Journal of Official Statistics
Vol. 5. No. 2, 1989, pp. 107–123

Adjustments for Quality Change in the U.S. Consumer Price Index

Paul A. Armknecht and Donald Weyback[1]

Abstract: The construction of the U.S. Consumer Price Index (CPI) assumes that commodities and services of constant quality are priced over time. When one product disappears and a new or slightly different product is substituted, it is difficult to distinguish quality change from price change. In 1983 and 1984 about 3.9% of the items for which prices were collected involved product substitutions, and by far, the highest rates of substitutions occurred among apparel items. This paper examines the effects of product substitutions on the CPI and provides an update on research that should help improve substitutions and the quality adjustment process for apparel items.

Key words: Consumer Price Index; quality adjustment; product substitution.

1. Introduction

The Consumer Price Index (CPI) for the United States is a measure of the average change in prices paid by urban consumers for a fixed market basket of goods and services. The theoretical framework for the U.S. CPI is based in the theory of consumer demand and presented by Gillingham (1974).

The U.S. CPI consists of samples for 85 geographic areas. The 29 largest areas are self representing because of their population size, while the remaining 56 are probability selected to represent medium and smaller size metropolitan areas and nonmetropolitan urban areas. The item structure for the U.S. CPI consists of 184 item strata for commodity groups such as white bread, carbonated drinks, women's apparel, etc. The market basket is the relative weight of each item stratum within geographic area and represents the relative spending patterns of U.S. urban consumers during the 1982–84 reference period. Price change in the all items U.S. CPI is a weighted average of the price change in each item stratum across all geographic areas. A detailed discussion of the sampling scheme for the U.S. CPI appears in the U.S. Department of Labor BLS Handbook of Methods (1988).

Ideally, we would like to be able to observe price changes for the same goods and services purchased during the reference period. The ability to obtain price information for the same bundle of commodities and services through time is almost impossible. Products frequently disappear from the market, for example, the Kodak Brownie camera. Existing products are replaced with new versions, as demonstrated by the annual model changeover in the U.S. auto

[1] Division of Consumer Prices and Indexes, Bureau of Labor Statistics, Washington, D.C. 20212, U.S.A.

industry, and new products such as video casette recorders emerge. If all replacement and emerging products were exact replacements for older and disappearing items in terms of providing consumers with the same satisfaction as the original item, there would be no problem in substituting the price of one item for that of the other when constructing our price index. However, many replacements and new products provide consumers with improved quality of performance or service, and as such, provide consumers with higher levels of satisfaction. The substitute products present problems for the construction of price indexes since they compromise the basic underlying assumption of measuring price change for items of constant quality (i.e., items yielding the same level of consumer satisfaction). A formal theoretical treatment of this problem along with theoretical solutions are presented in Fisher and Shell (1972). The basic issue for substitute products in measuring price change is to determine whether or not a quality difference exists and, if so, the value of that quality difference in terms of the product's price. With this information we can make adjustments for quality differences in the prices used to compute price change.

The problem of quality adjustment when product substitutions occur in the U.S. will be analyzed as follows. In Section 2, the types of adjustments in the U.S. CPI to keep quality constant when product substitutions occur are presented in some detail. In Section 3, information is provided on the frequency of product substitutions while in Section 4 estimates of the effects of substitutions in the U.S. CPI are presented.[2] Sections 5 and 6 contain descriptions of specific projects underway to improve the treatment of quality changes among apparel items using hedonic regression techniques. In the final sections, a brief summary and references are presented.

2. Methods of Adjustment for Quality Change

When a specific variety of good or service priced in a store is no longer available, a substitute must be found in order to continue the measurement of price change. Anytime a substitution occurs, the new product's price may reflect both pure price change and quality change.

In the U.S. CPI, the quality component in each substitution is handled in accordance with the information available. Most of the quality adjustments used are analogous to the technical factors discussed in Allen (1975). In instances where there is no appreciable difference in quality, the price of the substitute is used directly in the computation of the CPI. Thus, all of the observed price difference is counted as pure price change.

Where differences in quality exist, several options are available. These will be discussed in our order of preference based on the availability of information from the market place. We can estimate the quality change by observing the difference in market prices between the two varieties at the same point in time. For example, a manufacturer introduces a new product, B, that will eventually replace a current product, A. After B is introduced, we can observe the market prices of A and B. The current period is our overlapping month in which we collect prices for both items. The price change in current period is computed as the price change in the product A from its previous price. In the next period, we will collect only the price for product B and the

[2] Excluded from the study were price quotes for residential rent and homeowner's equivalent rent within Housing; used cars within Transportation; health insurance within Medical care; and magazines, periodicals, and books within Entertainment.

price change will be measured by the change in the price of B. The current period, where we have overlapping prices for the two products, provides the estimate of quality change. The quality difference is the difference in market price between A and B. We will refer to this technique as "overlap" pricing.

Another method for handling changes in quality is direct quality adjustment. Ideally, we would like to obtain the market value of the quality difference between an item and its substitute. Then we could adjust the price of the old item directly for the quality change. This rarely occurs, however. Usually quality changes are manifested in some differences in physical characteristics which are measurable. In such cases the value of the difference in quality can be estimated from production cost data. Changes in producer's cost are adjusted for normal profit margins at the wholesale and retail level. The resulting price change serves as a proxy for consumers' valuation of the price difference which cannot be observed directly in the market place. This approach is used most frequently in valuing quality change for new automobiles.

