Committee on National Statistics

The National Academies of SCIENCES · ENGINEERING · MEDICINE

Panel on Improving Cost-of-Living Indexes and Consumer Inflation Statistics in the Digital Age

Meeting #2
Wednesday, May 27, 2020

Sponsored by the Bureau of Labor Statistics

CNSTAT CPI Panel Roster
May 2020

Panel

Dan Sichel, Chair
Professor of Economics
Department of Economics, Wellesley College
dsichel@wellesley.edu

Ana Aizcorbe
Senior Research Economist
U.S. Bureau of Economic Analysis
ana.aizcorbe@bea.gov

Alberto Cavallo
Edgerley Family Associate Professor of Business Administration
Harvard Business School, Harvard University
alberto.cavallo@gmail.com

Jan De Haan
Senior Researcher
Statistics Netherlands
jandehaan1@gmail.com

Erwin Diewert
Professor of Economics
Affiliation: University of British Columbia and University of New South Wales
erwin.diewert@ubc.ca

Lisa Lynch
Provost and Maurice B. Hexter Professor
Brandeis University
lmlynch@brandeis.edu

Raven Saks Molloy
Assistant Director of the Division of Research and Statistics
Federal Reserve Board of Governors
raven.s.molloy@frb.gov

Brent Moulton
Senior Economist
Statistics Department, International Monetary Fund
bmoulton@imf.org

Marshall Reinsdorf
Senior Economist
International Monetary Fund
mreinsdorf@imf.org

Laura Rosner-Warburton
Senior Economist and Founding Partner
MacroPolicy Perspectives, LLC
lrosner@macropolicyperspectives.com

Louise Sheiner
Robert S. Kerr Senior Fellow; Policy Director
The Brookings Institution
lsheiner@brookings.edu

Project Staff

Chris Mackie
Study Director
National Academy of Sciences
cmackie@nas.edu

Anthony Mann
Program Associate
amann@nas.edu

Michael Siri
Associate Program Officer
msiri@nas.edu

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

Panel on Improving Cost-of-Living Indexes and Consumer Inflation Statistics in the Digital Age

Virtual Meeting with BLS – May 27, 2020

AGENDA

MEETING GOALS: BLS leadership will articulate their goals for the study and their motivation for commissioning it. Presentations by BLS experts will outline current CPI program activities and plans to reengineer price and expenditure input data used in the estimation of cost-of-living measures and consumer price indexes. Panel members and sponsor representatives will then revisit the project statement of task and revise as necessary. At the end of the meeting, the group will discuss the structure and content of an eventual workshop (public meeting) which will be the panel's main vehicle for gathering information from experts beyond the panel and BLS on various topics.

11:00 am Welcome, Introductions, Overview of the National Academies' Study Process
- **Brian Harris-Kojetin**, *Committee on National Statistics*

11:25 Study Charge, Meeting Goals
- **Dan Sichel**, *Panel Chair*

11:30 Opening Remarks from the Sponsor: Past CPI Studies; BLS's Vision and Strategy for the CPI; Goals for this Study
- **William Beach**, *Commissioner, Bureau of Labor Statistics*
- Open discussion

11:40 Improving Estimation of Elementary Indexes: Overview of high-level opportunities and challenges: CPI's two-stage aggregation and the alternative data initiative background and goals. Improving estimation of elementary item-area indexes broadly (beyond use of new data sources).
- **Rob Cage**, *DCPPI Assistant Commissioner*; **Anya Stockburger**, *Economist*
- Questions; Open discussion

12:00 pm *Break*

12:10 Difficult to Measure Categories: Overview: of BLS approach to measurement of goods and services—e.g., Health care/insurance, housing, high-tech—that raise unique measurement challenges.
- **Brett Matsumoto**, *Research Economist*
- Questions; Open discussion

12:30 CPI by Income and other Subpopulation Groups: BLS has done work over the past year on creating indexes for income quintiles using traditional methods while also researching methodological improvements.

- **Rob Cage**, *DCPPI Assistant Commissioner*
- Questions; Open discussion

12:50 Prioritization of Study Topics; Review and Revise (as necessary) the Project Statement of Task

Open Discussion: Content of Public Workshop; Prioritization of Information Gathering Tasks [time permitting; otherwise develop by email]

1:30 pm *Adjourn*

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

DIVISION OF BEHAVIORAL AND SOCIAL SCIENCES AND EDUCATION (DBASSE)

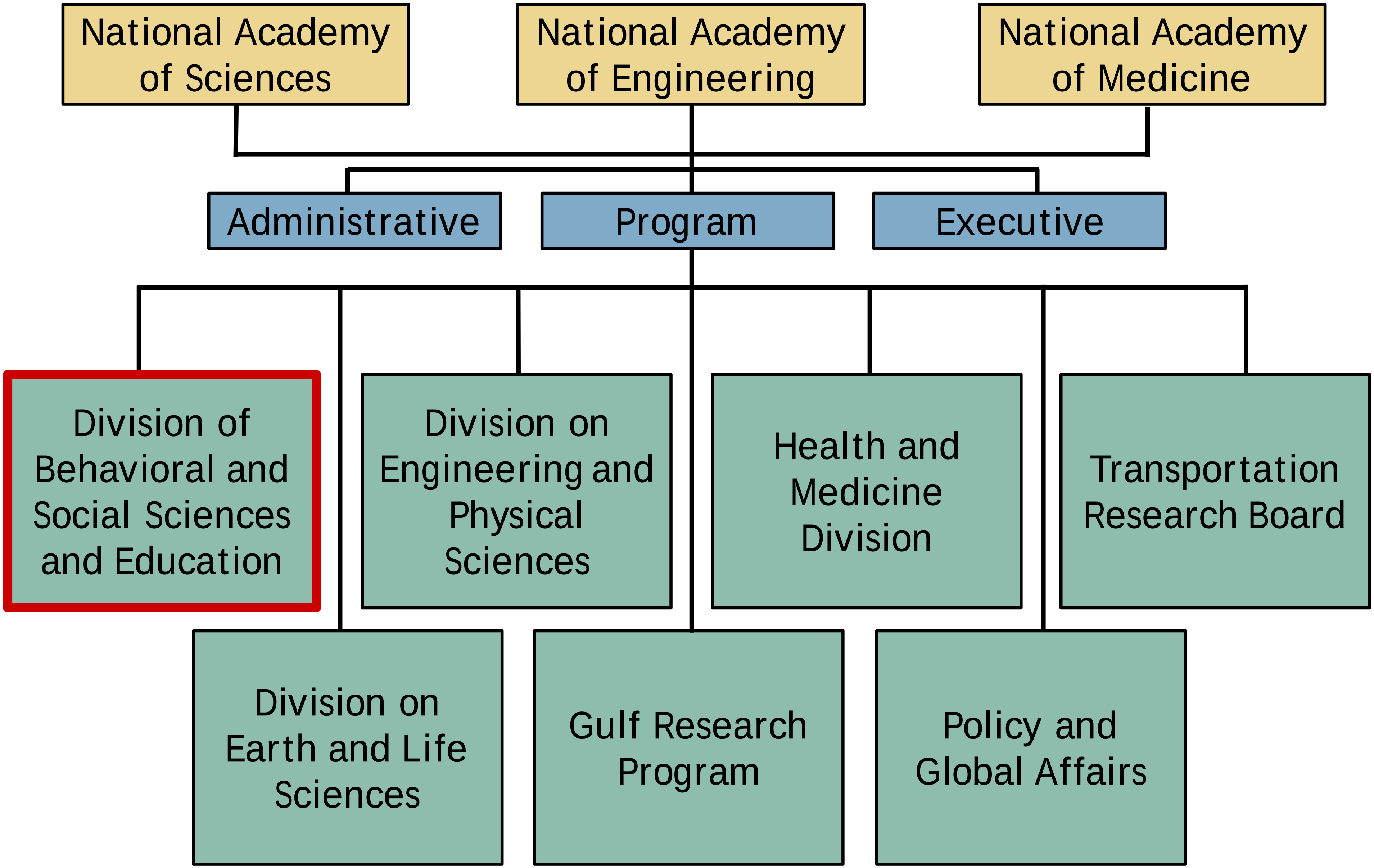
Panel on Improving Cost-of-Living Indexes and Consumer Inflation Statistics in the Digital Age

May 27, 2020

Brian Harris-Kojetin
Director, CNSTAT

U.S. National Academy of Sciences Charter (1863)

"…The Academy shall, whenever called upon by any department of the Government, investigate, examine, … and report upon any subject of science or art."


National Academy of Sciences
National Academy of Engineering
National Academy of Medicine
Administrative
Program
Executive
Division of Behavioral and Social Sciences and Education
Division on Engineering and Physical Sciences
Health and Medicine Division
Transportation Research Board
Division on Earth and Life Sciences
Gulf Research Program
Policy and Global Affairs

Division of Behavioral and Social Sciences and Education

(Work organized by Boards)

Board on Behavioral, Cognitive, & Sensory Sciences	Board on Children, Youth & Families	Committee on National Statistics	Board on Environmental Change & Society
Board on Science Education	Committee on Law & Justice	Committee on Population	Board on Human-Systems Integration

National Academies' Capabilities

Provide Consensus Advice	**Convene Experts**
Manage Original Research and Analysis	**Promote the Scientific Enterprise**

Consensus Studies

- ✓ Provide scientific evidence and advice to help policy makers and others make informed decisions
- ✓ Unbiased, authoritative advice
- ✓ All advice given in writing
- ✓ Provide neutral venue for open dialogue and discussion
- ✓ National Academies “Gold Standard”

Committee Membership

- ✓ Diverse, balanced range of interdisciplinary expertise related to the statement of task
- ✓ Potential nominees widely vetted
- ✓ Appointments made by President of National Academy of Sciences
- ✓ Before membership is finalized:
  - ✓ Formal public comment period
  - ✓ Hold discussion of balance, conflict of interest, and bias
- ✓ Volunteers serve without compensation and as individuals (not representatives of organizations or interest groups)

Consensus Study Process

Committee Process

- ✓ Statement of task: determines scope, focus and expertise
- ✓ Confidential deliberations are key
  - ✓ shield members from political or policy pressures
  - ✓ allows modulation of diverse viewpoints and types of expertise in order to achieve new insights and interpretations of evidence
- ✓ Draft report remains confidential throughout

Committee Meetings

Open and Closed Sessions

Closed Sessions:
Committee & Academies staff only

Open Sessions:
- Posted on website
- Open to public, invited visitors, sponsor
- Information-gathering

Consensus

Findings, Conclusions, & Recommendations

Findings

Observations based on facts and research evidence

Conclusions

Committee judgments about the findings taken together

Recommendations

Call for action or change

Based on evidence reviewed in report

What, if anything, should be done?

Who, what and how?

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

Report Review

- ✓ Independent, confidential review of draft report by outside experts
- ✓ Focus of review: does evidence support conclusions and recommendations?
- ✓ Report revised by committee in response
- ✓ Institution approves the report after revisions are complete

Report Release and Dissemination

- ✓ Sponsor briefed on final report just ahead of public release
- ✓ Report appears on National Academies Press website as free downloadable pdf
- ✓ Communications activities can include press release, webinar, public briefings, report highlights, social media, other tools

Special Thanks

- ✓ Our Sponsor
- ✓ Our Volunteers
- ✓ Our Staff

STATEMENT OF TASK

An ad hoc panel of the National Academies of Sciences, Engineering, and Medicine will review measurement issues in the Consumer Price Index program, which is overseen by the Bureau of Labor Statistics (BLS), and provide guidance for its continued modernization. The study will examine the potential to improve CPI methodology by incrementally transitioning from traditional survey-based data collection to an approach that blends multiple (survey and non-survey, government and commercial) data sources. The panel will consider opportunities to apply new data sources to improve the construction of specific elementary item-area indexes as well as to improve index aggregation along several dimensions.

Many data sources have emerged during the past 20 years (since the last CNSTAT review of the CPI). These sources could be used in the construction of the 7,000+ elementary item-area indexes in a way that improves the accuracy, timeliness, and detail of resulting price statistics, or reduces costs in the CPI program. The panel will identify specific areas where new kinds of data may be harnessed in a relatively straightforward way to improve price measurement of some goods such as food and electronics. The panel will also propose solutions for historically difficult-to-measure services such as health insurance and owner-occupied housing.

The panel will consider opportunities to use new data sources to improve aggregation of the elementary item-area indexes and also to mitigate upper-level substitution bias in the CPI-U and the CPI-W—for example, by taking advantage of the simultaneous availability of quantity and price information to update baskets and weights with shorter lags. As part of this task, the panel may revisit concerns about data sources used to estimate population item expenditure weights. The panel also will assess the prospects for creating new index aggregates that would present information about prices paid and expenditure weights for goods and services by households across the income distribution (by decile, or perhaps, by quintile). Work by BLS on this front would feed naturally into efforts by Bureau of Economic Analysis to provide statistics on how the nation's personal income is distributed across household income deciles.

Finally, the panel will offer forward-looking thoughts about what price measurement may look like in 20 years and what BLS can do to anticipate future research and policy needs. As part of its information-gathering activities, the panel will gather input from data users, stakeholders, and survey experts. The panel will produce a consensus report with conclusions and recommendations.

Improving Estimation of Elementary CPI Indexes

Rob Cage
Assistant Commissioner for CPI
Bureau of Labor Statistics

Anya Stockburger
Chief, Branch of Revision Methodology
Consumer Price Index Division

prepared for
National Academies of Sciences, Engineering, and Medicine
Committee on National Statistics
Virtual Meeting with BLS
May 27, 2020

CPI Measurement Objective

- Approximate a conditional cost-of-living index (COLI)
- Minimum expenditure in current period to achieve same level of satisfaction (utility) as base period
- The amount of expenditure change required for the consumer to be indifferent between current and past prices

Two-stage aggregation: consumption dimension

- Infinite set of eligible consumer goods and services bundled into 243 component groups (elementary item strata)

Housing

Food

Clothing

Transportation

Medical care

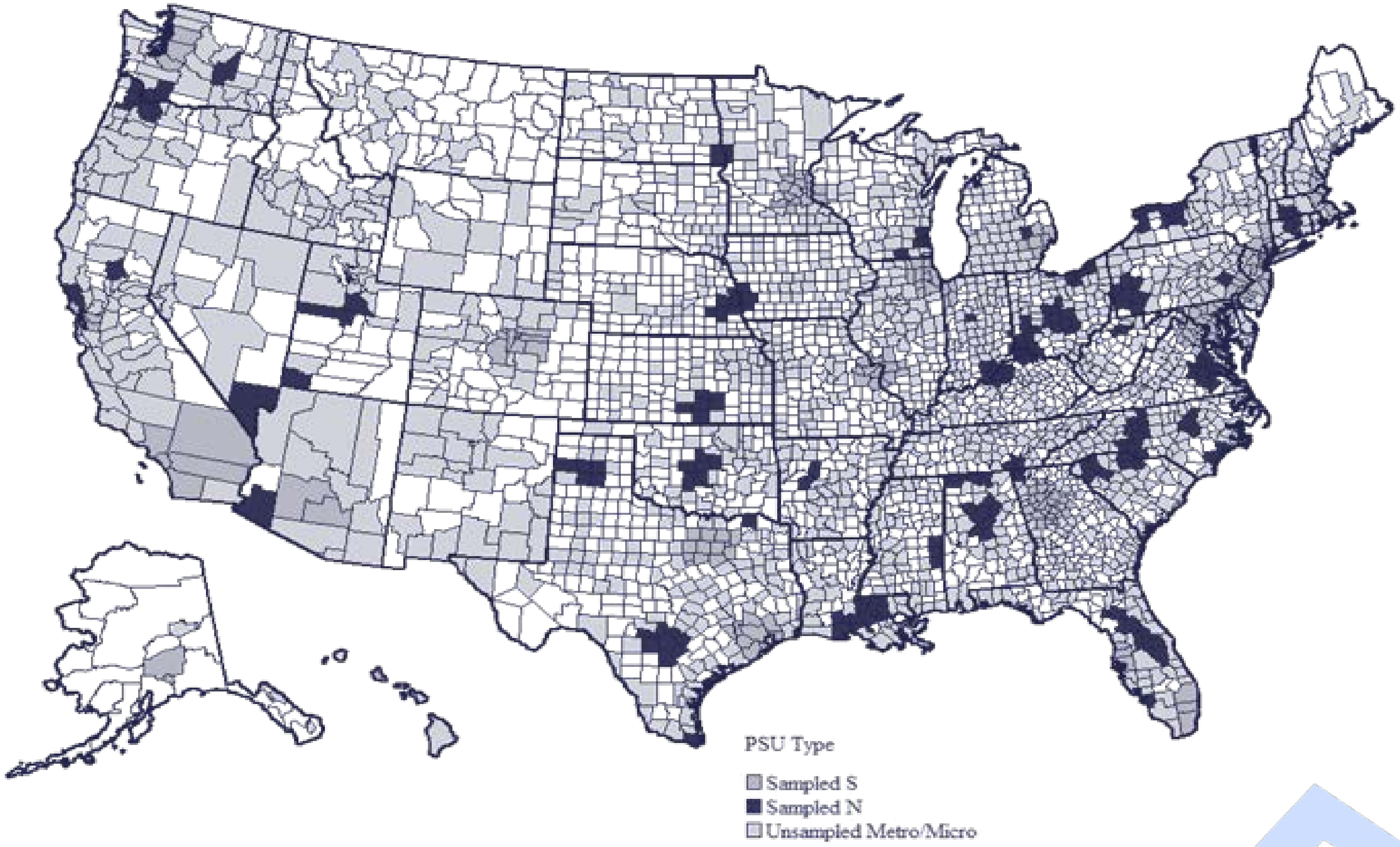
Education

Recreation

Other

Two-stage aggregation: geographic dimension

- Set of 929 core-based-statistical-areas bundled into 32 component groups (elementary areas)

