Evaluating the Consumer Price Index Using Nielsen's Scanner Data October 2013

Jenny FitzGerald and Owen Shoemaker
U.S. Bureau of Labor Statistics
2 Massachusetts Avenue, NE, Room 3655 Washington, D.C. 20212 U.S.A.
fitzgerald.jenny@bls.gov shoemaker_o@bls.gov

Abstract
The Consumer Price Index (CPI) estimates the change in prices over time of the goods and services U.S. consumers buy for day-to-day living based on price quotes selected from probability samples. The goal of our research is to determine how accurately the current CPI sample and indexes reflect reality. For our research, we calculate superlative price indexes from Nielsen's scanner data, which we assume to represent the real universe of commodities data. We then compare the Consumer Price Indexes for self-representing areas at the expenditure class level to our superlative price indexes derived from Nielsen's scanner data.

Key Words: Multistage Sample Design, Superlative Price Index, Wilcoxon Signed-Rank Test

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

1. Introduction

The Consumer Price Index (CPI) estimates the change in prices over time of the goods and services U.S. consumers buy for day-to-day living based on price quotes selected from probability samples. Since the CPI is based on a sample of commodities and services, there will always be questions as to how accurately the index reflects reality. In the past, there was no clear-cut method for evaluating the CPI's accuracy. However, the current availability of A. C. Nielsen scanner data provides us with an opportunity to evaluate the accuracy of some of the CPI indexes if we assume that the scanner data represent the real universe of commodities data.

In January 2012, the U.S. Bureau of Labor Statistics (BLS) purchased scanner data from the Nielsen Company summarizing the quantity and dollar amount of merchandise sold by Universal Product Code (UPC) and geographic area from September 2005 to September 2010. This paper attempts to evaluate how accurately the CPI reflects reality using the Nielsen scanner data for Expenditure Class FN (Juices and Non-alcoholic Drinks) and its lower level item strata. First, a brief overview of the CPI sample design and index estimation process is provided. Second, the study's methodology is described. Then, indexes are calculated from the purchased scanner data, using a "superlative" Tornqvist index estimator. Finally, the CPI's indexes are compared to the scanner data indexes at the Expenditure Class (EC) level and below.

2. Sampling in the CPI for Commodities and Services (C&S)

The CPI is calculated from a sample of price quotes, which are the ultimate outcome of several interrelated probability samples.

First, the BLS selects a sample of geographic areas, which are the primary sampling units (PSUs) for the CPI (Bureau of Labor Statistics, 2008). The BLS updates its CPI area sample once every ten years. To select its area sample, the BLS divides the entire U.S. into PSUs using the Office of Management and Budget's (OMB) definition of metropolitan statistical areas (MSAs). The BLS then classifies each PSU by its size. A PSU with a population greater than 1.5 million is a self-representing PSU and is given a class size of A. A PSU with a population less than 1.5 million is a non-self-representing PSU. A non-self-representing PSU can be a metropolitan area (with a class size of B) or a non-metropolitan area (with a class size of C).

The second classification variable for PSUs is Census region. After each PSU is mapped to its Census region and given a class-size, the BLS stratifies the PSUs in each region-class size into strata of similar PSUs. Self-representing PSUs are placed in a stratum by themselves; non-self-representing PSUs are stratified based on geographic variables correlated with price change and/or expenditure level. A program then selects one PSU per stratum using controlled selection to insure that the selected PSUs are well-distributed across states and to maximize the number of old PSUs selected in the new area sample. Currently, there are 87 PSUs that make up the CPI's 38 index areas.

Within each sampled PSU, the BLS selects a sample of outlets where consumers shop using the data collected via the Telephone Point-of-Purchase Survey (TPOPS). TPOPS (which is conducted by the U.S. Census Bureau for the BLS) uses random digit dialling to select a random sample of households. Eligible respondents are asked to provide information about where they bought items and how much they spent during a given recall period for a select group of items (Marsh, 2006). The reported outlets form the frame of outlets that the BLS uses to select its sample for the CPI. The BLS selects its sample of outlets from the frame independently for each PSU, replicate[1], and TPOPS category using a systematic probability proportional to size (PPS) sample design, where each outlet's measure of size (MOS) is its reported expenditure in the TPOPS category.

The outlet sample is then merged to an independent sample of entry level items (ELIs) that consumers buy. Specifically, the BLS selects a systematic PPS sample of ELIs for each PSU and replicate combination from the expenditure data collected by the Consumer Expenditure (CE) survey, which is aggregated by item stratum and region. An ELI's MOS is its expenditure total for the region compared to the region's total expenditure value for the item stratum. The CPI outlet sample and ELI sample is updated each year for 25 percent of the item strata in each PSU.