It is also feasible to measure the value of quality change using "hedonic" regression techniques. The price for item P is a function of the price, b_i, for each of its characteristics, X_i. In a linear form this would be

$$P = b_o + \sum_{i=1}^{K} b_i x_i.$$

The estimated parameters, b_i, from a linear regression provide implicit prices for each of the K characteristics. Whenever a quality characteristic changes, an estimated value for the amount of quality change can be calculated from the regression parameters. Griliches (1971), Triplett (1969, 1971a, 1971b), Triplett and MacDonald (1977), and Early and Sinclair (1983) have attempted to demonstrate the use of hedonic techniques to measure and adjust for quality change in price indexes.

Finally when the quality of a substitute item is not comparable with the old one and an overlap price or direct quality adjustment data are not available, the "link" technique is used. No price comparison is made between the two varieties when the substitution is brought into the index. The price change between the two observations is imputed as the average change for other items in that item stratum and geographic area. The remaining difference between the observed prices and the imputed change is assumed to reflect quality differences.

3. Rates of Product Substitution in 1983 and 1984

The rates of product substitution by major category and type of adjustment made in the U.S. CPI appear in Table 1. During 1983, 3.85% of all prices collected involved product substitutions compared to 3.95% in 1984.[3] The increase in the overall substitution rates occurred among comparable substitutes (an increase from 1.56% to 1.70%). Noncomparable and quality adjusted rates remained about the same. Among the major CPI categories, the housing component's rate rose from 4.25% to 4.83% reflecting a higher rate of comparable substitutions (2.21% vs. 2.67%). Higher rates of comparable substitutions in this group reflect shifts to new telephone services following the court-ordered divestiture of AT&T.[4] The other category showing

[3] Data for 1985 are not analyzed here because product redescriptions as the result of new collection documents introduced in preparation for the CPI revision created occassional operational situations that looked like substitutions but were not.

[4] In 1982 a federal court ordered AT&T to divest itself from ownership of regional telephone companies. The gradual divestiture involved transfer of many services to local companies from the former parent company.

Table 1. Product substitutions as a percentage of prices collected by major U.S. CPI category, 1983 and 1984

Category	Year	Comparable total	Noncomparable			Direct quality adjustment	Total
			Total	Use of overlap prices	Use of link prices		
All	1983	1.56	1.97	.23	1.74	.32	3.85
	1984	1.70	1.95	.23	1.71	.30	3.95
Food & beverages	1983	.52	1.29	.04	1.25	.00	1.81
	1984	.52	1.41	.08	1.33	.00	1.93
Housing	1983	2.21	1.89	.22	1.67	.15	4.25
	1984	2.67	1.93	.21	1.72	.22	4.83
Apparel & upkeep	1983	7.15	10.15	2.69	7.46	.03	17.34
	1984	7.80	9.70	2.43	7.27	.09	17.59
Transportation	1983	3.13	1.41	.06	1.35	2.18	6.72
	1984	3.02	1.03	.07	.96	1.74	5.80
Medical care	1983	.65	.94	.03	.91	.64	2.22
	1984	.80	1.02	.03	.99	.38	2.19
Entertainment	1983	1.92	2.51	.23	2.28	.18	4.61
	1984	2.85	2.96	.26	2.70	.27	6.08
Other goods & services	1983	1.44	1.69	.06	1.64	.17	3.30
	1984	1.94	1.64	.08	1.56	.40	3.99

Source: U.S. Bureau of Labor Statistics.

a substantial increase in rates of substitutions was the entertainment category (4.61% to 6.08%). The rate of comparable substitutions rose in other sporting goods and equipment. The increases in these categories were partially offset by the decline in the substitution rate in transportation (6.72% to 5.80%). The rate for noncomparable substitutions and quality adjustments declined largely due to less frequent substitutions among new vehicles.

The apparel and upkeep category registered the highest rate of substitution (17.59%) up slightly from the rate in 1983. Women's suits, women's dresses, and girl's coats, jackets, dresses and suits are strata within apparel that have substitution rates of 40% or greater. Entertainment (6.02%) and transportation (5.80%) demonstrated the next highest rates of substitution among CPI categories.

For the overall CPI the rate of comparable substitutions was 1.70%, up slightly from 1983. It is with this type of substitution that some quality changes may be counted as price changes and thus provide a bias (measurement error) in the CPI. The rate of comparable substitutions was higher in all categories except food and beverages and transportation. The rate of noncomparable substitutions changed little as did the rate of noncomparable substitutions with link prices. This latter type of substitution could result in some price change being counted as quality difference and excluded from the CPI. The rate of direct quality adjustments