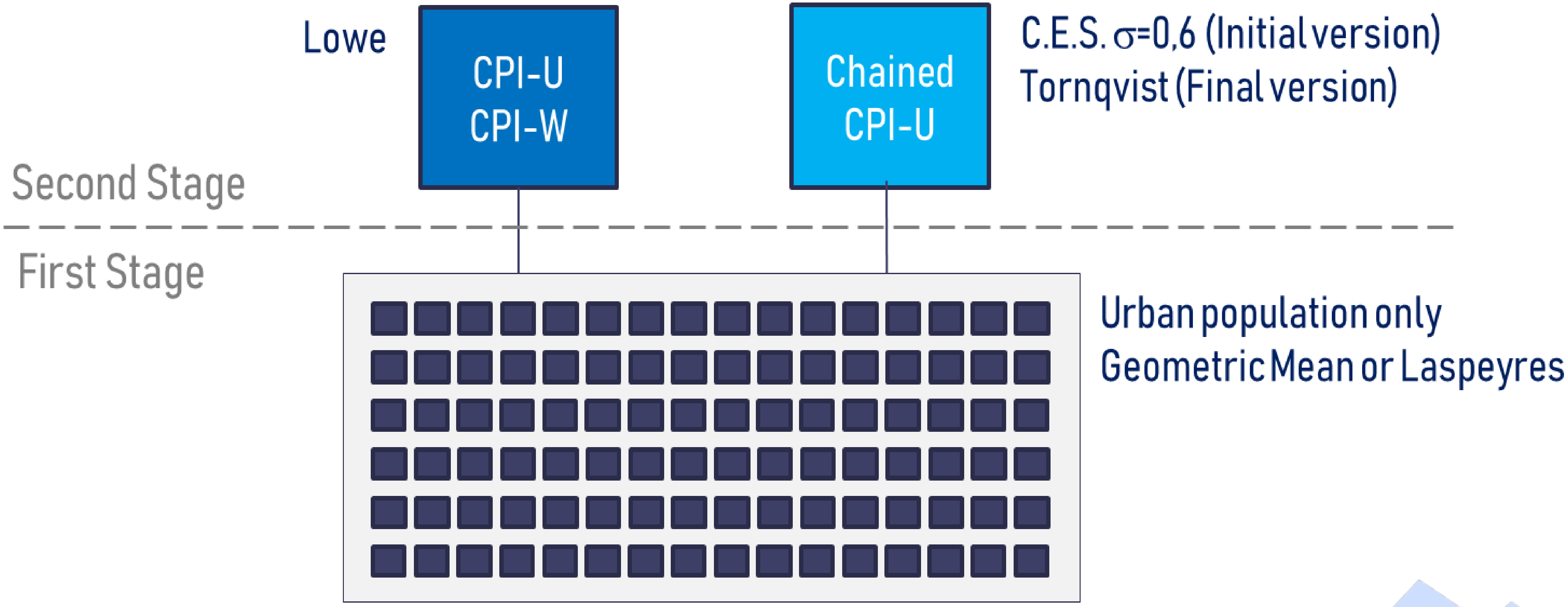


Two-stage aggregation: elementary indexes

- 243 items x 32 areas = **7,776** elementary indexes = basic building blocks

- Traditional survey collection:
  - Prices\rents:
    - In store via CAPI (tablet app), telephone, or manual keying of internet prices
    - Sample size: ~15 (goods\services); ~200 (rental units)
  - Weights:
    - In person recall interview or paper\online diary
    - Sample size: ~200 interviews, ~30 diaries
  - Estimation:
    - Fixed samples rotated every 4 years (goods\services) or 6 years (rental units)
    - Match-model pricing
    - Geometric Mean (90%) or Laspeyres (10%) aggregation

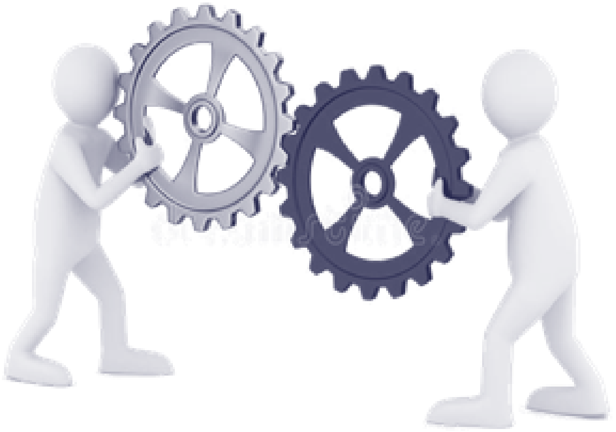
Two-stage aggregation: diagram



Continuous Improvement

- Maintain and improve accuracy, relevance, timeliness, transparency
- Over 75 improvements made since *1996 Boskin Commission* report

  - Relevance: more products
  - Timeliness: release sooner
  - Transparency: improve website
  - Accuracy: reduce measurement error



    - Variance: mitigate sources of sampling and nonsampling error
    - Bias: mitigate suspected biases

Elementary Indexes: improving accuracy

CHALLENGES

- Perennial issues
  - New goods bias
  - New outlets bias
  - Quality adjustment bias

IMPROVEMENTS

- Updating geographic structure
- Updating item classification
- Data review efficiencies
- Enhanced web-based collection forms
- Non-traditional data sources and collection modes

Opportunity: non-traditional data sources and collection modes

- **Goals**: improve accuracy of elementary indexes; improve efficiency of data collection
- **Strategic Objective:** convert a significant proportion of market basket from traditional collection to non-traditional sources and collection modes by 2024
- **Current Status:** 3.4% of market basket

Category	Data Source	Implementation Notes
Apparel, household goods	Corporate data	Implemented March 2019
Prescription drugs	Corporate data	Implemented May 2016
Postage	Publically available data	Oldest use of "alternative" data
Used cars	Purchased data	Another long-time alt data source

- **Other alternatives under investigation:** Respondent self-reporting, API, web-scraping, additional corporate sources or data acquisition via third party vendors

Alternative Data Initiative: data phases

Beginning Stages

Examples
- Food Away from Home
- Lodging Away from Home
- Apparel

Seek permission from corporations to obtain scanner data or scrape company websites

Work with corporation to **negotiate data** requirements (granularity, delivery mechanism, geographic coverage, etc.)

Collect data to build a time series for research to begin

Research

Examples
- Medical services
- Telecomm services
- Airfares

Use collected data to conduct **quality assessment** and fitness for use

Develop methodology to transform corporate data into measures of price change that can be blended with traditionally sampled and collected data. Methodology could be simple (data collection replacement) or complex (leveraging power of big data).

Attend research conferences, draft papers for publication, and publish experimental methods where needed to appropriately **vet methodology**

Implementation

Once the methodology is approved, **implement into production**. This includes modifying IT systems, finalizing contingency plans, and notifying the public.

Example
- Motor fuels, New Vehicles, Airline

Elementary Indexes: questions for the CNSTAT panel

CHALLENGES

- Perennial issues
  - New goods bias
  - New outlets bias
  - Quality adjustment
- Aggregation
  - Transaction data methods
  - Substitution bias
  - Chain-link bias
- Blending traditional data with non-traditional data
  - Fitting non-sampled data into traditional sampling framework

QUESTION & ANSWER

Contact Information

Rob Cage
Assistant Commissioner for CPI
202-441-2576
cage.rob@bls.gov

Anya Stockburger
Chief, Branch of Revision Methodology
202-691-5370
stockburger.anya@bls.gov

Difficult to Measure Items in the CPI

Brett Matsumoto
Research Economist
Division of Price and Index Number Research
CNSTAT Presentation
5/27/2020

Price measurement for hard-to-measure categories

Statement of Work:

As BLS has begun working with alternative data and/or collection methods, they are paying particular attention to various aspects of three hard-to-measure categories. What is the best method for estimation price change for health insurance, owner occupied housing, and durables?

Health Insurance

Statement of Work:

- *1. Should BLS change from indirectly pricing health insurance to directly pricing health insurance policies?*
- *2. And if so, is there a suggested approach to dealing with quality changes over time in the policies?*

Health Insurance

- HI premiums pay for medical goods and services plus services of the insurance company (processing claims, etc.).
- Indirect pricing: price insurance services using retained earnings method.
- Direct pricing: Premiums and out of pocket payments
- Implications for pricing medical goods and services.

Retained Earnings Method

- Currently the HI index is calculated as:

$$I_{t,t-1} = \left[\frac{\frac{RE_y}{B_y}}{\frac{RE_{y-1}}{B_{y-1}}}\right]^{1/12} * \left[\frac{\sum_i w^i MCPI_{t,b}^i}{\sum_i w^i MCPI_{t-1,b}^i}\right]$$

- The first term is brackets is the relative for the ratio of retained earnings (RE) to benefits (B). Collected annually, so 12^{th} root taken.
- The second bracket measures the aggregate change in the medical CPI. Each component weighted by the share of insurance payments to that service type.
- Base period b, y indexes year, t indexes month

Limitations

- Data
  - Only highly aggregated data
  - Data lagged over a year
- Quality Adjustment
  - Price of insurance services is retained earnings per real benefit dollar => any costly quality change will show up as a pure price change

Durables

Statement of Work:

Should the BLS adopt a flow-of-service approach to all durable goods, similar to the current methodology used for owner occupied housing?

Durables

- Durability means that a capital good is productive for two or more time periods, and this in turn, implies that a distinction must be made between the value of using or rental capital in any year and the value of owning the capital asset (CPI manual, ILO 2004)
- Methods for dealing with durability
  - Acquisitions approach
  - Payments (out-of-pocket)
  - Rental equivalence or leasing equivalence approach
    - Respondent provided (owner reports)
    - Renter-owner matching of characteristics
    - Hedonic approach to estimate market rents or leases (renter or leasee reports)
  - User cost approach (2 parts: costs of using capital + capital gains)
    - Capitalization rate
    - Full user costs

Treatment of Durables in the CPI

Category	Approach	Price
Owner Occupied Housing	Rental Equivalence	Rental prices of nearby units
New Vehicles	Acquisition Approach	Total transaction price
Used Vehicles	Acquisition Approach	Change in depreciated price
Other Durables (not included in housing)	Acquisition Approach	Total transaction price

Durable Issues

- Appropriate Method
- Consistency across the durable categories in the CPI
- Data needs and limitations of current data collection. E.g.,
  - estimate of depreciation rate, loan information, etc.
  - lack of rental equivalence data

Owner Occupied Housing

Statement of Work:

Under the rental equivalence approach to owner occupied housing, what is the best method to estimate monthly price change?

Currently, the CPI uses price change observed for rental housing units as a proxy for the hypothetical rental equivalent change associated with owner-occupied housing units.

Should the BLS directly measure the rental value of owner-occupied homes, and if so, what is the best approach?

Housing Data Issues for Rental Equivalence Approach

- Over-weighting of apartments in owners equivalent of rent.
- Use rents for all units or new leases only
- Lagged data due to infrequent repricing

Recommendations Needed

Recommendations for pricing health insurance

Durables – recommendations on how to implement flow of services approach.

Housing – recommendations on how to improve rental equivalence method.

Contact Information

Brett Matsumoto
Research Economist
Division of Price and Index Number Research
www.bls.gov/xxx
202-691-XXXX
matsumoto.brett@bls.gov

CPI by Income and Other Subpopulation Groups

Rob Cage
Assistant Commissioner for CPI
Bureau of Labor Statistics

prepared for
National Academies of Sciences, Engineering, and Medicine
Committee on National Statistics
Virtual Meeting with BLS
May 27, 2020

CPI Index Products

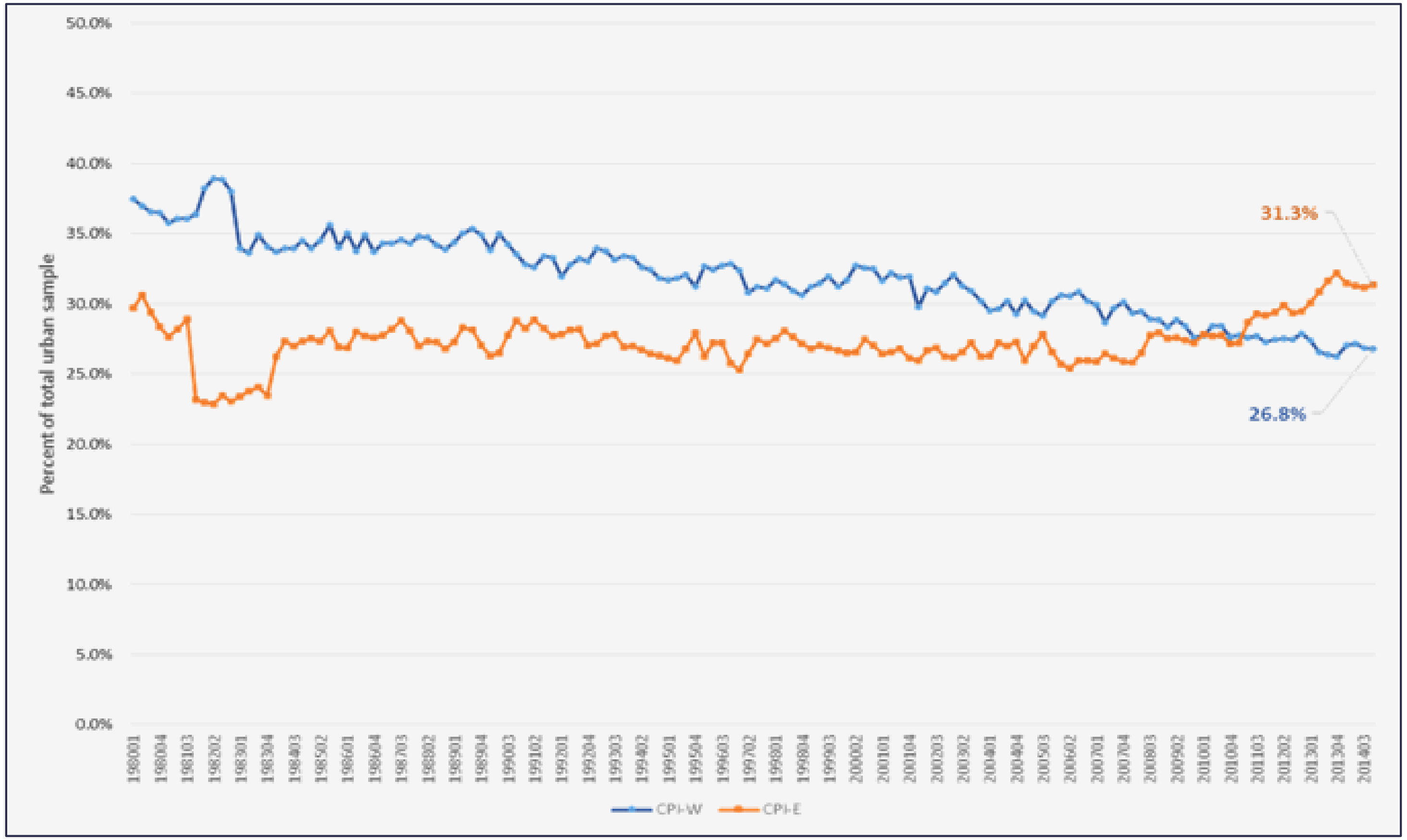
INDEX	COHORT	YEAR INTRODUCED	NOTE	WEBSITE HITS
OFFICIAL:				
CPI-W	Wage-earners and clerical workers	1921	Used for Social Security Cost of Living Adjustment (COLA)	6.0%
CPI-U	All urban consumers	1978	Headline index	93.1%
Chained CPI-U	All urban consumers	2002	Published with ~ 1-year lag; used for federal tax bracket adjustments	0.9%
RESEARCH:				
CPI-E	Elderly consumers	1988	Alternative to CPI-W for Social Security COLAs	0.0%
CPI-U-RS	All urban consumers	1999	Used by Census in calculation of real median income statistics	0.0%

Subpopulation group index

CPI-W versus CPI-E

- Both use set of n=7,776 CPI-U elementary indexes as input
  - No control for price change heterogeneity in outlet patronage nor specific goods and services purchased
- Both are reweighted Lowe aggregations using cohort elementary spending weights derived from household spending survey
  - Higher sampling error than CPI-U
- Chained Tornqvist versions not produced

Percent of urban sample, Consumer Expenditure Interview Survey

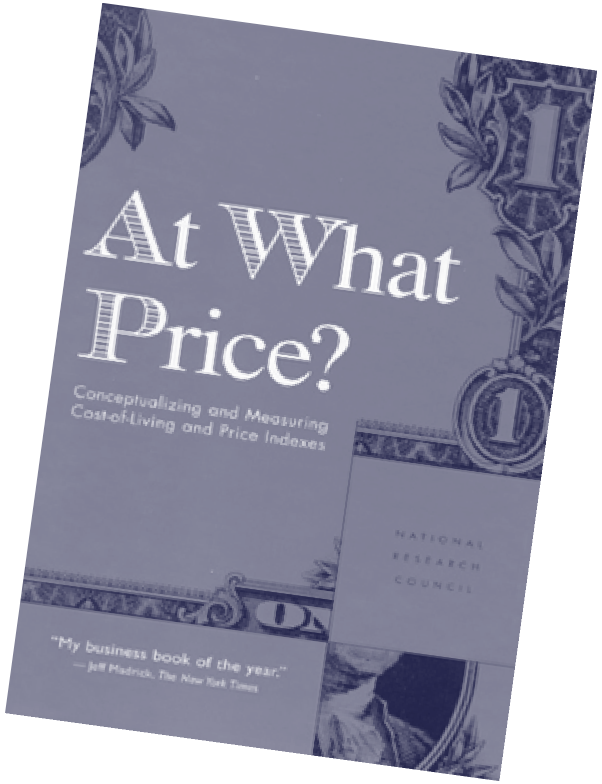


OMB Interagency Technical Working Group: Official Poverty Measure

- 13 members from eight different federal agencies
- Chartered February 2019
- Federal Register Notice published May 2019
- Primary objective:
  - Recommend index best suited for use in calculation of the U.S. Official Poverty Measure (OPM)
  - CPI-U has been in use since OMB Statistical Policy No. 14 was published in 1978
  - Evaluation of Chained CPI-U, among other existing index products
  - Public comment: poor experience inflation differently than total population at large

Improvement: Population subgroup index methodology

- **Date:** research started in 2018; implementation TBD
- **Goals:** improve accuracy of CPI-W; improve relevance by publishing new products, e.g. CPI for low-income households
- **Motivation:**
  - Declining sample size of CPI-W; contradictory classification of CPI-W relative to CPI-E
  - Potential user demand for CPI low-income product
  - 2002 CNSTAT *At What Price?* recommendations
    - Collect prices in a way that allows them to be associated with household characteristics (Recommendation 8.1)
    - 2019: point of purchase questions added to Consumer Expenditure Surveys (replacing TPOPS)



Phase 1: Research

- Focus on improvements that can be made given current funding level and data sources
  - Cohort definition
    - CPI-W urban wage-earners and clerical workers
    - Low-income
  - Treatment of owner-occupied housing
    - Payment approach instead of rental equivalence
  - Aggregation across households in cohort
    - Democratic weighting
    - Plutocratic weighting
  - Aggregation across items in basket
    - Examine feasibility of Chained Tornqvist

Phase 2: Research and improvements

- Seek budget initiative to elevate subpopulation index products to ‘gold standard’
  - Increase Consumer Expenditure Survey sample
    - Stratify household sample into 2x2 cross tabulation of income group and wage-earner status
    - Publish Tornqvist subpopulation indexes (Chained CPI-W and Chained CPI low-income)
  - Increase Commodities and Services Pricing sample
    - Capitalize on merger of TPOPS into CE as outlet frame source
    - Independent outlet selection for each population (U,W, low-income)
    - Potential for independent unique item selection
    - Begin to address elementary index heterogeneity issues

	SOCIAL SECURITY COHORT	
INCOME	A First quartile, SS household	B First quartile, non SS household
	C Upper 3 quartiles SS household	D Upper 3 quartiles non SS household

CPI-W = A + C
CPI Low-income = A + B
CPI-U = A + B + C + D

Contact Information

Rob Cage
Assistant Commissioner for CPI
202-441-2576
cage.rob@bls.gov

Population Subgroup Price Indexes:

Evidence of Heterogeneity or Measurement Error?

by

Robert Cage, Joshua Klick, William Johnson

Office of Prices and Living Conditions

Bureau of Labor Statistics | U.S. Department of Labor

Traditionally, national consumer price index (CPI) statistics are designed and engineered to represent a single population target: the noninstitutionalized resident civilian population of the nation. To that end, aggregation weights typically represent the experience of an average consumer or representative household. In reality, household budget shares can vary widely depending on the socio-economic characteristics of the household members, and the inflation experience of a given consumer cohort could vary far from the average. This results in some CPI users criticizing the official statistic as flawed, especially if the user's perceived personal experience with inflation differs from the published number. To address this issue, statistical agencies can produce cohort indexes tailored toward the specific indexation needs of various users or groups. However, this approach is encumbered by methodological challenges and practical limitations including available funding.