Finally, BLS field economists visit the sampled outlets and select individual items for each sampled ELI to be priced each month (or every other month) through a multistage

[1]Each geographic area of the CPI is made of two or more independent samples of items and outlets, called a replicate. A replicate is the basis of the CPI's variance estimates. Independent index estimates are calculated from the replicate samples, while the index produced from the full set of observed prices is called the full sample index estimate. CPI variance estimates are primarily computed using a stratified random groups (SRG) method.

probability sampling technique known as disaggregation. The single selection of a unique item is referred to as a price quote (Fuxa, 2010).

3. Index Estimation

Each month, the BLS calculates price relatives for all monthly and on-cycle bi-monthly elementary indexes for the CPI. An elementary index is an item stratum and index area combination. In the CPI, there are 211 item strata and 38 index areas. Thus, the CPI consists of 8,018 elementary indexes (211×38).

Most elementary indexes use an expenditure-share-weighted geometric average $PRX^{G}_{t,t-1}$ for price relative calculation; other elementary indexes use the Laspeyres formula average $PRX^{L}_{t,t-1}$ (Bureau of Labor Statistics, 2008). The formulas for $PRX^{G}_{t,t-1}$ and $PRX^{L}_{t,t-1}$ are as follows for each index area *a* and item stratum *i* combination:

$$PRX^{G}_{t,t-1} = \prod_{j \in a,i} \left(\frac{P_{j,t}}{P_{j,t-1}} \right)^{\frac{W_{j,POPS}}{\sum_{k \in a,i} W_{k,POPS}}} \quad (1)$$

$$PRX^{L}_{t,t-1} = \frac{\sum_{j \in a,i} \left(\frac{W_{j,POPS}}{P_{j,POPS}} \right) P_{j,t}}{\sum_{j \in a,i} \left(\frac{W_{j,POPS}}{P_{j,POPS}} \right) P_{j,t-1}} \quad (2)$$

Where:

$P_{j,t}$ = the price of the *j*th observed item in month *t* for area - item combination *a*, *i*;

$P_{j,t-1}$ = the price of the *j*th observed item in month $t-1$ for area - item combination *a*, *i*;

$P_{j,POPS}$ = item *j*'s price in the sampling period when POPS was conducted; and

$W_{j,POPS}$ = item *j*'s weight in POPS.

An elementary index value for area *a* and item stratum *i* is calculated by multiplying the previous month's index ($IX_{a,i,t-1}$) by the price relative for area *a* and item stratum *i* in month *t* ($PRX_{a,i,t}$):

$$IX_{a,i,t} = IX_{a,i,t-1} \times PRX_{a.i,t} \quad (3)$$

In the base month (where $t = 0$), the index for area *a* and item stratum *i* is set equal to 100.

$$IX_{a,i,t=0} = 100 \quad (4)$$

The CPI item structure has four levels of classification. That is, the CPI's 211 item strata indexes feed into 70 expenditure classes (ECs); the 70 ECs make up eight major groups;

and the eight major groups make up the entire CPI. To calculate the aggregated indexes at the EC level and above, elementary indexes are multiplied by an aggregation weight derived from tabulated CE data; the product is called a cost weight ($CW_{a,i,t}$). These cost weights are then aggregated to calculate the indexes for the three levels above the elementary index level. For example, equation five gives the formula to calculate an index for an EC for area *a* at time *t*:

$$IX_{a,EC,t,t-1} = \frac{\sum_{i \in a,EC} CW_{a,i,t}}{\sum_{i \in a,EC} CW_{a,i,t-1}} \quad (5)$$

Where:

$IX_{a,EC,t,t-1}$ = Index for area *a* for expenditure class *EC* at time *t*;

$CW_{a,i,t}$ = Cost weight *CW* for area *a* for item stratum *i* at time *t*; and

$CW_{a,i,t-1}$ = Cost weight *CW* for area *a* for item stratum *i* at time $t-1$.

4. Nielsen Scanner Data Project

The objective of this paper is: to evaluate how well the CPI indexes reflect reality for Expenditure Class (EC) FN (Juices and Non-alcoholic Drinks) and its lower level item strata. To accomplish this goal, parallel indexes were calculated from the Nielsen scanner data from October 2005 to September 2010 at the national level and at the city level for the CPI's self-representing areas. As mentioned previously, we assume the Nielsen scanner data is the best available representation of the real universe of commodities data for our research. However, the scanner data does have its deficiencies. Specifically, the scanner data excludes: drug stores with less than one million dollar in sales; grocery stores with less than two million dollars in sales; and currently a major retailer. For items from EC FN, grocery stores and mass supercenters account for 64 percent to 78 percent of the total market (Chanil 2012). Thus, the Nielsen scanner data represent at a minimum 64 percent of the total sales of items from EC FN (given that the scanner data also includes the sales figures from drug stores).

Before calculating the indexes from the Nielsen scanner data, BLS economists had to map each UPC from the scanner data to one and only one ELI from the CPI mapping structure. After the economists completed their concordance file of UPCs to ELIs, an average price for each UPC and market combination was calculated for each four week time period. Finally, a superlative Tornqvist index estimator was used to calculate price indexes from the Nielsen scanner data.