Table 2. Quality and pure price effects of substitutions during 1984

Major groups and selected indexes	Percent change in CPI–U[1]	Type of substitute									
		Total		Comparable		Noncomparable using overlap		Noncomparable using link		Direct quality adjustment	
		Quality	Pure price	Quality	Pure price	Quality	Pure price	Quality	Pure price	Quality	Pure price
All items studied[1]	3.40	1.23	3.26	0	1.37	.14	1.38	.99	.16	.10	.35
Food and beverages	3.73	.75	.59	0	.26	.07	.22	.68	.12	.01	0
Housing[1]	3.00	.85	1.10	0	.57	.06	.43	.80	.03	−.01	.06
Apparel and upkeep	1.95	9.66	41.57	0	17.49	2.18	22.73	7.35	1.45	.13	−.10
Apparel commodities	1.42	11.24	49.01	0	20.57	2.56	26.75	8.65	1.71	.03	−.02
Transportation[1]	2.09	.89	2.02	0	.62	0	.03	.56	.12	.33	1.24
New vehicles	2.51	3.75	8.62	0	1.04	.12	.04	1.44	.27	2.19	7.26
Medical care[1]	6.83	.24	.33	0	.25	−.03	.03	.21	.06	.05	−.02
Entertainment[1]	4.28	−.60	2.19	0	.68	−.23	.82	−.45	.47	.08	.22
Other goods and services	6.07	1.70	1.56	0	.92	.15	.13	1.37	.03	.18	.47

Source: U.S. Bureau of Labor Statistics

[1] Excluded from the study were price quotes for residential rent and homeowner's equivalent rent within Housing; used cars within Transportation; health insurance within Medical care; and magazines, periodicals, and books within Entertainment.

remained about the same (0.32%) with the preponderance of direct quality adjustments occurring among new vehicles in the transportation category.

4. Effects of Substitutions on the CPI in 1984

As has been indicated in the foregoing discussion, substitutions can result in some mixture of price and quality change. In Table 2 the effects of quality and price change from substitutions in 1984 are presented for the major CPI components. An earlier paper (Armknecht (1984)) presented results for 1983. The effects were measured by calculating the individual effects of each substitute. For example, when a new model car was introduced as a substitute for the older model, a price increase of 5% might have been observed. Of this price change 1% could have occurred due to improved quality and the remaining 4% was attributable to pure price change. The observed price change represents the combined effects of quality change and pure price change. Since each item included in the index has a specific weight within an item stratum and geographic area (estimating cell), the quality effect of each item can be estimated by multiplying the quality change (.01 in this example) by the quote weight. We can then compute a weighted average of these effects for each type of substitute (a direct quality adjustment in our example) at the estimating cell level. Since each of the estimating cells also has an expenditure weight associated with it, these quality effects can be aggregated further by computing a weighted average of any combination of estimating cells. The same calculation procedure was also used for averaging the pure price effects of each substitute at higher levels of commodity aggregation. In Table 2, the weighted average quality and price effects for each type of substitute have been calculated for U.S. all items, the major commodity groups and selected aggregations within groups.

The overall change in the CPI indexes of the items studied appears in the first column. The effects of quality change (which have reduced the observed price change) and pure price change (which represents the contribution of substitutes to the calculated CPI price change) are presented in subsequent columns. The total effects for all substitutes appear first, followed by those for each type of substitute. The CPI for the items studied rose 3.4% in 1984. The effect of quality adjustment (either implicit or explicit) on the change in the CPI during 1984 was to reduce it by 1.23 percentage points. That is, without quality adjustments of any kind, the CPI would have risen 1.23% more. Substitutes by themselves, after adjustment for quality change, contributed 3.26% to the overall increase.

As previously mentioned, the quality and pure price effects for each type of substitution in Table 2 were calculated for each individual substitute item. The estimates of the quality effect versus the pure price effect for each individual price quotation by type of substitution is as follows. For comparable substitutes there is no perceived quality difference between items so the quality effect of the individual substitute is zero. The pure price effect of the individual substitute is the change in observed prices: $(P_t^2/P_{t-1}^1) - 1$, where P is the price of the item, the subscript "t" represents the time period, and the superscript is the variety. For noncomparable substitutes using overlap prices (i.e., prices in the same time period for both varieties of the item) the quality effect is the price differential between varieties: $(P_t^2/P_t^1) - 1$, and the price effect is the

change in the observed price of the old variety: $(P_t^1/P_{t-1}^1) - 1$.

For noncomparable substitutes using link prices the quality effect is the differential between the price of the new variety and the imputed price for the old variety, $P_t^1{}^*$: $(P_t^2/P_t^1{}^*) - 1$. This imputed price $(P_t^1{}^*)$ is based on the price change of reported items in the same item stratum and geographic area. The imputation assumes the price change for the missing item, P_t^1, is the same as that for all other items which were available. The price effect is the differential between the imputed price and the previous price for the old variety: $(P_t^1{}^*/P_{t-1}^1) - 1$. Finally, the quality effect for items with direct quality adjustments is the proportion of the measured quality difference between the two varieties, (QA_{t-1}^1) to the price of the old item: (QA_{t-1}^1/P_{t-1}^1). The price effect is the differential between the price of the new variety and the price of the old variety adjusted for the quality difference: $(P_t^2/(P_{t-1}^1 + QA_{t-1}^1)) - 1$.

