104TH CONGRESS
2nd Session

COMMITTEE PRINT

S. PRT.
104–72

FINAL REPORT OF THE ADVISORY COMMISSION TO STUDY THE CONSUMER PRICE INDEX

COMMITTEE ON FINANCE
UNITED STATES SENATE

WILLIAM V. ROTH, JR., *Chairman*

DECEMBER 1996

Printed for the use of the Committee on Finance

U.S. GOVERNMENT PRINTING OFFICE
36–352—CC WASHINGTON : 1996

S362-1

For sale by the U.S. Government Printing Office
Superintendent of Documents, Congressional Sales Office, Washington, DC 20402
ISBN 0-16-054062-3

COMMITTEE ON FINANCE

WILLIAM V. ROTH, JR., Delaware, *Chairman*

JOHN H. CHAFEE, Rhode Island
CHARLES E. GRASSLEY, Iowa
ORRIN G. HATCH, Utah
ALAN K. SIMPSON, Wyoming
LARRY PRESSLER, South Dakota
ALFONSE M. D'AMATO, New York
FRANK H. MURKOWSKI, Alaska
DON NICKLES, Oklahoma
PHIL GRAMM, Texas
TRENT LOTT, Mississippi

DANIEL PATRICK MOYNIHAN, New York
MAX BAUCUS, Montana
BILL BRADLEY, New Jersey
DAVID PRYOR, Arkansas
JOHN D. ROCKEFELLER IV, West Virginia
JOHN BREAUX, Louisiana
KENT CONRAD, North Dakota
BOB GRAHAM, Florida
CAROL MOSELEY-BRAUN, Illinois

LINDY L. PAULL, *Staff Director and Chief Counsel*
MARK A. PATTERSON, *Minority Staff Director and Chief Counsel*

JOAN WOODWARD, *Economist and Chief Budget Analyst*
DAVID PODOFF, *Minority Chief Economist*

LETTER OF SUBMITTAL

DECEMBER 4, 1996

Hon. WILLIAM V. ROTH, JR., *Chairman,*
Hon. DANIEL P. MOYNIHAN, *Ranking Minority Member,*
Committee on Finance,
U.S. Senate, Washington, DC.

DEAR SENATORS ROTH AND MOYNIHAN:

The Advisory Commission to Study the Consumer Price Index herewith submits its Final Report in accordance with its charter based on Senate Resolution 73, Section 11b.

Sincerely,

MICHAEL J. BOSKIN,
Chairman
ELLEN R. DULBERGER,
Member
ROBERT J. GORDON,
Member
ZVI GRILICHES,
Member
DALE JORGENSON,
Member

CONTENTS

Executive Summary

1. The American economy is flexible and dynamic. New products are being introduced all the time and existing ones improved, while others leave the market. The relative prices of different goods and services change frequently, in response to changes in income and technological and other factors affecting costs and quality. This makes constructing an accurate cost of living index more difficult than in a static economy.

2. Estimating a cost of living index requires assumptions, methodology, data gathering and index number construction. Biases can come from any of these areas. The strength of the CPI is in the underlying simplicity of its concept: pricing a fixed (but representative) market basket of goods and services over time. Its weakness follows from the same conception: the "fixed basket" becomes less and less representative over time as consumers respond to price changes and new choices.

3. There are several categories or types of potential bias in using changes in the CPI as a measure of the change in the cost of living. (1) Substitution bias occurs because a fixed market basket fails to reflect the fact that consumers substitute relatively less for more expensive goods when relative prices change. (2) Outlet substitution bias occurs when shifts to lower price outlets are not properly handled. (3) Quality change bias occurs when improvements in the quality of products, such as greater energy efficiency or less need for repair, are measured inaccurately or not at all. (4) New product bias occurs when new products are not introduced in the market basket, or included only with a long lag.

4. While the CPI is the best measure currently available, it is not a true cost of living index (this has been recognized by the Bureau of Labor Statistics for many years). Despite many important BLS updates and improvements in the CPI, changes in the CPI will overstate changes in the true cost of living for the next few years. The Commission's best estimate of the size of the upward bias looking forward is 1.1 percentage points per year. The range of plausible values is 0.8 to 1.6 percentage points per year.

5. Changes in the CPI have substantially overstated the actual rate of price inflation, by about 1.3 percentage points per annum prior to 1996 (the extra 0.2 percentage point is due to a problem called formula bias inadvertently introduced in 1978 and fixed this year). It is likely that a large bias also occurred looking back over at least the last couple of decades.

6. The upward bias creates in the federal budget an annual automatic real increase in indexed benefits and a real tax cut. CBO estimates that if the change in the CPI overstated the change in the cost of living by an average of 1.1 percentage points per year over the next decade, this bias would contribute about $148 billion to

the deficit in 2006 and $691 billion to the national debt by then. The bias alone would be the fourth largest federal program, after social security, health care and defense. By 2008, these totals reach $202 billion and $1.07 trillion, respectively.

7. Some have suggested that different groups in the population are likely to experience faster or slower growth in their cost of living than recorded by changes in the CPI. We find no compelling evidence of this to date (in fact just the opposite) but further exploration of this issue is desirable.

8. The commission is making over a dozen specific recommendations to the BLS. These include the following:

i. The BLS should establish a cost of living index (COLI) as its objective in measuring consumer prices.

ii. The BLS should develop and publish two indexes: one published monthly and one published and updated annually and revised historically.

iii. The timely, monthly index should continue to be called the CPI and should move toward a COLI concept by adopting a "superlative" index formula to account for changing market baskets, abandoning the pretense of sustaining the fixed-weight Laspeyres formula.

iv. The new annual COL index would use a compatible "superlative-index" formula and reflect subsequent data, updated weights, and the introduction of new goods (with their history extended backward).

v. The BLS should change its procedure for combining price quotations by moving to geometric means at the elementary aggregates level.

vi. The BLS should study the behavior of the individual components of the index to ascertain which components provide most information on the future longer-term movements in the index and which items have fluctuations which are largely unrelated to the total and emphasize the former in its data collection activities.

vii. The BLS should change the CPI sampling procedures to de-emphasize geography, starting first with sampling the universe of commodities to be priced and then deciding, commodity by commodity, what is the most efficient way to collect a representative sample of prices from which outlets, and only later turn to geographically clustered samples for the economy of data collection.

viii. The BLS should investigate the impact of classification, that is item group definition and structure, on the price indexes to improve the ability of the index to fully capture item substitution.

ix. There are a number of additional conceptual issues that require attention. The price of durables, such as cars, should be converted to a price of annual services, along the same lines as the current treatment of the price of owner-occupied housing. Also, the treatment of "insurance" should move to an ex-ante consumer price measure rather than the currently used ex-post insurance profits based measure.

x. The BLS needs a permanent mechanism for bringing outside information, expertise, and research results to it. At the

request of the BLS, this group should be organized by an independent public professional entity and would provide BLS an improved channel to access professional and business opinion on statistical, economic and current market issues.

xi. The BLS should develop a research program to look beyond its current "market basket" framework for the CPI.

xii. The BLS should investigate the ramifications of the embedded assumption of price equilibrium and the implications of it sometimes not holding.

xiii. The BLS will require a number of new data collection initiatives to make some progress along these lines. Most important, data on detailed time use from a large sample of consumers must be developed.

9. The Commission is making several recommendations to the President and Congress. These include the following:

xiv. Congress should enact the legislation necessary for the Departments of Commerce and Labor to share information in the interest of improving accuracy and timeliness of economic statistics and to reduce the resources consumed in their development and production.

xv. Congress should provide the additional resources necessary to expand the CES sample and the detail collected, to make the POPS survey more frequent, and to acquire additional commodity detail from alternative national sources, such as industry surveys and scanner data.

xvi. Congress should establish a permanent (rotating) independent committee or commission of experts to review progress in this area every three years or so and advise it on the appropriate interpretation of then current statistics.

xvii. Congress and the President must decide whether they wish to continue the widespread substantial overindexing of various federal spending programs and features of the tax code. If the purpose of indexing is accurately and fully to insulate the groups receiving transfer payments and paying taxes, no more and no less, they should pass legislation adjusting indexing provisions accordingly.

This could be done in the context of subtracting an amount partly or wholly reflecting the overindexing from the current CPI-based indexing. Alternatively, a smaller amount would need to be subtracted from indexing based on the new revised annual index if and when it is developed and published regularly, to more closely approximate the change in the cost of living.

We hasten to add that the indexed programs have many other features and raise many other issues beyond the narrow scope of a more accurate cost of living index. We also wish to express our view that these findings and their implications need to be fully digested and understood by the BLS, the Congress, the Executive Branch and the public.

I. Introduction[1]

Accurate measures of changes in the cost of living are among the most useful and important data necessary to evaluate economic performance. The change in the cost of living between two periods, for example 1975 and 1995, tells us how much income people would have needed in 1975, given the prices of goods and services available in that year, to be at least as well off as they are in 1995 given their income and the prices of goods and services available then. For example, if a family with a $45,000 income in 1996 would have needed $15,000 in 1976, the cost of living has tripled in the interim.

If the American economy was quite static, with very few new products introduced, very little quality improvement in existing products, little change in consumers' income, and very small and infrequent changes in the relative prices of goods and services, measuring changes in the cost of living would be conceptually quite easy and its implementation a matter of technical detail and appropriate execution. Fortunately for the overwhelming majority of Americans, our economy is far more dynamic and flexible than that. New products are being introduced all the time and existing ones improved, while others leave the market. The relative prices of different goods and services change frequently, in response to changes in consumer demand, and technological and other factors affecting costs and quality. Consumers in America have the benefit of a vast and growing array of goods and services from which to choose, unlike consumers in some other countries or our ancestors many decades ago.

But because the economy is complex and dynamic is no reason to bemoan the greater difficulty in constructing an accurate cost of living index. Major improvements can and should be made to the various official statistics that are currently used as proxies for changes in the cost of living, such as the well-known Consumer Price Index (CPI).

The Consumer Price Index measures the cost of purchasing a *fixed* market basket of goods and services. Based on surveys of households from some base period, the index sets weights (expenditure shares) for different goods and services. The weights reflect average or representative shares for the groups surveyed.[2] Keeping these weights fixed through time, the CPI is then calculated by attempting to measure changes from one month to the next in prices of the same, or quite closely related, goods and services.

---

[1] We would like to thank the staffs of the Bureau of Labor Statistics, Congressional Research Service, Congressional Budget Office, Senate Finance Committee and other individuals in academe and the private sector too numerous to mention here for valuable assistance and advice during the Commission's work.

[2] The two most commonly used measures are the CPI-U and CPI-W. The former is for all urban consumers, roughly 80% of the population; the latter is for urban wage and clerical workers, about 32% of the population. Note that the expenditure shares may be quite different than the average for any particular household, and also on average for subgroups of the population. Also, the prices paid for some products may differ for some households from the prices actually sampled. In principle, if not practice, a separate cost of living index could be developed for each and every household based on their actual consumption basket and prices paid. The overall index is used to approximate this with the data reflecting representative consumers. Whether this is itself sufficiently misleading as to warrant separate price indexes for different population subgroups is discussed below.

But through time consumption baskets change, in part because of changes in the relative prices of goods and services, and therefore the weights from the base period no longer reflect what consumers are actually purchasing. Representative purchases also change as discount coupons, buyers' clubs and other marketing devices determine the best value and alter buying patterns. This failure to adjust for the changes in consumer behavior in response to relative price changes is called *substitution bias.* It is a necessary result of keeping the market basket fixed. Because the market basket is updated only every decade or so, as we get further away from the base period, there is more opportunity for relative prices to diverge from what they were in the base period, and for consumption baskets to change substantially.

Just as there are changes in what consumers purchase, there are also trends and changes in *where* purchases are made. In recent years, there has been a transformation of retailing. Superstores, discount stores and the like now comprise a large and growing fraction of sales relative to a decade or two ago. As important as keeping up with the basket of goods that consumers actually purchase is keeping up with the outlets where they actually purchase them, so that the prices paid are accurately recorded. The current methodology suffers from an outlet *substitution bias,* which insufficiently takes into account the shift to discount outlets.

Many of the products sold today are dramatic improvements over their counterparts from years ago. They may be more durable and subject to less need for repair; more energy efficient; lighter; safer; etc. Sometimes, at least initially, a better quality product replacing its counterpart may cost more. Separating out how much of the price increase is due to quality change rather than actual inflation in the price of a standardized product is far from simple, but is necessary to obtain an accurate measure of the true increase in the cost of living. To the extent quality change is measured inaccurately or not at all, there is a *quality change bias* in the CPI.

The same is true with the introduction of new products, which have substantial value in and of themselves—not many of us would like to surrender our microwave ovens, radial tires, and VCR's—as well as the value of greater choice and opportunities opened up by the new products. To the extent new products are not included in the market basket, or included only with a long lag, there is a *new product bias* in the CPI.

Finally, in a dynamic, complex economy like the contemporary United States, there are literally many thousands of goods and services consumed. Price data are collected at a considerable level of disaggregation and how the price changes are aggregated into an overall index involves quite technical issues that can lead to a *formula bias* in the CPI.

Even if no federal program on either the outlay or revenue side of the budget were indexed, it would still be desirable to improve the quality of measures of the cost of living from the standpoint of providing citizens a better and more accurate estimate of what was actually going on in the economy, a way to compare current performance to our historical performance or to that of other countries. For example, the most commonly used measure of the standard of living is real income or output per person. To measure changes in

real income requires the separation of nominal income changes from price changes. Obviously, that requires an accurate measure of price changes. The Commerce Department uses the component indexes of the CPI as inputs in estimating inflation and real GDP, and thus some of the bias from the CPI is transmitted to the national income accounts.

But numerous federal, state and local government programs and tax features are "indexed" for changes in the cost of living by the changes in the Consumer Price Index. The CPI is also used to index, formally or informally, a large number of private sector contracts, including wages in collective bargaining agreements and rents, to name obvious examples that affect millions of Americans. Currently, slightly under one-third of total federal outlays, mostly in retirement programs, are directly indexed to changes in consumer prices. Several features of the individual income tax, including the tax brackets, are indexed; the individual income tax accounts for a little under half of federal revenues.

Congress indexed these outlay programs and tax rules in order to help insulate or protect the affected individuals from bearing the brunt of increases in the cost of living. Yet the Bureau of Labor Statistics, the agency responsible for compiling and presenting the Consumer Price Index, has explicitly stated for years that the CPI is *not* a cost of living index, presumably for some of the reasons mentioned above. If the Consumer Price Index as currently produced, and as likely to be produced over the next few years, is not an appropriate cost of living index for the task Congress had in mind, then it is desirable to consider alternative measures.

The consequences of changes in the Consumer Price Index overstating changes in the cost of living can be dramatic. For example, if use of the CPI is expected to overstate the increase in the cost of living by one percentage point per year over the next dozen years, the national debt would be about $1 trillion greater in 2008 than if a corresponding correction were made in the indexing of outlays and revenues.

This report proceeds as follows: Section II discusses the historical and prospective budgetary implications of changes in the CPI *overstating* changes in the cost of living. Section III presents an overview of how the CPI is actually constructed. Section IV details why the CPI is not a true cost of living index and discusses substitution bias. Section V describes in greater detail the current procedures employed by the BLS to adjust for quality change and presents a survey of the studies and the Commission's judgment on the bias from quality change and new products. Section VI summarizes the Commission's findings on the size of the bias by type, plus the range of plausible overall bias. Section VII discusses the issue of separate price indexes for different groups and of aspects of the quality of life that fall primarily outside the market based consumption focus of cost-of-living measures. Section VIII presents the Commission's detailed recommendations of ways to produce and to use more accurate cost-of-living measures. The Conclusion offers a brief perspective and some cautionary notes on the use of the findings of the Commission.

II. Indexing the Federal Budget

The issue posed for fiscal policy makers by an upward bias in the CPI has been stated with admirable clarity by the Congressional Budget Office (1994):

> The budgetary effect of any overestimate of changes in the cost of living highlights the possibility of a shift in the distribution of wealth. If the CPI has an upward bias, some federal programs would overcompensate for the effect of price changes on living standards, and wealth would be transferred from younger and future generations to current recipients of indexed federal programs—an effect that legislators may not have intended.[3]

Social Security is by far the most important of the federal outlays that are indexed to the CPI. However, Supplemental Security Income, Military Retirement, and Civil Service Retirement are significant programs that are similarly indexed. Other federal retirement programs, Railroad Retirement, veterans' compensation and pensions, and the Federal Employees' Compensation Act also contain provisions for indexing. The Economic Recovery Tax Act of 1981 indexed individual income tax brackets and the personal exemption to the CPI.

How important have the budgetary consequences of upward bias in the CPI been *historically?* Obviously, a precise answer to this question would require extended study, taking into account the timing of the bias, the parallel development of indexing provisions in specific federal outlays and revenues, and interest on the accumulation of debt that has resulted. An indication of the potential size of these effects can be inferred from one important historical example of one clearly identified source of bias. A careful study of this type, which focuses on the most important federal program affected by indexing, namely, social security benefits, has been conducted by the Office of Economic Policy (OEP) of the Department of the Treasury.

On February 25, 1983, the Bureau of Labor Statistics (BLS) introduced an important technical modification in the Consumer Price Index for All Urban Consumers (CPI-U). This altered the treatment of housing costs by shifting the costs for homeowners to a rental equivalent basis. The new treatment of housing costs was incorporated into the Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W), used to index social security benefits, in 1985.

The rental equivalent measure of housing costs was a conceptual improvement and has been retained in subsequent official publications. However, housing costs in preceding years employed a "homeownership" measure " . . . based on house prices, mortgage interest rates, property taxes and insurance, and maintenance costs."[4] The treatment of housing costs prior to 1983 was not modified in publishing the revised CPI-U, so that the new treatment of housing introduced a discrepancy in the conceptual basis for the

---

[3] See Congressional Budget Office (1994).
[4] See Gillingham and Lane (1982).

CPI-U before and after 1983. Similarly, housing costs in the CPI-W prior to 1985 have not been modified.

BLS developed an "experimental" price index, CPI-U X1, based on a rental equivalent *treatment* of housing extending back to 1967. This provides the basis for the OEP assessment of bias in the CPI-W. The bias for 1975, the first year that social security was indexed to the CPI-W, was 1.1 percent. This bias mounted over subsequent years, reaching 6.5 percent by 1982 and then declining to 4.7 percent in 1984.[5]

Overpayments of social security benefits resulting from the bias in the CPI-W mounted through 1983, reaching a total of $8.76 billion or 5.55 percent of benefits paid in that year. These overpayments have resulted in a lower balance in the OASI trust fund and a larger federal deficit and debt. OEP estimates interest costs associated with these deficits at the rate of interest paid or projected to be paid on the OASI trust fund. Beginning in 1984 interest costs predominate in the total. In the current fiscal year the total cost is $21.79 billion, of which $17.64 billion is interest. The cumulative effect of just this one source of bias in the CPI-W via this one program on the federal debt amounts to $271.0 billion, as of 1996.