4.1 CPI Mapping Structure

The first step in calculating indexes to compare to the CPI's indexes was to create a concordance file mapping the UPC codes from Nielsen's scanner data to the ELIs from the CPI mapping structure. BLS Commodities Analysts (CAs) mapped each UPC from the Nielsen scanner data to one and only one ELI from the CPI mapping structure. In total, the Nielsen scanner data includes 1,463,373 unique UPCs.

The CAs mapped about 80 percent of the UPCs systematically without much effort to an ELI based on their product module description. The other 20 percent of UPCs had to be individually mapped to an ELI since their product module description did not match "nicely" with an ELI definition.

For UPCS that were mapped easily to an ELI, consider the UPCs that fell under the product module "Soft Drinks – Carbonated" from Nielsen. All of the UPCs under that product module were systematically mapped to ELI FN011 from the CPI for "Carbonated Drinks." For the UPCs that were more difficult to map, consider the UPCs that fell under the product module "Household Specialty Appliances." Some of those UPCs were mapped to ELI HM012 for "Power Tools;" others were mapped to ELI HM021 for "Powered Lawn and Garden Equipment and Other Outdoor Items." For the UPCs under the product module "Household Specialty Appliances," CAs reviewed descriptions of the items the UPCs represented. Then based on the item descriptions, the CAs mapped the UPCs to the appropriate ELIs from the CPI.

The UPC to ELI mappings were saved in a concordance file and then uploaded into an internal BLS database.

4.2 Average Price by Market and UPC

The second step in calculating indexes to compare to the CPI's indexes was to obtain prices from the Nielsen scanner data. The scanner data contain the variables: "Sales Dollars" and "Sales Units." The variable "Sales Dollars" gives the total sales of a UPC in U.S. dollars by market[2] for a one-week or four-week time period. "Sales Units" gives the total number of units sold of a UPC by market for a one-week or four week period. The four-week data run from September 2005 to September 2008, while the one-week data summarize scanner sales from September 2008 to September 2010. To get monthly price indexes from the Nielsen scanner data, the Nielsen's weekly data had to be aggregated by four-week intervals. Once the weekly data were summarized, average prices could then be calculated for a UPC and market combination for each four-week time period *t* from September 2005 to September 2010:

$$\text{AveragePrice}_{\text{UPC, Market, t}} = \frac{\text{Sales_Dollars}_{\text{UPC, Market, t}}}{\text{Sales_Units}_{\text{UPC, Market, t}}} \quad (6)$$

4.3 Nielsen Index Estimator

After average prices were calculated by UPC and market, a superlative Tornqvist index estimator was used to calculate the price relatives from the scanner data for the three item strata beneath expenditure class FN -- "Juices and Non-alcoholic Drinks":

[2] Market is a geographical area and store type. The store type might be: a drugstore (DRUG) with sales greater than one million dollars, mass merchandise store (MM) with sales greater than two million dollars, or the combination of food-drug-and mass merchandising stores (FDM).

$$\mathrm{PRX}_{t,t\text{-}1} = \prod_{UPC \in i, m \in a} \left(\frac{\mathrm{p}^{t}_{\mathrm{UPC,m}}}{\mathrm{p}^{t-1}_{\mathrm{UPC,m}}} \right)^{\left(\frac{1}{2} r^{t\text{-}1}_{\mathrm{UPC,m}} + \frac{1}{2} r^{t}_{\mathrm{UPC,m}} \right)} \quad (7)$$

$i =$ Item Stratum

$UPC =$ Universal Product Code

$a =$ Index Area

$m =$ Nielsen Market

$t =$ Current four-week time period

t-1 $=$ Previous four-week time period

$p^{t}_{UPC,m} =$ Average price for a UPC in market m at time t

$p^{t-1}_{UPC,m} =$ Average price for a UPC in market m at time $t-1$

$r^{t\text{-}1}_{UPC,m} =$ Expenditure of UPC in item stratum i from market m in index area a as a percentage of all UPCs in item stratum i in index area a in time period $t-1$

$r^{t}_{UPC,m} =$ Expenditure of UPC in item stratum i from market m in index area a as a percentage of all UPCs in item stratum i in index area a in current time period t

No item replacements or substitutions were made. That is, a UPC was included in the price relative in month *t*, if the UPC had an average price in the previous month *(t–1)*.

4.4 Obtaining Monthly Price Relatives from Nielsen's Scanner Data

The BLS publishes the CPI monthly. In order to evaluate the CPI using the Nielsen scanner data, exactly one price relative had to be calculated for each month from the Nielsen scanner data. As mentioned previously, the Nielsen scanner data summarize sales data by four-week intervals (and not by month). Consequently, five months of the reference period ended up with two price relatives: December 2005, December 2006, December 2007, November 2008, and October 2009. To deal with this issue, the price relatives for these five months were set equal to the product of their two four-week price relatives.