Currently, the U.S. Bureau of Labor Statistics (BLS) produces official consumer price indexes for two population cohorts: all urban consumers (CPI-U), and urban wage-earners and clerical workers (CPI-W). In addition, an experimental index is produced for a third cohort: urban elderly consumers (CPI-E). A storied legislative history, examined in Section 1 below, explains how the three indexes came about. Of note, the CPI-E was created to potentially address shortcomings of the CPI-U and the CPI-W for one specific indexation need: the annual cost-of-living adjustment (COLA) made to Social Security benefits by the Social Security Administration (SSA). Debate among policymakers and analysts over which index would be most appropriate for the COLA started in the late 1970s and amplified when a fourth index was added to the fray in 2002. That year, BLS began publishing a second index for the all-urban consumer cohort, the Chained CPI-U, that some argue is the most appropriate index for the COLA.

This paper highlights the practical issues and methodological problems BLS faces in estimating indexes for population subgroups, using the Social Security COLA indexation need as a focus. Public arguments for and against the CPI-E are summarized in Section 2. Methodological issues impacting the estimation of cohort indexes are summarized in Section 3. A model-based approach to constructing cohort indexes, within the confines of producing a national index and constrained by current survey data limitations, is

Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics.

proposed and outlined in Section 4 as a potential alternative to the current practice of varying expenditure weights only. Results are summarized in Section 5.

1. Historical Perspective

A review of seminal moments in BLS's history reveals why it currently produces CPIs for three different cohorts. The origins of the modern day U.S. CPI date back to World War I when rapid increases in prices created a demand for the statistic for use in employment contract negotiations.[1] To provide weighting patterns for the index, surveys of family expenditures were conducted in 1917 to 1919 exclusively in major industrial centers. The initial survey coverage was limited to married couples with at least one child, where the principal wage-earner in the family accounted for 75% or more of total family income. Periodic collection of prices started in 1919 and regular publication of a national index, commenced in 1921 with indexes estimated back to 1913. One of the first uses of the index was in cost-of-living escalator clauses built into the employment contracts of manual laborers, particularly shipbuilders working in industrial centers to support the war effort.

Over the next 50 plus years, the BLS produced the national CPI by tracking the buying habits of this specific, narrowly defined consumer cohort: families of manual laborers. During this time, however, the definition of the cohort and the geographic coverage expanded in scope.[2] By 1964, the target population had evolved to include all wage-earner and clerical worker households living in metropolitan areas of the country. To be considered a *wage-earner* household, at least one member of the household had to be employed full-time (i.e., 37 weeks out of the year) in an eligible occupation. Eligible occupation groups accounted for roughly 70% of the total U.S. workforce at that time, but excluded professionals, managers, and the self-employed. BLS's definition of *wage-earner and clerical worker* for the CPI has essentially remained unchanged since 1964.

1.1 Creation of the CPI-U

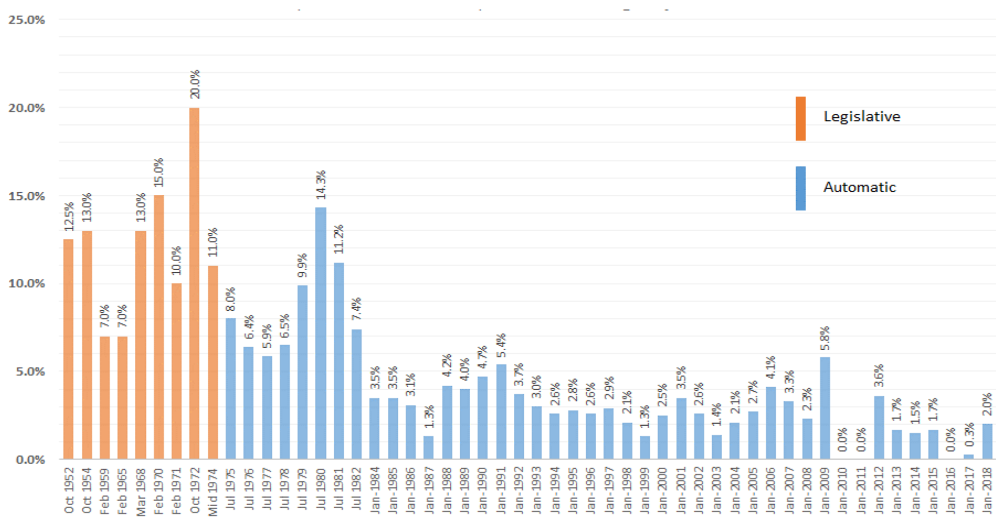
Then, in 1972, the passage of Public Law 92-336 changed the course of BLS's CPI history. This legislation, primarily aimed at increasing the national debt limit, contained an amendment to the Social Security Act of 1935. This 1972 amendment established an automatic annual cost-of-living adjustment (COLA) to Social Security benefits, to ensure the purchasing power of the otherwise fixed benefits would not be

[1] For a history of the first 100 years of the CPI, see Rippy, Darren. The first hundred years of the Consumer Price Index: a methodological and political history. *Monthly Labor Review*, Bureau of Labor Statistics. April 2014 at https://www.bls.gov/opub/mlr/2014/article/pdf/the-first-hundred-years-of-the-consumer-price-index.pdf.

[2] For a complete history of how the CPI target population evolved over time, see Exhibit 1 in Chapter 17 of the *BLS Handbook of Methods* at https://www.bls.gov/opub/hom/pdf/homch17.pdf.

eroded by the forces of inflation.[3] The 1972 statute mandated the Social Security Administration (SSA) to calculate the COLA using the CPI produced and published by the BLS. The law went into effect shortly thereafter, and the first automatic COLA was applied in 1975, an increase of 8%.[4] At that time, the only CPI published by the BLS was the index representing the *wage-earner and clerical worker* cohort. Hence, it was the index adopted by the SSA.

Figure 1. *History of Social Security cost-of-living-adjustments*



Note: The first COLA was applied in October 1950, an increase of 77%.

Meanwhile, the BLS was in the midst of a comprehensive overhaul and revision of the CPI program, partly in response to recommendations made by a committee commissioned by Congress to examine federal price indexes, commonly referred to as the Stigler Committee.[5] BLS planned to expand the population scope of the CPI to 'nearly all consumers' in 1978, engineering it to represent not just the existing urban wage-earner cohort, but all residents of metropolitan areas, including professionals, the self-employed, the poor, the unemployed, and retired persons.[6]

---

[3] Prior to 1975, cost-of-living adjustments to Social Security benefits required legislative action. Congress enacted public laws to increase benefits 10 times prior to the 1972 amendment taking effect: 1950, 1952, 1954, 1959, 1965, 1968, 1970, 1972, and 1974.

[4] For history of COLA increases see Whitaker, Julie. *Social Security: Cost-of-Living Adjustments*. Congressional Research Service, October 2017. https://fas.org/sgp/crs/misc/94-803.pdf

[5] See Stigler, George et. al. *The Price Statistics of the Federal Government. Report of the Committee before the Subcommittee on Economic Statistics of the Joint Economic Committee*, Congress of the United States. Sec 5(a) of Public Law 304, 79th Congress. January, 1961. Six recommendations included: (1) periodic weight updates at least once every 10 years; (2) probability sampling, (3) prompt introduction of new products, (4) funding for research, (5) extend population target to include single persons as well as families, and cover rural nonfarm as well as urban workers. Rural farm was excluded because a separate index produced by Department of Agriculture represented prices received and paid by farmers.

[6] Improvements made to the CPI in the 1978 Revision included: (1) updated weights using 1972-73 Consumer Expenditure Survey, (2) updated sample of items, (3) updated sample of outlets –introducing probability sampling, and (5) expanded population coverage. See The Consumer Price Index Revision-1978, Monthly labor Review. US Department of Labor, Bureau of Labor Statistics, 1978.

In 1974, BLS initially announced its intention to replace the urban wage-earner definition of the CPI with the broader all urban consumer definition, and publish just one index.[7] However, this plan was criticized by some members of Congress who feared the new index would, 'no longer be firmly grounded in the experience of low- and middle-income workers.'[8] As a result, Congress increased BLS's budget to allow both indexes to be calculated and published for at least three years. BLS introduced the all-urban index (CPI-U) in 1978, and relabeled the existing index as the CPI-W. From 1978 to 1980, separate outlet and item samples were selected to track price change for the separate CPI-U and CPI-W indexes. In 1981, Reagan era budget cuts to the BLS put the CPI-W on the chopping block once again. BLS could no longer afford to maintain two distinct outlet and item samples for the two indexes and could have terminated calculation of the CPI-W. However, because little difference between the CPI-U and CPI-W was observed during the 1978-1980 period, BLS decided to continue the CPI-W series using the CPI-U sample of areas, outlets, items, and prices. Hence, in 1981 the CPI-W became a reweighted version of the CPI-U, with the component weights used to calculate the aggregate index adjusted to reflect the spending habits of the urban wage-earner cohort. Since 1981, the CPI-U and CPI-W have differed only in the expenditure weights assigned to item categories and geographic areas.[9]

Had the BLS decided to terminate the wage-earner index in 1978, or 1981, the SSA would have had no option but to convert the calculation of the Social Security COLA to the lone index produced by the BLS: the CPI-U. Absent a legislative mandate identifying the specific index to use, the introduction of the CPI-U coupled with the continuation of the CPI-W presented the SSA with a rulemaking conundrum: which index was more appropriate for the stated purpose of the COLA?

In the late 1970s, Congress created a National Commission on Social Security to undertake 'fundamental, long-term, comprehensive consideration for changes to the Social Security system.'[10] The law establishing this commission expressly directed it to investigate the COLA calculation issue and which index would be most appropriate for said use. Notably, the directive called for an assessment of whether a special index for the *elderly* was most suitable. The report of the committee, issued in March 1981, recommended against the use of a CPI tailored toward any beneficiary group, such as the retired

---

[7] See Reed, Stephen and Kenneth Stewart. *Why does BLS provide both the CPI-W and the CPI-U?* Beyond the Numbers. BLS. February 2014. https://www.bls.gov/opub/btn/volume-3/why-does-bls-provide-both-the-cpi-w-and-cpi-u.htm.

[8] Joseph P. Goldberg and William T. Moyer, *First Hundred Years of the Bureau of Labor Statistics*, Bulletin 2235 (U.S. Bureau of Labor Statistics, September 1985).

[9] Between 1983 and 1985, the indexes also differed in the treatment of the shelter component of the index. A rental equivalence approach was adopted in the CPI-U in 1983, and in 1985 for the CPI-W. See https://www.bls.gov/opub/btn/volume-2/pdf/owners-equivalent-rent-and-the-consumer-price-index-30-years-and-counting.pdf for more information on why this change was made.

[10] See https://www.ssa.gov/history/reports/80commission.html.

elderly, primarily due to a lack of analytical data and research on how to properly construct such an index. The committee did, however, recommend use of the CPI-U over the CPI-W for the Social Security COLA, primarily because the CPI-U included retired persons in its population target while the CPI-W did not.

Neither Congress nor the SSA acted on the recommendation. Inaction was likely due, in part, to the financial crisis enveloping the Social Security program at the time. By the late 1970s, the Old-Age and Survivors Insurance Trust Fund (OASDI) was projected to become insolvent as early as 1983. A second National Commission on Social Security Reform was created by Congress in 1981, commonly referred to the Greenspan Commission, to solve the financial problems facing Social Security.[11] The Greenspan Commission report, issued in 1983, *inter alia* recommended against the use of the CPI for automatic benefit increases, preferring instead to base them on changes in average wages, when the latter was lower.[12]

So within the span of two years, two separate national commissions had issued two conflicting opinions on the appropriate indexation method for Social Security COLAs. One, in effect, favored an index representing the inflation experience of a retired cohort, and the other favored an index representing the experience of workers.

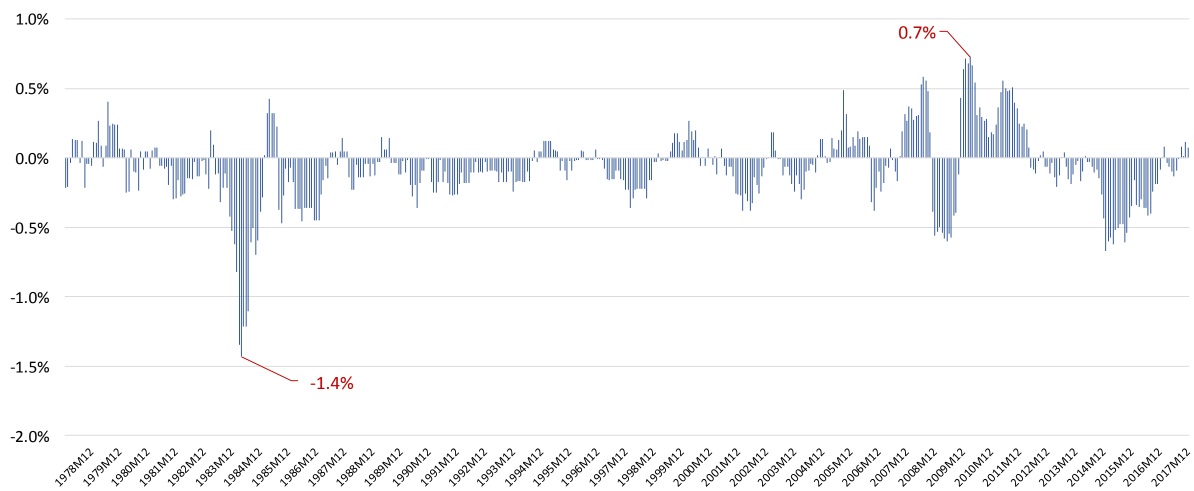
Complicating matters was the double-digit levels of inflation experienced by the American economy at the beginning of the decade. In early 1983, the CPI-U was yielding higher estimates of inflation compared to the CPI-W, a gap which would turn out to be the largest difference in the history of the two series to date (see Figure 2). Perhaps fearing a switch from the CPI-W to the CPI-U would hasten the woes of the OASDI, both Congress and SSA remained idle. Automatic Social Security COLAs have been tied to the CPI-W ever since, while many other COLAs cited in federal legislation are tied to the CPI-U or Chained CPI-U.[13] Public debate over the appropriate index for COLA indexation has not only endured, but has arguably amplified over the past 30 years.

[11] See https://www.ssa.gov/history/reports/gspan.html. Alan Greenspan was the chair of National Commission on Social Security Reform.

[12] See Greenspan Commission report, recommendation 15.

[13] In 2017, legislation was enacted to index federal income tax brackets to changes in the Chained CPI-U. See Public Law 115-97, The Tax Cuts and Jobs Act of 2017.

Figure 2. *Difference in annual inflation estimates, CPI-W minus CPI-U*



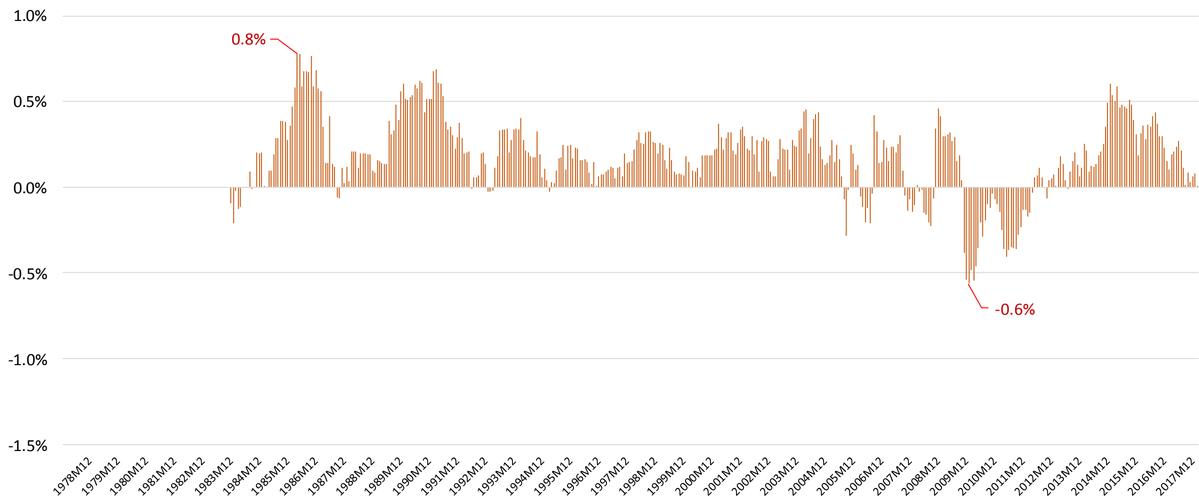
1.2 Creation of the CPI-E

With over 67 million beneficiaries, Social Security is by far the largest U.S. government outlay directly adjusted using a CPI. Hence, SSA's use of the CPI-W instead of the CPI-U has remained a topic of discussion since the initial debates of the early 1980s. The primary criticism of the CPI-W is its cohort definition: the index is calculated using expenditure weights of wage-earners. This population is employed and the spending habits are reflective of people who work. In contrast, most Social Security recipients are retired and not in the labor force.

Suitability of the CPI-W came into Congress's crosshairs again in 1987, when it directed the BLS to develop an index reflecting the spending patterns of seniors, so that it could be evaluated by SSA as a possible replacement for the CPI-W for Social Security COLAs. However, this Congressional mandate came without any additional funding. In turn, BLS released the experimental CPI for Americans 62 Years of Age and Older (CPI-E) in 1988, with estimates calculated back to 1982. The index was labeled experimental because, like the CPI-W, it differed from the CPI-U only in the expenditure weights used to aggregate component indexes together. The index was not constructed by selecting a dedicated sample of outlets and items based on the spending choices of the elderly, and it did not track the prices paid exclusively by the cohort, which would not have been possible without additional funding. For these reasons, and because the sample size used to estimate the weights for the elderly cohort were smaller than the wage-earner cohort, the BLS deemed the CPI-E less accurate than the CPI-W, reinforcing its experimental moniker.