In summary, the BLS made two decisions in revising the treatment of housing costs in the CPI-W in 1985. The first decision was to change the treatment of housing costs to a rental equivalent basis beginning in January 1985. The second was not to revise the treatment of housing costs for 1984 and earlier years. As a consequence of these two decisions the level of the CPI-W is 4.7 percent above the CPI-U X1, a measure of the cost of living based on the same primary data sources and similar methodology, but with a consistent treatment of housing costs.

The increases in federal outlays resulting from the bias in the CPI-W cannot be justified as cost of living adjustments. These increases are the consequence of an inappropriate treatment of housing costs before 1985 and have resulted in large transfers to beneficiaries of the OASI program that are devoid of any economic rationale. The overpayments have continued up to the present, but are declining in importance. However, the resulting decline in the OASI trust fund continues to mount due to rising interest costs and now contributes more than two hundred billion dollars to the federal debt.

Of course, nobody would suggest retroactively undoing the overindexing due to this or any other source of bias. The point of this discussion is to demonstrate how important it is to correct biases in the CPI as quickly and fully as possible before their consequences mount, indeed compound.

What would be the effect of an upward bias in the CPI on future budget deficits? More than half of federal spending of $1.5 trillion is now attributable to entitlements and mandatory spending programs. In January 1995 the annual Congressional Budget Office (CBO) outlook for the economy and the federal budget showed that this proportion is projected to rise to almost two-thirds of federal spending during fiscal year 1998. Cost-of-living adjustments at a projected rate of 3.0 percent will contribute $43 billion to total

[5] See Duggan, Gillingham, and Greenlees (1995).

spending on mandatory programs in that year and $80 billion in fiscal year 2000.[6] This is 6.8 percent of projected spending on mandatory programs in fiscal year 2000.

Testimony presented by the CBO to the Committee on Finance shows the impact of a hypothetical correction (reduction) of 0.5 percentage point in cost of living adjustments for fiscal years 1996–2000. Federal outlays would decline by $13.3 billion in fiscal year 2000, while revenues would rise by $9.6 billion. The decline in debt service resulting from reduced deficits in fiscal years 1996–2000 would be $3.3 billion, yielding a total contribution to deficit reduction of $26.2 billion in fiscal year 2000.[7] This is more than ten percent of the deficit projected by CBO in that year.

The CBO has provided the Commission with updated projections of the impact of hypothetical corrections (reductions) of 0.5 and 1.0 percentage point in cost of living adjustments for fiscal years 1997–2006.[8] With a reduction of 0.5 percentage point the total contribution to deficit reduction rises to $67.5 billion in 2006. Of this amount, an increase in revenue accounts for $22.3 billion and reductions in outlays, including debt service, amounts to $45.3 billion (of which debt service is $13.1 billion).

CBO projections for the impact of a hypothetical correction (reduction) in cost of living adjustments of 1.0 percentage point are, of course, even more dramatic. The total change in the deficit in the year 2006 is $134.9 billion. Federal revenues would be increased by $44.5 billion and federal outlays reduced by $90.5 billion; of the reduction in outlays $26.1 billion can be attributed to lower debt service and $64.4 billion to lower outlays on indexed programs.

Stated differently, if the change in the CPI overstated the change in the cost of living by an average of one percentage point per year over this period, this bias alone would contribute almost $135 billion to the deficit in the year 2006. That is one-third the projected baseline deficit (which assumes no policy changes such as the current balanced budget proposals). More remarkably, the upward bias by itself would constitute the fourth largest federal outlay program, behind only social security, health care and defense. By 2008, the increased deficit would be $180 billion and national debt $1 trillion. (See Appendix Figures A–1 and A–2 for related detail).

In summary, an upward bias in the CPI would result in substantial overpayments to the beneficiaries of federal entitlements and mandatory spending programs. In addition, such a bias would reduce federal revenues by overindexing the individual income tax. In short, the upward bias programs into the federal budget every year an automatic, real increase in indexed benefits and a real tax cut. Correction of biases in the CPI, while designed to adjust benefits and taxes for true changes in the cost of living more accurately, would also contribute importantly to reductions in future federal budget deficits and the national debt. These reductions can be attributed to higher revenues, lower outlays, and less debt service. Lower outlays—cuts in indexed federal spending programs and re-

[6] See Congressional Budget Office (1995).

[7] See O'Neill (1995). These CBO budget estimates are relative to CBO's January 1995 baseline and do not include the small adjust assumed in the out-years of the budget resolution.

[8] These estimates are relative to the CBO's May 1996 Baseline. See CBO (1996).

duced interest payments—account for over two-thirds of the long-run deficit reduction, while higher revenues account for the rest.

III. How the CPI is Constructed: A Brief Introduction

Knowledge of how the CPI is constructed is needed to understand the reasons that biases occur and the rationale for our recommendations for improvements and changes. This section provides a brief description of the BLS methodology highlighting the places where biases and key issues are likely to arise. We refer the reader to BLS documentation for more detail on data collection procedures and index construction methodology, as well as to recent articles by Armknecht (1996), and Shapiro and Wilcox (1996b).[9]

As could be inferred from the discussion above about the complexities of a modern dynamic capitalist economy, the CPI program is a complex and difficult undertaking. To make it manageable, the BLS applies a simplified view of the marketplace and consumer behavior. This simplified view is reflected throughout the CPI approach. It takes expenditures for a *fixed* market basket of goods and services at some point in the past, called the base or reference period, and estimates what it would cost today to purchase the same market basket. The formula used to construct the CPI, called Laspeyres, assumes that purchases are made in fixed quantities based on decisions from some previous period's experience. In other words, the CPI attempts to answer the question, "what is the cost, at this month's market prices, of purchasing the same market basket actually purchased in the base period?" Since the Laspeyres formula does not allow for the substitution of products or services in response to current prices and choices, it is an "upper bound" to a cost of living.

The market basket consists of total expenditures on items directly purchased by all urban consumers, that is, food, clothing, shelter and fuels, transportation, medical services and other goods and services that people buy for day-to-day living. The BLS uses scientific sampling techniques to select specific items. The BLS measures the price changes in these items over time. The sample design involves a multistage process for sampling by geographic area, retail outlet, item category, and individual goods and services within an outlet and category.

Several samples are used to try to make the CPI representative of the prices paid by consumers: urban areas selected from all U.S. urban areas, consumer units within each selected area, outlets from which these consumer units purchased goods and services, specific items—goods and services—purchased by these consumer units, and housing units in each urban area (for the shelter component of the CPI). The key sources of information used to determine the items which comprise the market basket and the outlets at which prices are to be collected are the Consumer Expenditure Survey (CES) and the Point-of-Purchase Survey (POPS).

[9] Chapter 19 of BLS Handbook of Methods. We are especially grateful to John Greenlees and Brent Moulton of BLS for clarifying several of these issues for the Commission.

Each month, prices for approximately 71,000 goods and services are collected from 22,000 outlets, in 44 geographic areas.[10] Separately, information is collected each month from about 5,000 renters and 1,000 homeowners for the housing components of the CPI.[11] The price quotations are combined, that is, aggregated, into the overall CPI. The determination of representative items to be priced, the procedure for collecting prices at the outlets, and the levels at which the prices are combined into indexes and the indexes are combined into higher aggregates, are all based on a fixed structure or system in which a number of key assumptions are embedded.

The item structure has four levels of classification beginning with major groups such as food and beverages, transportation, and medical care. The seven major groups are made up of 69 expenditure classes (EC's), for example fresh fruits (EC 11) and hospital and other related services (EC 57). The expenditure classes are in turn divided into 207 groupings called item strata, the lowest level at which indexes are constructed.[12] Two examples of item strata are apples, and nursing and convalescent home care. It is important to note that while the item categories are mutually exclusive and exhaustive for all consumer expenditures, this does not mean that new goods and services are automatically brought into the sample if they were not available during the reference period. It just means that every good and service can be classified within an existing stratum and there is no need to create a new stratum for when a new good or service is introduced. (This is made possible in part by numerous item categories called "other.")

Within each item stratum, entry level items, called ELIs for short, are defined. Indexes are not constructed at this level. Many stratum have only one ELI for example, Apples. The ELIs are the lowest level sampling units for items. They are the level of item definition at which the data collectors begin item sampling within each sample outlet. For example, prices for Brand "X" fever thermometers for babies, model 41303 41082, 4-3/10 inches long with plastic case, sold by "Y" Foods, Inc. in West Terre Haute Indiana, might be collected for Medical Equipment for General Use, ELI 55032, within Non-prescription Drugs and Medical Supplies, item stratum 5503, within Non-prescription Drugs, EC 55, within the Medical Care Commodities component of the Medical Care major expenditure category.[13] A schematic of the item structure is shown below in Figure 1.

---

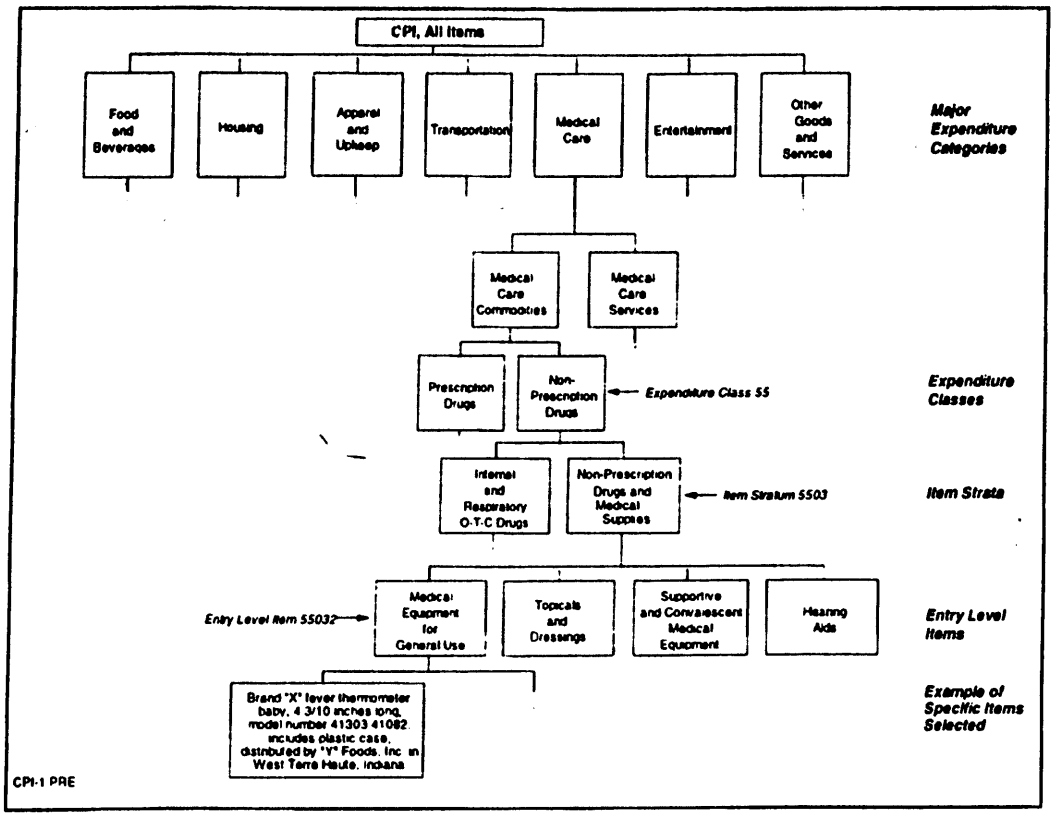
[10] Prices are actually collected in 88 locations, called primary sampling units, or PSUs. In eight PSUs (the five largest urban areas), prices are collected for all items every month. In other areas, prices are collected monthly for food, fuels and a few other items, and bi-monthly for all other items. Of the 44 areas which go into the index every month, 32 are self-representing because of their size and 12 are composites constructed from 56 PSUs which provide representation for smaller and mid-size cities across the country.

[11] The housing component properly captures the multiple period consumption accrued subsequent to the purchase of a house. This same approach is warranted but not used in measuring durable goods such as automobiles and refrigerators.

[12] Only 184 of the item groups are actually priced. The other 23 strata account for less than 2 percent of the weight of the overall index. Price indexes for these groups are moved with changes in the indexes for the item groups which are priced.

[13] This example is current and real, provided by BLS. The names of the brand and store are withheld to adhere to confidentiality requirements.

FIGURE 1
CPI HIERARCHY STRUCTURE



Outlets within geographic areas are sampled, too. The probability of selection for any given outlet is proportional to that outlet's share in total expenditures in the survey area for the item category in question. This is done so that the price quotes for selected items are obtained at outlets which are representative of the places that consumers made their purchases and also because the outlet is assumed to be an important characteristic of the purchase and component of price change. It follows from this assumption that differences in prices of the same item in different outlets must represent differences in aspects of the purchase such as quality of service or convenience of location and that consumers will pay the same proportional difference over time for these other aspects. When this assumption does not hold, such as when some outlets grow faster than others, the methodology will prevent adequate accounting in two ways: the current methodology will not adequately provide for obtaining more price quotes or give more importance to the more favored outlets, nor does it provide for direct comparison of the quality differences in purchasing the same item at two different places.

There is a process to "refresh" items and outlets sampled, called sample rotation, which generates a sample of specific items each of which had a probability of selection into the sample proportional to its share in recent consumer expenditures. Approximately 20 percent of the sample is rotated every year such that full rotation

takes 5 years. The items rotated in are not directly compared to those they replaced. The procedure assumes that at the time of rotation, the original item and the one rotated in have the same quality adjusted price.

BLS procedure provides for selecting alternative items to be priced when the previously priced item is sold out, discontinued or otherwise permanently unavailable. The field agent is given guidelines to use in selecting the replacement or substitute item within the same ELI and a judgment is made as to the comparability of the specifications. (However, there is no provision to assure that the replacement is the product which has taken market share from the one that has disappeared.)[14] When the substitute is determined to be non-comparable, BLS most often assumes that the quality difference accounts for the price difference, net of the price change since the last pricing period for similar items.[15] In some cases, attempts are made to measure the quality differences. Notice that it is the disappearance of an item which triggers the mechanism to price a substitute.

Prices of new goods not falling within an established stratum, which are introduced after the base period and therefore not in the reference market basket, are not given special preference in item substitution and sample rotation, and consequently are often not included in the index until the subsequent decadal revision.[16] (Moreover, the impact of new goods is not measured retrospectively because the CPI is not revised historically.[17]) Frequently cited examples of important new products which were not introduced until many years after their introduction are air conditioners and VCRs.[18] Cellular telephones will be included in the 1998 revision of the CPI.

While the methodology does not ensure the introduction of new products until the market basket is updated, improving the timeliness through more frequent updates of the market basket solves only part of the problem. Direct comparisons of the quality of new products with those with which they compete is often difficult. Furthermore, proper accounting of the impact of new products often requires comparisons with products in other item groups. The current item structure prevents the CPI from fully capturing the effects of a drug replacing surgery, of electronic information services replacing newspapers, of automobile leasing competing with purchases, of video rentals replacing cinema attendance. Over time, price changes in successful products will be given greater weight in the CPI, but full measurement of the price impacts across item groups is not possible when close substitutes are in different item

---

[14] There are other reasons that may result in the disappearance of a specific item from a specific outlet. When the reason is loss of competitive market share, the BLS replacement procedures are likely to result in upward bias.

[15] This procedure is based on the assumption that the marketplace adjusts fully and instantaneously to price differences among competing products. This can happen in two ways: prices of substitutes change immediately to make them equal (quality-adjusted), or quantities of what would be higher priced products fall to zero making them disappear from the market.

[16] The market basket has been updated once per decade historically and introduced with a several year lag. For current data, it represents an average derived from surveys for 1982–1984. The next revision of the index is scheduled for introduction in 1998 at which time the base period will be updated to 1993–1995.

[17] Only seasonal factors are revised historically.

[18] Sometimes a new product is introduced with sample rotation. An example might be a new variety of apples.

groups.[19] Although the item structure has several purposes, index estimation is the most important.

Prices for specific goods and services at specific outlets in specific locations are combined into item group-area indexes and these indexes are further aggregated by weighting them together either up through the item classification structure or by geographic area, to form a national CPI. The weights are derived from the Point-Of-Purchase Survey, the Consumer Expenditure Survey (which contains only modest detail) and from the statistical approach used in initiating specific commodities or services at the selected outlets. The design does not provide for collecting changes in quantities over time (since the market basket is assumed to be unchanged, this is not necessary to construct the CPI).

The use of arithmetic means to combine price changes within item groupings, for example different types of apples, implements the restriction that quantity weights do not change when prices change.[20] The arithmetic mean fails certain common sense tests, as discussed in the next section.

The greater the substitutability of the items whose prices are combined this way, the greater the resulting substitution bias in the index. An alternative to assuming no change in quantities is to assume no change in expenditure shares. This can be accomplished through the use of geometric means, which effects a price increase that is proportionally offset by a quantity decrease. For example, if a ten percent increase in the price of granny smith apples were associated with a ten percent reduction in the quantity purchased, geometric means would be the appropriate way to capture the market response. If there were no quantity change associated with the price increase then arithmetic means would be appropriate. In the case of granny smith apples, the availability of other varieties of apples may yield a market response to quantities that more than offsets the price increase. When this happens, the use of geometric means *understates* the market response.

It is worth noting that the published geographic indexes do not provide comparisons of the price level across geographic areas; rather, they provide comparisons of rates of change in the CPI. Clearly, if the rates of change are different, then the levels must also differ at some point. Indeed, the differences in levels would be of significant interest as a comparison of the cost of living across geographic areas. Yet the methodology does not provide such comparisons. Geographic areas play an important role in the sampling design, however, geographic area indexes as they are constructed today serve no other purpose than a step in aggregation, en route to a national CPI.

In summary, sampling techniques are used to determine which items are priced at which outlets. The methodology requires allegiance to the concept of a fixed market basket which by design does

[19] The ability of the CPI to fully capture price impacts such as these depends on the degree to which the classification structure is consumption-based. That is, items which are the closest substitutes for each other in terms of how they are used, must be in the same item group, the lowest level at which indexes are constructed. The item structure is updated with the decadal revisions of the CPI. The new item structure which will be introduced with the 1998 revision will make some improvements toward placing close substitutes together. Much more is needed.

[20] Another way to state this is that the elasticity of substitution among items within the lowest grouping, say types of apples, is assumed to be zero.

not change item category weights until the market basket is updated, historically every ten years or so, and hence fails to capture some new products. Neither does it make direct comparisons of the purchase experience at different outlets, by assumption not capturing the lower prices to which consumers respond making some outlets grow faster than others. The most detailed level at which price indexes are constructed is for 207 item groups for each geographic area. Geographic area price indexes are constructed to provide estimates of price change in specific geographic areas en route to the national CPI, but they do not provide inter-area comparisons of the cost of living. Price indexes are successively combined into broader categories until a national CPI is reached.