4.5 Significance Tests

Because the CPI is a chained index, the differences between the CPI indexes and the scanner data indexes are compounded over time. To see if the differences between the two sets of price change data are significant, Wilcoxon Signed-rank tests and paired t-tests were conducted on the two sets of twelve month percent changes[3] derived from the CPI and Nielsen index estimates. Explicitly, a twelve-month percentage change is calculated as follows:

[3] Significance tests were originally conducted on the CPI and Nielsen monthly price relatives. All of the significance tests, however, indicated that no significant differences exist between the two sets of monthly price relatives at both the city level and the national level. This was alarming when one pair of indexes clearly diverged. Seemingly, the monthly price changes are too small to make any conclusion. Consequently, the significance tests were conducted on the twelve-month percentage changes instead.

$$PC_{12m} = \left(\frac{IX_t}{IX_{t-12}} - 1 \right) \times 100\% \quad (8)$$

$PC_{12m} =$ Twelve month percent change

$\mathrm{IX}_t =$ Index in month t

$IX_{t-12} =$ Index 12 months earlier in month $t-12$

The Wilcoxon Signed-rank test is a nonparametric alternative to the paired student's t-test. To be precise, data must be normally distributed for a paired t-test but not for the Wilcoxon Signed-rank test. From Hollander and Wolfe (1999), the sign-rank test is designed for analyses in which the primary interest is centered on the median of a population. It was hypothesized that there would be no difference between the CPI's twelve-month percentage changes and the twelve-month percentage changes calculated from the Nielsen scanner data. Thus, the median of the differences was expected to be zero.

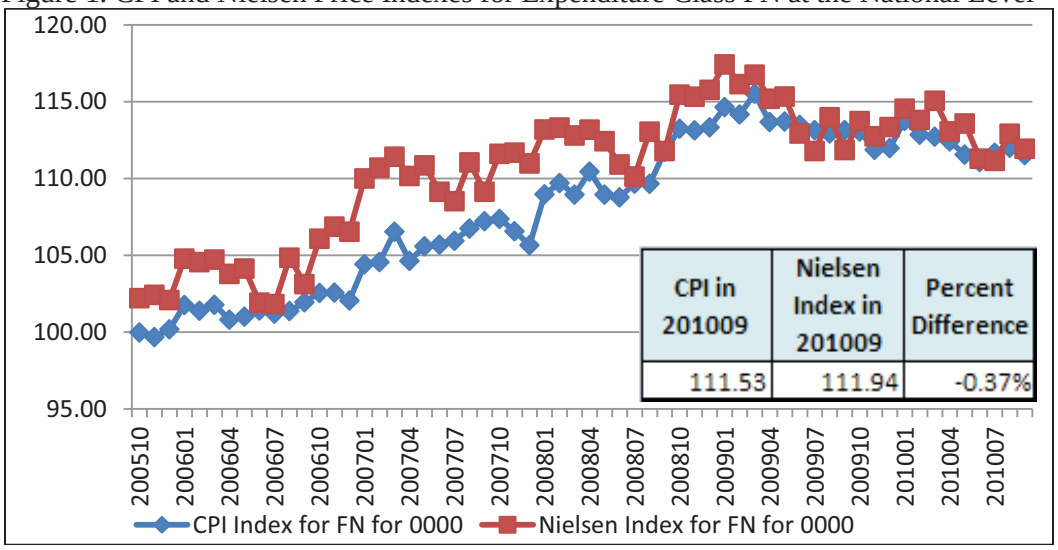
5. Findings

Figure 1 below compares the CPI index for EC FN (Juices and Non-alcoholic Drinks) to the index estimate calculated from the Nielsen scanner data for EC FN. The base period is 200509 for both the CPI and Nielsen indexes. If we assume that the index estimate calculated from the Nielsen scanner data is the best estimate of reality, the CPI index estimate slightly underestimated the actual price increases of "Juices and Non-alcoholic Drinks" from 200510 - 200906. The greatest difference between the CPI and Nielsen indexes occurs in 200702 when the CPI index (104.56) is about 5.6 percent less than the Nielsen index (110.71). From 200907 – 201009, the CPI trend line appears to run right through the middle of the Nielsen index estimates. At the end of five years, the CPI index for expenditure class FN differs by less than one percent from the Nielsen index. Table 1 below gives the results of the paired t-test and Wilcoxon Signed-rank test that were conducted on the two sets of twelve-month percentage changes from 200610 – 201009 derived from the CPI and Nielsen index estimates. The paired t-test indicates that no significant difference exists between the CPI and Nielsen twelve-month percentage changes for expenditure class FN, while the Wilcoxon Signed-rank test does show a significant difference at the $\alpha = 0.05$ level.