BLS continues to calculate the experimental CPI-E and release it upon request. Historically, the index has yielded annual inflation estimates about 0.2 percent higher than the CPI-U, and 0.3 percent higher than the CPI-W (see Figure 3). Had the SSA used the CPI-E instead of the CPI-W for the most recent COLA, the average recipient would have received an additional $1 in average monthly benefits in 2018. This seemingly *de minimis* difference would equate to an additional $1.2 billion in outlays from the OASDI trust fund. Over the years, bills have been periodically proposed in Congress to mandate the use of the CPI-E or a similar index product targeting the elderly cohort for Social Security COLA indexation. In 2017 alone, as many as five different bills were sponsored but to date none have passed into law.[14]

Figure 3. *Difference in annual inflation estimates, CPI-E minus CPI-U*



1.3 Creation of the Chained CPI-U

In the early 1990s, concerns over potential biases in the CPI and subsequent impacts on the Social Security COLAs and financial solvency of the OASDI Trust Fund, led Congress to once again appoint a commission to study the CPI. The Advisory Commission to Study the Consumer Price Index, commonly referred to as the Boskin Commission, issued its report in 1996 which concluded the CPI overstated inflation by about 1.1 percentage points per year.[15] The Boskin Commission summarily identified four specific sources of upward bias in the CPI aggregating to a plausible range of 0.80 to 1.60 percent per annum:

[14] See http://www.ncpssm.org/PublicPolicy/SocialSecurity/Documents/ArticleID/1159/The-CPI-E-%E2%80%93-A-Better-Option-for-Calculating-Social-Security-COLAs.

[15] See Boskin, Michael, et al., *Toward A More Accurate measure of the Cost of Living*, December 4, 1996. Michael Boskin was the chair of the Advisory Commission.

Figure 4. *Estimated biases in the CPI, Boskin Commission, 1996*

Upper Level Substitution Bias	0.15
Lower level Substitution Bias	0.25
New Products/Quality Change	0.60
New Outlets	0.10

Among the CPI market basket categories, the Commission concluded the quality change bias was most problematic for medical care, estimating an upward bias in prescription drug, hospital service, and physician service indexes as high as 3 percent per annum.[16]

Meanwhile, BLS was already considering measures to improve the handling of some of these biases. To mitigate lower level substitution bias, BLS adopted a hybrid approach to estimating lower-level or component indexes in 1999, whereby the Geometric Mean index formula was adopted for the majority of market basket strata.[17] This index assumes consumers respond to relative price change by holding the budget shares of the specific items purchased constant over time (i.e., a unitary elasticity of substitution). The Laspeyres index, which assumes consumers maximize utility by holding the quantity mix of specific items purchased constant over time (i.e., zero elasticity of substitution) was retained for a few categories.[18] The CPI-U, CPI-W, and CPI-E each use this hybrid combination of Geometric Mean and Laspeyres indexes at the component level.

To address the upper level substitution bias, BLS began publication of the Chained CPI-U (C-CPI-U) in August 2002 with data going back to 2000. The C-CPI_U utilizes a Tornqvist index to aggregate component indexes. This index uses observed monthly expenditure changes made by consumers, ostensibly capturing marketplace substitutions brought about by changing relative prices.[19] The formula was not – and has not been -- adopted for the flagship CPI-U index, because the requisite monthly expenditure data needed for index calculation become available with roughly a one year lag. Failing to meet BLS's timeliness standard of a 2-week lag for the headline national inflation index, BLS decided to

[16] See Boskin Commission, Table 2, pp. 39-40.

[17] See Dalton, Kenneth, John Greenlees and Kenneth Stewart. Incorporating a geometric mean formula into the CPI. *Monthly Labor Review*, October 1998. https://www.bls.gov/mlr/1998/10/art1full.pdf.

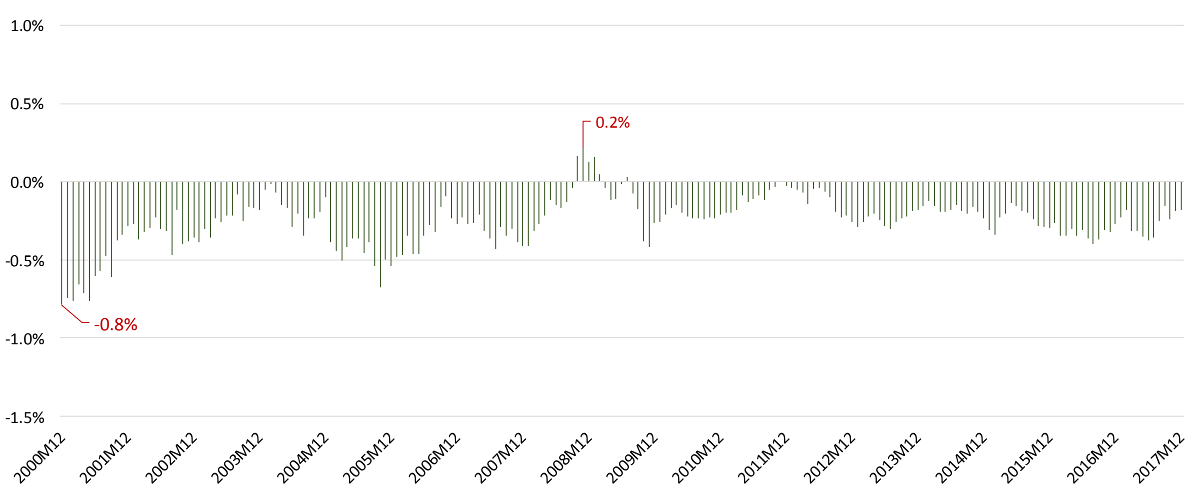
[18] Elementary items using the Laspeyres formula include rent, owner's equivalent rent, electricity, natural gas, housing at school, water and sewerage maintenance, most, medical care services, and vehicle registration and licensing fees.

[19] See Cage, Robert, John Greenlees and Patrick Jackman. Introducing the Chained Consumer Price Index. Meeting of the International Working Group on Price Indexes, Paris France, May 2003. https://www.bls.gov/cpi/additional-resources/chained-cpi-introduction.pdf.

release the Chained CPI as a new, separate index product. Preliminary versions are published on the same schedule as the CPI-U, but are subject to revision.[20]

Over its history, the Chained CPI has yielded annual inflation estimates that are on average approximately 0.25 percentage points less than the CPI-U (see Figure 5). As a consequence, some observers, notably fiscal conservatives, argue it is most suitable for the Social Security COLA.[21] As part of the 2012 and 2013 "fiscal cliff" negotiations, President Obama proposed switching the COLA to the Chained CPI-U as a way to address budgetary shortfalls.[22] Congress did not adopt his proposal. But five years later, with the passage of the Tax Cuts and Jobs Act of 2017, Congress changed the indexation of federal income tax brackets from the CPI-U to the Chained CPI-U.[23] The new legislation, however, did not mandate this change for the Social Security COLAs which remain indexed to the CPI-W.

Figure 5. *Difference in annual inflation estimates, Chained CPI-U minus CPI-U*



2 Criticisms of population subgroup indexes and the current COLA debate

2.1 Population Targeted

Proponents of the CPI-E argue it is the more appropriate index for Social Security COLAs precisely because the cohort definition, elderly consumers, more properly aligns with the targeted group

[20] See Klick, Joshua. Improving initial estimates of the Chained Consumer Price index. Monthly Labor Review, February 2018. https://www.bls.gov/opub/mlr/2018/article/improving-initial-estimates-of-the-chained-consumer-price-index.htm.

[21] See, e.g., Rosenberg, Adam and Marc Goldwein. Measuring Up: The Case for the Chained CPI. Moment of Truth Project, December 2012. http://www.momentoftruthproject.org/sites/default/files/Measuring%20Up.pdf.

[22] The term fiscal cliff was popularized by Ben Bernanke former chairman of the US Federal Reserve, describing the point when Bush era tax cuts would expire and spending cuts would be triggered in 2013.

[23] See Public Law 115-97, The Tax Cuts and Jobs Act of 2017.

addressed by the adjustment. Research has shown that spending patterns differ between the elderly and the general population, especially in the medical care category. Seniors 65 and older spend more than twice on health care, and those 75 and older spend nearly three times more on health care than younger consumers.[24] Moreover, price change for medical care items have increased at a higher rate than other items in the CPI market basket. Critics of using the CPI-W argue its primary flaw is its failure to take these differences for the elderly population into account, and is therefore biased downward for the COLA purpose.[25]

Figure 6. *Relative importance for select CPI categories, 2013-2014*

	Category	CPI-U	CPI-W	CPI-E
A1	Adult clothing	1.7%	1.6%	1.3%
A2	Children and infant clothing	0.4%	0.5%	0.1%
E0	Telephone and electronics	6.0%	6.9%	5.5%
E1	Tuition	2.7%	2.0%	1.2%
F1	Food at home	10.9%	12.6%	9.9%
F2	Food away from home	5.1%	5.1%	4.4%
F3	Alochol and tobacco	1.4%	1.6%	1.1%
H0	Shelter	32.1%	30.6%	35.5%
H1	Household utilities	5.5%	6.0%	5.7%
H2	Housefurnishings and operations	3.7%	2.6%	4.1%
M0	Medical Care	8.3%	6.5%	11.1%
R0	Entertainment and recreation	4.5%	3.6%	4.3%
T0	Vehicles	7.0%	7.9%	6.4%
T1	Gasoline and vehicle maintenance	9.5%	11.4%	8.0%
T2	Public transportation	1.3%	0.9%	1.3%

Note: Estimates from CE Interview Survey only, research sample data.

Opponents of the CPI-E are just as numerous. Arguments against its use for indexing public transfer payments generally fall into the following categories:

1. Medical Care. Out-of-pocket medical care expenses account for the majority of the growth rate differences between the CPI-E and the other indexes. But, as the 1996 Boskin Commission hypothesized, failure to adequately capture quality change in the medical sector could bias these component indexes upward as much as 3 percent per year. Switching the COLA to CPI-E, where medical care has a higher weight, would exacerbate this perceived bias.
2. Substitution Bias. BLS does not currently produce a chained version of the CPI-E utilizing the Tornqvist version to mitigate upper level substitution bias. For this reason, some commentators

[24] See http://www.ncpssm.org/PublicPolicy/SocialSecurity/Documents/ArticleID/1159/The-CPI-E-%E2%80%93-A-Better-Option-for-Calculating-Social-Security-COLAs.

[25] *Ibid.*

argue the CPI-E is flawed due to its upper-level aggregation formula. Accordingly, they maintain the Chained CPI-U is a more accurate measure of overall inflation, making it a better candidate for the COLA.[26]

3. Accuracy of Aggregation Weights. CPI-E aggregation weights are derived from smaller CE survey sample sizes, leading to larger sampling error in the weights and a less accurate index compared to the CPI-U and Chained CPI-U. Households classified as CPI-E comprised about 30% of the CE sample in 2015.

4. Outlet and Item Samples. The CPI-E fails to adequately incorporate outlets patronized by the elderly, nor track the prices paid by the elderly, and therefore the index falls short of accuracy standards. Due to known third degree price discrimination for some items, e.g., senior citizen discounts for food and cinema admissions, the CPI-E cohort faces different prices for the same goods and services to some extent. To the extent that inflationary change for the set of actual prices paid by the cohort for the subset of specific goods and services purchased at the subset of retail establishments differs from price change experienced in the overall economy, then the CPI-E is of unknown accuracy.

5. Cohort Definition. Some argue if the dedicated goal of the COLA is to protect the purchasing power of Social Security income, as opposed to other sources of income, then the adjustment should be narrowly defined to represent the spending of this income by Social Security recipients. This viewpoint essentially argues the current cohort definition of the CPI-E is ill-fitting. In 2007, Burdick and Fisher of the SSA argued the most practical objection to using the CPI-E for Social Security COLAs was its cohort definition.[27] They posited one-fifth of OASDI beneficiaries were under the age of 62, and therefore excluded from the CPI-E. Likewise, one-fifth of persons age 62 or older were not beneficiaries but were included in the index.[28] In effect, Burdick and Fisher suggested the targeted cohort of the COLA should be all OASDI recipients, regardless of age. Some argue the intent of the COLA is to guard against the erosive effects of *overall* inflation, and hence a national index measuring the macro economy is best suited, such as the CPI-U or Chained CPI-U. These commentators argue a switch to a cohort-

[26] See Rosenberg and Goldwein, *supra*.

[27] See Social Security Bulletin, Vol. 67, No. 3, 2007 at https://www.ssa.gov/policy/docs/ssb/v67n3/v67n3p73.html.

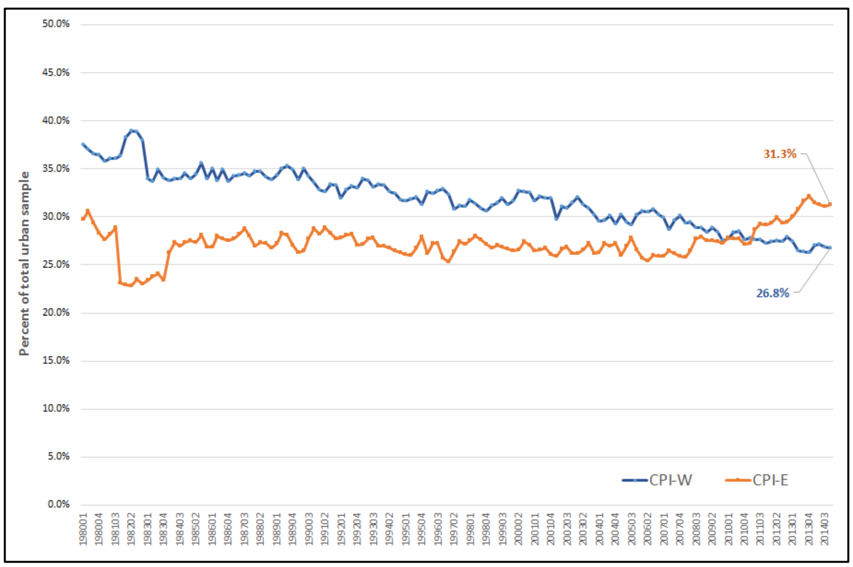
[28] In 2013-2014, approximately 7% of non CPI-E households in the Consumer Expenditure Interview Survey had the majority of their income sourced from Social Security, and 20% of CPI-E households had the majority of total household income from non-Social Security sources.

specific index targeting a population subgroup would constitute a public policy change, rather than a change necessitated to improve the accuracy of the COLA.[29]

2.2 Accuracy of the CPI-W

Although largely overlooked by analysts, each of these potential criticisms levied at the CPI-E can also attach to the index actually used for the COLA: the CPI-W. Of particular concern is the declining sample size of the CPI-W in the CE household expenditure survey. Once constituting 100% of the sample prior to the expansion to urban population coverage in 1978, the CPI-W sample has been in decline. In fact, it has been smaller than the CPI-E sample since mid-2011. The CPI-W accounted for a little over a quarter of all survey respondents by the end of 2015 (see Figure 7). This phenomenon is partly explained by higher survey non-response rates among worker households compared to elderly households, but also indicative of a labor force transitioning into the modern digital economy. The out-of-date delineation of occupations eligible for the CPI-W cohort is shrinking as a percentage of the total workforce.

Figure 7. *CPI-U and CPI-W sample size, as proportion of total sample*
Consumer Expenditure Interview Survey



Furthermore, BLS and SSA currently utilize different customized definitions of the *wage-earner* cohort. Historically, the term has been used pejoratively to describe the poor, underprivileged, and propertyless proletariat.[30] It has also been used to describe a class of workers, usually manual laborers, distinguished by method of recompense: paid by the hour earning a fixed hourly wage rate, rather than an annual

[29] See Rosenberg and Goldwein, *supra*.
[30] See George Bernard Shaw, *What Socialism Is*. The Fabian Society, 1890.

salary.[31] Neither of these connotations, low-income and earning method, are technical criterion used in the current BLS nor SSA wage-earner definition.

The CPI-W cohort is currently defined as urban households with a majority of income sourced from earnings in a limited list of specific occupation types.[32] This cohort represented roughly 29% of all U.S. households in 2015-2016, and only 40% of full-time labor force participant urban households. In contrast, the SSA defines wage-earner as an individual who earns Social Security credits while working for wages or for self-employment income.[33] Initially excluding the self-employed, federal and state government employees, and other classes of workers, Social Security eligibility has expanded since 1935 and over 90% of the U.S. labor force is now covered under the program.[34] Hence, there is a large gap between the CPI and SSA wage-earner definitions. If the targeted population of the index used for the Social Security COLA is the cohort of Social Security tax payers in the labor force, then the CPI-W represents only a fraction of this group.

Moreover, CPI-W households have lower incomes, on average, than non CPI-W households with earnings. Stores and retail establishments patronized by lower-income households, and the specific variety and brands of products purchased within CPI component categories, may be different from the all-urban population at large – the same concern associated with the CPI-E. Any bias associated with using the all-urban component indexes to construct the CPI-W likely increases as the cohort shrinks in size. Nonetheless, these aforementioned shortcomings of the CPI-W have yet to gain much attention in the public debate over the COLA.

3 Methodological Issues in Cohort Index Construction

3.1 The Schultze Advisory Panel

The difficulty of linking the actual prices observed on a monthly basis in the CPI, in the sampled set of retail establishments, to the characteristics of the consumers paying the prices, hampers the ability of statistical agencies to construct accurate indexes for population subgroups. This was a major finding of a National Research Council panel of the National Academy of Sciences sponsored by BLS to investigate conceptual, measurement, and other statistical issues in the development of cost-of-living indexes,

---

[31] In most English dictionaries, wage-earner is defined as 'a person who works for wages, especially as distinguished from one paid a salary.' American English dictionaries typically omit the wage-salary distinction, defining wage-earner as 'a person who works for wages *or* salary.'

[32] Delineated occupations included in the CPI-W include construction workers, mechanics, forestry, fishing, grounds keeping, machine or transportation operators, protective service, retail sales workers, administrative support including clerical, and other service workers such as janitors and food industry workers.

[33] See https://www.ssa.gov/agency/glossary/.