In conclusion, improving the CPI as a measure of the cost of living requires addressing a range of issues beginning with revisiting critical assumptions, adjusting resource optimization criteria, and abandoning the Laspeyres index formula. The Commission's recommendations are presented in Section VIII.

IV. The Consumer Price Index and a Cost of Living Index: Measurement Issues

A cost of living index is a comparison of the minimum expenditure required to achieve the same level of well-being (also known as welfare, utility, standard-of-living) across two different sets of prices. Most often it is thought of as a comparison between two points of time. As with any practical application of theory to index number production, estimating a cost of living index requires assumptions, a methodology, data gathering processes and index number construction.

There are two sets of potential biases in the CPI: biases relative to an "ideal" cost of living index and biases which arise within its own terms of reference. The strength of the CPI is in the underlying simplicity of its concept: pricing a fixed (but representative) market basket of goods and services over time. Its weakness follows from the same conception: the "fixed basket" becomes less and less representative over time as consumers respond to price changes and new choices.

Consumers respond to price changes by substituting away from products that have become more expensive and toward goods whose prices have declined relatively. As the world changes, they are faced with new choices in shopping outlets, varieties, and entirely new goods and services, and respond to these as well. These changes make the previously "fixed basket" increasingly irrelevant.

In trying to keep true to its concept in a rapidly changing world, the current CPI procedures encounter difficulties. Biases result when they ignore some of these changes such as the appearance of discounters, and also when they try to do something about them such as when items are rotated out of the sample and replaced with new items. Attempting to capture the changes in a way that tries to mimic the pricing of a "fixed basket" within a rather patchwork framework just cannot be done without introducing other problems into the resulting index. These different biases overlap and have been discussed under a number of headings: substitution

bias; formula bias; outlet substitution bias; quality change; and new product bias.

The "pure" substitution bias is the easiest to illustrate. Consider a very stylized example, where we would like to compare an initial "base" period 1 and a subsequent period 2. For simplicity, consider a hypothetical situation where there are only two commodities: beef and chicken. In period 1, the prices per pound of beef and chicken are equal, at $1, and so are the quantities consumed, at 1 lb. Total expenditure is therefore $2. In period 2, beef is twice as expensive as chicken ($1.60 vs. $0.80 per pound), and much more chicken (2 lb.) than beef (0.8 lb.) is consumed, as the consumer substitutes the relatively less expensive chicken for beef. Total expenditure in period 2 is $2.88. The relevant data are presented in Table 1. How can we compare the two situations? Actually, there are several methods, each asking slightly different questions and therefore, not surprisingly, giving different answers.[21]

Table 1
Hypothetical Example of Substitution Bias

	Price in Period 1	Quantity in Period 1	Price in Period 2	Quantity in Period 2	Price Relatives		Relative Weights	
					P2/P1	P1/P2	1	2
Beef	1	1	1.6	0.8	1.6	0.63	0.5	0.43
Chicken	1	1	0.8	2.0	0.8	1.25	0.5	0.57

The simplest comparison is to ask "How much more must I spend in my current situation (period 2) to purchase the same quantities that I purchased initially (in period 1)?"[22] This is the question asked by the CPI. The price index for period 2 relative to period 1 uses the initial period 1 basket of consumption as the weights in the computation. To buy 1 lb. of beef and 1 lb. of chicken in period 2 costs $2.40. The price index for period 2 relative to period 1 is 1.20 (2.40/2.00), that is a 20% increase.

Intuitively, it is easy to understand why such a computation imparts an upward (substitution) bias to the measure of the change in the true cost of living. It assumes the consumer does not substitute (cheaper) chicken for beef. In the real world, as in the hypothetical example, consumers change their spending patterns in response to changes in relative prices and, hence, partially insulate themselves from price movements.

An alternative approach would be to ask the question "How much more am I spending in my current situation (period 2) than I would have spent for the same goods and services at the prices that prevailed initially (in period 1)?"[23] This price index compares expenditures in period 2 ($2.88) with what it would cost to buy the current (period 2) market basket at the initial prices ($0.80 for the beef plus $2.00 for the chicken equals $2.80). This price index is 1.03, that is only a 3% increase. This approach understates the rise in the true cost of living as it overstates substitution.

[21] Each method has come to be named for its inventor. See below.
[22] This index is called the Laspeyres index.
[23] This index is called the Paasche index.

The idea underlying a cost of living index is to allow for the substitution that follows relative price changes. The question answered by a cost of living index is: "How much would we need to increase (or decrease) initial (period 1) expenditure in order to make the consumer as well off as in the subsequent period (period 2)." Although the answer to this question might appear to require detailed knowledge of a consumer's preferences, an excellent approximation can be obtained by using a "superlative" index formula instead of the traditional fixed weight index employed in the CPI.

The concept of a superlative index number was introduced by the American economist, Irving Fisher (1922), to describe index numbers that met certain reasonable criteria and thus agreed closely with his "ideal" index, described below.[24]

A major difficulty with a fixed weight index is the failure of time reversibility. This simple and intuitive requirement or "test" for an index number is that the index should remain the same if the underlying prices undergo a reversal. For example, suppose that the price of beef in Table 1 rises from 1.0 in Period 1 to 1.6 in Period 2, but then falls back to 1.0 in Period 3, reversing the change that took place between Periods 1 and 2. A fixed weight index increases by 60% between periods 1 and 2, but decreases by only 37.5% between periods 1 and 3, so that the increase in the "beef" index between periods 1 and 3 is 22.5% or 11.25% per period, rather than zero, as required for time reversal.

A geometric average satisfies the time reversal test, since it is based on the square root of the product of the price ratios between periods. In the example of the beef price from Table 1, the price ratio between Period 1 and Period 2 is 1.6, while the price ratio between Period 2 and 3 is .625. The product of these two price ratios is one, as required for time reversal, so that the average price increase is zero per period.

This concept was generalized by the Canadian economist, Erwin Diewert (1976), and used to describe any index number formula that provides a satisfactory approximation to an underlying economic index, such as a cost of living index.[25] The CPI is based on a fixed weight index formula that does *not* provide such an approximation, fails to meet these sensible criteria and worse yet is known to be biased upward. A superlative index requires the same information on prices and quantities as a fixed weight index, but involves interpolating between the two periods rather than treating one of them as the "base" period. There are two ways of doing this.

The first approach to interpolating between time periods is to use the geometric mean of the two fixed weight indexes—using the initial period and the subsequent period as "base" periods. The geometric mean is the square root of the product of the two indexes. This is the ideal index originated by Irving Fisher (1922) and now called the "Fisher ideal index" in his honor. In our example, this comes to 1.11, an 11% increase. By comparison the CPI-type fixed weight index, treating period 1 as the "base" period, is biased upward by 9% (1.20 minus 1.11). Alternatively, a fixed weight index with period 2 as the "base" period is biased downward by 8% (1.03

[24] See Fisher (1922).
[25] See Diewert (1976).

minus 1.11). The Fisher ideal index is employed by the Bureau of Economic Analysis in compiling data on the U.S. national income and product accounts.

An alternative approach to interpolation is to use a weighted average of the growth rates in prices with relative weights equal to the average of the weights in the two periods. This is called the "Tornqvist" index in honor of one of its originators, the Finnish statistician Leo Tornqvist (1936).[26] In our example, this is 1.10, a 10% increase. We conclude that the two superlative index formulas yield very similar approximations to the cost of living index. Estimates of the biases of the two fixed weight indexes are also similar. The BLS has compared a fixed weight index with the Fisher ideal and Tornqvist indexes to assess the bias in the fixed weight index as a measure of changes in the cost of living.

How large are substitution biases in the CPI? To answer this question we must take into account the hierarchical nature of the construction of the CPI described above. It is useful to focus initially on Upper Level Substitution Bias, which occurs when indexes for the 207 item groups and 44 areas are aggregated to form the CPI. The BLS uses a fixed weight index for this purpose (with weights derived from the Consumer Expenditure Survey (CES), a survey of household expenditure patterns), and hence ignores substitutions of chicken for beef, apples for oranges, etc. The BLS has measured this form of substitution bias by comparing a fixed weight index with an index generated by one of the interpolation methods we have described. Estimates are presented in Section VI.

The second type of substitution bias in the CPI is Lower Level Substitution Bias, which occurs when prices for the approximately 71,000 goods and services and information on housing costs are used to form indexes for the 207 items and 44 areas. This part of the index construction involves probability sampling with probabilities derived from the CES and the Point-of-Purchase Survey (POPS) of retail establishments in order to reflect the likelihood of purchases of individual items at specific retail outlets. It is useful to think of this as an alternative fixed weight index with probabilities playing the role of expenditure weights.

Since 1978 the formula at the lower level of index construction has been closely analogous to the fixed weight index at the upper level and assumes no substitution between commodities in response to price changes within the lowest level. For example, a change in the relative price of delicious and Granny Smith apples would lead consumers to substitute one for the other, but this is ignored with a fixed weight index at the lower level. In constructing a measure of the cost of living a fixed weight index at the lower level results in a bias. To assess the magnitude of this bias a geometric average based on the square root of the product of the price ratios between periods provides an appropriate standard for comparison. For example, Moulton and Smedley (1995) have compared the BLS fixed weight approach at this level with a weighted geometric approach.[27]

---

[26] See Tornqvist (1936).

[27] See Moulton and Smedley (1995).

Diewert (1995) has provided a detailed review of the properties of alternative approaches to index number construction at the lower level.[28] These include time reversal, as well as other reasonable requirements for index numbers at the lower level. Shapiro and Wilcox (1996b) have provided an elegant rationale for the geometric approach based on the correlation of relative prices over time.[29] Provided that this correlation is small, a modification of the geometric mean is approximately unbiased for the underlying cost of living index, and this characterization does not require information about the underlying system of consumer's preferences.

Modified geometric means have been widely used as a standard for evaluating methods for index number construction at the lower level. Diewert (1995) gives a useful review of the empirical studies. In addition to the work of Moulton and Smedley (1995), Carruthers, Sellwood, and Ward (1980) have conducted a study of this type for the U.K., Schultz (1994) for Canada, Dalen (1994) for Sweden, and Woolford (1994) for Australia.[30] These studies show that fixed weight indexes, like those used by BLS, are biased upward; the order of magnitude of the bias is similar to that suggested by the study of Moulton and Smedley (1995) for the U.S. These problems have led an increasing number of statistical agencies, such as Statistics Canada, to follow Irving Fisher's (1922) advice and jettison the arithmetic mean in favor of the geometric mean.

A relatively subtle problem developed in implementing the fixed weight index at the lower level. When sample items are replaced by substitute items for which no previous price observations are available, base period prices for the substitute items must be "imputed" to fill this gap. The procedure adopted by BLS for doing this had the effect of linking the weights for the substitute items to the prices used in the CPI and produced a bias that is an important component of Lower Level Substitution Bias. This problem also arises during rotations of items included in the sample of 70,000 prices for goods and services and the sample of housing costs.

An estimate of the overall Lower Level Substitution Bias is given by the difference between the fixed weight index and a geometrically weighted average, where the fixed weight index is based on the methods for price imputation introduced by BLS in 1978. In 1995 and 1996 BLS introduced new procedures based on "seasoning" the price estimates. Seasoning involves lengthening the period between a price imputation and the period when an item is actually introduced into the CPI. By lengthening this period the link between weights and prices for individual items can be broken and the bias reduced. However, the bias associated with the fixed weight formula remains.[31]

Our Interim Report anticipated that what we called "formula bias" and now refer to as Lower Level Substitution Bias would be eliminated by BLS. The BLS did alter its procedures by introducing "seasoning" where appropriate; while this eliminated bias due to methods for price imputation, it did not affect the bias due to the

---

[28] See Diewert (1995).

[29] See Shapiro and Wilcox (1996b).

[30] See Carruthers, Sellwood, and Ward (1980), Schultz (1994), Dahlen (1994), and Woolford (1994).

[31] For detailed discussion of Lower Level Substitution Bias, see: Moulton (1996).

use of a fixed weight formula at the lower level. Accordingly, we have recommended below that BLS should replace the fixed weight formula by a geometrically weighted average. This has been tested and found to be feasible in an important study by Moulton and Smedley (1995).[32] Additional work is currently underway to extend the period of this study.

The introduction by BLS of a fixed weight index at the lower level of aggregation was viewed at the time as introducing consistency of indexing at both upper and lower levels of aggregation. However, the disadvantages of the fixed weight approach at the upper level carry over to the lower level. A superlative index formula is required to provide a satisfactory approximation to the underlying cost of living index at the upper level. This avoids the bias associated with the fixed weight index formula employed in the CPI. Similarly, lacking quantity or expenditure information at the lower level, a good approximation to the underlying cost of living index is obtained from a geometrically weighted average formula.

Just as consumers change the goods they purchase in response to changes in relative prices, as in the beef and chicken example, so do they change the location where they make their purchases. The opening of a new discount store outlet may give consumers the opportunity to purchase at a lower price than before. At present, the CPI procedures ignore such reductions that occur when consumers change outlets. However, if consumers cared only about obtaining goods at the lowest price, then we would observe all goods sold at the same price at all outlets. Instead, we observe low prices at discount stores and warehouse clubs at the same time as medium prices at supermarkets and higher prices at convenience stores. Evidently, consumers care not only about prices, but the level of services such as availability of clerks, wrapping services, and the distance between home and alternative outlets.

Current procedures in the CPI ignore price changes when consumers switch outlets. This incorporates into the CPI the implicit assumption that price differentials among outlets entirely reflect the differences in service quality. This approach would be legitimate if the economy stood still with a stable set of outlets providing alternative levels of service quality. However, there has been a continuous increase in the market share of discount stores as more efficient technologies of distribution allow low price outlets to expand while older, higher priced outlets have contracted and in some cases gone out of business. This shift in market share indicated that many consumers respond to price differentials and do not consider them to be fully offset by differences in service quality. Completely ignoring all differences in service quality by incorporating all such price reductions into the CPI would err in the opposite direction. Further research is required to disentangle true changes in prices from changes in service quality. This problem is analogous to the need to disentangle the changes in prices from changes in product quality.

Quality change and new goods present the most difficult problems for measurement. They include capturing the introduction of new products in a timely manner; making direct quality compari-

[32] See Moulton and Smedley (1995).

sons of new products with existing ones; making direct quality comparisons of new products with other products against which they compete (in other classification groupings such as a new drug and the surgical treatment it replaces); and capturing the combined impact of quality and substitution as these new products displace others within and across their classification grouping.

A well-known expert on price indexes has stated the general issue clearly: " . . . heterogeneity in economics pertains to transactions, and not just the physical description of the product. Whenever two transactions involve different bundles of explicit or implicit attributes, they differ qualitatively. Differences in terms of sale, services provided with the sale, . . . are exactly identical from the economics of the matter, to physical changes that we normally think of as "quality change" (Triplett (1990)).

For example, it is not just what is purchased where (and how), but possible also when that matters. There may also be a time of week bias. The BLS does not collect prices on weekends and holidays when certain items and types of outlets disproportionately run sales.[33] There appears to have been a sizeable increase in the fraction of purchases made on weekends and holidays perhaps reflecting the increased prevalence of two earner families. We know of no systematic study of this issue and urge the BLS to conduct the research necessary to examine it thoroughly, perhaps with scanner data.

A full treatment of these issues reinforces the problem of focusing on the "average" or "representative" consumer. Different consumers have different tastes and time costs, and hence value the appearance of new outlets and new products differentially, with some (the majority) becoming better off with supermarkets and others losing out as the corner grocery store disappears. The CPI is not equipped to account for special characteristics of different consumers or groups of consumers.[34] [35]

There are still other issues that would in principle apply to obtaining a true cost of living index (COLI). Consider two examples: the negative effects of higher crime rates and the concomitant purchases of security devices and higher insurance premiums and the positive effects of improvements in information technology that permit a parent to work at home when a child is ill. Surely these would enter a calculation of "the minimum expenditure necessary to be at least as well off." Section VII below explores some of these problems.

V. Quality Change and New Products

Introduction

The difficult questions posed by quality change and the continuing arrival of new products have been called the "house-to-house combat of price measurement." In this section we will treat new product bias as a component of quality change bias and will not at-

---

[33] This undoubtedly arose both to save overtime pay and to accommodate retailers who prefer BLS price-takers to be in their establishments when they are less crowded.

[34] See Jorgenson and Slesnick (1983).

[35] See Fisher and Griliches (1995).

tempt to break down our overall bias estimate into the separate contributions of quality change bias and new product bias.

Quality changes have occurred at a rapid rate for some products but not others. The CPI has done a better job capturing the effect of quality change for some products than others. The CPI has introduced some new products faster than others. Because the magnitude of quality change bias differs so much across product categories, any overall evaluation of the magnitude of quality change bias must be conducted "down in the trenches," taking individual categories of consumer expenditure, assessing quality change bias for each category, and then aggregating using appropriate weights.

Further complicating the analysis is that quality change bias, assessed at the level of individual products, appears to have changed significantly over time. For instance, important improvements in BLS methodology largely or entirely eliminated an upward bias in the CPI for new automobiles prior to the mid-1960s and a downward bias for apparel after the mid-1980s. Likewise, an important source of downward bias in the CPI rent index was eliminated in the late 1980s.[36]

Previous evaluations of quality change bias, e.g., Shapiro and Wilcox (1996c) and Lebow, Roberts, and Stockton (1994) have tended to take bias estimates from earlier research on particular products, e.g., consumer appliances or automobiles, apply that bias estimate with the weight of those products in the CPI, and assumes that in the rest of the CPI the rate of quality change is zero. We do not view that approach as likely to emerge with a neutral evaluation of the bias, simply because the evaluation that the rest of the CPI is unbiased represents an extreme one-sided answer to the question as to whether the components of the CPI subject to relatively little research are biased. They may be as likely to be subject to the average rate of bias of those components which have been subject to careful research as to no bias at all. In this section we evaluate the CPI component-by-component and extrapolate research on bias from one category to another when the categories seem related. Nevertheless, we attribute bias estimates of zero to a number of categories which seem quite dissimilar to those categories subject to intensive research, or where unmeasured quality change and new products have been relatively unimportant.