In 1984 the greatest effect of product substitutions in terms of both quality and price change was among the apparel indexes. Apparel commodities increased in price by 1.4%. Had there been no allowance for quality differences among substitutes, prices would have risen by 11% more. When substitutions occurred for apparel items, large price increases were recorded which alone would have resulted in a 49% rise in the apparel indexes. However, the price effects of substitutes were mitigated by the large number of items that are periodically sold at reduced prices during the year. These offsetting effects are the result of the pricing practices in U.S. retail markets for most apparel commodities, and the BLS procedures for pricing substitutions.

The general approach to price setting in U.S. retail apparel markets is to introduce items at high profit margins in the beginning of their selling season, and then take continual price reductions, referred to as "sales" in the U.S., until the stock is depleted and new lines are introduced (see Pashigian (1988)). Thus an item will be brought into the CPI at its initial price and then proceed to be reduced in price until it is no longer available. At this point a substitution occurs, usually to the new line of apparel. When there is a comparable substitution, a large price increase can occur. These comparable substitutes contributed 20.6% toward apparel commodity price increases in 1984.

Quite often the new item is considered noncomparable so that it would be linked into the index. Linking in such cases would result in a downward bias in the index. As just mentioned, the new line is being sold at its high introductory price. But linking results in no price comparison being made between items at the time of substitution. So in subsequent months, as price reductions occur, the index will show continuous price declines. The following example provides a heuristic treatment of this situation.

In January through March we were

Table 3. Prices in US$ for hypothetical apparel item

	Jan.	Feb.	Mar.	Apr.	May	June
Old item in US$	100	75	50	–	–	–
New item in US$	–	–	–	150	125	100
% Change	–	−25	−33	1[a]	−17	−20

[a]Imputed from other apparel items.

pricing the old item which showed large declines. In April the old item is not available and we substitute to a new item. The two are declared noncomparable and we have no overlap price. If the new item is linked into the index, the price change is imputed as the average of all other items in that item stratum/geographic area. Assume this to be 1%.

In May's pricing the new item shows a decline of 17% and in June there is another decline of 20%. Thus, the index would show an almost continual downward movement. To overcome this inherent bias, there is a procedural rule used that prohibits an item from going out of the index following these continuous price reductions. When the item is no longer available, and a noncomparable substitute is selected, an artificial overlap price is used. This estimated price is the item's last regular price before the reductions began. In the above example, when a noncomparable substitute occurred in April, an artificial price of US$100 would be used as an estimated overlap for computing the index so that the March–April change would be 100% (US$50 to US$100), thus bringing the index back to its level before the price reductions began. This procedure of estimated overlaps accounts for the large price effect (27%) among noncomparable substitutes with overlap prices in apparel commodities. In most other areas of the U.S. CPI, the effect of this procedure is minimal.

The type of action which has the greatest effect in terms of quality in the CPI is the noncomparable substitute where the linking technique is used. In such instances most of the price differential between the old and new varieties is implicitly treated as quality change. These actions could have affected the studied areas of the CPI by 1%. In 1984 the apparel commodities (8.7%) and household furnishings (3.6%) were the components with the largest quality effects. By contrast, direct quality adjustments resulted in only a 0.2% reduction of the change in the overall CPI. The adjustments were concentrated in the new vehicles component (2.2%).

5. Efforts to Improve the Treatment of Price and Quality Changes in Apparel Commodities

Several research projects are currently underway to identify and resolve problems that affect price versus quality change in the CPI. As the foregoing analysis indicates, the area where the greatest opportunity for improving quality adjustment may exist is with apparel commodities. The index for apparel commodities had shown an increase of only 25.5% over the nine year period ending in December 1986, while all the items CPI increased by 77.9%.

Seven of the 16 strata that comprised men's, women's, boy's and girl's apparel had increased by less than 10%, and one stratum had actually shown a decrease of 13% over that same time period. The large number of items that require substitutions each month and the large percentage of those substitute items that are noncomparable and, therefore, not used in the calculation of the index probably contribute heavily to this result. Because the price changes for these are imputed from the average of changes for other items that have either the same variety or comparable substitutes in the specific month, the price changes associated with the introduction of a new noncomparable variety may be understated. For example, in men's suits for those quotes in which the same item was priced during 1983, the price change was 0.3%. When comparable substitute items were priced, the average price change was 15%, suggesting that the introduction of

new varieties may be accompanied by larger pure price changes than the price changes for continuing varieties.

In order to reduce the number of apparel substitutes that are noncomparable, a research project was undertaken to develop a hedonic regression model to derive the importance of individual specifications in determining the price of various apparel commodities.

The utility of hedonic regressions on apparel items can be seen on three levels. First, parameter estimates from the regression models can enhance, or "upgrade," the existing comparability decision making capabilities of economic analysts by providing more clear cut (and statistically supported) distinctions between significant and insignificant price factors. At a more general level, the models can be used to identify those "major" price factors which contribute the bulk of the quality makeup of that item. These major price factors were used to redesign the collection documents used by CPI field staff. Limiting noncomparability to changes in these major price factors allows analysts to increase the number of comparable versus noncomparable substitutes while still maintaining control over the level of quality of the sample.

Finally, hedonic regression models may allow analysts in the future to use the parameter estimates as implicit price values for direct quality adjustments. Using this approach most substitutions could be used for price comparisons, with the chief error in the overall quality of the sample being the degree to which the model did not explain price variability.