[34] See https://www.ssa.gov/policy/docs/ssb/v48n4/v48n4p33.pdf .

commonly referred to as the Schultze Panel.[35] The panel's report, issued in 2002, contained a comprehensive examination of population group indexes, from conceptual basis to practical estimation methods[36]

Among the notable findings and recommendations of the Schultze Panel:

- Group indexes should account for two sources of heterogeneity: across stratum product categories and within stratum product categories; within stratum differences arise from idiosyncratic tastes and preferences, but also from differences associated with age, income, family composition, and geographic location. Most experimental group indexes control for the former, but there is a lack of evidence on the latter to enable adequate exploration.[37]
- In the absence of an index that can capture differences in price or qualities of good purchased by a subgroup, there is no rationale for switching to an index (e.g., the CPI-E) for purposes of Social Security COLA indexation. Differences between the CPI-U and CPI-W are not large enough to warrant a change.[38]
- If inflation rates differ among population cohorts, then use of an overall CPI to index Social Security COLAs may overcompensate some groups and undercompensate others, in a way most people would deem unfair or unjust.[39]
- BLS should explore collecting prices in a way that allows them to be associated with household characteristics, and Congress should fund it.[40]

The panel seemed to suggest the preeminent weakness in cohort indexes, such as the CPI-E and CPI-W, is their failure to incorporate cohort-specific outlet and item samples, and to track cohort-specific transaction prices. They went as far to proclaim, 'the current BLS data collection system cannot produce group indexes.'[41] Most consumer population subgroup indexes produced by national statistical agencies suffer from this problem.

Since the issuance of the Schultze and Mackie advisory report in 2002, research has begun to emerge using scanner and large transaction data, seeking evidence of within-category differences across cohorts. Diamond, Watanabe, Watanabe (2016) used individual level purchase and demographic household scanner data, linked to 33 million transactions, to decompose group indexes into weight and

[35] See Schultze, Charles and Christopher Mackie. *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes.* National Research Council, 2002. https://www.nap.edu/read/10131/chapter/1. Professor Schultze was the chair of the panel.
[36] Schultze, Chapter 8.
[37] Schultze, p. 224.
[38] Schultze, p. 198.
[39] Schultze, p. 226.
[40] Schultze, Finding 8.1.
[41] Schultze, p. 226.

price effects in Japan– essentially controlling for heterogeneity across and within category groupings.[42] They concluded price and weight effects do not vary across age group, except at the in the oldest group, but warn measures of variation are highly dependent on how the groups are defined.

Beyond the within product category heterogeneity issue, two additional conceptual problems plague the estimation of cohort indexes: treatment of owner-occupied housing and aggregation method. However, a consensus approach among researchers and practitioners appears to be emerging for these.

3.2 Treatment of Owner Occupied Housing

The definition of the weight for owner occupied housing is controversial. BLS switched from a user cost to rental equivalence approach for this component in 1983.[43] Most elderly consumers in the Consumer Expenditure Survey sample are homeowners, and the majority of these own their home without a mortgage, having disposed of the debt obligation earlier in life.[44] Hence, the current expenditure needed to maintain the home, or user cost, is limited to property taxes, insurance, utilities, and maintenance. Insurance, utilities, and household maintenance are explicitly captured in the CPI in different component categories. In effect, the rental equivalence approach discards the mortgage payment and property tax expense, in favor of an estimate of the market rental value of the home. In particular, this approach imputes a hypothetical expenditure typically larger than the carrying cost of the home, especially for homeowners not holding any mortgage debt. For these consumers, this lowers the relative importance of other CPI items – items actually purchased in the current period. When shelter inflation outpaces inflation for other market basket items, the rental equivalence approach may overstate the actual inflation of the elderly cohort.

While many economists agree the rental equivalence approach is appropriate for an index needed to gauge overall national inflation, many argue it may not be appropriate for indexation needs of targeted population subgroups. For example, Hobijn and Lagakos questioned its use for indexing a cash benefit program like OASDI because rental equivalence measures an *opportunity cost* rather than actual out-of-

[42] See Diamond, Jess, Watanabe, Kota and Tsutomu Watanabe, Estimating Consumer Prices Inflation by Household at http://www.price.e.u-tokyo.ac.jp/img/event/2015SWET-Jess.pdf.
[43] See Gillingham, Robert and Walter Lane. Changing the treatment of shelter costs for homeowners in the CPI. *Monthly Labor Review*. June 1982.
[44] In 2013-2014, 77% of CPI-E households were classified as homeowners, and 67% of these owned without a mortgage.

pocket expenses for many elderly.[45] Boskin argues the rental equivalency method, in effect, compensates homeowners for capital gains on their homes.[46]

3.3 Aggregation across Households

The choice of proper aggregation method consistently arises in the price index literature. To aggregate household specific indexes into aggregate measures for a group, one can weight each household by their total expenditures on the CPI market basket items, a plutocratic weighting approach, or assign each household an equal weight, referred to as a democratic approach. The plutocratic approach tends to skew the aggregate index toward the behavior of high-income households, who get a larger relative weight. The CPI-U, CPI-W, CPI-E, and Chained CPI-U are all plutocratically weighted. Deaton (1998) suggested the representative household suggested by the plutocratic CPI-U was around the 75th percentile of income. Schultze argues, 'it is hard to imagine anyone would deliberately make decisions about public pensions by tracking households at the 75th percentile.'[47] The Schultze panel suggested the optimal aggregation choice for a national index may differ from the choice for subgroups, essentially recommending the plutocratic approach for the index used as a barometer for the macroeconomy, and a democratic approach for subgroup indexes. The U.K. Royal Statistical Society concurred with this recommendation in 2015.[48]

4 Improving population subgroup index estimates

To understand the dynamics generating current differences in CPI-U, Chained CPI-U, CPI-W, and CPI-E inflation estimates, and to seek evidence of current inflation heterogeneity across cohort groups in the United States, we scrutinized the sample unit level data used to compile the U.S. official indexes from 2014 to 2017. Then, we utilized a model-based machine learning clustering algorithm to control for heterogeneity across households, in an experimental cohort aggregation approach.

---

[45] See Hobihn, Bart and David Lagakos. Inflation Inequality in the United States. *Review of Income and Wealth*, December 2005.

[46] See Boskin, Michael et al. Consumer Prices, the consumer price index, and the cost of living. *Journal of Economic Perspectives*. Vol.12, No. 1. Winter 1998. https://pubs.aeaweb.org/doi/pdf/10.1257%2Fjep.12.1.3.

[47] Schultze, p 240.

[48] See Bentley, Alan. *Household-group inflation: methods to combine expenditure patterns*. Meeting of the Group of Experts on Consumer Price Indices. Geneva, Switzerland. May 2016.

4.1 Data set

The experimental research data set started with all households (consumer units) participating in the 2013 and 2014 CE Quarterly Interview Survey, the expenditure reference period of the CPI-U, CPI-W, and CPI-E in 2016 and 2017 (n=24,080). This data set was then restricted to the subset of respondents participating in four quarterly interviews during this period, to directly obtain annual expenditure estimates for each household (n=6,606).[49] Due to the reduced sample size, expenditure data were summed across the 243 component CPI item categories to form household (*j*) shares (*w*) at a broader, expenditure class (*k*) level (n=70) for the arbitrarily selected base-period of Θ = 2013-2014 .[50] See Equation 1 and 2.

$$(1)\quad E_{k,j,\theta} = \sum_{i \in k,j}^{i} (PQ)_{i,j,\theta} \qquad\qquad (2)\quad w_{k,j,\theta} = \frac{E_{k,j,\theta}}{\sum_{k \in j}^{k} E_{k,j,\theta}}$$

The CE Interview Survey, a 3-month recall survey, is not designed to capture detailed expenses for all of these categories. Detailed food expenses are only collected in the CE Diary Survey. Official CPI aggregation methods integrate data from the Diary and Interview survey components to estimate expenditures at the component cell level (n=9,234) by calendar month. Hence, the official method neither requires the collection nor calculation of household specific budget shares for each component item. This research jettisoned the Diary data, and used expenditures collected in the Interview for the categories officially sourced from the Diary, e.g., food at home and some apparel items. This reduced the set of component cells to 52 item categories.[51] Then, an initial Laspeyres and Geometric Mean index for each household was computed using corresponding component (*k*) national average price change from December 2014 to December 2017. See Equations 3 and 4.[52]

$$(3)\quad I^{L}{}_{j;\theta \to t} = \sum_{k \epsilon j}^{k} w_{k,j,\theta} \left(\frac{P_{k,t}}{P_{k,\theta}} \right) \qquad\qquad (4)\quad I^{G}{}_{j;\theta \to t} = \prod_{k \epsilon j}^{k} \left(\frac{P_{k,t}}{P_{k,\theta}} \right)^{w_{k,j,\theta}}$$

---

[49] Using households with partial year reported data would bias household budget shares toward items with strong seasonal spending behavior. Use of 4-panel households eliminated any seasonal bias in the household specific shares, at the expense of jettisoning 75% of the data used in the constructing of the official BLS indexes.

[50] This results in a loss of precision in measuring heterogeneity across household and groups of households, notably variation in spending behavior that might exist across component items within expenditure class.

[51] Garter, et al. demonstrate a model-based approach to link Diary households to Interview, in order to estimate item-specific weights, sourced from the Diary, to Interview households. Adoption of this method would allow for a complete matrix of spending for each Interview household at the CPI component stratum level. See https://www.bls.gov/ore/pdf/ec020060.pdf .

[52] Using national level indexes sacrifices precision and variation attributable to geographic variation. Future enhancements of this research will control for geographic variation.

Descriptive statistics for the experimental research sample are shown in Figure 8. The restricted research sample has slightly more elderly and retired households, and fewer wage-earning households, compared to the full CPI-U sample.

Figure 8. *Descriptive statistics for research sample*

Descriptive Statistics 2013 and 2014 Consumer Expenditure Interview Survey	4-Interview CPI-U Households		All CPI-U Households	
Sample size	6,606		24,080	
Sample with:	n	%	n	%
Income > 0	6,596	*99.8%*	22,488	*93.4%*
One or more members age 16 or under	2,571	*39.0%*	8,772	*39.0%*
One or more owned motor vehicles	5,756	*87.3%*	18,782	*83.5%*
One or more members age 64 or over	2,134	*32.4%*	5,821	*25.9%*
One or more members earning income	4,832	*73.3%*	17,308	*77.0%*
Only one member (single consumer)	1,894	*28.7%*	7,235	*32.2%*
Living in center city of metropolitan area	1,906	*28.9%*	7,574	*33.7%*
50% or more of income sourced from Social Security	1,424	*21.6%*	3,726	*16.6%*
50% or more of income sourced from wage earnings	3,671	*55.7%*	14,224	*63.3%*
Total household annual income before taxes				
First quintile	$ 26,889		$ 21,720	
Third quintile	$ 89,520		$ 85,015	
Average	$ 68,058		$ 65,223	

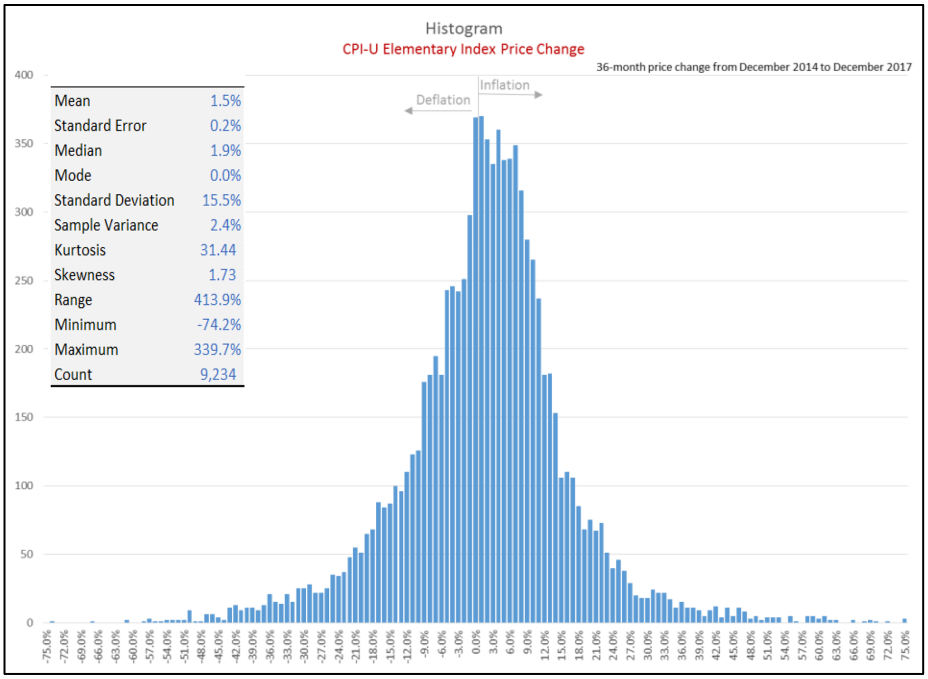
4.2 Evidence of price change dispersion across categories

Overall inflation estimates from the study period, December 2014 to December 2017, are shown below. The research example exhibits the same ordinal ranking of estimates as the official indexes, CPI-E > CPI-U > CPI-W for this time period.

		RESEARCH SAMPLE		
	Official BLS Index	Sample Mean	Democratic	Plutocratic
CPI-U	4.988	4.912	4.845	4.796
Chained CPI-U	4.220	na	na	na
CPI-W	4.618	4.504	4.440	4.333
CPI-E	5.611	5.530	5.498	5.385

If there is minimal variation in price change (1) within the component items in the CPI market basket, and (2) across the component items, combined with (3) homogeneous budget share patterns across subgroups, then the four CPI products will tend to produce similar estimates of inflation. Component price change for CPI elementary items exhibited significant variation during the study period, ranging from a minimum of -74.2% (car rentals in St. Louis) to a maximum of 339.7% (moving expenses in Houston). See Figure 9. The distribution of price change at the broader expenditure class (*k*), nationwide level is narrower, ranging from a low of -19.1% for recreational goods to a high of +22% for vehicle insurance. Categories exhibiting above average inflation during this period include medical care, shelter, food away from home, and college tuition. Households with disproportionately large budget shares for these items will have higher than average household specific inflation. Categories exhibiting the most deflation during this period include apparel (except jewelry and watches), household goods, and telephone and internet services. Households with large budget shares for these categories will have lower household-specific inflation experience.

Figure 9. *Distribution of CPI component item stratum and area price change*



4.3 Evidence of inflation heterogeneity across households

The distribution of the household specific Laspeyres index estimates for the 2013-2014 experimental research sample is shown in Figure 10. The distribution exhibits less variation than the component indexes, with a min of -2%, a max of 11.9%, and a coefficient of variation of 33%. The distribution is slightly skewed toward the higher inflation tail, with 4% of the sample having inflation rates beyond two standard deviations from the mean, compared to 2% in the corresponding lower tail. Scatterplots of price change by age indicate a significant upward trend, although age does not explain much of the household inflation variation. See Figure 11. Similarly, a scatterplot of price change by income shows a downward linear trend, with a heteroskedastic range at lower incomes. See Figure 12. Note the households with the highest inflation values are the lower income elderly households. Hence, this evidence is consistent with prior BLS research suggesting the poor tend to face higher inflation rates compared to the rich, and the elderly face higher rates than the young.[53]

Figure 10. *Distribution of household specific price change*

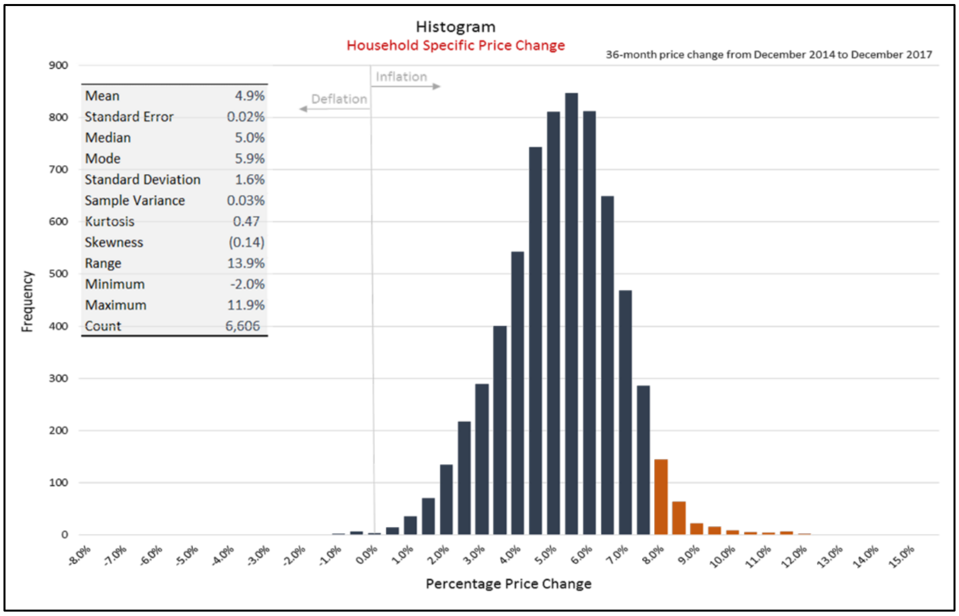
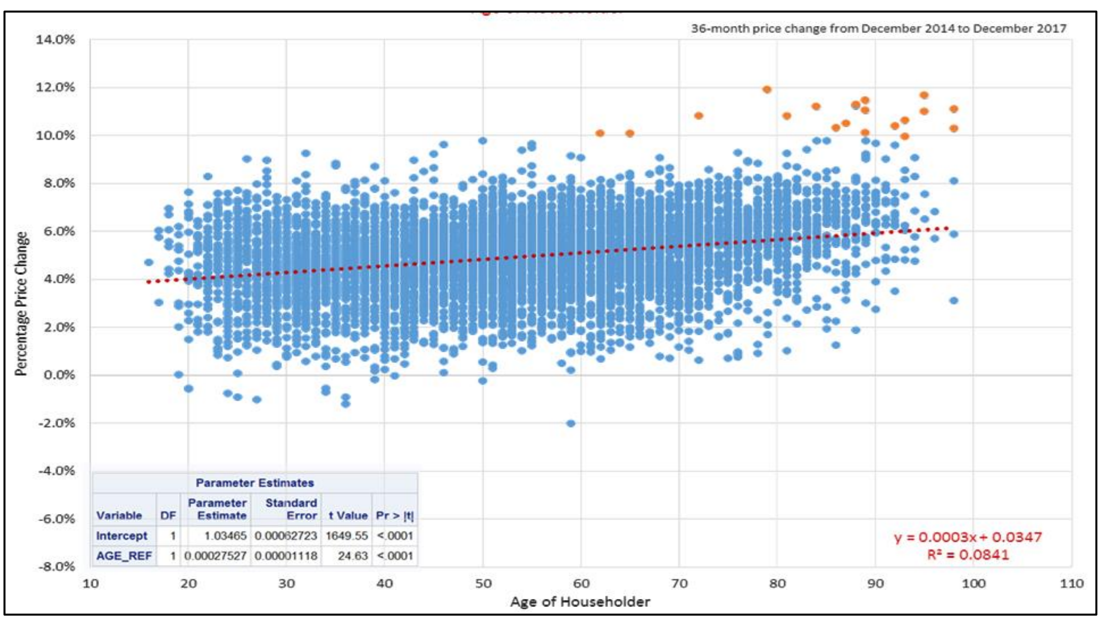
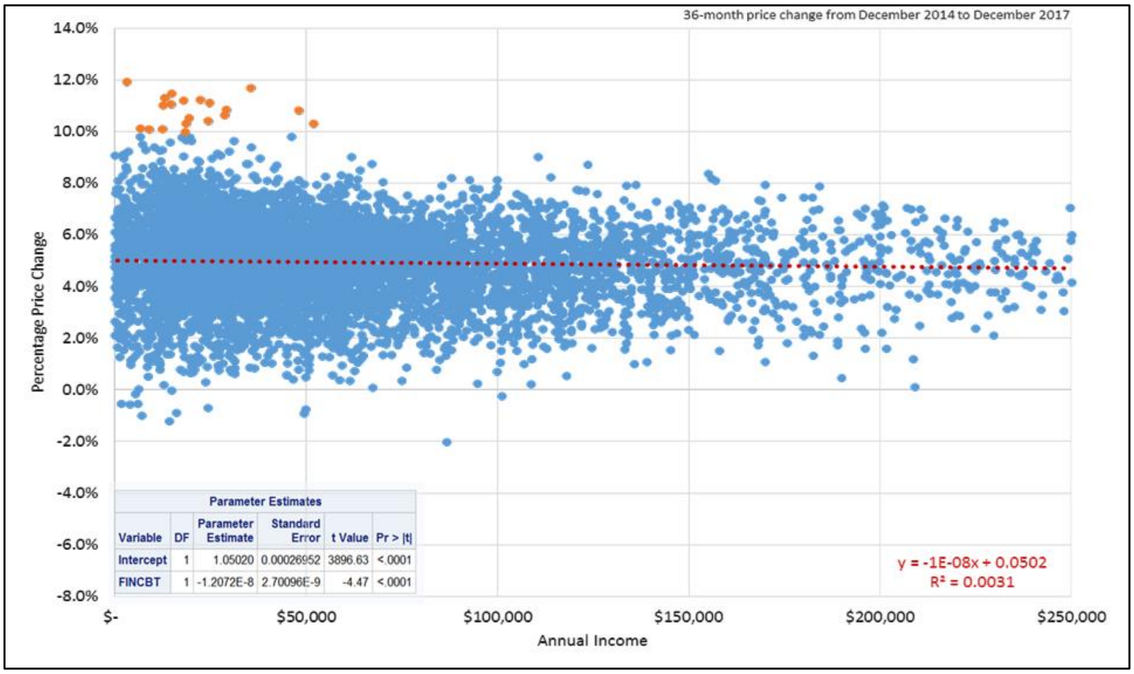