Table 1. Paired T-Test and Wilcoxon Signed-Rank Tests at the National Level for EC FN

CPI Area	Expenditure Class	Ave Diff b/w CPI and Nielsen 12 Month Pct Changes	N	DF	T-Value	P-Value	Wilcoxon Statistic	P-Value
All U.S. (0000)	FN	0.3748	48	47	1.52	0.1350	749	0.0493

Figure 1. CPI and Nielsen Price Indexes for Expenditure Class FN at the National Level



From the graph in Figure 1, the CPI index for expenditure class FN at the national level performs well in measuring the actual price change of all "Juices and Non-alcoholic Drinks" if we assume that the Tornqvist index estimate calculated using the Nielsen scanner data is the best estimate of reality. At lower levels, however, the CPI index may not do as well. To test this theory, the item strata beneath expenditure class FN were investigated. Specifically, expenditure class FN is made up of three item strata: FN01 (Carbonated Drinks), FN02 (Frozen Noncarbonated Juices and Drinks), and FN03 (Non-frozen Noncarbonated Juices and Drinks). FN01 accounts for about 42 percent of EC FN; FN02 makes up about two percent; and FN03 contributes about 56 percent. Figures 2 - 4 below compare the CPI index estimates to the Nielsen index estimates for item strata FN01, FN02, and FN03, respectively, at the national level. At the end of five years, the CPI's index estimates for FN01 and FN02 are about four percent greater than the Nielsen index estimates; the CPI's estimate for FN03, on the other hand, is about five percent less than the Nielsen index estimate.

Figure 2. CPI and Nielsen Price Indexes for Item Stratum FN01 at the National Level

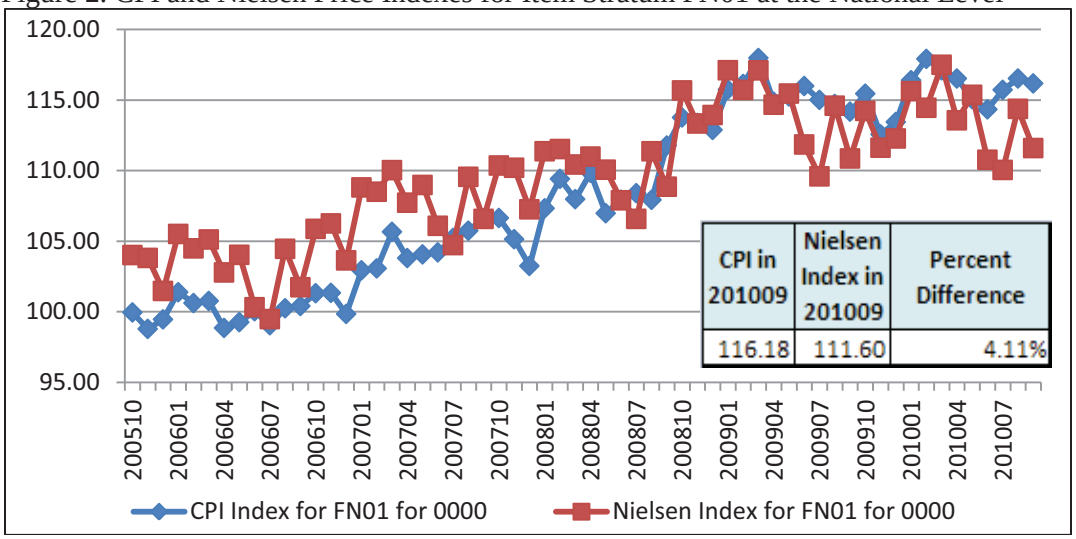


Figure 3. CPI and Nielsen Price Indexes for Item Stratum FN02 at the National Level

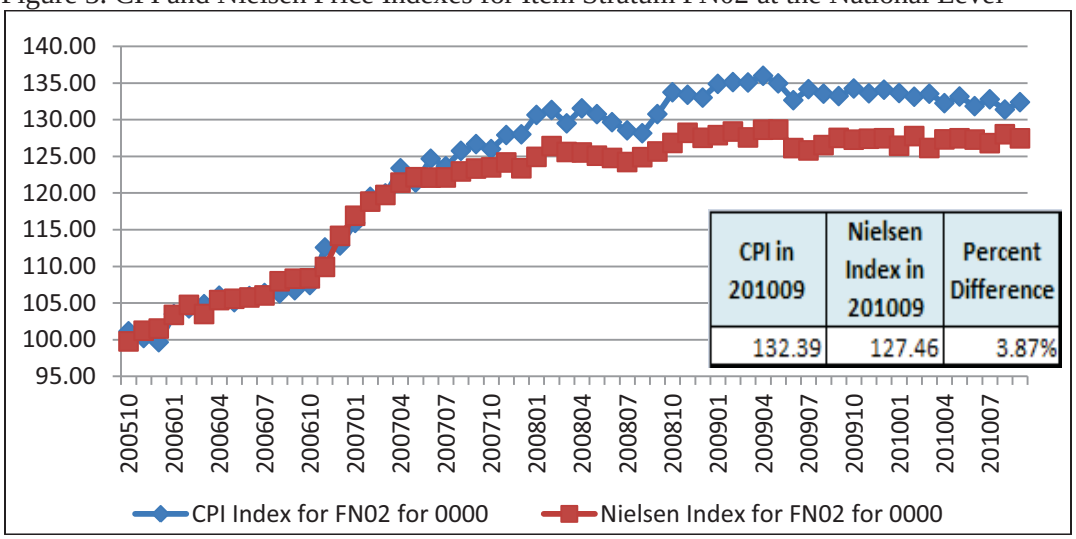


Figure 4. CPI and Nielsen Price Indexes for Item Stratum FN03 at the National Level