While the problem of bias due to quality change and new products can be largely separated from the other forms of bias considered above, this is not entirely possible. Evidence on quality change bias developed in other studies, for instance Gordon (1990), is based on an attempt to measure prices directly from sources independent of BLS price quotations, using such sources as mail order

[36] There is no presumption that the magnitude of upward quality bias has declined over time. One consideration is that the growing importance of such hard-to-measure categories as consumer electronics and medical services may have increased the significance of quality change bias in the past decade. Another problem is suggested by a "thought experiment" recently conducted by Nordhaus (1996), who extrapolated backward substantial upward bias in the CPI over a long period of 190 years and arrived at implausibly low estimates of the standard of living of the average U.S. citizen in the year 1800. The implausibility of continuous upward bias in earlier decades at the rate suggested for recent decades in this report implies that in some earlier era the upward bias in the CPI was substantially less. This, of course, is natural. Long ago more was more important than better, e.g., enough to eat was more important than variety. As incomes rise beyond some point it is natural to expect increased demand for quality in many goods and services. We return to this issue below in our discussion of housing prices.

catalogues and Consumer Reports.[37] However, any differences between these independent indexes and the CPI for the same goods may reflect not just quality change and new product bias, but also traditional substitution bias (since the mix of products and models shifts faster in the alternative source than the CPI), outlet substitution bias (since alternative price quotes are often an average of market prices which adjusts for the changing mix of discount stores), and formula bias (since the alternative indexes are free from the formula bias problems discussed previously).[38]

Conceptual Issues

The difficulty created by quality change in existing products, and by the introduction of new products, is highlighted by returning to the definition of a cost of living index: a comparison in two time periods of the minimum expenditure required to achieve the same level of well-being. What does the "same level" mean when the models of a given product available in the second time period embody different quality attributes than in the first time period? And, an even more profound difficulty, what does the "same level" mean when entirely new products are introduced that were unavailable in the first time period?

A pervasive phenomenon called the "product cycle" is critical in assessing the issue of new product bias in the CPI and applies as well to new models of existing products. A typical new product is introduced at a relatively high price with sales at a low volume. Soon improvements in manufacturing techniques and increasing sales allow prices to be reduced and quality to be improved. For instance, the VCR was introduced in the late 1970s at a price of $1,000 with clumsy electromechanical controls; by the mid 1980s the price had fallen to $200 and controls were electronic, with extensive preprogramming capabilities. Later on in the product cycle, the product will mature and eventually will increase in price more rapidly than the average product of its class. The sequence is easily visualized as a "U"-shaped curve—the price of any given product relative to the consumer market basket starts high, then goes down, is flat for a while, and then goes back up. To the extent that the CPI overweights mature products and underweights new products, it will tend to have an upward bias. Some recent academic research, notably Berndt, Griliches, and Rosett (1993) and Berndt,

---

[37] Consumer Reports since 1959 has based its quotation for the average price of a particular model on an average of prices obtained in a market survey, typically of 15–20 price quotations obtained across the country. Thus a shift to discount stores would show up in the Consumer Reports price quotations and account for part of the difference between Gordon's indexes and the CPI index for the same category.

[38] As a further example of the difficulty of disentangling quality change from other sources of bias, one important fact to be assessed is the relatively large recent difference between the growth rates of the CPI and deflator for Personal Consumption Expenditures (hereafter PCE deflator). The PCE deflator is part of the National Income and Product Accounts, which is the responsibility of the Bureau of Economic Analysis, a division of the Department of Commerce. Over the 12 months ending in August, 1996, the CPI rose at 2.9 percent per year, while in the four quarters ending in 1996: Q3, the implicit PCE deflator rose at 2.0 percent per year. An alternative PCE deflator excluding expenditures on personal computers and medical care increased at 2.7 percent per year, suggesting that much of the difference between the PCE deflator and CPI over this period may involve their differing treatments of computers and medical care. This calculation was made by Bob Arnold of the CBO by subtracting nominal and real expenditures on medical care and consumer purchases of computers from total personal consumption expenditures. But that is not just an issue of differing treatment of quality change, but also substitution bias, since consumer expenditures on both personal computers and medical care have increased greatly since the 1982–84 base year of the CPI.

Cockburn, and Griliches (1996), computes alternative price indexes with the mix of prescription drugs actually sold and the limited and older sample contained in the CPI, and this research attributes a significant upward bias to the CPI on the grounds of its lateness in introducing the mix of models and varieties actually sold. Another recent study, Dulberger (1993) demonstrated the upward bias in price indexes produced as a result of delayed introduction of new products, and further showed that the magnitude of the bias increased with the length of the delay.[39]

An important criterion for the assessment of quality change and new product bias is the evolution of market shares for particular models and products. When a new model is introduced that is more expensive than an old model, but it gains market share, we can conclude that it was superior in quality to the old model by more than the differential in price between the two.

The same criterion helps us deal with outlet substitution bias. When consumers shift from traditional supermarkets to new, more expensive specialized food markets offering an improved selection or variety of produce, we can deduce that consumers are better off. The fact that Wal-Mart both charges lower prices and has become by far the largest retail chain over the past 15 years indicates that consumers do not consider the lower Wal-Mart prices to be offset by inferior service, as implicitly assumed by the CPI, but rather that consumers view Wal-Mart to offer a superior combination of prices and service to the previously available mix of outlets. The fact that convenience stores like 7–11 both charge higher prices and have gained market share indicates that consumers view convenience stores as providing a value of extra convenience that is worth more than the extra price that they charge. Many consumers shop at both Wal-Mart and convenience stores, paying both lower and higher prices on particular items than with the previous mix of stores, and the shift in market share suggests that the new mix is an improvement. The same evaluation can be made of restaurants, where consumers have shifted toward low-priced fast food outlets like McDonalds, medium-priced franchises like Olive Garden and Red Lobster, and in some urban areas, sophisticated high-priced restaurants specializing in Tuscan, Thai, and other ethnic food specialties. An important strand of academic research on such diverse products as medical imaging devices (Trajtenberg, 1990) and breakfast cereal (Hausman, 1996) attributes substantial value to increases in product variety. Thus, the "value of variety" is critically important in our assessment both of outlet substitution bias and, in this section, of quality change and new product bias.

BLS Methodology

Our discussion of quality change and new product bias begins with a review of the methods used by the CPI to handle quality changes in existing products and then turns to problems posed by new products. The BLS has five different methods to cope with a model change for an existing product.

[39] Dulberger studied electronic components which are inputs to many consumer electronics products including personal computers, automobiles and home security systems.

- The "direct comparison" method treats all of the observed price change between the old model and the new model as a change in price and none as a change in quality. There is no necessary bias, because quality can decrease as well as increase. But in practice most goods tend to undergo steady improvement, and often a better model is introduced with no change in price, causing the quality change to be missed entirely.
- The "deletion" method makes no comparison at all between the prices of the old and new model. Instead, the weight attributable to this product is applied to the average price change of other products in the same commodity classification. To the extent that the deletion method is used, the CPI consists disproportionately of commodities of constant quality which may be further along in the product cycle.
- The "linking" method can be used if the new and old model are sold simultaneously. In this case the price differential between the two models at the time of introduction of the new model can be used as an estimate of the value of the quality differential between the two models. As indicated above, this can lead to an understatement of quality change if the new model gains market share. Also, a quality improvement in the new model can occur even if it costs less or the same as the old model, as in the case of the VCR where the price fell continuously while programming capability and reproduction quality improved.
- The "cost estimation" method attempts to establish the cost of the extra attributes of the new model. Problems in practice with the costing method have been its infrequency of use, and the fact that it has been applied disproportionately in the case of automobiles relative to other products. This raises the possibility that there is a spurious upward "drift" in the relative price of other products relative to automobiles due to an uneven application of the costing method. An emerging source of upward bias is that products like automobiles are benefitting from the improved quality of materials like steel (which does not rust as it once did) and tires (which last many more miles). To the extent that some of these inputs to the auto production process are experiencing quality improvements of their own in excess of differences in cost, these will not be picked up by the BLS cost-based quality estimation procedure.
- Thus far, the CPI has introduced only in its apparel category an alternative methodology called the "hedonic regression method" for estimating the value of quality change.[40] The hedonic approach can be viewed as an alternative method to manufacturers' cost estimates in making quality change adjustments. It assumes that the price of a product observed at a given time is a function of its quality characteristics, and it estimates the imputed prices of such characteristics by regressing the prices of different models of the product on their differing embodied quantities of characteristics. Thus the hedonic approach is less a new method than an alternative to cost esti-

[40] Dulberger compared the rates of change in price indexes constructed from the same sample, using hedonic techniques and conventional linking. Her results, for the period 1972–1984 showed an average annual price decline for computing equipment that was 9 percentage points lower using the hedonic approach.

mates to be used when practical factors make it more suitable than the conventional method.

By their very nature hedonic indexes require large amounts of data. Given the thousands of separate products that are produced in any modern industrial society, the need to collect a full cross-section of data on each product presents a substantial obstacle to the full-blown adoption of the hedonic technique. But in many cases the data already collected by CPI field agents can be used for hedonic regression analysis; this is already done in the case of apparel.

Another possible objection is that it is impossible to construct a hedonic index in the timely fashion required by the CPI, with its orientation to producing within a few weeks an estimate of month-to-month price changes that can never be revised. But this ignores the fact that coefficients can be estimated on the basis of historical data, and these previously estimated coefficients can be used to evaluate quality change when a new model is introduced. This approach would be particularly suitable for product categories subject to a rapid succession of new model introductions, notably TV sets and personal computers.

This list of BLS methods reveals at least four potential sources of upward bias: the use of the direct comparison method that does not address the quality issue at all, the use of the deletion method that bases price change on models that are unchanged in quality and may be further along in the product cycle, the use of the linking method when quality improvements are greater than the price differential across models, and the use of the cost method which may miss quality improvements achieved by those firms which supply better materials and inputs to producers of final goods.

A potentially greater difficulty is that the CPI makes no attempt to create systematic estimates of the value of quality improvements which increase consumer welfare without raising the price of products. For instance, many consumer electronic products and household appliances have experienced a reduction in the incidence of repairs and in electricity use, and few if any of these improvements have been taken into account by the CPI. The increased longevity of automobiles (cited below), appliances, and other products introduces a similar source of bias.

New Products Bias

We turn now to the issue of new product bias. There is no debate regarding the reality of the product cycle, and nobody debates the fact that the CPI introduces many products late, thus missing much of the price decline that typically happens in the first phase of the product cycle. An extreme example involves room air conditioners, which were widely sold in 1951, but not introduced into the CPI until 1964, 13 years later. More recently, the microwave oven was introduced into the CPI in 1978 and the VCR and personal computer in 1987, years after they were first sold in the marketplace. As an even more contemporaneous example, there are currently 36 million cellular phones in use in the United States, but as yet the CPI has no price index for cellular phones. Thus none of the benefit to consumers of being able to keep track more easily of children, spouses, or of aged parents has yet received any credit

in our national measures of inflation, real output, or productivity. Even more recently, there are more than 40 million cellular phone subscribers in the U.S., but the cellular phone has yet to be introduced into the CPI.[41]

A second aspect of new product bias results from a narrow definition of a commodity. When a new product is finally introduced into the CPI, no comparison is made of the price and quality of the new product with the price and quality of an old product that performed the same function. For instance, people flock to rent videos, but the declining price of seeing a movie at home, as compared to going out to a theater, is not taken into account in the CPI. Similarly, the CPI missed the replacement of electric typewriters by electronic typewriters and then PCs with word-processing and spell-checking capability, or CD-ROM encyclopedias that cost far less than old-fashioned bound-book versions and eliminate many trips to the library. Inevitably, however, many new products embody genuinely new characteristics that have no previous counterpart. Electronic mail that provides a new set of bonds and communication between parents and their children who are off at college and cellular telephones that make possible virtually continuous contact with a sick child or aged parent are but two examples.

This discussion of new products leads inevitably to deeper questions about changes in the standard of living of the average American. Positive changes made possible by consumer electronics need to be weighed against increasing crime, pollution, and other "bads." We return to these issues in Section VII below.

Quality Change and New Product Bias by Product Category

Because quality change bias differs in magnitude, direction, and timing across product categories, the only way to narrow the range of uncertainty of the magnitude of quality change bias is to examine the available evidence, category by category. Table 2 is designed to provide a guide to this assessment. The left-hand column lists each major product category within the CPI next to its "relative importance," i.e., percentage weight, in December, 1995. In this section we review the available evidence on bias related to quality change and new products, by category.

In some categories there is little if any published evidence that allows us to reach a determination. However, we do not follow previous research by assuming that in these categories the overall bias due to quality change and new products is necessarily zero. Instead, we discuss the likely direction of bias in the context of the definition of a cost of living index: a comparison in two time periods of the minimum expenditure required to achieve the same level of well-being.

1. Food and beverages. The most dramatic evidence of upward bias in the food and beverages category was produced by Reinsdorf (1993), who found during the 1980–90 period an annual rate of change of average price paid for 50 narrowly defined commodities that was fully 2.0 percent per annum slower than the CPI for the same product categories. While Reinsdorf thought at the time that

[41] The number of cellular phone subscribers is taken from Business Week, December 2, 1996, p. 104.

this difference reflected outlet substitution bias, in fact he later concluded that the difference represented a mix of formula bias and outlet substitution bias. Whatever the interpretation of Reinsdorf's study, it does not represent evidence on quality change, since his commodities were chosen to be identical to those priced in the CPI.

Besides his study, there is little if any published evidence on the food category, other than Hausman's (1996) attempt to establish the value for the introduction of a new variety of breakfast cereal. Perhaps more important than new varieties of packaged goods has been a wave of technological improvements that has greatly increased the variety of fresh fruits and vegetables available in the typical supermarket during the winter months, and a trend toward more services provided in supermarkets, eliminating the need to travel to small specialty shops, especially fresh fish markets and deli counters preparing fresh-cooked food. How much would a consumer pay to have the privilege of choosing from the variety of items available in today's supermarket instead of being constrained to the much more limited variety available 30 years ago? A conservative estimate of the value of extra variety and convenience might be 10 percent for food consumed at home other than produce, 20 percent for produce where the increased variety in winter (as well as summer farmers' markets) has been so notable, and 5 percent for alcoholic beverages where imported beer, microbreweries, and a greatly improved distribution of imported wines from all over the world have improved the standard of living. Increased variety and convenience in food away from home, in every price category from McDonalds to luxury restaurants (as discussed above), can also be credited with a 10 percent premium. The annual rates of bias in Table 2 are calculated by converting these assumed premia to annual geometric growth rates over the past 30 years.

2. Housing. By far the largest single weight in the CPI is given to housing component, and within that to shelter. The shelter component shifted to a rental equivalence approach in 1983, and the CPI-U-X1 index represents an attempt to provide a consistent treatment of housing using the rental equivalence concept back to 1967. The annual rate of change of the CPI shelter index exceeds that of the CPI residential rent index by 2.33 percent per annum from 1967 to 1983, and correspondingly the annual rate of change of the official CPI-U exceeds that of CPI-U-X1 by 0.52 percent per annum over the same interval.[42] The BLS has also shifted methodology in 1995 to correct formula bias and in 1988 to correct an "aging bias" that resulted from pricing in successive periods housing units that were becoming progressively older. Randolph (1988) estimates this pre-1988 aging bias at 0.3 percent per annum, a concept that represents the effect of depreciation net of any maintenance and renovation expenditures.

First, we register our skepticism that the Randolph aging bias should be considered a bias in its entirety. Older units rent for less than new units for two reasons. First, they may physically deteriorate by more than is offset by repairs and maintenance. But, second, they may lose value as newer units come on the market containing amenities such as central air conditioning. Such economic

[42] All annual growth rates calculated in this report are logarithmic.

obsolescence does not represent a decline in the quality of the service provided by the older apartments, but rather represents the result of the fact that the income elasticity of demand for shelter amenities is positive, and people expect higher quality in apartments and houses as the nation's per capita income increases. An exact analogy is the introduction of the jet plane, discussed in detail by Gordon (1990). The quality of the ride on a propellor-driven DC-7 did not decline when the pure-jet DC-8 was introduced in 1958. Rather, consumers valued the ride on the jet plane so highly that the demand for flights on the DC-7 vanished. The DC-7 was scrapped prematurely, within five to ten years after the introduction of the jets. Consumers gained the entire surplus from the transition from propellor to jet planes for long-distance air travel, and the declining rents of older apartments represent a less dramatic example of the same phenomenon. Thus far there has been little investigation into quality change in the apartments included in the CPI rent survey. The "CPI methods hold most housing quality constant by measuring rent changes longitudinally for a cross-section of housing units" (Randolph, 1988, p. 359). That is, rent changes on a given unit are followed through time, and alternative units are rotated in, with the overlap handled by deletion. If there is a general tendency for more recently constructed units to have more and better appliances, central air conditioning, and other amenities that were not present in previous decades, there is the possibility of an upward bias in the CPI rental index if consumers value these amenities at more than their extra cost. The continuous movement of households to newer apartment complexes in suburbs and in the Sunbelt may be part of a process by which housing quality steadily improves. The "market share" test suggests that many households prefer new sunbelt apartments to older types of apartment in central cities in the north central and northeastern states.

The U.S. Census Current Housing Reports report median monthly rent of all rental occupied units. The ratio for 1993 to 1976 is 2.92 ($487/$167). The CPI rental index ratio (not adjusted for formula or aging bias) for the same years is 2.46. The implied annual difference in growth rates for the CPI is −1.00 percent per year. An alternative comparison for 1973–88 yields a difference of −1.10 percent per year.[43]

While only limited data are available on the quality of rental units, there is evidence that rental units have improved in quality at approximately the pace of owner-occupied units, for which more data are available. Two key measures have persuaded us of the comparability of rental and owner units (the CPI uses rent indexes for both the rental and owner-occupied segments of housing, so these findings support the CPI choice). First, between 1970 and 1993 the mean number of rooms increased by 9.7 percent in all occupied units (of which about 1/3 were rental units), while the mean number of rooms in rental units increased by a similar 7.8 percent. Perhaps more important, the number of rooms per person increased by 30.2 percent in all units and 27.0 percent for rental

[43] Brown's (1994) compilation of annual rent paid by a working class household yields $5160 for 1988 (Table 7.8a) and $1803 for 1973 (Table 6.8a), for a ratio of 2.86. The ratio of the CPI unadjusted rent index for the two years is 2.43.

units.[44] This set of comparisons supports the view that quality has improved at approximately the same rate in rental and owner-occupied units, and that we can use some of the available data on the totality of occupied units to reach a judgment on the extent of quality change.

While the best data are available for newly constructed units, some important data are available for the entire stock of existing units. For the entire stock of existing rental units alone, the mean number of bathrooms increased by 23.3 percent between 1970 and 1993. And for the entire stock of all units, the fraction containing central air conditioning increased from 10.8 to 41.7 percent.