Figure 11. *Inflation scatterplot by age of householder*



53 See Garner, Thesia, et al. An experimental consumer price index for the poor. Monthly Labor Review. September 1996.

Figure 12. *Inflation scatterplot by household income*



4.4 Evidence of variation within categories

To address the failure of the cohort indexes to account for within stratum variation, we stratified the CPI sample of rents and goods and services into three price level groups based on the price of each item in the sample as of December 2014: lowest quintile prices, highest quintile prices, and middle 60%. The stratification was executed at the CPI elementary item level. Then, the stratified samples were pooled by expenditure class (*k*) and component indexes were recalculated over the 36-month study period. Due to sample rotation, the sample of usable rents and price quotes with available price data in both periods was roughly 50% of the full CPI sample (n=67,268). Descriptive statistics for this restricted sample are shown in Figure 13. Price change in the study period for the low price items, on average, is higher compared to high price items.

Assuming lower income households have income to purchase only the lower price items within each category, and higher income households purchased the higher priced items, the stratified component indexes were matched to households by income quintile. This experiment is a heuristic way to capture potential variation existing in the marketplace, lacking the actual prices paid for specific brand and variety of items purchased by each household. In effect, the price level of the items proxy for the hypothesized limited vector of all marketplace prices paid by income cohort groupings.

Figure 13. *Descriptive statistics for CPI sample data, by price level strata*

	ALL	*LOWEST*	*MIDDLE*	*HIGHEST*
Mean	1.07102	1.0824	1.0709	1.0600
Standard Error	0.00149	0.0036	0.0019	0.0031
Mode	1	1	1	1
Standard Deviation	0.38661	0.4150	0.3854	0.3595
Sample Variance	0.14947	0.1722	0.1485	0.1293
Kurtosis	347.8	270.3	422.2	175.0
Skewness	11.7	10.9	12.9	8.5
Range	19.95	17.34	19.95	12.17
Minimum	0.05	0.05	0.05	0.05
p10	0.3983	0.4555	0.3901	0.3762
Q1	0.9542	0.9479	0.9574	0.9515
Median	1.0308	1.0213	1.0337	1.0313
Q3	1.1393	1.1438	1.1410	1.1298
p90	1.3017	1.3333	1.3011	1.2800
Maximum	20	17.39	20.00	12.22
Count	67268	13540	40196	13532

4.5 A model-based clustering aggregation approach

Lacking data that connects the prices and quantities of marketplace transactions to the household characteristics of the consumers making the purchases due to lack of funding, how can statistical agencies improve subgroup index estimation methods using the survey data available to them? Using a subset of the household, rent, and price sample data, we developed an experimental approach to constructing a hypothetical subgroup index targeting Social Security, within the confines of estimating a national index. The experimental approach follows the guidance of the Schulze advisory panel, and differs from the official CPI methods in four ways: (1) treatment of owner-occupied housing, (2) cohort definition, (3) component prices, and (4) aggregation method.

Step 1. Stratify the starting or base period sample of households into a limited number of mutually exclusive but exhaustive cohorts of most interest to users. This could be income quintiles or CPI market basket expenditure quintiles, to control for hypothesized inflation variation across the income distribution. In the U.S., the prominent use of the CPI for Social Security COLA might suggest stratifying the sample by Social Security recipient status. We stratified the research sample into three income (lowest quintile, medium quintiles, highest quintile) and two Social Security (recipients, non-recipients) groups.[54]

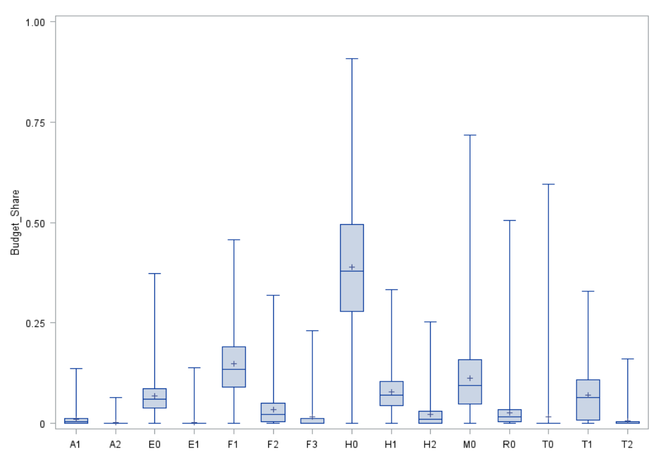
Figure 14. *Initial population strata sample sizes*

Code	Social Security households:	n	%
	Social Security households:	1424	*22%*
SL	Low income	648	*10%*
SM	Middle Income	776	*12%*
	Non Social Security households:	5171	*78%*
NL	Low income	671	*10%*
NM	Middle income	3181	*48%*
NH	High income	1319	*20%*

Note: No households classified as Social Security recipients were in the high income quintile group, resulting in five initial population strata instead of six.

In order to apply the economic approach to index number theory, it is necessary to assume that each household *(j)* within each of the five initial strata (s) defined in Step 1, behave as a single household.[55] That is, the member households in each (s) should be homogeneous with respect to their CPI budget shares such that the assumption of homothetic preferences applies. However, there remains a great deal of variation in budget shares within these initial strata (see Figure 15). To strengthen the assumption of homothetic preferences, we cluster the households within each strata (s) into cohorts (c) based on their revealed spending shares.

Figure 15. *Boxplot of household budget shares for select CPI categories, Social Security low-income strata*



[54] To be classified as a Social Security household, 50% or more of total household income had to be sourced from Social Security benefits.

[55] See International Labour Office. Consumer Price Index Manual, Theory and Practice. Chapter 18, Section 18.1 p. 337.

Step 2. Calculate household specific budget shares. See equations (1) and (2) in section 4.1. Calculate household specific indexes. See equations (3) and (4) in section 4.1. We calculated two sets: one using a rental equivalence approach for owner-occupied housing, and one using a payment approach as outlined in section 4.5. Household expenditures were summed to a broader 15-category level (used for clustering in subsequent steps) as shown in Figure 16.

Figure 16. *Budget shares for select CPI categories, by experimental cohort*

RENTAL EQUIVALENCE

	Category	NL	NM	NH	SL	SM
A1	Adult clothing	1.2%	1.6%	2.1%	1.1%	1.0%
A2	Children and infant clothing	0.6%	0.4%	0.4%	0.1%	0.1%
E0	Telephone and electronics	6.8%	6.6%	5.3%	6.2%	5.5%
E1	Tuition	2.6%	1.9%	4.9%	0.3%	0.5%
F1	Food at home	15.7%	11.9%	8.8%	12.3%	10.2%
F2	Food away from home	4.2%	5.1%	5.8%	3.6%	3.9%
F3	Alochol and tobacco	1.8%	1.5%	1.3%	1.3%	1.0%
H0	Shelter	35.2%	31.7%	30.2%	39.4%	34.5%
H1	Household utilities	7.2%	5.8%	4.3%	7.1%	6.3%
H2	Housefurnishings and operations	2.2%	3.0%	4.9%	2.6%	3.4%
M0	Medical Care	5.7%	7.6%	7.3%	12.0%	13.9%
R0	Entertainment and recreation	3.3%	3.9%	5.9%	2.8%	3.7%
T0	Vehicles	3.8%	7.0%	8.0%	3.2%	6.9%
T1	Gasoline and vehicle maintenance	8.9%	10.7%	8.8%	7.2%	8.3%
T2	Public transportation	0.9%	1.0%	1.8%	0.7%	0.9%

PAYMENTS APPROACH

	Category	NL	NM	NH	SL	SM
A1	Adult clothing	1.4%	1.7%	2.3%	1.4%	1.3%
A2	Children and infant clothing	0.7%	0.5%	0.5%	0.2%	0.2%
E0	Telephone and electronics	7.5%	7.2%	5.7%	7.8%	7.1%
E1	Tuition	2.8%	2.0%	5.2%	0.4%	0.6%
F1	Food at home	17.2%	13.0%	9.5%	15.6%	13.1%
F2	Food away from home	4.6%	5.6%	6.3%	4.6%	5.0%
F3	Alochol and tobacco	2.0%	1.7%	1.4%	1.7%	1.2%
H0	Shelter	28.8%	25.6%	25.1%	23.1%	16.1%
H1	Household utilities	7.9%	6.4%	4.7%	9.0%	8.1%
H2	Housefurnishings and operations	2.4%	3.3%	5.3%	3.3%	4.3%
M0	Medical Care	6.3%	8.3%	7.8%	15.2%	17.8%
R0	Entertainment and recreation	3.6%	4.3%	6.3%	3.5%	4.7%
T0	Vehicles	4.2%	7.6%	8.6%	4.1%	8.9%
T1	Gasoline and vehicle maintenance	9.8%	11.7%	9.5%	9.1%	10.6%
T2	Public transportation	1.0%	1.1%	2.0%	0.9%	1.1%

Figure 17. *Difference in rental equivalence and payment approach shares, by experimental cohort*

	Category	NL	NM	NH	SL	SM
A1	Adult clothing	0.1%	0.1%	0.2%	0.3%	0.3%
A2	Children and infant clothing	0.1%	0.0%	0.0%	0.0%	0.0%
E0	Telephone and electronics	0.7%	0.6%	0.4%	1.7%	1.6%
E1	Tuition	0.3%	0.2%	0.4%	0.1%	0.1%
F1	Food at home	1.5%	1.1%	0.6%	3.3%	2.9%
F2	Food away from home	0.4%	0.5%	0.4%	1.0%	1.1%
F3	Alochol and tobacco	0.2%	0.1%	0.1%	0.3%	0.3%
H0	Shelter	-6.3%	-6.1%	-5.0%	-16.2%	-18.4%
H1	Household utilities	0.7%	0.5%	0.3%	1.9%	1.8%
H2	Housefurnishings and operations	0.2%	0.3%	0.4%	0.7%	1.0%
M0	Medical Care	0.6%	0.7%	0.5%	3.2%	3.9%
R0	Entertainment and recreation	0.3%	0.4%	0.4%	0.7%	1.0%
T0	Vehicles	0.4%	0.6%	0.6%	0.9%	1.9%
T1	Gasoline and vehicle maintenance	0.9%	1.0%	0.6%	1.9%	2.3%
T2	Public transportation	0.1%	0.1%	0.1%	0.2%	0.2%

Step 3. Using the sample of all households, identify budget shares of most significance in explaining variation in the household specific inflation estimates. The objective of the clustering problem is to group households into like-kind, homogenous groups based on budget shares. In the U.S. CPI, with 243 component strata, this becomes a high dimensional data problem, with vector *n* number of budget shares equal to the number of component strata in the CPI classification scheme. To reduce the number

of dimensions, we summed the shares into 15 groupings.[56] We than executed the following LASSO model:

$$(5)\quad {I^L}_{j;\theta\to t} = v(w_{k_i})$$

to determine the optimal set of variables for clustering. LASSO modifies OLS regression by adding a penalty term to the error. Mathematically the problem is to find $(\beta_1, \beta_2, \dots, \beta_{15})$ which minimizes the error of:

$(6)\quad {I^G}_{j;\theta\to t} = \sum_{i=1}^{15} \beta_i w_{k_i}$ where the error is $\sum_{observations}\left({I^G}_{j;\theta\to t} - \sum_{i=1}^{15} \beta_i w_{k_i}\right)^2 + \lambda \sum_{i=1}^{15} |\beta_i|$

The penalty term increases the error if there are large coefficients. This has a tendency to reduce or drive to zero the coefficients of shares which are not strongly predictive of the response. The value of λ is determined by cross validation.[57]

We selected the following shares as most explanatory: f1- food at home, f2-food away from home, h0-rent or owner's equivalent rent, m0-medical area, and t1-gasoline and vehicle maintenance.

Step 4. In order to reduce the variation in spending behavior *within* the income-Social Security strata, we executed a hierarchical clustering algorithm to form clusters of homogeneous households based on the revealed budget shares.[58] This approach, in effect, superimposes homothetic utility preferences inside the cluster as the vector of budget shares on CPI items will be approximately equal across all households within the cluster. If the clustering algorithm performs well, then a democratic average and plutocratic average index will converge for the cluster, eliminating the difficult choice of aggregation method.

---

[56] A1=Adult apparel, A2=Infant and children apparel,E0=Telephone, internet, audio, video, electronics, E1=Tuition, F1=Food at home, F2=Food away from home, F3-Alcohol and tobacco, H0=Shelter, H1=Household utilities, H2=Housefurnishings and operations, M0=Medical care, R0=Entertainment and recreation, T0=Vehicles, T1=Gasoline and vehicle maintenance, T2=Public transportation.

[57] Typically 10-fold cross validation is used. This means the data are divided into 10 blocks, the betas are fit using 9 blocks of data and used to predict the 10th block. The combined error of predicting the 10 blocks in turn is used to select the best value of λ and the coefficients kept are the ones corresponding to this value of λ. If a coefficient is exactly zero then the variable should be dropped from the model. If the coefficients are close to zero relative to other coefficients then it is up to the modeler's judgement whether or not to keep these variables.

[58] Clustering is an unsupervised machine learning method to partition a set of data into mutually exclusive clusters where data points within a cluster are similar. There are many ways to cluster data but here we use hierarchical clustering.

Hierarchical clustering is a two-step process. First a tree is formed and then the tree is cut at a level chosen by the researcher to provide a partition into the desired number of clusters. It starts with every one of N data points being its own cluster. Then it combines the two nearest points into a cluster resulting in N-1 clusters at the next higher level of the tree. Again the two nearest clusters are combined so there are N-2 clusters at the next highest level of the tree. This is repeated until all points are in a single cluster. The main choice in hierarchical clustering is the choice of a distance function for two clusters. We chose Ward's minimum variance method, which minimizes the increase in total within cluster variance from combining two clusters.[59]

The formula for the distance between two clusters K and L is $D(K,L) = \frac{\|\overline{X_K} - \overline{X_L}\|^2}{\frac{1}{N_K} + \frac{1}{N_L}}$ where the numerator is the squared distance between the centroids of clusters K and L and N_K is the number of points in cluster K (likewise for N_L).

Once the tree is constructed then one has to determine how many clusters to divide the data into. There is no method that is guaranteed to select the optimal number of clusters so one is left to rely on diagnostic criteria.[60] After deciding on the number of clusters, Proc Tree reads in the hierarchical tree and cuts it at the appropriate stage to produce the desired number of clusters.

To visualize the clustering problem, the t-distributed Stochastic Neighbor Embedding method was used to visualize the structure of the high dimensional budget share data in two dimensions.[61] An example is shown in Figure 18 below.

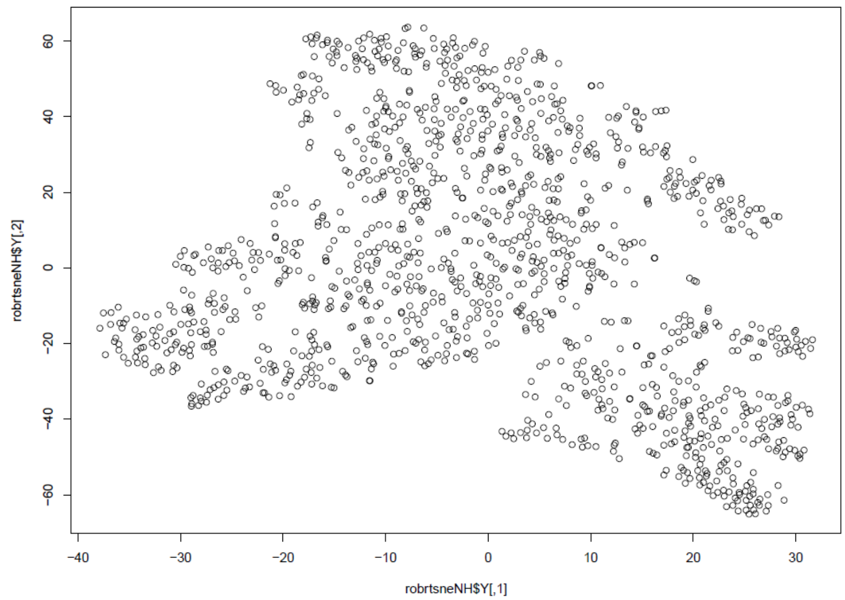
---

[59] See https://arxiv.org/pdf/1111.6285.pdf.