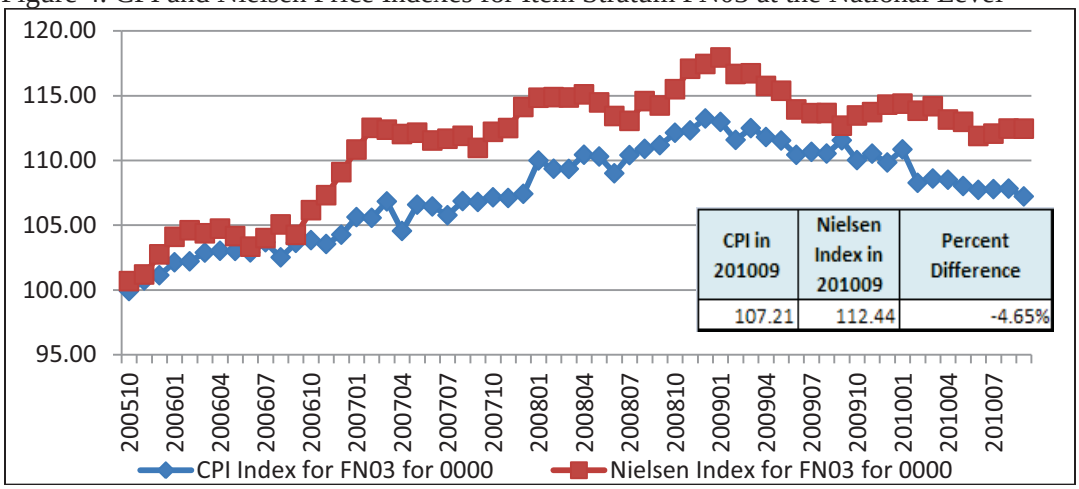


Table 2 below gives the results of the paired t-test and Wilcoxon Signed-rank test that were applied to the average twelve-month percentage changes calculated from the CPI and Nielsen index estimates for item strata FN01, FN02, and FN03 at the national level. Both the paired t-tests and the sign-rank tests indicate that a significant difference exists between the CPI and Nielsen twelve-month percentage changes for all three item strata at the $\alpha = 0.05$ level. This finding is not surprising given the smaller sample sizes of the three item strata beneath EC FN.

Table 2. Paired T-Tests and Wilcoxon Signed-Rank Tests at the National Level for FN01, FN02, and FN03

CPI Area	Item Stratum	Ave Diff b/w CPI and Nielsen 12 Month Pct Changes	N	DF	T-Value	P-Value	Wilcoxon Statistic	P-Value

CPI Area	Item Stratum	Ave Diff b/w CPI and Nielsen 12 Month Pct Changes	N	DF	T-Value	P-Value	Wilcoxon Statistic	P-Value
All U.S. (0000)	FN01	1.3082	48	47	5.77	<0.0001	1036	<0.0001
All U.S. (0000)	FN02	1.2773	48	47	4.25	0.0001	942	0.0001
All U.S. (0000)	FN03	-0.7273	48	47	-2.28	0.0273	420	0.0424

If we assume that the Nielsen index estimates are the best estimates of price change for the three item strata, the CPI's twelve month percent changes for FN01 differ the most from reality. To see what cities contribute to this difference, price relatives were calculated from the Nielsen scanner data for item stratum FN01 at the city level for all of the CPI's self-representing cities. The CPI's self-representing cities together account for about 50 percent of the total weight of FN01; the other 50 percent comes from the CPI's non-self-representing cities. Table 3 gives the results of the paired t-tests and Wilcoxon Signed-rank tests that were ran on the CPI and Nielsen twelve-month percentage changes for item stratum FN01 at the city level.