Further indication of the change in quality standards is indicated by changing characteristics of new single-family houses completed in 1993 compared to 1976: median square feet increased by 30 percent, bathrooms from 2.0 to 2.4, percentage with central air conditioning from 49 to 78, percentage with one or more fireplaces from 45 to 63, and percentage with a garage from 72 to 84.[45]

We have already determined that between 1976 and 1993 the average rent paid in the U.S. increased 1.0 percent faster than the CPI rent increase. To conclude that the CPI is unbiased, we would have to determine that the quality of the average rental unit increased by 1.0 percent per year over that period, or 18 percent over the entire period. From the evidence we have examined, we believe that 20 percent is a low-end estimate of the increase in the average size of apartments, which would support the conclusion that the average rent per square foot has increased no faster than the CPI. But also, we find convincing evidence that the average quality of apartments per square foot has increased as well. The transition to central air conditioning proceeded at a rapid rate during the past two decades. Other amenities were added which increased the average quality of apartments, particularly swimming pools, health clubs, on-site free parking, and climate (since the mix of apartments shifted toward southern climates which reduced the impact of winter weather on tenants, particularly older tenants).

For the period since 1970 we find it plausible that the CPI accurately measures rent per square foot of apartment space, but its measure of shelter rent is upward biased by neglecting the increase in the quality of apartments per square foot. It is entirely natural that an increase in per-capita income would spill over into increased quality of housing, because there is no reason why housing size and quality should have an income elasticity of zero. The improved quality of appliances documented by Gordon (1990) applies to the shelter sector, since most apartments are now provided with relatively recent refrigerators, stoves or oven/cooktop combinations, dishwashers, and garbage disposals. The rental equivalent of these appliances must be substantial and they have been included in both new and older apartments mainly since 1955–60. A conservative estimate is that the total increase in apartment quality per square foot, including the rental value of all appliances, central air conditioning, and improved bathroom plumbing, and other amen-

[44] All citations regarding housing quality in this section refer to the U.S. Census of Housing, 1970 and the U.S. Census Bureau, American Housing Survey, 1993.

[45] Data from Statistical Abstract of the United States. 1978 issue, Tables 1371 and 1373. 1995 issue, Tables 1214 and 1230.

ities, amounts to 10 percent over the past 40 years, or 0.25 percent a year. Accordingly, Table 2 records an upward bias in the CPI of 0.25 percent per year for the shelter component, and this well may be an understatement.

For years before 1973, there is some evidence that the CPI rent index may be biased downward by more than can be explained by changes in quality. For instance, average annual rental expenditures for working class families in the CES increased from $444 in 1950 to $1803 in 1973, a ratio of 4.06, while the equivalent ratio for the unadjusted CPI rent index is only 1.93. This translates into a slower annual growth rate of the CPI of −3.24 percent per year. The same comparison for 1918 to 1950 yields an annual difference of −2.82 percent per year.[46] Without a measure of annual quality change per year, we cannot make a judgment on the magnitude of the bias, but the possibility that the CPI rental index incorporates a substantial downward bias prior to 1973 may help to explain the "Nordhaus thought experiment problem" identified above, namely that backward extrapolation of substantial CPI bias for a century or more yields implausibly low levels of the standard of living during the 19th century. Further judgment on this issue must await the development of quantitative measures of the change in apartment quality between 1918 and 1973, although we note that there has obviously been a major improvement in quality since 1918, when only 36 percent of apartments had bathrooms and only 61 percent had inside water closets (Brown, 1994, Table 3.6A).

Turning now to other components of housing expenditure, there is no reason to suppose that the CPI has measured the price of fuel or electricity inaccurately, since these commodities are homogeneous and among the easiest to measure of any goods or services. However, when we think of why people prefer to live in the modern age and would (in most cases) not willingly choose to go back to the conditions of 70 years ago, the change in the nature of household heating fuel surely enters the calculation. In 1918, 80 percent of American homes were heated with coal and wood, which had to be stored and carried, and produced a fire that had to be tended, used a stove that had to be cleaned, and smoke that polluted the air.[47] Because the transition from coal and wood heat to other sources of fuel had been largely completed by the early 1970s, we do not include this major improvement in the quality of life as a source of recent bias in the CPI.

The rest of the weight in the CPI on housing is applied to a myriad of expenditures, each having a relatively small weight, including telephone service, refuse collection, cable TV, curtains, furniture, bedding, video and audio products, major household appliances, and a large number of miscellaneous items. Most of the CPI weight on "other utilities" is applied to local and long distance telephone service and cable television. Even if the CPI correctly tracked the prices of each of these items, quality change would be missed. There has been continuous improvement in the quality of telephone service (e.g., reduction of static and improvement in clarity), improved convenience (credit card pay phones, itemized bill-

[46] The 1918 and 1950 data are for annual rental expenditures of working class households, from Brown (1994), Tables 3.6A and 5.10.

[47] Brown (1994), p. 63.

ing), and a great increase in picture quality and consumer choice achieved by cable television viewed as a new product. The fact that more than 60 percent of American households are now wired for cable TV, despite substantial monthly program fees, suggest that the development of cable TV has created a product yielding substantial consumer surplus. We conservatively estimate the quality bias connected with this category as 10 percent per decade, or 1.00 percent per year.[48]

The appliance and radio-TV category has been subject to more extensive research than any other category of consumer spending. Over the full period 1947–83 Gordon's detailed study (1990, p. 552), based on model-by-model comparisons from Consumer Reports, found an upward bias in the PCE deflator (which in turn is based on the CPI) of 3.22 percent per year for appliances and 5.94 percent for radio-TV. For the 1973–83 subperiod, the respective rates are 2.83 percent and 4.69 percent. These rates are applied in the CPI to a remarkably small fraction of consumption, just 0.8 percent according to Table 2. Consumer electronics alone, i.e., excluding electric appliances, recorded annual factory sales (i.e., net of retail markups) of $55.9 billion in 1994, which amounted to 1.25 percent of nominal personal consumption expenditures.[49] The 1995 share in PCE of final sales to consumers of audio and video equipment, including TV sets and VCRs, was also 1.25 percent, appliances contributed an additional 0.55 percent and personal computers an additional 0.33 percent, for a total weight in PCE of 2.13 percent, well over double the weight of the same products in the CPI.[50]

This small slice of personal consumption is the source of the largest annual rate of bias, with the possible exception of medical care. Our overall estimate of bias, based on Gordon's research, incorporates both quality-change bias and also new product bias, since his estimates of the overall bias take account of the fact that the quality-adjusted price of the VCR was declining at 30–40 percent per year in the early 1980s, prior to the introduction of this product into the CPI in 1987. Similarly, in recent years the price of personal computers purchased by consumers has been declining by at least 25 percent per year, but this has no impact at all, because home purchase of PCs were negligible in the CPI base period of 1982–84.

Our estimate of overall bias in this sector is 3.0 percent for appliances, 4.0 percent for radio-TV, including VCR's and camcorders, and 15 percent per year for personal computers.[51] Applying respective current nominal weights of 0.8 percent for appliances, 1.0 percent for consumer electronics, and 0.4 percent for personal computers, this category contributes an annual rate of quality change and

---

[48] This estimate of 1.0 percent per year multiplies an estimate of 1.5 percent per year for telephone service by its weight in this category of roughly two-thirds.

[49] The U.S. Consumer Electronics Industry In Review 1995, Electronic Industries Association, p. 13.

[50] Data provided by the BEA. Audio-video equipment includes VCRs, camcorders, and videotapes but excludes "audio media," i.e., cassettes and CDs. By 1996:Q3 computers accounted for 0.43 percent of nominal PCE and 1.34 percent of real PCE in 1992 dollars.

[51] While the CPI incorporates a matched-model index that records a decline in computer prices of 10–20 percent per year, this is applied to a negligible weight (based currently on 1982–84 weights) and so has no practical importance for assessing the magnitude of the quality change bias in the personal computer category. Berndt, Griliches, and Rappaport (1995) have estimated an annual difference between "matched model" indexes for personal computers (of the type used by the CPI) and hedonic indexes (used by the BEA) of roughly 15 percent per year.

new product bias of 0.10 percent per year to the total CPI. The figure entered into Table 2 for this category is a weighted average of the bias estimate, but the bias figure for the total CPI is based on weights corresponding to current nominal expenditures, not the CPI weights displayed in Table 2. Also, prior to 1994 the bias figure is based only on appliances and radio-TV, since personal computers did not emerge as a significant product until that date.

Regarding house furnishings other than appliances and video-audio products, there is no available research to provide guidance. The available range of furniture, draperies, etc., allows consumers to substitute among products, fabrics, and outlets along dimensions that are not captured by the CPI. There have been many new products in this area, including furniture and fabrics that are much less susceptible to damage by stains and children's accidents than was previously possible. This category also includes soap and cleaning products, where substantial progress has been made. We view a bias rate of 0.33 percent per year, or 10 percent over the past 30 years, as conservative.

3. Apparel. It is often assumed that there has been no quality change in apparel. But new apparel products are constantly introduced that improve consumer welfare, including denim jeans and shorts, advanced varieties of running shoes, iron-free synthetic fabrics, and lightweight but water-resistant raingear. Despite this, apparel is the other major area where the CPI is thought to have incorporated a downward bias. One source of downward bias occurred when the CPI price quotations followed the decline in price of an old model placed on sale, and then (using the deletion technique) made the transition to a new model without accurately recording the corresponding increase in price. Reforms in the CPI in the mid-to-late 1980s eliminated this source of downward bias and shifted to the hedonic price technique for some quality adjustments within the apparel component.[52]

The CPI apparel index is relatively easy to assess by accumulating outside evidence from such sources as mail-order catalogues. While style changes in fashion goods are frequent, quality changes in utilitarian apparel products purchased by average urban consumers are sufficiently infrequent to allow careful price comparisons across identical models from mail-order catalogues. By limiting itself to a month-to-month measurement framework, without cross-checks based on yearly or decadal comparisons, the CPI is vulnerable to persistent drift that emerges from measurement flaws such as the treatment of products on sale, as discussed above.

In a new project Gordon (1996) has compiled an apparel price index from the Sears catalogue based on thousands of year-by-year comparisons of identical apparel items over the intervals 1965–93.[53] The ratio of the CPI relative to the Sears apparel index rose at an annual rate of +1.92 percent per year during 1985–93.[54] The rapid rate of increase of the CPI apparel index after 1985 relative to Sears is surprising, because Sears in those years was losing market share to Wal-Mart and other discounters. Thus there is reason to think that the Sears catalog index might overstate the

[52] Recent changes in the CPI treatment of apparel are discussed in Liegey (1990, 1994).
[53] A total of 1,769 matched price comparisons were made for 1914–47 and 4,640 for 1965–93.
[54] Recall that the CPI treatment of apparel changed in 1985.

increase in true apparel prices faced by the average American consumer. Nevertheless, we shall take the conservative approach of cutting the implied bias rate from +1.92 percent suggested by the Sears index to a smaller 1.0 percent bias rate.

4. Transportation. The transportation component of the CPI consists of a wide variety of heterogeneous goods, including new vehicles, used vehicles, motor fuel, vehicle repairs, auto insurance and registration, and public transportation, mainly airline fares.

The most important questions to be addressed in the transportation sector are the valuation of mandated safety and anti-pollution devices, and the treatment of used cars relative to new cars. As documented by Gordon (1990, p. 364) for the period 1947–83, the actual price of new cars increased much faster than the CPI for new cars, and after 1967 almost none of this relative increase could be explained by increases in the dimensions included in the traditional hedonic regression equations for new cars. The key ratios of 1983 to 1967 prices were that actual prices had increased by a ratio of 289.9, the CPI for autos had increased by a ratio of 202.6, and that the difference had been more than explained by the contribution of CPI adjustments for safety and environmental quality and Gordon's adjustment's for fuel economy.[55] The resulting upward bias in the CPI relative to Gordon's final auto index is 0.44 per year from 1967 to 1983.

However, Gordon accepted the CPI's treatment of anti-pollution devices as a quality improvement rather than a price increase. We are persuaded that mandated anti-pollution devices are analogous to an indirect tax. Gasoline taxes may be used to provide a benefit in the form of better highways, but a tax increase is treated correctly by the CPI as an increase in the cost of living. Anti-pollution devices provide a benefit in the form of cleaner air, but by analogy to taxes should be treated as an increase in the price of the car. Using the detailed information given by Gordon, we can calculate an alternative to his index that converts the CPI adjustment for anti-pollution devices from a quality change to a price change, and this results in the finding that the CPI for new cars was downward biased during 1947–83 by 0.94 percent per year. We do not make a similar adjustment for the value of quality change taking the form of safety devices such as seat belts and crash-resistant bumpers, since our feeling is that consumers see the connection between their own safety and the devices more directly than they do between anti-pollution devices and air quality. Subsequently we will adjust the 0.94 percent downward bias for an offsetting increase in automobile durability.

What has happened since 1983, the terminal year of Gordon's study? Berry, Kortum, and Pakes (1996) show that emission control standards for automobiles became markedly tighter in two stages, 1975 and 1979–81, but did not change thereafter through the conclusion of their study in 1990. They also develop a fuel efficiency index that adjusts for changes in the horsepower and weight of cars; this rises by 67 percent between 1972 and 1983 and then drifts down by 13 percent by 1990. An alternative study of fuel

[55] The 1967–83 annual rates of growth were for the "raw" price of new cars 6.65 percent, the safety-pollution adjustment −2.03 percent, the fuel economy adjustment −1.10 percent (totalling 3.52 percent), and the new car CPI 4.42 percent.

economy (Sykuta, 1996) extends the data to 1994 and concludes that "new car fuel efficiency reached a plateau by 1986 and has not since changed by more than 1 percent in any direction" (p. 12). He attributes the absence of a decline in fuel economy in response to lower fuel prices to the high cost of "switching back to older, less efficient designs and technologies." Thus it appears that neither changes in anti-pollution equipment nor in fuel economy have been important potential factors in creating a bias in the new car CPI since the mid-1980s.

However, neither Gordon's study nor the CPI have taken into account the increased service lifetime of the typical new car. The average age of automobiles in use increased at an annual rate of 2.1 percent per year during 1970–83, part of the period covered by the Gordon study, and at 1.3 percent per year during 1983–93.[56] The CPI should be pricing automobiles on a rental equivalent basis, parallel to their treatment of housing. If the useful life of a car is improved by technological change that raises quality, reduces maintenance requirements, and minimizes rusting and corrosion, then consumers benefit. The cumulative increase in the useful life of a car from 1970 to 1993 amounted to 48 percent. Consider an automobile costing $10,000 with a life of 10.0 years; this is equivalent to an annual cost of $1,000 per year. Now consider the same car with a life extended by 48 percent to 14.8 years; its annual cost has been correspondingly reduced to $676.

The reduction in automobile depreciation is only part of the user cost of owning a car. Many of the additional elements of cost, e.g., insurance and repairs, are priced separately by the CPI in the major category "Other Private Transportation," where we impute no bias at all. The remaining component of user cost, in addition to depreciation, is real interest expense. Balancing alternative methods of paying for cars, including cash, installment finance, and leasing, we think that 10 percent is a reasonable estimate of the real interest cost. The extension of automobile longevity has reduced user cost from roughly 20 percent per year (consisting of a 10 percent real interest cost plus 10 percent depreciation rate) to 16.7 percent (the same real interest cost plus a 6.7 percent depreciation rate), for an overall reduction in user cost of 16.7 percent, or a geometric rate of −0.79 percent per year over the period 1970–93.

Since the rate of increase of average age slowed after 1983, we distribute this effect accordingly, as contributing to an upward bias in the CPI at a rate of 0.95 percent per year during 1970–83 and 0.59 percent per year after 1983. For the 1970–83 period, we add the upward bias in the true price of cars due to increasing durability of 0.95 percent per year to our previous estimate of a downward bias of 0.94 percent per year, exactly cancelling out and yielding a net bias for zero. For the subsequent post-1983 period, we found no other reason to suppose the CPI is biased in either direction, so the durability adjustment is converted to a net upward bias of 0.59 percent per year.

[56] The growth rates for trucks are similar. Data for 1970 and 1983 come from Statistical Abstract, 1990, Table 1027, and for 1993 from Statistical Abstract, 1995, Table 1025.

In assessing these calculations, we recognize that some fraction of the durability effect may not represent a pure quality change, but rather may reflect other factors, such as an increased relative price of cars that induces users to hold onto cars longer, or an improved quality of highways. This might suggest that our durability adjustment is too large. However, an offsetting error may be more important and may represent an important source of quality improvement that is not taken into account in the CPI or in our adjustments—namely, the marked decrease in the incident of defects on both imported and domestic new cars, as measured by the J. D. Power survey and other evidence. This additional source of quality improvement, which we have not taken into account, suggests that our estimate of the CPI bias for new cars is probably conservative.

The CPI index for used cars has long been known to be upward biased, simply because no quality adjustments were applied to this category at all. The upward bias over the period 1967–87 is very large, amounting to 2.44 percent per year, if we take the difference in the growth rate between the new car and used car index to represent a measure of the bias.[57] Adding this to the new car upward bias of zero, we arrive at a total bias for used cars for 1967–87 of 2.44 percent per year. For the period since 1987 we apply the CPI durability bias to used cars as well.

Regarding other components of the transportation category, we regard motor fuel as homogeneous and easy to measure. However, numerous improvements in fuels and related products have contributed both to cleaner air and improvements in autos, such as increased durability. While the cleaner air is valuable in its own right (see Section VII), we treat the mandate for cleaner gasolines like an indirect tax, not quality change per se. The improvements to autos are counted there, but must be understood to result in part from improvements in fuels, lubricants, etc. The CPI treats full-service and self-service fuel as two different commodities, as in the case of full-service department stores and no-frills discount stores. The shift in market share from full-service to self-service motor fuel provides evidence of a type of outlet substitution bias, which we do not count here as quality change bias. For the past decade, when the transition to self-service was largely complete, we attribute a small upward bias of 0.25 percent per year to the CPI for ignoring the increased convenience and time-saving contribution of automatic credit-card readers built into gasoline pumps.[58]

"Other private transportation" expenditures, an important CPI category, consist primarily of automobile maintenance and repair and automobile insurance. We are not aware of any evidence of price index bias involving these areas and thus assign a zero bias. However, in the public transportation category, which is dominated by airline travel, it is well documented that prior to 1982 the CPI incorporated a substantial upward bias due to a failure to take discount fares into account. We take our estimate of this bias from

[57] In 1987 the BLS began to adjust for quality changes in used cars in a parallel manner to its adjustments for new cars.

[58] The 1982 Census of Retailing reports that the number of gas pumps in the United States declined 28 percent between 1972 and 1982, and that the fraction of self-service pumps (which was not reported in 1972) had reached 54 percent by 1982.

Baily-Gordon (1988, p. 416) and multiply it by the weight (about 2/3) of airline fares in the public transportation category.

5. Medical Care. The medical care category may be the location of substantial quality change bias at a rate as rapid or more rapid than in appliances and radio-TV products, but its weight in the CPI is much greater. That weight is controversial in itself, since the 7.4 percent weight is based only on out-of-pocket expenses by consumers and does not include expenses paid for medicare, medicaid, or by employer-financed payroll deductions. Since one of the main uses of the CPI is to convert indexes of real income, compensation, and wages into "real terms," the current weight is wrong. Changes in medical care prices or technology should influence the CPI with the same weights that are relevant for total income and compensation, that is, they should include the total medical care bill, which in National Income and Product Accounts version of consumption amounts to 16 percent, not the much smaller 7.3 percent weight in the CPI.