[60] We used the CCC statistic output from the PROC CLUSTER procedure in SAS. SAS Technical Report A-108 describes the formulas and empirical testing behind the CCC statistic. SAS Proc Cluster provides a plot of the CCC for each number of clusters. One should look for a peak in the CCC curve with a value greater than 2.

[61] Laurens van der Maaten and Geoffrey Hinton, Visualizing Data using T-SNE, Journal of Machine Learning Research volume 9(Nov):2549—2578, 2008. To define the structure of the 15-dimensional data, a measure of similarity is constructed. For a point P_i use a Gaussian distribution centered at P_i. Then using the distance between P_i and P_j one can calculate the probability that P_j is the neighbor of P_i. This measure is not symmetric for P_i and P_j so the similarity of P_i and P_j is the average of the similarity of P_j to P_i and the similarity of P_i to P_j. This defines a distribution of similarities in the original 15 dimensional space. The goal of t-SNE is to find a set of points in 2-dimensional space such that the distribution of similarities among points is as similar as possible to the distribution of similarities in the original space. The construction of similarities in 2-dimensional space uses a Student's t distribution with one degree of freedom instead of a Gaussian distribution. The t-distribution has thicker tails than the Gaussian distribution which results in points separated by a moderate distance in the 15-dimensional space being somewhat closer together in the 2-dimensional space. The purpose of this is to hold together any cluster structure in the original 15-dimesional space and to improve the separation of clusters.

Figure 18. *t-SNE plot of stratum NH = Non Social Security, High Income Households*



The final clustering algorithm optimized on 40 clusters for each strata, except the NM (Non Social Security, Middle Income) strata which required 75, for a total of 235 strata-cluster combinations over the entire sample. The average household sample size per strata-cluster was roughly 30, with a min=2, q1=13, q3=37, max=126. The strata-cluster groups form the component building blocks for subsequent index aggregation.

Step 5. Calculate strata-cluster indexes using the Laspeyres formula (to support CPI-U calculation), and a superlative formula (Tornqvist) to mitigate substitution bias, by aggregating across clusters within stratum. Lacking current-period weights for the research sample, we estimated a preliminary version of the superlative using the base-period weights and base-period household clustering, by applying the CES formula and an estimate of substitution elasticity (here σ=0.6) across all clusters.[62] This step can be performed using democratic or plutocratic weights, as the two schemes should converge. We found an average difference of 0.015% between the two weighting approaches across all 235 clusters.

$$(7)\quad I^{L}{}_{s,c;\theta\to t} = \sum_{k,j\epsilon s,c}^{k} w_{k,j,\theta}\left(\frac{P_{k,t}}{P_{k,\theta}}\right) \qquad (8)\quad I^{C}{}_{s,c;\theta\to t} = \prod_{k\epsilon j}^{k}\left(\frac{P_{k,t}}{P_{k,\theta}}\right)^{w_{k,j,\theta}}$$

[62] Ideally, the value of σ should be optimized for each stratum-cluster and could vary. This will be a future enhancement of this research.

Step 6. Democratically average the strata indexes into cohort specific indexes needed by specific users for tailored indexation needs. Here, we calculated a democratic average of the low-income and middle-income Social Security recipient strata to estimate an index for all Social Security recipient households.

Step 7. Plutocratically average the strata indexes into the national index.

Step 8. In order to calculate Tornqvist indexes at the strata-cluster level, a mechanism to classify the sample of current period CE survey households into the base-period within stratum clusters is needed. Since the longitudinal period of the CE survey is only four quarters, households in the survey constantly rotate and we lack current-period spending for the exact set of household as in the base period.

We plan to use the PROC DISCRIM procedure available in SAS on demographic variables to identify current period households with the same characteristics of base-period households, in a future enhancement of this research.[63]

By using classification for the new current-period households, each household can be assigned to base-period strata-cluster based on different variables than were used to perform the original clustering (i.e., the budget shares) as the classification models make no assumptions about how the values of the response in the training dataset were produced. Ostensibly, this will allow the vector of budget shares of the current-period members of the cluster to vary from the vector of the original members, where the variance can be assumed to be response to changing relative prices. This could be an improvement to the current method used to construct Tornqvist indexes, as the clustering should ostensibly control for

[63] Classification is a supervised machine learning method. There is a response variable, which is the cluster number based of the clustering using the 15 budget shares described in step 4. Supervised machine learning means that there is a training dataset with the values of the response filled in. The classification model fitted to the training data (here the clustering in step 4) will predict which cluster a new point should belong to. There are a great many classification models. Most classification models in SAS are part of the SAS Enterprise Miner, such as neural networks, support vector machines, and decision trees. However SAS Proc Discrim is capable of classifying new households based on demographics and the clustering of the initial dataset. Proc Discrim by default assumes a multivariate normal distribution of the demographic variables for points in the training data (the households which were clustered). Since this assumption is unlikely to hold, we use a non-parametric classification method. This method is called K-nearest neighbors. The K-nearest neighbors classification algorithm works as follows: Choose the value of K, the number of neighbors to be used in classification. For a new point P, find the K closest points in the training dataset to the point P. The cluster number containing the most points from the K-nearest neighbors is assigned to the point P. Choosing an odd value for K reduces the likelihood of ties. One can use cross validation to find a very good value for K. In cross validation, the training data is broken into say 10 blocks. Then the K-NN procedure is used to classify the 10th block on the basis of the other 9. Each of the 10 blocks is classified in turn and the error rate or AUC (area under curve) are used to select the value of K which most accurately reproduces the original clusters.

exogenous factors that also influence spending behavior, such as changes in household composition and demographics.

5 Results

We computed aggregate indexes for each stratum using (1) a rental equivalence and (2) payments approach for owner occupied housing. Laspeyres and CES indexes for each stratum were computed using three different sets of component indexes: (1) official CPI-U expenditure class (*k*) national indexes; (2) a restricted sample of CPI-U (*k*) indexes, and (3) a restricted paired price sample of (k) indexes, pairing low-price items to the low income quintile strata, low and medium prices to middle income strata, and all prices to high incomes strata. Results are in Figure 19.

Figure 19. *Index results, rental equivalence approach*

Formula	Owner Occupied Housing	Component Indexes	Low Income Quintile	High Income Quintile	Social Security Cohort	All Urban Consumers
Laspeyres	Payment	Full sample, urban	4.24%	4.50%	4.33%	4.19%
Laspeyres	Payment	Restricted price vector	6.12%	4.44%	5.47%	4.71%
Laspeyres	OER	Full sample, urban	5.06%	4.86%	5.56%	4.80%
Laspeyres	OER	Restricted price vector	7.05%	4.75%	6.81%	5.34%
CES σ=0.6	Payment	Restricted price vector	5.98%	4.32%	5.31%	
CES σ=0.6	OER	Restricted price vector				5.21%

Conclusion

In this paper, we examined micro-level CPI source data from 2014 through 2017 seeking evidence of inflation heterogeneity across households, and experimented with an alternative technique to calculate cohort specific indexes that controls for the heterogeneity. To control for the presence of heterogeneity in budget shares across households, we clustered an experimental research sample into homogeneous groups defined by the similarity of spending patterns on select CPI items. Using these homogenous clusters as building blocks, we constructed a national all-urban consumer index and a hypothetical population subgroup index for Social Security recipient households, with alternative treatments of owner-occupied housing. To incorporate hypothesized within category heterogeneity for low-income groups, we calculated separate component indexes for the lowest-priced products and services in the CPI sample, and imputed those indexes to the low-income households.

Due to the underlying sample size of the research data set, we do not draw statistical inferences from the exercises performed herein or sanction the accuracy of the index estimates. However, we note the following observations:

- The experimental indexes calculated herein are within the 95% confidence interval of the official CPI-U for the corresponding time period, with two exceptions. Low income quintile and Social Security cohort indexes using the restricted, low-price vector and rental equivalence definition for owner occupied housing were larger.
- The treatment of owner-occupied housing and within-item inflation variation have a large impact. The payment approach for shelter yields a lower inflation estimates during the study period, notably for the Social Security cohort. Use of restricted samples for component indexes (i.e., the low price level quintile) for low income cohorts yields higher overall inflation estimates.
- Stratified clustering of households based upon budget shares may improve the accuracy of superlative index estimation

Statistical agencies are responsible for producing accurate, reliable, timely statistics and should resolve to meet the needs of all index users. In the case of the index needed by the Social Security Administration to index COLAs, if the index is biased too high, there is a heightened risk of hastening the insolvency of the OASDI trust fund absent a commensurate increase in taxes levied on workers to populate the fund. If it is biased too low, it undermines the intent of the law and the purchasing power of recipients, especially the poorer subgroup that depends on the income for livelihood. Due to the sheer number of beneficiaries in the program, a difference of one tenth of one percent in the CPI translates into more than a billion dollars of outlays. Accuracy is paramount.

When COLAs are provided in public transfer payments, such as Social Security, an overarching question is whether the adjustment should compensate the recipients for changes in the overall cost-of-living for the nation as a whole, or should it account for any significant inflation difference experienced by the intended recipients? This is a question for policy makers to resolve. If the latter, the ability of statistical agencies to produce accurate subgroup indexes is hampered by a lack of sufficient household spending data, both with respect to survey sample size and consumption detail.

The BLS is currently redesigning its household consumption survey. BLS plans to add point of purchase questions to the Consumer Expenditure Survey and outlets will then be linked to households. This will enable BLS, potentially, to add population groups to its stratification of items and areas. This could, partially at least, enable BLS to control for within-category heterogeneity currently not accounted for in the CPI-E and CPI-W index estimates. Clustering households into homogeneous groups, and scientifically selecting samples independently for each group, could improve cohort index estimation.

CNSTAT Panel Planning

1 Basic index calculation	BLS References	External References
Basic index methodology	CPI Handbook of Methods https://www.bls.gov/opub/hom/pdf/cpihom.pdf	
Blending traditionally collected data with alternatively collected and sourced data	Friedman, D., Konny, C., Williams, B. (2019). Big Data in the U.S. Consumer Price Index: Experience & Plans. *NBER CRIW Conference Proceedings* https://conference.nber.org/conf_papers/f111229/f111229.pdf	
New products		Broda, C., and Weinstein, D. E. (2010). Product creation and destruction: Evidence and price implications. *American Economic Review*, 100(3), 691-723. Diewert, W.E. and Feenstra, R. C. (2018). Estimating the Benefits of New Products. Discussion Paper 18-04. Vancouver School of Economics, University of British Columbia. Feenstra, R. C. (1994). New product varieties and the measurement of international prices. *The American Economic Review*, 157-177. Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In *The Economics of New Goods* (pp. 207-248). University of Chicago Press.

	Quality change	Erickson, T. and Pakes, A. (2011). An Experimental Component Index for the CPI: From Annual Computer Data to Monthly Data on Other Goods. *American Economic Review* 101, pp. 1707-1738.	Melser, D. and Syed, I. A. (2016). Life Cycle Price Trends and Product Replacement: Implications for the Measurement of Inflation. *Review of Income and Wealth* 62.3, pp. 509-533. Pakes, A. (2003). A Reconsideration of Hedonic Price Indexes with an Application to PC's. *American Economic Review* 93.5, pp. 1578-1596. Nevo, A. (2003). New Products, Quality Changes, and Welfare Measures Computed from Estimated Demand Systems. *Review of Economics and Statistics* 85.2, pp. 266-275.
	New outlets		
2	**Upper-level substitution bias in the CPI-U and CPI-W**	Cage, R., Greenlees, J., and Jackman, P. (2003). Introducing the Chained Consumer Price Index. *Seventh Meeting of the International Working Group on Price Indices*, Paris France. https://www.bls.gov/cpi/additional-resources/chained-cpi-introduction.pdf Shoemaker, O. (2006). Comparison between Chained CPI-U and Regular CPI-U All-US Indexes in the Housing Sector (2000-2004). BLS working paper.	Diewert, W. E. (1976). Exact and Superlative Index Numbers. *Journal of Econometrics*, *4*(2), 115-145. Diewert, W. E. (1978). Superlative Index Numbers and Consistency in Aggregation. *Econometrica* 46.4, pp. 883-900. Hill, Robert J. (2006). Superlative Index Numbers: Not All of Them Are Super. *Journal of Econometrics* 130.1, pp. 25-43.

	https://www.bls.gov/osmr/research-papers/2006/pdf/st060220.pdf	
3 Methodology for estimating price change		
Health insurance	Measuring Price Change in the CPI: Medical Care https://www.bls.gov/cpi/factsheets/medical-care.htm#A2	
Owner occupied housing	CPI Handbook of Methods Measuring Price Change in the CPI for Rent and Rental Equivalence https://www.bls.gov/cpi/factsheets/owners-equivalent-rent-and-rent.pdf Garner, Thesia I. and Verbrugge, Randal(2009). Reconciling User Costs and Rental Equivalence: Evidence from the U.S. Consumer Expenditure Survey. BLS Working Paper 427, 1-46. https://www.bls.gov/cex/research_papers/pdf/reconciling-user-costs-and-rental-equivalence-evidence-from-the-us-consumer-expenditure-survey.pdf	W. Erwin Diewert, Erwin W. and Nakamura, Alice O. (2009). ACCOUNTING FOR HOUSING IN A CPI. Research Department Federal Reserve Bank of Philadelphia WORKING PAPER NO. 09-4 https://pdfs.semanticscholar.org/1163/288c9f32b1182e74c2eb94512abef5d54440.pdf Kudlyak, Marianna (2012). Housing Services Price Inflation. Economic Quarterly. Volume 98 (3), 185-207. https://www.richmondfed.org/~/media/richmondfedorg/publications/research/economic_quarterly/2012/q3/pdf/kudlyak.pdf
Durables		Osborne, M. (2018). Approximating the Cost-of-Living Index for a Storable Good. *American Economic Journal: Microeconomics* 10(2), pp. 286-314.

UNSW Business School
Centre for Applied Economic Research

Measuring Real Consumption and CPI Bias under Lockdown Conditions

W. Erwin Diewert and Kevin J. Fox

7 May 2020

The Challenge

1. Millions of good and services now unavailable.

2. Unprecedented situation – methods haven't been developed for this situation. Advice defaults to standard treatment of non-available products.

3. Consumer expenditure patterns have clearly changed dramatically yet statistical agency practice is to use expenditure weights from a previous period; these weights are likely to be irrelevant during lockdown conditions.

4. This situation risks the public and policy makers losing confidence in key economic statistics.

International Advice to National Statistical Offices

<u>Advice from Eurostat</u> to European Union countries on how to calculate the EU's Harmonized Index of Consumer Prices (HICP):

"The compilation of the HICP in the context of the COVID-19 crisis is guided by the following three principles:

- **Stability of the HICP weights,**
- **Compilation of indices covering the full structure of the European version of the Classification of Individual Consumption According to Purpose (ECOICOP),**
- **Minimizing the number of imputed prices and sub-indices."**

The weights reflect "household consumption expenditure patterns of the previous year."

Advice is effectively to carry on as usual, as if nothing has happened.

International Advice to National Statistical Offices

UNECE advice is similar, but notes:

"In all cases, it is important to apply imputation methods that ensure the index reaches the correct level when again it becomes possible to collect prices and include them in the index."

Hence, it provides an explicit explanation for the carry on as usual methodology; i.e., when things return to "normal", the post lockdown CPI indexes will be comparable to the pre-lockdown CPI index.

IMF advice is consistent with Eurostat and UNECE, but it is more explicit in one respect in that it rules out simple carry forward pricing and endorses inflation adjusted carry forward prices.

Empirical Evidence of Changing Expenditure:
From Credit Cards in Spain

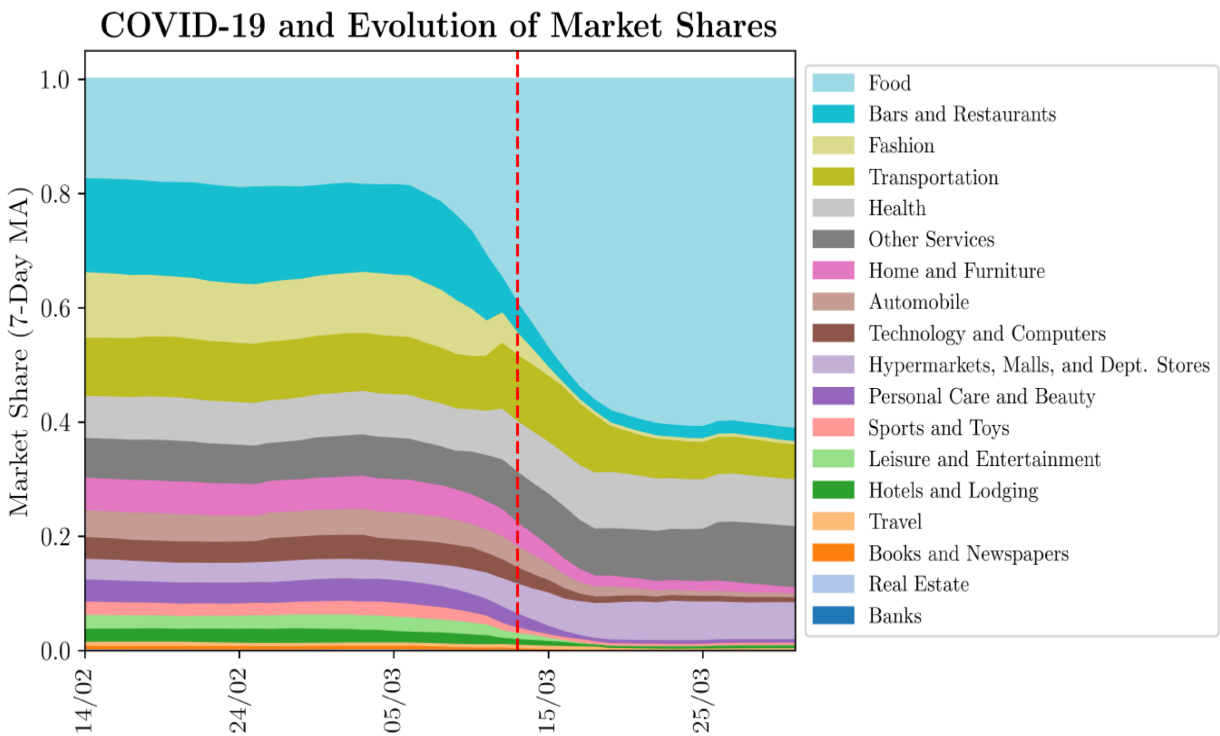
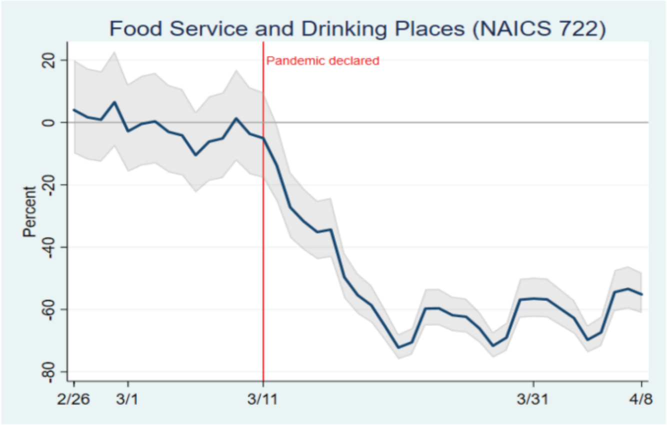


Fig. 7: The evolution of market shares for broad expenditure categories. Categories are stacked top to bottom in order of pre-crisis shares. The red dash indicates the announcement of the lockdown. Shares are expressed as a seven-day moving average.