The Women's suits price index was selected for initial study because it has exhibited the most severe downward price movement (a decrease of 13.4% over a nine year time period) among apparel items, and the collection document was the most complex in terms of the number and correlation of specification elements. Useful hedonic regression results from this stratum would be a strong indication that similar information could be achieved for most, if not all of the "less complex" apparel items.

July and August of 1984 were selected as the initial study time period. Two months are needed to include all local areas in the CPI because apparel is priced bimonthly in many areas. Specifications and prices were integrated into a database containing the entire women's suits sample. Any imputed prices as well as "sale" prices were then replaced by the item's last regular reported price using historical price data. This approach enables both fall/winter and spring/summer seasonal items to be fully and equally represented. Those items which had never shown a reported price as well as those "sale" items which had no previous regular prices were removed, reducing the database to 314 observations.

Continuous variables were then created for each of the fiber specification elements; dummy variables were created for each of the other specification elements, as well as various other combinations of variables (e.g., items which had a lining and were also 100% wool).

Similar databases of the women's suits sample in previous years were then created using the same set of procedures – specifically July/August of 1983, July/August 1981, and July/August 1979. These data were then combined to form an aggregate database for women's suits of 1584 observations.

Preliminary regressions were run on a pooled 1983–1984 database to determine the "base" variables for each of the specification categories. Polyester was found to be the base variable for the fiber category with all other fibers having positive correlations with price.

Forward stepwise regressions were then run to determine the relative importance and significance of those characteristics represented by the set of continuous and dummy variables. During this time the relationships between the variables were being analyzed through the use of a correlation matrix; guarding against excessive collinearity between pairs of variables at this stage reduces the likelihood of excessive multicollinearity occurring in the regression models. Using the preliminary "identifier" regressions, the stepwise results, and the correlation matrix, full linear regression models were run on that set of characteristics which combined substantial explanatory power and acceptable collinearity levels.

Results from these procedures isolated 14 distinct characteristics presented in Table 4. These characteristics, all of which are seen to be statistically significant as measured by their individual F values (see Table 4), explained a substantial amount (R^2 = 82%) of the price variability in the sample. Multicollinearity (as measured by the tolerance factors, which equal one minus the R^2 that results from the regression of the other variables in the model on that regressor), was significant, but deemed to be within acceptable levels for the purpose of identifying the important price factors. One point to be considered here is that in the stepwise regressions, as each variable was added to the model, those variables already in the model remained statistically significant. In high multicollinearity situations, additional variables tend to alter substantially the number and composition of statistically

Table 4. Women's suits – 1983 and 1984 – regression coefficients and F values

Characteristics	Regression coefficient	F value	Tolerance factor
Intercept	16.87	–	–
Exclusive brand	320.67	913.40	.620
National/Regional brand	102.23	292.09	.381
Private label brand	39.94	77.49	.465
Silk	0.63	15.35	.768
Wool	0.56	93.16	.443
Rayon	0.36	17.80	.676
Cotton	0.31	6.47	.920
Top, shirt or blouse	33.26	33.99	.752
Set on waist (pants)	28.02	6.91	.910
Lining	16.26	9.30	.404
Short sleeve jacket	14.67	2.84	.962
Skirt style	12.91	6.49	.951
Set on waist (skirt)	7.86	3.25	.642
Misses size range	6.67	3.61	.881

Adjusted R^2 = 82%
Chow test F value = 2.34 Threshold F value = 2.04

Note: All variables other than those corresponding to fiber are dummy variables, whose coefficients can be interpreted as direct quality contributors to the price of the item, in dollar amounts (i.e., the presence of a blouse contributes US$33.26 to the price of the suit). Fiber content variables are continuous; their coefficients can be interpreted as the contribution to price per unit change in fiber content (50% wool content = 50 × 0.56 = US$28.00 contribution to price over the price of a polyester suit).

significant variables already included in the model.

Of considerable importance is the fact that each of the regression coefficients make logical sense, from both an ordinal standpoint and from a more general quality perspective. For example, the presence of silk is more important with respect to price than the presence of wool, which has a greater coefficient than rayon, etc. Also the characteristics that one would expect, a priori, to contribute most to the quality of a woman's suit have positive values and are statistically significant.

These results were used to revise the collection document (called a checklist) for women's suits. A new "tiered" approach to the checklist design was used. In the new design the most important price factors that determine quality and comparability appear in the first tier. (See Exhibit 1.) The set of first tier price factors were selected from the hedonic model as those factors which contribute the bulk of the quality make up of women's suits. When a substitution is necessary, our field agents are instructed to match all first tier elements, if possible. If they can, the substitution will be comparable.

The second tier, as identified by the hedonic models, contains minor product/quality characteristics; the third tier contains product identifiers. Limiting noncomparability to changes in first tier price factors allow us to increase the number of comparable versus noncomparable substitutions while still maintaining control over the level of the quality variability of the sample.

The substitution procedures to be used in conjunction with the tiered checklists have also been oriented towards maximizing comparable substitutions. The differences basically involve two modifications to the timing of substitutions. The constraint for immediate substitution is relaxed. The collection staff can delay substitutions until such time as an item with matching first tier specifications to the previous item is available (if future availability of such an item can be reasonably assured). Secondly, the collection staff has freedom to substitute immediately from the previous item if a substitute item with matching first tier specifications is available (unless it can be assured that the previous item will be available next pricing period).