The CPI weight of 7.4 percent is applied to three primary categories—drugs, professional medical services (i.e., doctors and nurses), and hospitals. There has been substantial recent research on prescription drugs. For instance, Berndt, Cockburn, and Griliches (1996) have studied prices of anti-depressant drugs. They find a substantial difference between an index based on BLS methods, which rises during 1993–96 at 4.6 percent per annum, and an alternative index based on their preferred methodology which rises at 1.1 percent per year. This alternative index uses an alternative (Divisia) weighting scheme, introduces new items much more promptly, and treats generic drugs as fully equivalent to proprietary brand-name drugs. In July, 1995, the CPI shifted its treatment of generics, so that any decline in price when generic equivalent drugs become available is taken fully into account as a price change rather than being treated as a different good and hence "linked out" of the index. Based on the Berndt et al. research, and related research by Griliches and Cockburn (1994), we conclude that prior to 1995 there was a 3.0 percent annual bias in the CPI for prescription pharmaceuticals. The new CPI treatment of generic drugs after July 1995 reduces this annual rate of bias from 3.0 to 2.0 percent, and the remaining bias can be attributed (based on the previously cited research) to the late introduction of particular drugs into the CPI and the failure to attribute a positive value to newly introduced drugs that gain market share and thus appear to add value for consumers.

The major weight of the CPI medical care component is applied to medical care services, e.g., physician fees, and hospital costs. No attempt is made in the CPI to value health care "outcomes" as contrasted to "inputs." A hybrid approach is taken by the Producer Price Index (PPI), which also prices inputs but with different weights and increases by roughly 2.0 percent per year more slowly than the CPI in both the doctor and hospital category in the recent past (1995–96). Thus government indexes already provide important evidence that the CPI is upward biased by at least two percent per year, but that ignores many changes in medical care practice and technology that suggest a higher rate of bias. Cutler et al. (1996, Table 1) contrast input price indexes for the cost of heart-

attack treatment with alternative "outcome " indexes that take account of the cost of achieving a given increase in life expectancy. The authors contrast an "input" index of the type currently compiled by the CPI, which rises by 3.3 percent per year over 1983–94, with a final outcomes index that takes into account a conservative valuation for the extension of life expectancy attributable to new heart attack treatments and declines by 1.1 percent a year, for a net bias of 4.4 percent per year. Shapiro and Wilcox study cataract surgery for the period 1969 and 1993. Their "protypical" index that duplicates CPI methods increases by a factor of 9 while their preferred alternative index increases by a factor of 3, implying an annual rate of upward bias of 4.57 percent. The closeness in the Cutler et al. and Shapiro-Wilcox studies of quite different medical procedures is striking. In view of the fact that the CPI has been rising relative to the PPI for medical care (taking the input rather than the outcomes approach) by 2 percent per year in 1995–96, the rate of upward bias that we have assigned to the medical care category, 3.0 percent per year, is probably conservative. There is probably additional quality change taking into account the wide variety of new diagnostic and test equipment, the reduction of pain of routine procedures, the shortening of hospital stays, and also the quality of hospital rooms.

This new research by Cutler and others opens up the potential for a major improvement in our understanding of the economics of medical care. This category should receive a substantial component of the CPI's future research investment, and we strongly endorse a move in the CPI away from the pricing of health care inputs to an attempt to price medical care outcomes.[59]

6. Entertainment. The entertainment category is divided roughly equally between commodities and services. Commodities consist of newspapers, magazines, sporting equipment, and toys. We assume that there is no bias in newspapers and magazines (although electronic news services provide convenience and timeliness to some) but that sporting equipment and toys are subject to a somewhat smaller bias than Gordon found for appliances, namely 2.0 percent per year as contrasted to 3.0 percent for appliances. This may represent an average of the rate of bias in electronic toys, e.g., Nintendo games, which may be close to the 15 percent rate we have applied to personal computers, and a bias rate of zero for other toys, including stuffed animals and non-electronic board games. Entertainment services consist of club memberships, admission fees, and lesson fees. There has undoubtedly been an improvement in the comfort of attending sports events, with domed stadiums and more comfortable seats, but we have not assigned any bias to the CPI measure of entertainment services prices.

7. Other Goods and Services. While purchases of such products as cigarettes, toiletries, and cosmetics may have been subject to outlet substitution bias, there is no reason to think that quality change has been missed or that new product bias has been important. However, this category includes a small weight for small personal care appliances, e.g., hair dryers, and it is reasonable to sup-

[59] The BLS is moving to price and reprice a hospital bill every month, a potential improvement, although not something which per se gets at quality change.

pose that these items are subject to the same magnitude of bias as large appliances, i.e., 3.0 percent per year. Much of the rest of the weight in the "other" category consists of school books and fees, primarily college tuition. However, about one-tenth of the weight in this category consists of personal financial services, which have been subject to rapid technological change, particularly with the widespread diffusion of ATM machines and all-in-one cash-management accounts. Taking a conservative 2.0 percent estimate of personal financial services and applying that to one-tenth of this category yields our 0.2 percent annual rate of bias.

Conclusion on Quality Change Bias

Our final estimate of quality change bias, taking the weights and bias rates in Table 2, aggregates to approximately 0.6 percent per year. This is slightly higher than the rate of 0.5 percent per year for the combined categories of quality change and new product bias estimated in our interim report or 0.35 percent in Shapiro and Wilcox (1996c). Our higher estimate results from a much more extensive examination of the existing literature and consideration of factors that affect each of the 27 separate categories included in Table 2.

Table 2

Relative Importance of Components in the CPI-U, U.S. City Average, December 1995, and Estimates of Quality Change and New Product Bias for Selected Time Intervals

Major and Selected Components	*Relative Importance in Percent*	*Estimated Quality Change Bias at Annual Rate for Selected Time Intervals*	
1. Food and beverages	17.332		
Food at home other than produce	8.543	0.30 (1967–96)	
Fresh fruits and vegetables	1.337	0.60 (1967–96)	
Food away from home	5.886	0.30 (1967–96)	
Alcoholic beverages	1.566	0.15 (1967–96)	
2. Housing	41.346		
Shelter	28.289	0.25 (1976–96)	
Fuels	3.792	0.00	
Other Utilities, incl. telephone	3.222	1.00	
Appliances incl. electronic	0.806	[1]3.6 (1973–94)	[1]5.6 (1994–96)
Other housefurnishings	2.639	0.30	
Housekeeping supplies	1.116	0.00	
Housekeeping services	1.482	0.00	
3. Apparel and upkeep	5.516	−0.95 (1965–85)	1.00 (1985–96)
4. Transportation	16.953		
New vehicles	5.027	0.00 (1970–83)	0.59 (1983–96)
Used Cars	1.342	2.44 (1967–87)	0.59 (1987–96)
Motor Fuel	2.908	[2]0.0 (1974–84)	[2]0.25 (1984–96)
Other Private Transportation	6.153	0.00	

Table 2—Continued
Relative Importance of Components in the CPI-U, U.S. City Average, December 1995, and Estimates of Quality Change and New Product Bias for Selected Time Intervals

Major and Selected Components	*Relative Importance in Percent*	*Estimated Quality Change Bias at Annual Rate for Selected Time Intervals*		
Public Transportation	1.523	2.66 (1972–77)	4.60 (1977–82)	0.00 (1982–96)
5. Medical Care	7.362			
Prescription drugs	0.891	3.00 (1970–95)	2.00 (1995–96)	
Nonprescription drugs and medical supplies	0.391	1.00		
Professional medical services	3.465	3.00		
Hospital and related services	2.257	3.00		
Health insurance	0.358	0.00		
6. Entertainment	4.367			
Commodities	1.975	1.20		
Services	2.392	0.00		
7. Other Goods and Services	7.123			
Tobacco, smoking products	1.610	0.00		
Personal care	1.170	0.90		
Personal and educational expenses	4.342	0.20		
TOTAL	100.000	0.612 (for 1996)		

[1] Applied to the weight of this category in 1995 annual nominal personal consumption expenditures, 2.13 percent, as contrasted to the December, 1995 CPI relative importance of 0.8 percent.
[2] See text.

Has quality change bias increased or decreased? Table 2 provides some insights to that issue. To take the specific example of 1980, we can aggregate the rates of bias given in Table 2 and come up with an upward CPI bias due to quality change and new products of 0.488 percent per year, compared to 0.613 for 1996. The main differences come from the reversal of the previous downward bias for apparel and the increased upward bias in the appliance/radio/TV component due to the growing role of personal computers. Partially offsetting these sources of increased upward bias are reductions in the extent of upward bias in used cars, airline fares, and prescription drugs.

VI. Estimates of Biases by Type and in Total

The CPI is not a cost of living index, but rather a fixed weight index, implemented by means of a modified Laspeyres formula. This creates an immediate conflict between the objectives of the CPI and many of the purposes for which it is intended. For example, the CPI is used to index private contracts, tax brackets and government transfer programs, such as Social Security, in order to compensate beneficiaries for changes in the cost of living. A fixed weight index exaggerates the effect of price changes on the cost of living, because it fails to allow for substitutions that enable consumers to avoid the full impact.

The assessment of biases in the CPI requires a cost of living index as a point of reference. An approximation to a cost of living index can be generated by combining the results of research on dif-

ferent types of biases in the CPI. The purpose of this section is to summarize this research and assess the importance of these biases in the CPI. Our estimates depend primarily on studies produced prior to the convening of our Commission, many of them by the BLS.

Our Interim Report of September 15, 1995, presented initial estimates of biases in the CPI. We estimated that the overall bias had been 1.5 percentage points per annum in recent years, but changes in the CPI methodology then in prospect from BLS would eliminate as much as 0.5 percentage point per annum of this bias, reducing the bias going forward to 1.0 percentage point per annum. We have now revised our estimates to reflect changes in the CPI announced by BLS on March 29, 1996, and new estimates of the impact of biases due to the introduction of new products and changes in the quality of existing products. The BLS has eliminated some of this bias totaling 0.24 percentage point per year, raising our estimate of the bias going forward by one-quarter of one percentage point. In addition, we have revised our estimates of new products/quality change bias upward by 0.10 percentage point per year.[60]

In assessing biases in the CPI it is essential to separate two types of substitution bias. First, BLS uses a fixed weight index based on the modified Laspeyres formula to combine price indexes for 207 items for 44 areas into a national CPI. The weights are derived from the Consumer Expenditure Survey (CES) and reflect surveys of individual households. We refer to the substitution bias at this level as Upper Level Substitution Bias. This bias is measured as the difference between the modified Laspeyres formula used by BLS and a Tornqvist index, which is (approximately) free of substitution bias.[61] Most estimates cluster around 0.2 to 0.25, including numerous estimates from BLS. The latest estimates available to the Commission reflecting just produced unpublished corrections of previous research by BLS show an average bias over the period 1988–1995 of about 0.15 percentage point per year. While we have not had time to analyze these new results, to be conservative, we adopt this figure.

The second type of substitution bias results from combining price observations for approximately 71,000 goods and services and information on prices for housing components of the CPI into indexes for the 207 items and 44 areas. We refer to the substitution bias at this level as Lower Level Substitution Bias. The prices to be collected are selected by probability sampling. The probabilities are derived from the CES and the Point-of-Purchase Survey (POPS) of retail establishments and are intended to reflect the share of items and areas in consumer expenditures in the base period.

---

[60] In addition to the comprehensive bias estimates presented in our Interim Report, estimates have been made by Erwin Diewert (1996), Shapiro and Wilcox (1996), and others. An especially valuable earlier survey was presented by Lebow, Roberts, and Stockton (1994).

[61] Aizcorbe and Jackman (1993) have estimated Upper Level Substitution Bias, but these estimates have been updated and revised by BLS.

Beginning in 1978 BLS introduced a sample rotation procedure in the estimation of price indexes within each item and area category. At the same time prices were combined in a way that reflected the modified Laspeyres index formula.[62] These two changes had the effect of introducing considerable Lower Level Substitution Bias, a fact discovered by Reinsdorf (1993) after a lapse of fifteen years.[63] This issue has been intensively studied by BLS and steps have been taken to deal with it, beginning in January 1995.[64]

For data from the CPI covering the thirty month period from June 1992 to December 1994, Moulton and Smedley (1995) have estimated that the difference between the modified Laspeyres formula used at the lower level of aggregation by BLS and a weighted geometric mean formula for non-shelter components of the CPI to be 0.49 percentage point per annum. This difference is an estimate of the bias of the Laspeyres formula, since Shapiro and Wilcox (1996) have shown that the geometric mean provides an unbiased estimate of the underlying cost of living index. Armknecht, Moulton, and Stewart (1995) have estimated the bias for owners' equivalent rent to be 0.50 percentage point per annum.[65] BLS is currently testing the feasibility of producing indexes based on this methodology for price data beginning in 1990. Substitution of the weighted geometric mean formula for the modified Laspeyres formula at the lower level of aggregation is an important step in the direction of a cost of living index.

In January 1995 the BLS introduced an improved method for the imputation of price changes for food at home, owners' equivalent rent, and prescription drugs. These changes are described by Armknecht, Moulton, and Stewart (1995). For food at home items the changes included the introduction of a procedure called "seasoning." The seasoning period is the time to obtain the data needed to weight each new sample observation before introducing it into the index. This period was lengthened to three months for food at home items, breaking the link between the weights for these items and prices eventually used in the CPI.

On March 29, 1996, the BLS announced that the seasoning procedure would be extended to all non-shelter items, effective with the CPI for June 1996.[66] The announcement pointed out that residential rent and owners' equivalent rent were no longer subject to the bias associated with sample rotation procedures, as a consequence of the changes introduced in January 1995. In addition, the BLS stated that (with rare exceptions) the weight for a substitution item would be kept constant throughout the life of the item. BLS estimated the reduction in bias due to the January 1995 changes to be 0.14 percentage point per annum, while the 1996 changes for the non-shelter items reduced the bias by a further 0.10 percentage point per year. This reduces the Moulton-Smedley estimate of the remaining Lower Level Substitution Bias to 0.25 percentage point per annum. On July 16, 1996, BLS introduced changes in the classification and definition of the hospital and re-

[62] For details, see Gillingham (1974.)
[63] See Reinsdorf (1993), pp. 227–254.
[64] The results of BLS research on Lower Level Substitution Bias are summarized by Moulton (1996).
[65] See Armknecht, Moulton, and Stewart (1995).
[66] See Division of Consumer Prices and Price Indexes (1996b).

lated services component of the CPI.[67] These were intended to improve the measurement of this important component of the CPI, but were not accompanied by an estimate of bias reduction. Thus, our point estimate of Lower Level Substitution Bias is 0.25 percentage point per annum.

The goal of BLS is to measure goods and services of constant quality; however, the disappearance of products from the marketplace necessitates the substitution of other products. Armknecht and Weyback (1989) summarized the methods used by BLS to adjust the CPI for quality change.[68] Elimination of new goods bias and quality change bias are essential steps in measuring the cost of living. Important empirical research on new goods bias has been done for breakfast cereals by Hausman (1996), prescription drugs by Griliches and Cockburn (1994), new cars by Pakes, Berry, and Levinsohn (1993), and many others as described above in Section V.[69] We have estimated the total bias due to new products and quality change of existing products to be 0.6 percentage point per year.

Reinsdorf (1993) has provided the principal empirical evidence on New Outlet Substitution Bias. This is based on comparisons between prices for certain food and fuel items for outlets rotating into the sample covered by the CPI and outlets rotating out. He estimated the bias to be 0.25 percentage point per annum. Lebow, Roberts, and Stockton (1994) have extrapolated this estimate to the CPI as a whole by identifying components of the index that would be affected by outlet substitution bias. These amount to 40% of the CPI, so that Outlet Substitution Bias is 0.1 percentage point per annum.

These separate biases are approximately additive and likely to be independent of modest swings in the true inflation rate. Thus, a bias of 1 percentage point implies that when changes in the CPI show inflation rising from 3% to 5%, it is likely actually to be rising from 2% to 4%. Note the bias primarily affects the level, not the change, in the inflation rate. At very high rates of inflation, the bias may increase (one might assume greater outlet and commodity substitution), but we currently have no evidence regarding this issue.

Table 3 summarizes our evaluation of biases in the CPI. This includes point estimates based on the best available evidence as well as a plausible range for the overall bias. The average of our estimates of the overall bias in the CPI is 1.1 percentage point per annum with a range of 0.8 to 1.6 percentage point.

---

[67] See Division of Consumer Prices and Price Indexes (1996a).

[68] See Armknecht and Weyback (1989).

[69] See Hausman (1966), Griliches and Cockburn (1994), and Pakes, Berry, and Levinsohn (1993).

Table 3
Estimates of Biases in the CPI-Based Measure of the Cost of Living (Percentage Points Per Annum)

Sources of Bias	*Estimate*
Upper Level Substitution	0.15
Lower Level Substitution	0.25
New Products/Quality Change	0.60
New Outlets	0.10
Total	**1.10**
Plausible range	(0.80–1.60)

The BLS is preparing for a benchmark revision in January 1998, when the CPI will incorporate new expenditure weights from the 1993–1995 Consumer Expenditure Surveys. However, BLS will retain the modified Laspeyres formula, so that our estimates of bias will carry over to the revised CPI. In addition, BLS has continued to introduce important modifications in the CPI in order to improve measurements and remedy deficiencies that have come to light. However, these revisions, like the forthcoming benchmark revision, employ the modified Laspeyres framework, so that important differences between the CPI and a cost of living index will remain.