Carvalho, V.M., J.R. Garcia, S. Hansen, Á. Ortiz, T. Rodrigo, J.V. Rodríguez Mora and J. Ruiz (2020), “Tracking the COVID-19 Crisis with High-Resolution Transaction Data”, Cambridge-INET Working Paper Series No: 2020/16, University of Cambridge.

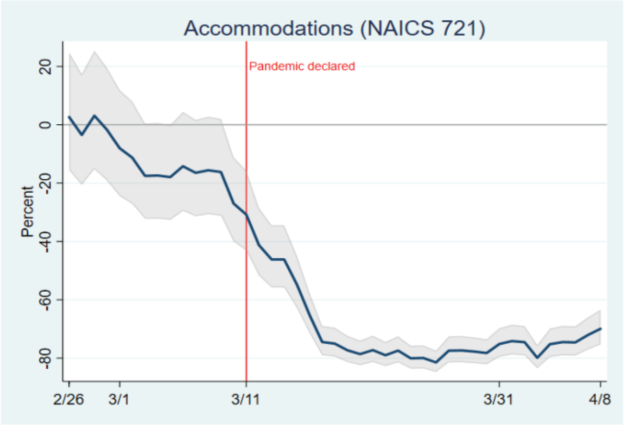
Empirical Evidence of Changing Expenditure: US

Figure 3. Event Study for Restaurants



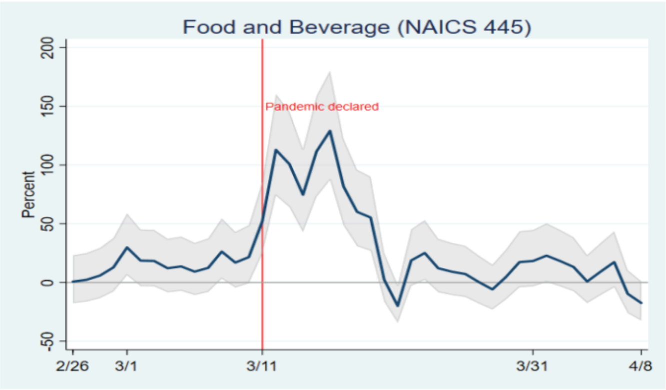
Notes. The estimates shown here have been transformed from log scale to percentages by using the exponential of the point estimate minus one, multiplied by 100. The vertical red line represents March 11, the date on which WHO declared a global pandemic. Deviations away from 0 indicate the change in the sector associated with the timing of the event. The bars represent the 95 percent confidence interval bands around the point estimate.

Figure 4. Event Study for Accommodations



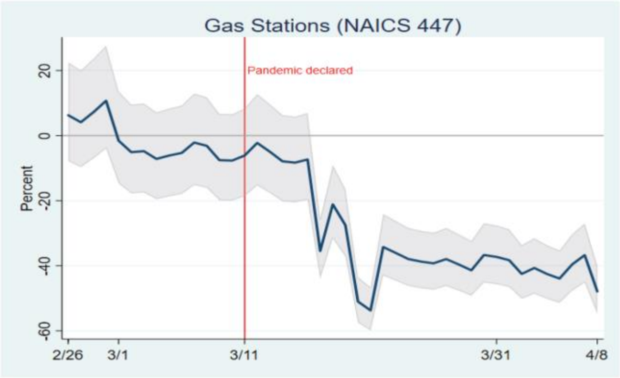
Notes. The estimates shown here have been transformed from log scale to percentages by using the exponential of the point estimate minus one, multiplied by 100. The vertical red line represents March 11, the date on which WHO declared a global pandemic. Deviations away from 0 indicate the change in the sector associated with the timing of the event. The bars represent the 95 percent confidence interval bands around the point estimate.

Figure 5. Event Study for Food and Beverage



Notes. The estimates shown here have been transformed from log scale to percentages by using the exponential of the point estimate minus one, multiplied by 100. The vertical red line represents March 11, the date on which WHO declared a global pandemic. Deviations away from 0 indicate the change in the sector associated with the timing of the event. The bars represent the 95 percent confidence interval bands around the point estimate.

Figure 6. Event Study for Gas Stations



Notes. The estimates shown here have been transformed from log scale to percentages by using the exponential of the point estimate minus one, multiplied by 100. The vertical red line represents March 11, the date on which WHO declared a global pandemic. Deviations away from 0 indicate the change in the sector associated with the timing of the event. The bars represent the 95 percent confidence interval bands around the point estimate.

Dunn, A., K. Hood and A. Driessen (2020), "Measuring the Effects of the COVID-19 Pandemic on Consumer Spending using Card Transaction Data," U.S. Bureau of Economic Analysis Working Paper WP2020-5.

Our Advice

1. **In the short run, collect whatever prices are available and supplement these from scanner data and web scraped prices to make up for missing prices due to changes in price collection methodology. For prices which are still missing, use inflation adjusted carry forward prices, consistent with current advice from the international agencies.**

2. **At the same time, put in motion some method for getting current expenditure weights for the consumption basket. This would require either a continuous consumer expenditure survey or the use of new sources of data. These new sources could include credit card companies and companies that produce household expenditure data from households scanning barcodes of purchased items ("homescan" data).**

3. **Once the new consumer expenditure information becomes available, produce a new analytic CPI. This would be revisable while the new methodology was developed further. This would supplement the existing CPI, which would likely be heavily compromised due to the treatment of missing prices and use of out-of-date expenditure weights.**

Issues Addressed

1. Measurement of real consumption.
2. Measurement of the CPI.
3. Advantages and disadvantages of using various "practical" approaches that NSOs are likely to implement, taking into account different levels of data constraints.
4. Construction of elementary indexes with a lack of matching product prices.
5. Other practical measurement problems facing NSOs in CPI construction under pandemic conditions.

Key Findings

1. Using carry-forward prices (either unadjusted or adjusted for inflation) will lead to:
   - An overestimation of real consumption growth.
   - An underestimation of changes in consumer inflation.

2. Fixed basket indexes, such as the Lowe index used in most countries to construct the CPI, are inadequate when there are dramatic changes in consumer expenditure.

3. Need new expenditure weights for the lockdown period. Once the lockdown ends, price change comparisons should be made with the pre-lockdown period using pre-lockdown weights.

4. A revisable CPI is needed.

Real Consumption and CPI Biases

The problem is how to think about missing prices.

From the economic approach to index number theory, a price index is a ratio of expenditure functions with changing prices but fixed utility.

That is, consumers must have preferences over the same set of products in both periods being compared.

In the context of new goods, Hicks (1940) proposed reservation prices: the prices that drove demand to zero in the period before they are observed.

We adapt this to the disappearing goods context. This approach allows us to identify biases from the carry-forward prices approach.

Real Consumption and CPI Biases

Why reservation prices?

A lockdown is like being sent to jail – deprived of products and confined to a particular place.

People are prepared to pay a lot of money to avoid jail, indicating that lockdowns decrease welfare.

To capture such declines in welfare, need the prices for unavailable products to be much greater than the corresponding prices in the previous period.

Real Consumption and CPI Biases

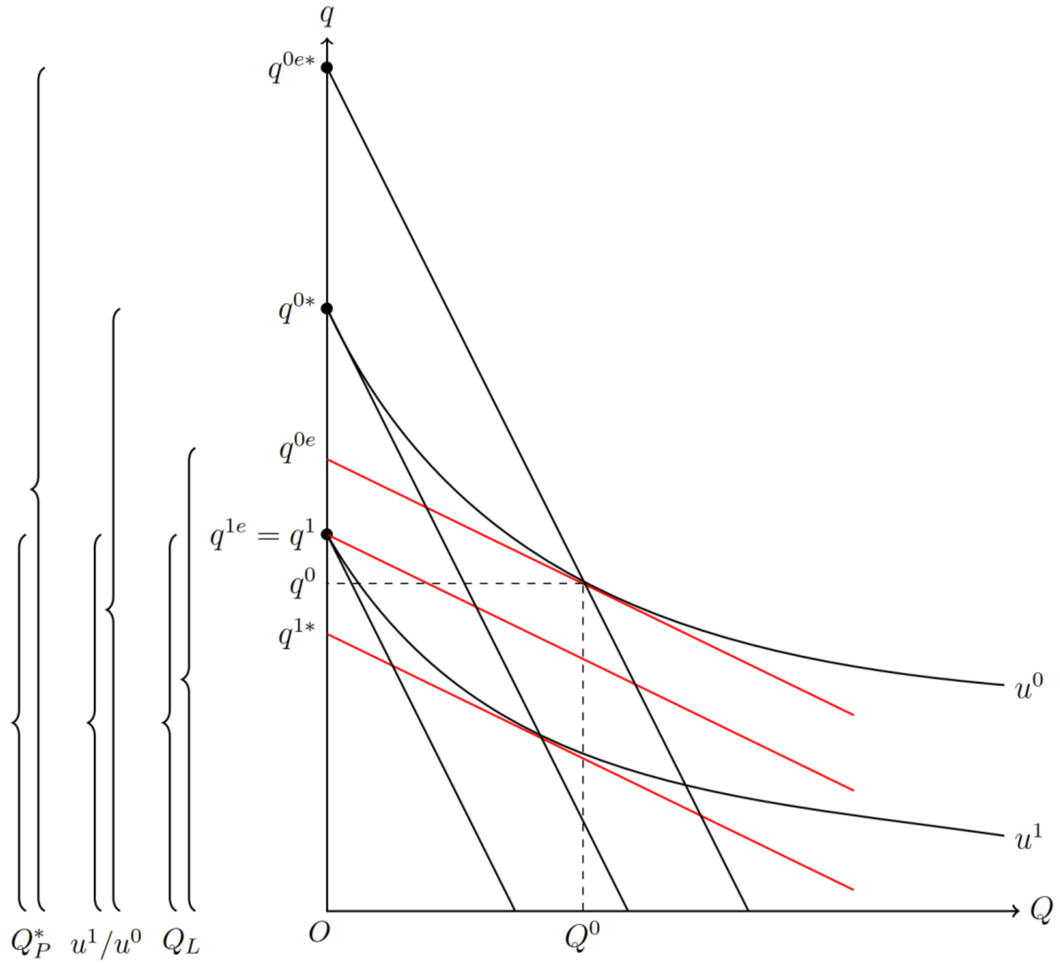
1. A theoretical quantity index is the ratio of two expenditure functions with prices held constant and utility allowed to change. Hence, it is a measure of welfare change.

2. As (in ratio terms) we want value change = price change x quantity change, then for a fall in welfare (i.e. the quantity index) we need an increase the corresponding price index (as Q = V/P).

3. Inflation adjusted carry forward prices will not give this increase.

4. Hence, need reservation prices for the lockdown period, which will be much higher than carry forward prices.

Real Consumption and CPI Biases



q available in both periods, Q available only in period 0

Some Other Practical Problems Considered

No NSO employee price collection:

- **Use web scraping and other non-traditional methods, but need to make sure that only collect prices for products that were actually consumed by any household.**

Stockpiling Problem:

- **Look at what to do about goods that enter and exit the consumption basket as supply-chain issues/lockdown rules change. That is, products drift in and out of scope. Response may depend on information available.**
- **CPI is constructed (mainly) on an acquisitions approach rather than a consumption approach – should it be changed to reflect consumption not taking place in the period of acquisition? NSOs will likely stick with acquisition approach. But the assumption of a constant basket equal to a pre-lockdown basket for all post lockdown periods is going to be a rather poor assumption.**

Some Other Practical Problems Considered

No NSO employee price collection:

- **Use web scraping and other non-traditional methods, but need to make sure that only collect prices for products that were actually consumed by any household.**

Stockpiling Problem:

- **Look at what to do about goods that enter and exit the consumption basket as supply-chain issues/lockdown rules change. That is, products drift in and out of scope. Response may depend on information available.**
- **CPI is constructed (mainly) on an acquisitions approach rather than a consumption approach – should it be changed to reflect consumption not taking place in the period of acquisition? NSOs will likely stick with acquisition approach. But the assumption of a constant basket equal to a pre-lockdown basket for all post lockdown periods is going to be a rather poor assumption.**

Some Other Practical Problems Considered

Free Dwelling Rent:

- **If there is a policy of rent forgiveness, then these nonpaying tenants are getting free rent.**

- **Propose a methodology for getting reservation prices (Appendix C).**

- **It will typically be difficult to estimate these reservation prices, but either carry forward prices or inflation adjusted carry forward prices can be taken as approximate reservation prices.**

- **In this case, show how changes in price indexes will be dependent on depreciation of properties (i.e. declines in quality).**

Conclusions

NSOs should:

- **Collect whatever prices are available, including from non-traditional sources. For missing prices, use inflation adjusted carry forward prices. (We prefer reservation prices, but acknowledge that NSOs will not be able to calculate these in a timely fashion.)**

- **Urgently start a program to obtain current expenditure weights for the consumption basket.**

- **Produce a revisable CPI as an analytical series that can be updated as new methodology is developed and new data sources are exploited.**

- **Explain to the public and policy makers that the usual measures of real consumption and inflation are compromised due to the pandemic…and seek a big increase in budget to address this!**

CNSTAT CPI Panel: Potential Workshop Topics and Participants

This list includes more topics and presenters than can be done in our open meetings. We envision perhaps 5 sessions with 2 presenters and discussion in each. We also have the BLS presentations (listed in the old April 3 meeting agenda) which could be integrated into the topic-organized sessions here.

Core Topics

Big Data I: Use of alternative (nonsurvey) data sources in official statistics (possible panel leaders/discussants: Lynch, Sichel).

- **Shapiro/Johnson**, University of Michigan: health care, use of organic data, housing. Using transactions data directly from companies. Re-engineering Key National Economic Indicators, https://www.nber.org/papers/w26116 (agreed in principle to participate).
- **Van Rompaey**, Statistics Canada: big data. Strong reputation.
- **Mehrhoff,** Deutsche Bundesbank: big data in official statistics.

Big Data II: Use of high frequency data in price measurement (possible leaders/discussants: Cavallo, De Haan) BLS is looking for practical guidance on how to shift methodology to take advantage of alternative data sources, and to blend those sources with traditional survey data (often at the basic item/location cell level).

- **Nakamura,** UC Berkeley; new outlets, using high frequency data, studying outlet effects, and in macro policy. Has agreed in principle to participate in workshop.
- **Van Loon**, Statistics Belgium: scanner data from supermarkets in the calculation of the CPI since 2015. The applied method is a version of the so-called "dynamic method".
- **Nevo**, Northwestern University: demand estimation, substitution patterns, scanner data.
- **Chernozhukov**, MIT: econometrics, nonparametric methods, hedonic price indexes, big data.
- **Diewert and Fox's** paper discussing the need for high frequency data (COVID economy)?
- **Hitsch, Hortacsu** and **Lin** at UChicago. Prices and Promotions in US Retail Markets: Evidence from Big Data.

Price indexes for population (income) subgroups (possible leaders/discussants: Moulton, Sichel).

- **Jaravel**, London School of Economics and Political Science: Also has an interesting paper that uses scanner data to document the differences in price changes across different parts of the income distribution. Could be in session with Rob Cage.
- **Cage, BLS,** about their project to estimate demographic subgroup price indexes (by income group)

https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.22/2018/BLS_4.pdf.

- **Kaplan**, University of Chicago: price dispersion (people paying different prices for the same item), inequality, macroeconomics.

Housing prices (possible leaders/discussants: Molloy).

- **Vavra**, University of Chicago Booth: price distributions, macroeconomic implications of inflation, housing.
- **Moulton**, University of North Carolina, housing services. Valuing Housing Services in the Era of Big Data: A User Cost Approach Leveraging Zillow Microdata: Marina Gindelsky, Jeremy Moulton, Scott A. Wentland.
- **Kudlyak,** Federal Reserve Bank of San Francisco: Housing prices.
- **Hill**, University of Graz in Austria: owner occupied housing.

Medical care/health insurance session (possible leaders/discussants: Aizcorbe, Sheiner).

- **Ho**, Princeton University: health insurance, industrial organization [suggested by BLS].
- **Krsinich,** Statistics New Zealand: time dummy hedonics, non-survey data, housing hedonics.
- **Cutler**, Harvard University.

International Perspectives (possible leaders/discussants: De Haan, Diewert).

- **Schreyer**, OECD. Overview of what are other countries doing in the key topic areas identified in the panel's statement of task.
- **Oulton**, London School of Economics: productivity, official statistics.
- **Watanabe**, University of Tokyo.
- **Payne**, ONS.

Additional "time permitting" Topics

Data Users perspective (possible leaders/discussants: Rosner-Warburton). BLS just completed a data user survey, the results of which they can share.

- **Lebow, Rudd**, Federal Reserve. Use of CPI in monetary policy. Priorities for improving price measurement.
- **Haver,** Haver Analytics: BLS data user.
- **Hatzius,** Goldman-Sachs: economic forecaster.

Measuring prices in the digital economy/the price of information (possible leaders/discussants: Reinsdorf).

- **Byrne**, Federal Reserve (digital services, information, new outlets, quality/tech change).
- **Corrado**, Conference board. Prices and productivity in the knowledge economy.
- **Timmer**, University of Groningen: internet prices, online prices, productivity.
- **Fox**. (k.fox@unsw.edu.au). The Digital Economy, New Products and Consumer Welfare.
- **Weinstein** Columbia: ecommerce, big data sources, trade and prices, new products).

<u>New Products</u> (possible leader/discussant: Diewert).

- **Feenstra**, UC Davis (new goods, housing, international comparisons, etc.).
- **Klenow**, Stanford University; new products, sticky prices, macroeconomic implications of inflation.
- **Bils,** University of Rochester: new products, quality change.

<u>Other Durable goods (beyond housing)</u>. BLS is interested in pursuing a discussion of the flow of service approach, including on related categories such as vehicle leasing, that aren't durables per se.

- **Jorgenson**, Harvard University. Service flow approach.

Other Experts to Involve

Groshen, Cornell (former BLS Commissioner); **Lin**, Chicago-JD.com Silicon Valley Research Center (ad measurement, pricing, and consumer demand modeling; innovative and scalable advertising measurement products and intelligent pricing products drawing methods from marketing science, economics, and machine learning); **Verbrugge, Bradley, Greenlees**, and **Armknecht** are all former BLS staff; **Chevalier**, Yale University.