To illustrate these differences, consider the following example. A women's suit containing 100% wool is priced in September. For pricing in November, suits containing 100% wool are sold out, and will not be sold again in the store until they restock for the next fall season. Under former substitution rules, since the current item itself is not available and is not expected to be available in the immediate future, the field representative would substitute, in all likelihood to a noncomparable item. Using the revised substitution procedures, the field representative would determine if the first tier specification elements are temporarily or permanently not available. Since suits containing 100% wool (and which match the other first tier specifications) will be sold in the outlet next September, the field representative would wait to substitute until that time, in all likelihood to a comparable item. In the meantime the item's price is imputed.

The preceding example illustrates the constraining influence on the substitution decision; the following illustrates the new price collection policy which allows an immediate substitution when identical first tier specifications are available. Consider again an example of a suit priced in September. Upon returning to collect a price in November, if the previous item is unavailable and the respondent is unsure of the future availability of the item, previous

Exhibit 1

BUREAU OF LABOR STATISTICS **U.S. DEPARTMENT OF LABOR**
CONSUMER PRICE INDEX - ELI CHECKLIST

collection period: _ _ _ _ outlet number: _ _ _ _ _ _ _ _ quote code: _ _ _ arranging code: _ _ _ _

ELI No./ title 38051 WOMEN'S SUITS cluster code 01C

item availability: AVAILABLE ELI NOT SOLD INIT INCOMPLETE

purpose of checklist: INIT INIT COMPLETION SPEC CORR SUB REINIT CHECKLIST REV

CURRENT PERIOD

price _ _ _ _ _ _ . _ _ _

type of price: REG SALE

SALES TAX

applicable: YES NO

included: YES NO

rate: _ _ . _ _ _

description: _____

_____

_____

seasonal: YES NO in-season: JAN FEB MAR APR MAY JUN JUL AUG SEP OCT NOV DEC

respondent: location:

field message:

FIBER CONTENT

- A1 100% Silk
- A2 50-99% Silk
- A3 1-49% Silk
- A4 100% Wool
- A5 50-99% Wool
- A6 1-49% Wool
- A7 50-100% Cotton
- A8 1-49% Cotton
- A9 100% Polyester
- A10 All manmade fiber (except 100% polyester), and all man made fiber blends
- A99 Other, _____

BRAND/LABEL CATEGORY

- B1 Store brands
- B2 Private label
- B3 National/regional brand
- B4 Exclusive brand
- B5 Uncertain

COMPOSITION

- C1 Jacket or coat
- D1 Vest
- E1 One skirt or one pair pants
- F1 Top, shirt, or blouse
- G1 Sweater
- H99 Other, _____

CLEANING METHOD

- I1 Machine washable
- I2 Dry clean only
- I99 Other, _____

SIZE RANGE

- J1 Juniors/misses
- J2 Womens/half sizes
- J99 Other, _____

SPECIFY ADDITIONAL INFORMATION ON PAGE 38051-2

SPECIFIC FIBER CONTENT

- K99 Silk, _____%
- L99 Wool, _____%
- M99 Cotton, _____%
- N99 Rayon, _____%
- P99 Nylon, _____%
- Q99 Polyester, _____%
- R99 Other, _____%

SPECIFIC BRAND/LABEL NAME

- S99 _____

ZZ99

BLS 3400B (Rev. May 1986) 38051-1 Revised May 1986

Exhibit 1 (Cont.)

38051 - 01C WOMEN'S SUITS - CONTINUED

JACKET OR COAT DESCRIPTION

LINING/INTERLINING
- T1 Lined
- T2 Unlined
- U1 Interlined
- U2 Not interlined

SLEEVES
- V1 Long
- V2 Short

LENGTH
- W1 Hip length
- W2 Waist length
- W99 Other, _____

BELT
- X1 No belt
- X2 Self belt, tied or with buckle
- X99 Other, _____

SKIRT DESCRIPTION

WAISTBAND
- Y1 Pull-on elasticized
- Y2 Set-on with zipper
- Y99 Other, _____

STYLE
- AA1 A-line
- AA2 Flared or full
- AA3 Straight
- AA4 Gathered (dirndl)
- AA5 Pleated
- AA99 Other, _____

LENGTH
- AB1 Short (approx. knee length)
- AB99 Other, _____

LINING
- AC1 Lined
- AC2 Unlined

PANTS DESCRIPTION

WAISTBAND
- AD1 Pull-on elasticized
- AD2 Set-on with zipper
- AD99 Other, _____

LINING
- AE1 Lined
- AE2 Unlined

TOP, SHIRT OR BLOUSE DESCRIPTION

SLEEVES
- AF1 Long or 3/4
- AF2 Short
- AF3 Sleeveless

STYLE
- AG1 Open front
- AG2 Pullover
- AG99 Other, _____

NECK
- AH1 No collar
- AH2 Shirt style
- AH3 Stock tie
- AH99 Other, _____

FIBER
- AI99 Cotton, _____%
- AJ99 Rayon, _____%
- AK99 Polyester, _____%
- AL99 Other, _____%