The Upper Level Substitution Bias in the CPI will persist after the forthcoming benchmark revision of the CPI, since BLS plans to retain the modified Laspeyres formula. Second, BLS has reduced so-called formula bias, the part of Lower Level Substitution Bias resulting in substantial measure from the introduction of sample rotation procedures and the modified Laspeyres index at the Lower Level of aggregation in 1978. However, the objective of the changes in January 1995 and those announced in March 1996 was to improve the implementation of the modified Laspeyres formula, not to eliminate the Lower Level Substitution Bias quantified by Moulton and Smedley (1995). Finally, New Item, New Outlet, and Quality Change Biases are unaffected by the changes already announced by BLS or the benchmark revision.[70]

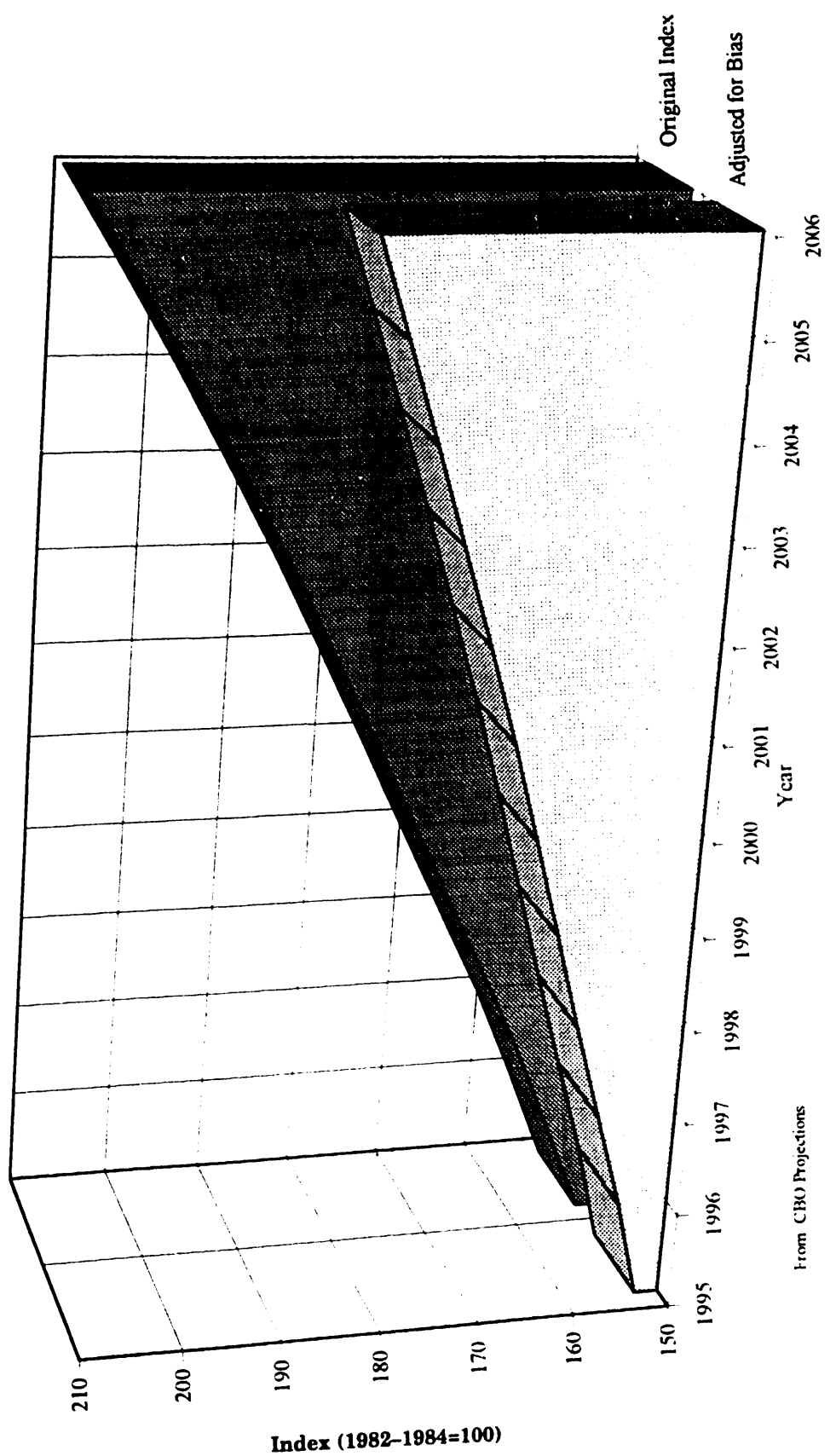
To summarize we have revised the estimates presented in our Interim Report to reflect BLS revisions of the CPI and the accumulation of new findings on the magnitude of biases. Our main conclusion is that the limitations imposed by the modified Laspeyres formula make the CPI unsuitable for cost of living measurement. By combining a Tornqvist formula at the Upper Level of aggregation and a weighted geometric formula at the Lower Level, BLS could eliminate both types of Substitution Bias. However, these changes alone would fail to encompass adjustments for New Item, New Outlet, and Quality Change Biases. Adjustments for these biases are essential for measurement of the cost of living.

Figure 2 illustrates the compounding effect of a 1.1 percentage point per annum bias on CBO projections of the CPI-U through 2006. While 1.1 percentage point may seem to be a small amount in any given year, cumulatively year after year it adds up to a sizable difference, 14% over a dozen years.

[70] Although the BLS is continuously attempting to improve the price measures, for example moving to price and reprice hospital bills.

Figure 2

Effect of a 1.1 Percentage Point Upward Bias in the CPI-U Through 2006



VII. Other Issues

A. A Separate Price Index for the Elderly?

In principle, if not practice, a separate cost of living index could be developed for each and every household based upon their actual consumption basket and prices paid. As noted above, the aggregate indexes use data reflecting representative consumers. Some have suggested that different groups in the population are likely to have faster or slower growth in their cost of living than recorded by changes in the CPI. We find no compelling evidence of this to date, and in fact two studies suggest that disaggregating by population group, for example by region or by age, would have little effect on measured changes in the cost of living.[71] Further, work on this subject remains to be done. In particular, the prices actually paid, not just expenditure shares, may differ.[72]

Beyond the different consumption baskets, it is important to understand our analyses of the sources of bias are applied to representative or average consumers. Some consumers will substitute more than others, and the substitution bias may be larger for some, smaller for others. Likewise, some are more likely to take advantage of discount outlets; others less so. Perhaps most importantly, the benefits of quality change and the introduction of new products may diffuse unevenly throughout the population. Some will quickly gain the benefits of cellular telephones, for example, while others may wait many years or decades or never use them. This is yet another reason why we have been very cautious in our point estimates for these particular sources of bias.

B. Broader Considerations on the Quality of Life

Not all change is positive and not all change is positive for everyone. In making the case for the importance of quality change, we have also to look at the other side of the ledger. There are at least three types of change to consider here:

1. New goods may drive out older goods which are still valued by a subgroup of the population, or what is equivalent, the loss of economies of scale may drive up their price significantly. To the extent that it is measured, it does not represent a new problem of the price index construction.

2. An existing good or service may deteriorate in its quality. That is a less frequent phenomenon, in spite of the mantra that "they don't make them the way they used to." The most significant recent example is probably the HMO-induced tightening-up of the rules of access to medical care which is likely to be perceived as a significant deterioration in the expected services that had been contracted for by the purchase of medical insurance. It is unlikely, however, to have outweighed the

---

[71] See Boskin and Hurd (1985); Jorgenson and Slesnick (1983). However, very preliminary unpublished work suggests that for the period 1982–91 the larger fraction of expenditures on out-of-pocket healthcare by the elderly combined with the more rapid rise in healthcare prices than overall prices for this period might lead to a slightly faster rise in a price index for the elderly. The rate of healthcare price inflation has slowed substantially of late, so it is unlikely this result will be reproduced for the mid-1990s. Also, as noted above in Section V, healthcare inflation is seriously overstated because of the substantial uncounted quality change.

[72] Seniors get special discounts, for example, and their geographic distribution, and other factors might cause the prices they pay to differ slightly from those recorded in the CPI.

medical advances of recent years. Few would trade today's restricted access to medical care, for a more free access to the technologies of yesteryear, foregoing the improvements in bypass operations, ulcer treatments, or cataract surgeries.

3. The largest effect may come from change in our physical, social, and economic environment which impose on us higher expenditures necessary to keep up with our previously achieved utility levels.

It is not clear, however, whether events such as a colder winter, the appearance of AIDS, or a rise in the crime rate should be included in the definition of a price index. A change in expenditures due to an unanticipated change in weather should raise the price index only to the extent that energy prices go up, not quantities consumed. The latter, if the event persists, will ultimately affect the commodity weights in the index, but that is a different matter. The rise in AIDS would drive up the price index of health, if we define it as the expenditure necessary to achieve an equivalent base-period health level. But while this component represents a real rise in the "cost-of-living," it may not be an appropriate component in an indexing formula, since there are no "gains" among the young which could be called to "compensate" the retirees for such price increases.

The appearance of AIDS did raise the cost of living. Counting subsequent improvements in its treatment as a positive quality change will be inappropriate, if the original deterioration in the environment was not taken into account in the measurement framework in the first place. Similarly, counting quality improvements in locks and other security devices, may overestimate the "gains" from such defensive consumer investments.

While it is impossible to provide a full and accurate accounting for such changes, it is extremely unlikely that the rise of such "bads" out-balances the "good" that is contained in the developments alluded to above. In the major areas of concern and poor measurement, environment, health, and crime, there have been gains in the first two and we have come close, in recent years, to holding steady on the last one. Thus, while we do recommend extending measurement efforts beyond the current concept of what constitutes the consumption "basket," we see no strong reason to temper our conclusions about the measurement issues in the areas we did examine: the more traditional components of consumption as defined in the current content of the CPI. The following presents some brief background for our view.

The industrial revolution caused widespread air and water pollution, and this was indeed a negative factor up until the 1950s and 1960s. However, since then a shift from coal to natural gas as the dominant energy source for home heating, cleaner fuels and cars and environmental regulation have caused a major decline in the presence of many types of contaminants in the air and water. The shift in heating fuels also brought about a major increase in cleanliness and convenience around the house, as the coal scuttle was replaced by the silent and automatic transmission of natural gas. The CPI implicitly values the improvement in air quality made possible by mandated anti-pollution devices in automobiles, since it treats the cost of mandated anti-pollution devices as an improve-

ment in quality rather than an increase in price. We have recommended that the CPI treat changes in price due to additional anti-pollution devices as a price rather than quality increase. But that concept, incorporated in our quality change bias estimates in this report, leaves the improvement in air quality unmeasured. This is a source of understatement in our final estimate of CPI bias. Further, the CPI is inconsistent, since a portion of the higher cost of electricity, steel, paper, and other products is also due to environmental regulation, and the benefits of higher air and water quality made possible by regulation of products other than automobiles is not taken into account.[73]

Crime is another type of externality. The rate of crime victimization increased in the past, e.g., from 0.096 incidents per capita in 1970 to 0.098 incidents in 1981. Since then, however, the victimization rate has fallen to 0.074 incidents in 1992 (the latest year available).[74] The share of violent crimes increased from 0.025 in 1970 to 0.028 in 1981 and decreased slightly to 0.026 in 1992. Since 1992 there has been widespread attention to a sharp decline in homocide rates in many major American cities.

Looking further for negative factors, perhaps the most important are such social issues as divorce, illegitimacy, and the reduced role of the nuclear family. The divorce rate increased by 50 percent between 1970 and 1980 but since then has been stable.[75] The suicide rate was stable between 1980 and 1992 but the age-adjusted death rate declined by 14 percent, while the infant mortality rate *fell* by 58 percent between 1970 and 1992. Perhaps most importantly, life expectancy at birth increased from 70.8 in 1970 to 75.5 in 1993. The major negative has been that births to unmarried women have increased, from 18.4 percent to 30.1 percent. There may be other intangible negatives, such as perceived increased job insecurity, possible increased inequality, and the decreased job opportunities for workers with only a high school education.

On the positive side, there is no question that goods have improved in ways that our discussion of quality change cannot fully incorporate. Gordon (1990, p. 38) notes a number of dimensions of quality which his measures did not capture, some of which are the faster speed and reduced vibration of jet planes, improved reliability of appliances and automobiles, improved sound quality of audio equipment in homes and automobiles, improved safety devices on home power tools and power lawn mowers, reduction in the noise, weight, and installation cost of room air conditioners, and "immeasurably better picture quality of color TV sets."

Many aspects of the change in the American standard of living from the rural horse-drawn economy of 1870 to today's modern economy occurred many decades ago and are not current sources of CPI bias. The elimination of animal waste was a major contribution of the motor car, in addition to its speed and flexibility, but this achievement was largely accomplished before World War II. The achievement of electrical appliances in reducing household

---

[73] Detailed case studies in Repetto et al. (1996) demonstrate substantial understatement of output and productivity growth, and hence an overstatement of price growth, in the electricity, pulp-paper, and agricultural industries.

[74] All data in the paragraph come from Statistical Abstract, 1995, Table 317.

[75] Data in this paragraph come from Statistical Abstract, 1995, Tables 114, 124, 136, and 141.

drudgery was largely accomplished by the 1960s. But some improvements have been continuous. There has been a steady transition in the quality and variety of home entertainment options, from the player piano, to the radio, to black-white and then color TV, to the VCR, and now cable TV, satellite TV, and the World Wide Web and the other features available with personal computers. Further, the rapid spread of central air conditioning has made possible a substantial movement of households to the southern and southwestern states. Millions have chosen to take this option, voting with their feet to enjoy milder winters with artificially cooled summers.

Overall, we find that the presumed negatives (pollution, crime, suicide, divorce), the worsening of which may have increasingly detracted from the quality of life at one time, have reached a plateau and in the case of pollution and crime seem actually to have reversed direction, thus recently improving the quality of life. The remaining negatives are important but seem to us to have been more than offset by increased quality and variety of goods, services, and choice of outlets along dimensions that are partly but not entirely captured by our measures of bias, but most importantly by the major increase in longevity which perhaps swamps everything else. Accordingly, our estimate of the current bias in the CPI is, if anything, probably understated.

VIII. The Commission's Recommendations

Introduction

Our recommendations are directed at three different audiences: 1. Our elected officials (the President and the U.S. Congress), who are the providers of funds, the supervisors, and also major consumers of the information contained in the CPI; 2. The BLS, which is the producer of the CPI; and 3. The community of professional economists and statisticians, who ultimately provide both the manpower and the knowledge base required for a successful operation of this major measurement and observational tool for our economy. The role of the BLS in this can be viewed as running one of our major National (Economic) Observatories, producing both timely information about the state of our economy and providing the inputs for advancing our knowledge of and understanding about how it functions and the interrelationships among its various components.

Recommendations

1. The BLS should establish a cost of living index as its objective in measuring consumer prices.

All of our specific recommendations are aimed toward this goal.

2. We recommend the development and publication of two indexes:

One which is published monthly on a timely basis and is designed to maintain the spirit of the cost of living index yet accommodate the inconsistent timing schedules of the required information; and a second index which is published and updated annually

and revised historically to introduce improvements arising from new information and new research results. The purpose of having two indexes is to accommodate the complex issues that must be addressed and the time delay in obtaining all of the necessary data.

The recommendations to the BLS are divided into three parts:

1. Short run: Those we think can be implemented immediately, with little additional resources or new data collection initiatives. These center on changing the current CPI computation, primarily to make it more current, and second, on computing an annually updated and subsequently revised COL index;

2. Intermediate run: Reforms that are feasible within the current state-of-the-arts, but would require new data collection, reorganization of activities, and changes in the detail of the various sub-indexes produced by the CPI; and

3. Longer-run recommendations, emphasizing topics and areas that need additional research and attention.

Short Run

3. The timely, monthly index should continue to be called the CPI and should move toward a COLI concept by adopting a "superlative" index formula to account for changing market baskets, abandoning the pretense of sustaining the Laspeyres formula.

To accommodate the delay in obtaining information on quantities needed to combine the price changes of items in the lowest groupings, BLS should move away from the assumption that consumers do not respond at all to price changes in close substitutes. We thus recommend BLS move to a "trailing Tornquist" Index (weighted geometric mean of price relatives), at the stratum and ELI level, and also, concurrently, to geometric means of price relatives at the elementary aggregation level. Both of these moves would alleviate the problem of the growing irrelevancy of "baskets" based on decade-old consumption patterns, reduce significantly the substitution and formula bias, and facilitate the speedier introduction of new goods and services into the index.[76]

a. Because of the lag in collecting up-to-date information on consumer spending patterns, the weights will have to be based on a trailing two- or three-year average of past expenditures, e.g. 1993–4 weights for the 1996 price changes. They should be changed every year.[77] This implies that

b. The BLS should organize itself for "permanent" rather than decadal revisions in the CPI. Both the weights and the priced commodity and services assortment need more frequent updating. Also,

c. Wherever possible, scanner data and other "outside" data should be used both to reduce the cost of data collection and

[76] A parallel trailing chained-link Laspeyres formula, with revisions producing a comparable Fisher Ideal Index with a 4-year lag might be useful as well. Even so, we still recommend the move to geometric means at the elementary aggregates level.

[77] What needs more rapid changing are the commodity strata weights and the assortment of items priced. City weights could be held constant and changed much less frequently. Here decades would do. See Shapiro and Wilcox (1996a) for additional discussion of the mechanics of such revisions.

(primarily) to expand the assortment of goods and services priced concurrently, to provide current item weights, and to introduce new items as quickly as they enter the market. Whether this will result in a net reduction in the cost of data collection is an open question.

d. As subsequent data become available, the weights are updated, and new goods are introduced and their history extended backward, the information incorporated in the published CPI should undergo retroactive revision, as far back as the new information warrants, in the form of a new annual COL index, using a compatible "superlative-index" formula. This "revised" COL index would be published annually, with a lag of a year or two, and would be subject to additional revisions after new information emerges and new methodology is introduced. The published versions of this index need never be "final."

4. **The BLS should move to geometric means at the elementary aggregates level.** We believe it to be the closest approximation to a full implementation of the COLI concept, which could be ultimately implemented also at this level, as scanner data become available for most of the currently sampled commodities. Changing to geometric means will not only solve the "formula bias" problem, much of which has been recently eliminated by a switch to "pre-seasoning," but will also alleviate the below-stratum-level substitution bias. It will not solve, however, the "outlet bias." To aid in updating the required weights, the BLS should be able to acquire the detailed commodity-level shipments data currently collected by the Census but not accessible to the BLS.

Intermediate Run

5. **The BLS should study the behavior of the individual components of the index to ascertain which components provide most information on the future longer-term movements in the index and which items have fluctuations which are largely unrelated to the total and emphasize the former in its data collection activities.**

This could result in the down-weighting or even elimination of data collection for certain cities and a revision of the commodity structure of the index which would consider some goods as having a national market, sampling a larger number of items but with less regard to geography, focusing on geographical differences only for more "local" commodities, such as fuel costs, rent, personal services, and fresh produce.

Currently, the BLS collects a large number of price quotes on bananas, because they are inexpensive to collect and their prices are quite variable, even though these variations are not related systematically to the underlying trend-movements in the CPI. At the same time, less attention is paid to less variable but more likely to change (disappear or be redesigned) and harder to measure commodities, such as surgical treatments, consumer electronics, and communication services.

6. The BLS should change the CPI sampling procedures to de-emphasize geography, starting first with sampling the universe of commodities to be priced and then deciding, commodity by commodity, what is the most efficient way to collect a representative sample of prices from which outlets, and only later turn to geographically clustered samples for the economy of data collection.