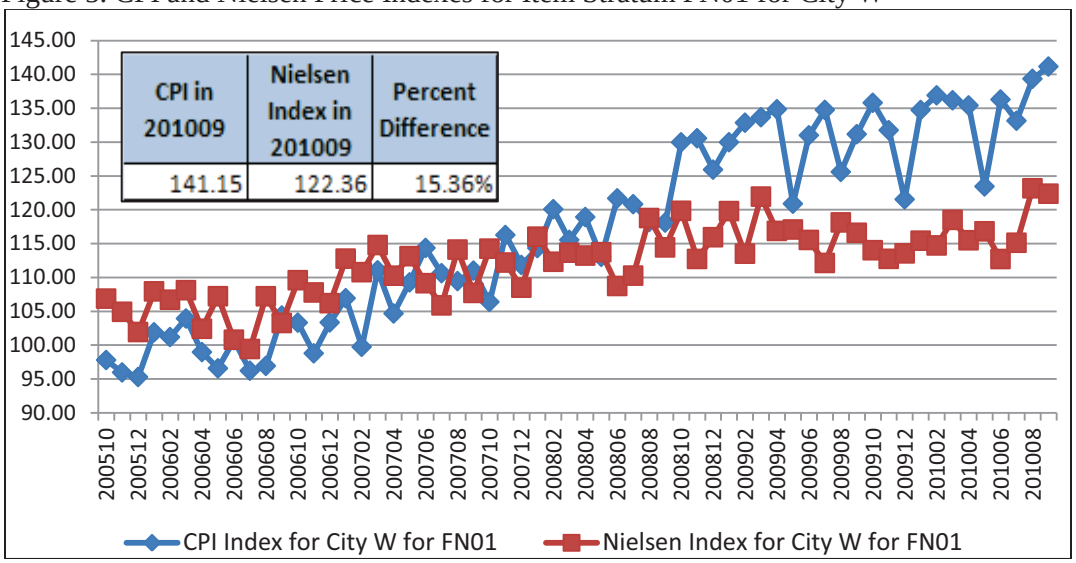
Table 3. Paired T-Test and Wilcoxon Signed-Rank Test for FN01 for Each Self-Representing City

PSU Description	Ave Diff b/w CPI and Nielsen 12M Pct Changes	N	DF	T-Value	P-Value	Wilcoxon Statistic	P-Value
City A	1.3226	48	47	2.69	0.0098	873	0.0017
City B	-0.4439	48	47	-0.33	0.7420	565	0.4068
City C	0.8391	48	47	0.78	0.4384	616	0.3870
City D	-0.0248	48	47	-0.03	0.9970	528	0.2691
City E	2.2429	48	47	2.84	0.0066	840	0.0049
City F	0.9924	48	47	2.07	0.0439	797	0.0160
City G	0.2714	48	47	0.29	0.7739	655	0.2460
City I	2.9229	48	47	1.68	0.1002	742	0.0571
City J	0.1767	48	47	0.23	0.8218	587	0.4959
City K	0.7085	48	47	0.83	0.4095	685	0.1599
City L	-1.7931	48	47	-1.67	0.1021	407	0.0317
City M	3.3930	48	47	4.60	<0.0001	981	<0.0001
City N	3.1250	48	47	2.68	0.0101	831	0.0063
City O	-0.8687	48	47	-0.79	0.4329	490	0.1574
City P	2.1952	48	47	2.41	0.0199	804	0.0134
City Q	1.8021	48	47	2.28	0.0270	804	0.0134
City R	1.0622	48	47	0.93	0.3563	654	0.2492
City S	2.5890	48	47	3.42	0.0013	957	<0.0001
City T	3.0849	48	47	4.40	<0.0001	950	0.0001
City U	2.7131	48	47	3.75	0.0005	933	0.0002
City V	1.2696	48	47	1.04	0.3059	735	0.0658

PSU Description	Ave Diff b/w CPI and Nielsen 12M Pct Changes	N	DF	T-Value	P-Value	Wilcoxon Statistic	P-Value
City W	5.2305	48	47	7.27	<0.0001	1110	<0.0001
City X	2.7472	48	47	3.12	0.0031	856	0.0030
City Y	0.8772	48	47	1.78	0.0820	748	0.0504
City Z	1.1535	48	47	1.31	0.1980	687	0.1550
City AA	3.5147	48	47	2.98	0.0046	834	0.0058