FIBER DESIGN
- AM1 No design/solid color
- AM2 Printed
- AN1 Multicolor, fiber or yarn dyed
- AN2 Jacquard or dobby, solid color
- AN3 Jacquard or dobby, multicolor
- AN99 Other, _____

SWEATER DESCRIPTION

FIBER
- AP99 Wool, _____%
- AQ99 Acrylic, _____%
- AR99 Polyester_____%
- AS99 Other, _____%

CONSTRUCTION
- AT1 Cut and sewn
- AT2 Full fashioned

**** STYLE NUMBER**

AU99 _____

**** OTHER DESCRIPTIVE OR IDENTIFYING INFORMATION**

AV99 _____

procedure called for the field representative to enter a code of unavailable. In many cases, the field representative would then return in January to find the entire line of suits either sold out or out of season. The new procedures allow the field representative to substitute in November to a suit with matching first tier specifications and show a price comparison.

The degree to which this new methodology improves the accuracy of the apparel indexes depends in large part upon its ability to maintain the quality of the apparel samples to within acceptable levels. In effect, the new methodology comprises a more lenient comparison structure for substitutes, which carries with it the potential danger of introducing biases due to quality change. Our concern is not so much for levels of quality of individual items, but rather for higher levels of aggregation – specifically the entire sample of a particular stratum. It is important to keep this change in orientation from the individual item level to the aggregate level in mind.

It has been demonstrated, however, that the change in quality will not be adversely affected. Three separate internal studies investigated the degree to which the quality of apparel items have changed over time. The first study involved a comparison of the percentage frequencies of specification elements of four women's apparel strata over a three year time period. Since specification elements are indications of quality, differences in their relative frequencies should point to the existence and location, if any, of specific changes in quality for each stratum. The other two studies involved comparisons of individual quotes over a recent three year time period using cardinal and ordinal "comparison scales." These studies concluded that there had not been large changes in quality in the four strata under study.

6. Quality Adjustment Factors From Hedonic Regression Models

The coefficients from hedonic regression models can be used to derive implicit quality adjustments. From an operational perspective it would be desirable for the parameter estimates to remain stable over time. The standard Chow test, which is a modified F test used to determine the statistical reliability between regression equations, can be used to provide evidence of parameter stability. Initial F values for the regression model previously mentioned rejected (at the .01% level of significance) the assumption that the model was stable over a one year time period (1983 to 1984). A Chow test using paired years (1979/1981 to 1983/1984) for this model also rejected the null hypothesis of stability.

Further analysis of the parameter estimates over varying time periods revealed that some coefficients were less stable than others. This led to the belief that, if these "unstable" parameters were among the minor price contributors, their removal from the model would reduce its instability without significantly reducing its explanatory power, or R^2.

A series of regressions were then tested using different combinations of variables and other techniques to search for an optimal model which also (1) is stable over, at minimum, a one year time span, and (2) meets a more stringent multicollinearity threshold than in effect for the tiered checklist regressions—specifically, that the tolerance factor for each variable is greater than 0.50. The form of the regression model was also respecified to a semi-log model in which the natural logarithm of price is regressed over the independent variables. This form of the model is one commonly suggested in the literature on hedonic models, e.g., Griliches (1971) and Triplett (1971a, 1971b). The

Table 5. Women's suits – 1983 and 1984 – regression model for quality adjustment

Characteristics	Regression coefficient	*F* value	Tolerance factor
Intercept	3.19	–	–
Silk	.0140	59.79	.949
Wool	.0084	177.52	.568
Cotton	.0057	14.56	.929
Rayon	.0056	28.88	.690
Lining	.6883	154.88	.580
Top, shirt or blouse	.5244	64.59	.826
Set on waist (pants)	.3531	7.55	.973
Jacket	.3050	7.96	.928
Misses size range	.2113	24.23	.916
Skirt style	.1710	7.51	.979
Adjusted R^2 = 62%			
Chow test *F* value = 2.14 Threshold *F* value = 2.25			

Note: The independent variables in the regression were run on the natural logarithm of price; the regression estimates can be interpreted as the estimated percentage change in price to a unit change in the particular characteristic.

semi-log model yielded a slightly better statistical result. The regression parameters are interpreted as percentage contributions to price rather than actual dollar contributions. The resulting regression model listed in Table 5 meets these criteria. Worthy of note is the fact that the brand categories, removed due to their significant instability, bore primary responsibility for the reduction of R^2 to 62%.

Table 6. Original and re-estimated CPI-U indexes using hedonic regression parameters for quality adjustments[1]

Month and year	Women's coats and jackets			Women's suits		
	Actual	Test	Diff.	Actual	Test	Diff.
July '87	95.2	94.9	.3	91.8	92.0	−.2
August '87	89.6	89.3	.3	89.9	89.9	0
September '87	97.8	98.3	−.5	110.0	111.6	−1.6
October '87	108.3	108.0	.3	118.6	120.6	−2.0
November '87	111.7	110.7	1.0	119.4	123.5	−4.1
December '87	110.8	109.4	1.4	117.5	122.2	−4.7
January '88	104.0	102.4	1.6	113.7	116.4	−2.7
February '88	100.3	97.6	2.7	109.6	112.0	−2.4
March '88	102.7	98.9	3.8	107.0	110.0	−3.0
April '88	108.9	103.8	5.1	122.7	126.7	−4.0
May '88	108.3	103.6	4.7	123.5	127.7	−3.2
June '88	104.9	101.0	3.9	119.5	124.0	−4.5

[1]The test indexes presented here were produced by Efthemia Georges and Paul Liegey in an internal BLS report entitled, "An Examination Using Hedonic Regression Techniques to Measure the Effects of Quality Adjustment on Apparel Indexes", November 1988.

Further Chow test results over longer time periods, revealing at what point the coefficients become statistically unstable, indicate that the regression parameters become unstable for time periods of two years or more. However, this does not appear to be a substantial limitation to the use of the regression coefficients for quality adjustment. We will use implicit prices from previous year's apparel lines to estimate the value of quality differences with current apparel lines.

To measure the potential effects of using direct quality adjustments for apparel items, studies are underway to recalculate a few women's apparel indexes over a one year time period. First, an attempt was made to quality adjust all noncomparable substitutions using estimated coefficients similar to those in Table 5. These implicit characteristic prices were used to adjust directly individual price observations for quality differences. The first two strata tested were women's coats and jackets and women's suits. The tests indicate a few operational problems in using hedonic methods. A large proportion of noncomparable substitutions could not be quality adjusted because one of the first tier specification elements was not reported by the field collection staff. Because of design problems, some confusion occurred in recording the composition specification elements on the tiered checklist (Exhibit 1). Also, a number of items with the A99 or H99 "other" specification had no regression parameter from the hedonic model. As a result only about 20% of the noncomparable substitutes could be quality adjusted using the regression estimates. The actual CPI and the re-estimated test CPI using the hedonic quality adjustment factors are presented in Table 6. The one year period covers June 1987 to May 1988. For women's coats and jackets, the test index was lower than the published index indicating a potential upward bias in existing linking procedures for this stratum. For women's suits, the test index was higher indicating a possible downward bias for the stratum. However, these results are inconclusive because of the small proportion of noncomparable substitutes that could be quality adjusted.

Actions have been taken to remedy this situation. Revised checklists have been issued to field staff and procedures for headquarters staff to deal with the "other" specifications in the first tier of the checklist are being developed. In addition more control variables (e.g., region of country, city size and type of outlet) are being added to the model. Another set of test indexes will be produced with data for mid-1988 to mid-1989.

7. Summary

Product substitutions in the U.S. CPI make measurement of pure price change more difficult because of the uncertainty surrounding changes in quality between old and new varieties of items. Efforts are continuing to identify techniques and procedures that reduce the possible biases that changes in quality have on the U.S. CPI. Apparel commodities are one component within the U.S. CPI where such bias could be substantial. The use of hedonic regression techniques to identify significant quality factors and their implicit contribution to price differences appears very promising and should lead to future improvements in measuring price change among apparel items.

9. References

9.1. References cited in the text

Allen, R.G.D. (1975): Index Numbers in Theory and Practice. Aldine Publishing, Chicago.

Armknecht, P.A. (1984): Quality Adjustment in the CPI and Methods to Improve It. American Statistical Association, Proceedings of the Business and Economic Statistics Section, pp. 57–63.

Early, J.F. and Sinclair, J.H. (1983): Quality Adjustment in the Producer Price Indexes. In the U.S. National Income and Product Accounts: Selected Topics. Murray F. Foss, ed., Chap. 2. University of Chicago Press, Chicago.

Fisher, F.M. and Shell, K. (1972): The Economic Theory of Price Indices. Academic Press, New York.

Gillingham, R.F. (1974): A Conceptual Framework for the Revised Consumer Price Index. American Statistical Association, Proceedings of the Business and Economic Statistics Section, pp. 246–252.

Griliches, Z. (1971): Introduction: Hedonic Price Indexes Revisited. In Price Indexes and Quality Change, Z. Griliches, ed., Harvard University Press, Cambridge, pp. 3–15.

Pashigian, B.P. (1988): Demand Uncertainty and Sales: A Study of Fashion and Markdown Pricing. American Economic Review, 78, pp. 936–953.

Triplett, J.E. (1969): Automobiles and Hedonic Quality Measurement. Journal of Political Economy, 77, pp. 408–417.

Triplett, J.E. (1971a): Quality Bias in Price Indexes and New Methods of Quality Measurement. In Price Indexes and Quality Change, Z. Griliches, ed., Harvard University Press, Cambridge, pp. 180–214.

Triplett, J.E. (1971b): The Theory of Hedonic Quality Measurement and Its Use in Price Indexes. BLS Staff Paper No. 6, U.S. Department of Labor, Bureau of Labor Statistics, Washington, D.C.

Triplett, J.E. and McDonald, R.J. (1977): Assessing the Quality Error in Output Measures: The Case of Refrigerators. Review of Income and Wealth, 23, pp. 136–157.

U.S. Department of Labor, Bureau of Labor Statistics (BLS) (1988): Consumer Prices. BLS Handbook of Methods, Chapter 19, Bulletin 2285, pp. 154–208.

9.2. Reference not cited in the text

U.S. Department of Labor, Bureau of Labor Statistics (BLS) (1983): Problems in Measuring Consumer Prices, Report 697.

Received March 1988
Revised July 1989