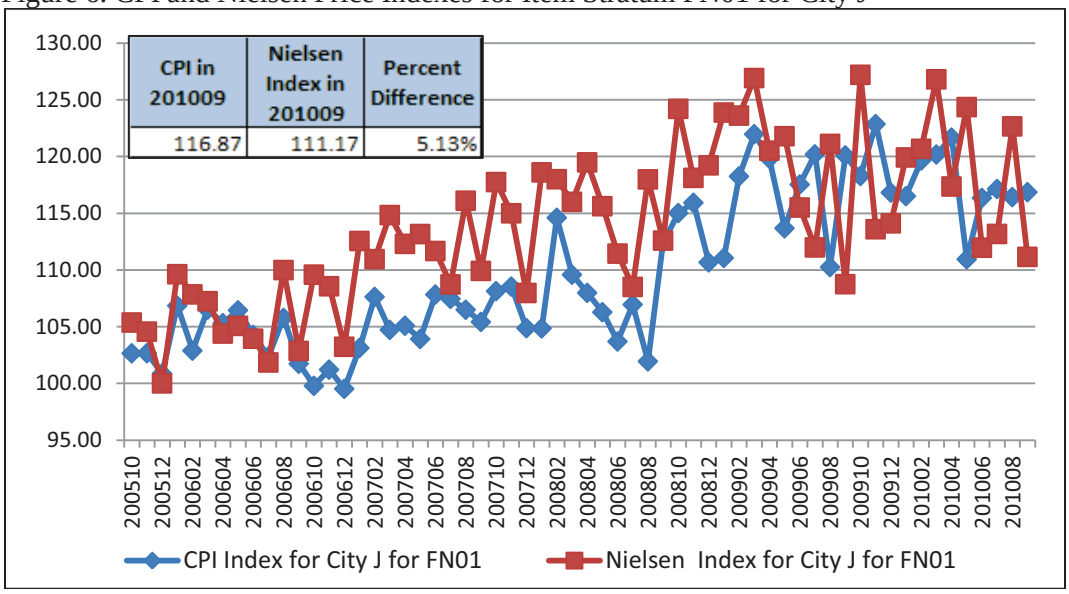
About half of the paired t-tests and Wilcoxon Signed-rank tests indicate that there is a significant difference between the CPI and Nielsen twelve-month percentage changes at the $\alpha = 0.05$ level. The CPI indexes, however, appear to be more precise for some cities than other cities for FN01 if we assume that the Nielsen index is the best estimate of the real market. To get a visual depiction of how much of a difference exists between the CPI index and Nielsen index for a city with low p-values, the CPI and Nielsen indexes for City W for FN01 are shown below in Figure 5.

Figure 5. CPI and Nielsen Price Indexes for Item Stratum FN01 for City W



Both the CPI and Nielsen indexes for FN01 indicate inflation from the base period of 200509 to 201009. However, in 201009, the CPI index indicates about a 41.2 percent increase in price from 200509, whereas the Nielsen index indicates only a 22.4 percent increase. Thus, the CPI index is about 15.4 percent greater than the Nielsen index. In contrast, the CPI and Nielsen price indexes were also compared for a city with high p-values from the significance tests. Figure 6 below shows the difference between the CPI and Nielsen indexes for FN01 for City J (which had p-values of 0.8218 and 0.4959 for the t-test and signed-rank test, respectively). The Nielsen index line tends to run above the CPI's index line up to 200907. From 200907 until 201009, the CPI's trend line appears to run through the middle of the Nielsen indexes. At the end of five years, the CPI and Nielsen indexes differ by about five percent and both signify inflation.

Figure 6. CPI and Nielsen Price Indexes for Item Stratum FN01 for City J



6. Conclusion

In summary, the CPI index estimates for expenditure class FN and its three lower level strata (FN01, FN02, and FN03) perform well at the national level if we assume that the Nielsen price indexes are the best estimates of reality. The Nielsen and CPI index estimates for expenditure class FN differ by less than one percent after five years at the national level, while FN's three lower level item strata all differ by less than five percent. Additionally, the paired t-tests and Wilcoxon Signed-rank tests conducted on the CPI and Nielsen monthly price relatives for FN, FN01, FN02, and FN03 did not show a significant difference between the two sets of price relatives at the $\alpha = 0.05$ level. The significance tests, however, did show that significant differences do exist between the twelve-month percentage changes of the three lower level item strata.

As expected, larger differences exist between the CPI and Nielsen price indexes at the city level. For example, the CPI index for self-representing City W for item stratum FN01 indicates a 41.15 percent increase over the five year reference period, whereas the Nielsen index signifies only a 22.36 percent increase. For indexes that perform poorly, future work should attempt to identify the causes of those differences. Possible areas to be examined are the disaggregation of items under those particular item strata and the types of outlets that contribute price data to the problematic strata.

References

Bureau of Labor Statistics (2008). "Chapter 17 of the BLS Handbook of Methods: Consumer Price Indexes." Accessed February 4, 2013 from the Bureau of Labor Statistics http://www.bls.gov/opub/hom/homch17.htm.

Chanil, Debra and Meg Major (2012). "2012 Consumer Expenditure Study: Deconstructing the Market Basket." The Progressive Grocer. Accessed April 23, 2013. http://www.progressivegrocer.com/inprint/article/id3153/deconstructing-the-market-basket/ .

Fuxa, Mary Lee (2010). "Sampling and Weighting of Commodity and Service Units for the Elementary Level of Computation of the U.S. Consumer Price Index." *Proceeding Business and Economic Statistics Section, American Statistical Association* [CD ROM], 5014-5025. Alexandria, VA.

Hollander, Myles and Douglas A. Wolfe, (1999). Nonparametric Statistical Methods, Second Edition, New York: John Wiley, pp. 35-42.

Marsh III, Fred (2006). "Comparison between Newly Proposed Response Rates and Current Response Rates for the TPOP Survey." *Proceedings on Consumer Prices and Expenditures Section on Government Statistics, American Statistical Association* [CD ROM]. Alexandria, VA.