The current city level price indexes are useless for geographical comparisons of levels and misleading as measures of rates of change, since they are not based on any clearly defined levels. To do an adequate job of describing the geography of price levels in the U.S. will require the collection of prices for the same commodities and services in different cities. To study differential changes in the price levels across cities, arising from different competitive and population trends, it may prove adequate to sample the "national" commodities in specific cities only once a year or so, on a rotating basis. More generally, one could design a model consisting of an underlying "national" trend level of the CPI, which would be the primary focus of monthly estimation, and more slowly changing city differentials, which would be based on less frequently collected data.

This would allow the CPI to concentrate resources on expanding the sample and analysis in rapidly changing areas of the commodity and services spectrum, such as health services, communication services, and food away from home, where quality change and commodity turnover is endemic.

Moving to a notion of a new "basket" each year will allow for a faster introduction of new items and new outlets. Moving to a national sample for most of such items would allow expansion of the number of specific items (models, varieties, types) sampled within a particular ELI and reduce thereby the number of forced substitutions. Also, this would allow for the use of new sources of data, such as scanner data on prices, and industry-wide information on sales of specific items (for more detailed weights), leading to a quicker identification of new goods and their faster incorporation into the index. This is also the level at which more extensive quality adjustments and "comparable" substitutions could be made, recognizing the appearance of new outlets and new versions of services which provide consumers, effectively, with cheaper sources for the same or similar items consumed previously.

7. The BLS should investigate the impact of classification, that is item group definition, on the price indexes, to improve the ability of the index to fully capture item substitution.

In addition, a classification rule should be implemented for new products that groups them within the same low-level group (stratum) as those for which consumers are most likely to substitute for them. On-line news services which compete with newspapers, automobile purchases with leases, and drugs with the surgical procedures they replace, are examples of products for which direct comparisons are needed so that the full substitution effect can be accounted for.

8. There are a number of additional conceptual issues that require attention. The price of durables, such as cars, should be converted to a price of annual services, along the same lines as the current treatment of the price of owner-occupied housing. Also, the treatment of "insurance," should move to an ex-ante consumer price measure rather than the currently used ex-post insurance profits based measure.

9. The BLS needs a more permanent mechanism for bringing outside information, expertise, and research results to it.

This commission did not have the resources or the time to investigate all the various aspects of the CPI in adequate depth. Nor would a subsequent similar group if it were again assembled ad hoc. A more permanent body should be created, at the request of the BLS, organized by an independent public professional entity, such as the American Economic Association, the NRC-NAS or NBER, with a significant resource commitment. Such a group could pursue more fundamental research in cooperation with the BLS and provide a framework for experimentation with various alternative data collection and estimation approaches. It would also provide the BLS with a more permanent channel for access to a range of professional and business opinions on the statistical, economic and current market issues arising in the normal process of data collection, on index number construction, and on the implementation of some of the reforms suggested here.

Longer Run

10. The BLS should develop a research program to look beyond its current "market basket" framework for the CPI.

In the longer run, the big issues are new commodities and new services and the changing economic, social, and environmental climate within which the consumer is operating. This program should explore measuring the value of time saved by new medical procedures and communication devices, the value of life extended and its associated quality, the losses experienced, in terms of longer distances to new shopping centers by the closing of some neighborhood stores, and the "consumption" increases forced on consumers, by rising crime, new diseases, or changes in taxation.

11. BLS should investigate the ramifications of the embedded assumption of price equilibrium.

This assumption, which means that prices or quantities adjust immediately to quality changes or the introduction of substitutes is fundamental to many elements of the methodology and its failure to hold sometimes is at the heart of many of the issues discussed in this report. We recommend that BLS identify the methodological changes required to relax this major assumption as research warrants.

12. The BLS should develop a number of new data collection initiatives to make some progress along these lines.

First and foremost, BLS or a companion agency will need to collect data on detailed time use from a large sample of consumers. We would also need to extend the current health status survey to include more information on various "quality of life" issues. Progress should also be made (perhaps jointly with the BEA, which is already doing work in this area such as environmental satellite accounts) on incorporating data from victimization surveys and from various measures of the status of our physical environment into an experimental set of national satellite social-economics accounts, accounts that value not only the market consumption basket, but also the resulting leisure and quality of life experienced by the average individual. Such accounts could also provide information on the distribution of these measures across different age and social groups. It will be difficult to integrate these into the main cost of living framework, but over time, progress on these fronts should provide useful supplementary information to policy makers and the public.

Suggestions for Congress

13. Congress should enact the legislation necessary for the Departments of Commerce and Labor to share information in the interest of improving accuracy and timeliness of economic statistics and to reduce the resources consumed in their development and production.

14. Congress should provide the additional resources necessary to expand the CES sample and the detail collected, to make the POPS survey more frequent, and to acquire additional commodity detail from alternative national sources, such as industry surveys and scanner data.

While the Commission has identified some potential areas of cost savings, and it sympathizes with the recent trend to use private business sector methods to make the federal government more efficient, it notes the overwhelming trend for private businesses to be investing heavily on information technology, from tangible capital such as hardware to intangibles such as increasingly important software, to human capital.

15. Congress should establish a permanent (rotating) independent committee or commission of experts to review progress in this area every three years or so and advise it on the appropriate interpretation of the then current statistics.

This would be useful in its own right, but especially so to smooth the transition to a new index.

16. Congress and the President must decide whether they wish to continue the widespread overindexing of various federal spending programs and features of the tax code. If the purpose of indexing is fully and accurately to insulate the groups receiving transfer payments and paying taxes, no more and no less, they should pass legislation adjusting indexing provisions accordingly.

This could be done in the context of subtracting an amount partly or wholly reflecting the overindexing from the current CPI-based indexing. Alternatively, a smaller amount could be subtracted from the new revised annual index if and when it is developed and published regularly, to more closely approximate a true cost-of-living index.

We hasten to add that the indexed programs have many other features and raise many other issues beyond the narrow scope of a more accurate cost of living index. We also wish to express our view that these findings and their implications need to be digested and understood by the BLS, the Congress, the Executive Branch and the public.

To the Economists and Statisticians

These professions should treat training in data collection, data analysis, and interpretation more seriously and give it more space and attention in the standard curriculum. There should be more emphasis on measurement and sampling issues in the training of economists and statisticians. Effort should also be put into improving the ties between professionals in government and their academic and business colleagues. The academic world needs to be cognizant of the important work done by its colleagues in government who provide them with much of the "raw material" for their subsequent analyses and show more appreciation of their efforts and understanding of the constraints under which they are laboring.

IX. Conclusion

While the CPI is the best measure currently available, it is not a true cost of living index. It suffers from a variety of conceptual and practical problems as the vehicle for measuring changes in the cost of living. Despite important BLS updates and improvements in the Consumer Price Index, it is likely that changes in the CPI have substantially overstated the actual rate of price inflation. Moreover, revisions have not been carried out in a way that can provide an internally consistent series on the cost of living over an extended span of time. More importantly, **changes in the Consumer Price Index are likely to continue to overstate the change in the true cost of living for the foreseeable future.** This overstatement will have important unintended consequences, including overindexing government outlays and tax brackets and increasing the federal deficit and debt. If the intent of such indexing is to insulate recipients and taxpayers from changes in the cost of living, use of the Consumer Price Index has

in the past, and will in the future, substantially overcompensate (on average) for changes in the true cost of living.

This report has laid out a variety of issues to be addressed in developing a more accurate measure of the cost of living. It has also presented a series of recommendations to the agency responsible for the nation's price statistics and to the elected officials who are funders, supervisors, and consumers of those statistics. We have no doubt that implementation of our recommendations would greatly improve the accuracy of the nation's price statistics. This in turn would lead to more accurate measurement of everything from cost-of-living adjustments in private contracts and public programs to information for the Federal Reserve to improved inputs to the nation's national income and product accounts. These improvements in turn would better enable citizens and policy makers alike to measure economic progress over time, among groups, and across countries.

While the commission's best estimate of the overstatement of changes in the cost of living based on changes in the consumer price index is a little over one percent, our broader point is that even small differences compound over time and matter a lot. This was evidenced in Section II when the improvement in the treatment of owner-occupied housing was introduced early in the 1980's. The same is true of the recently corrected formula bias issues which added an additional bias of about 0.24 percent per year for 1979–95. The cumulative ramifications are substantial.

While subsequent analysis, research, and economic events may result in a slight change in these estimates—at least as likely to be up as down in our opinion—some care in their use is warranted. While the analysis in this report represents our best judgment, this Commission did not have the substantial resources that the previous major effort to examine the nation's price statistics, the so-called Stigler Commission, had in 1961. The Stigler Commission was able to commission and produce substantial original research, while this Commission did not have the time or resources to do so. Nevertheless, this report incorporates new information from a wide variety of sources, both within the government and from outside the government. We are gratified by the tremendous outpouring of suggestions, advice, and assessment of individual issues that have arisen in the course of the committee's investigation.

The readers of this report need some time to digest and understand the results, analysis, and recommendations. This includes the BLS, the Congress, the Executive Branch, the private sector, and academe. We very much hope that careful and thoughtful consideration will be given to the findings presented in this report in the spirit in which they are offered: an attempt to provide some guidelines on how to improve the production and use of the nation's price statistics and the continuous process of improving them in a complex, dynamic economy.

REFERENCES

Advisory Commission To Study the Consumer Price Index. 1995. "Toward a More Accurate Measure of the Cost of Living, Interim Report," September 15, 1995.

Aizcorbe, Ana M. and Patrick C. Jackman. 1993. "The Commodity Substitution Effect in CPI Data, 1982–1991," *Monthly Labor Review*, Vol. 116, pp. 25–33 (December 1993).

Armknecht, Paul A. 1996. "Improving the Efficiency of the U.S. CPI in the Future," paper presented at NBER Productivity Workshop, April 19, 1996.

Armknecht, Paul A. and Donald Weyback. 1989. "Adjustments for Quality Change in the U.S. Consumer Price Index," *Journal Of Official Statistics*, Vol. 5, No. 2, pp. 107–123 (1989).

Armknecht, Paul A., Brent R. Moulton, and Kenneth J. Stewart. 1995. "Improvements to the Food at Home, Shelter, and Prescription Drug Indexes in the U.S. Consumer Price Index," Working Paper 263, Bureau of Labor Statistics (1995).

Berndt, Ernst and Zvi Griliches. 1993. "Price Indexes for Microcomputers: An Exploratory Study," in *Price Measurements and Their Uses*, edited by M. Foss, M. Manser, and A. Young; NBER Studies in Income and Wealth, Vol. 57. Univ. of Chicago Press, 1993, pp. 63–99.

Berndt, Ernst R., Iain Cockburn, and Zvi Griliches. 1996. "Pharmaceutical innovations and market dynamics: Tracking effects on price indexes for anti-depressant drugs," *Brookings Papers on Economic Activity* (1996).

Berndt, Ernst R., Zvi Griliches, and Joshua G. Rosett. 1993. "Auditing the Producer Price Index: Micro Evidence from Prescription Pharmaceutical Preparations," *Journal of Business and Economic Statistics*, 1993, 11(3), 251–264.

Berndt, Ernst R., Zvi Griliches, and Neal Rappaport. 1995. "Econometric Estimates of Price Indexes for Personal Computers in the 1990's," *Journal of Econometrics*, 68 (1995) 243–68.

Berry, Steven, Samuel Kortum, and Ariel Pakes. 1996. "Environmental Change and Hedonic Cost Functions for Automobiles," NBER working paper 5746, September 1996, forthcoming in Proceedings of the National Academy of Sciences, 1996.

Boskin, M. and M. Hurd. 1985. "Indexing Social Security Benefits: A Separate Price Index for the Elderly," *Public Finance Quarterly*, Vol. 13, No. 4, pp. 436–449 (October 1985).

Brown, Claire. 1994. *American Standards of Living*, Oxford, UK: Basil Blackwell Ltd., 1994.

Carruthers, A. G., D. J. Sellwood, and P. W. Ward. 1980. "Recent Developments in the Retail Price Index, *The Statistician*, Vol. 29 (1980), pp. 1–32.

Congressional Budget Office. 1994. "Is the Growth of the CPI a Biased Measure of Changes in the Cost of Living?" *CBO Papers*, Washington. Congress of the United States, October, p. 32.

Congressional Budget Office. 1995. *The Economic and Budget Outlook: Fiscal Years 1996–2000*, Report to the Senate and House Committees on the Budget, Washington, Congress of the United States, January 1995, Table 2–8, p. 43.

Cutler, David and Mark McClellan. 1996. "The Determinants of Technological Change in Heart Attack Treatment," NBER working paper No. 5751, Cambridge, MA, September 1996.

Cutler, David M., Mark McClellan, Joseph P. Newhouse, and Dahlia Remler. 1996. "Are Medical Prices Declining?" NBER working paper No. 5750, September 1996.

Dahlen, J. 1994. "Sensitivity Analysis for Harmonizing European Consumer Price Indices," paper presented at the International Conference on Price Indices (October 31–November 2, 1994), Ottawa, Statistics Canada.

Diewert, Erwin. 1976. "Exact and Superlative Index Numbers," *Journal Of Econometrics*, Vol. 4, No. 2, pp. 115–145 (1976).

Diewert, Erwin. 1996. "Comment on CPI Biases," *Business Economics*, Vol. 31, No. 2, pp. 30–35 (April 1996).

Diewert, W. Erwin. 1995. "Axiomatic and Economic Approaches to Elementary Price Indexes," Working Paper No. 5104, Cambridge, National Bureau of Economic Research, Inc. (May 1995).

Division of Consumer Prices and Price Indexes. 1996a. "Changing the Hospital and Related Services Component of the Consumer Price Index," Bureau of Labor Statistics, Press Release (July 16, 1996).

Division of Consumer Prices and Price Indexes. 1996b. "Extending the Improvements in the CPI Sample Rotation Procedures and Improving the Procedures for Substitute Items," Bureau of Labor Statistics, Press Release (March 29, 1996).

Duggan, James, Robert Gillingham, and John Greenlees. 1995. "Housing Bias in the CPI and its Effect on the Budget Deficit and the Social Security Trust Fund," Office of Economic Policy, U.S. Department of the Treasury, June 30, 1995, p. 6.

Dulberger, Ellen R. 1989. "The Application of an Hedonic Model to a Quality-Adjusted Price Index for Computer Processors," in Technology and Capital Formation, edited by Dale W. Jorgenson and Ralph Landau, Cambridge, Mass: MIT Press.

Dulberger, Ellen R. 1993. "Sources of Price Decline in Computer Processors: Selected Electronic Components," in Price Measurements and Their Uses, edited by Murray F. Foss, Marilyn E. Manser and Allan H. Young, Chicago, Ill: The University of Chicago Press.

Fisher, F.M. and Zvi Griliches. 1995. "Aggregate Price Indices, New Goods, and Generics," *Quarterly Journal of Econometrics* (1995).

Fisher, Irving. 1922. *The Making of Index Numbers,* Boston, Houghton-Mifflin, 1922.

Gillingham, Robert and Walter Lane. 1982. "Changing the Treatment of Shelter Costs for Homeowners in the CPI," *Monthly Labor Review,* June 1982, p. 9.

Gillingham, Robert. 1974. "A Conceptual Framework for the Consumer Price Index," *Proceedings of the Business and Economics Section,* American Statistical Association, pp. 246–252 (1974).

Gordon, Robert J. 1990. *The Measurement of Durable Goods Prices.* University of Chicago Press for NBER, 1990.

Gordon, Robert J. 1996. "The Sears Catalog Revisited: Apparel and Durable Goods," mimeo, Northwestern University, April 1996, revised August 1996.

Griliches, Zvi and Iain Cockburn. 1994. "Generics and New Goods in Pharmaceutical Price Indexes," *American Economic Review,* 1994, 84(5), 1213–1232.

Griliches, Zvi and Iain Cockburn. 1996. "Generics and the Producer Price Index for Pharmaceuticals," in *Competitive Strategies in the Pharmaceutical Industry,* edited by Robert Helms, Washington, DC: AEI Press, 1996, pp. 19–34.

Hausman, Jerry A. 1966. "Valuation of New Goods Under Perfect and Imperfect Competition," in Timothy Bresnahan and Robert J. Gordon, eds., *The Economics Of New Goods,* Chicago, University of Chicago Press (1966).

Hausman, Jerry. 1996. "Valuation of New Goods Under Perfect and Imperfect Competition," in T. Bresnahan and R. J. Gordon, *The Economics of New Goods,* Studies in Income and Wealth (University of Chicago Press for NBER, 1996).

Jorgenson, Dale W. and Daniel T. Slesnick. 1983. "Individual and Social Cost-of-Living Indexes," *Price Level Measurement,* W.E. Diewert and C.Montmarquette (eds.), Ottawa, Statistics Canada, pp. 241–336 (1983).

Lebow, David E., John M. Roberts, and David J. Stockton. 1994. "Monetary Policy and 'The Price Level'," Washington, D.C., Board of Governors of the Federal Reserve System (1994).

Lebow, David E., John M. Roberts, and David J. Stockton. (no date) "Monetary Policy and 'The Price Level'," mimeo, Federal Reserve Board, Washington DC.

Liegey, Paul F., Jr. 1990. "Adjusting Apparel Indexes in the Consumer Price Index for Quality Differences," in Murray F. Foss, Marilyn E. Manser, and Allan H. Young (eds.), *Price Measurements and Their Uses.* Chicago: University of Chicago Press for NBER.

———. 1994. "Apparel Price Indexes: Effects of Hedonic Adjustment," *Monthly Labor Review* (May), pp. 38–45.

Moulton, Brent R. 1996. "Estimation of Elementary Indexes of the Consumer Price Index," Bureau of Labor Statistics (May 1996.)

Moulton, Brent R. and Karin E. Smedley. 1995. "A Comparison of Estimator for Elementary Aggregates of the CPI," Bureau of Labor Statistics (June 1995).

Nordhaus, William D. 1996. "Do Real-Output and Real-Wage Measures Capture Reality? The History of Light Suggests Not," in T. Bresnahan and R. J. Gordon, *The Economics of New Goods,* Studies in Income and Wealth (University of Chicago Press for NBER, 1996).

O'Neill, June. 1995. "Prepared Statement," *Consumer Price Index,* hearings before the Committee on Finance, U.S. Senate, 104th Congress, First Session, Washington, U.S. G.P.O., Table 1, p. 146.

Pakes, Ariel, Steven Berry, and James A. Levinsohn. 1993. "Applications and Limitations of Some Recent Advances in Empirical Industrial Organization: Price Indexes and the Analysis of Environmental Change," *American Economic Review,* Vol. 83, No. 2, pp. 241–246 (May 1993).

Randolph, William C. 1988. "Housing Depreciation and Aging Bias in the Consumer Price Index," *Journal of Business and Economic Statistics* (July 1988), Vol. 6, No. 3, pp. 359–71.

Reinsdorf, Marshall B. 1993. "The Effect of Output Price Differentials on the U.S. Consumer Price Index," in Murray F. Foss, Marilyn E. Manser, and Allan H.

Young, eds., *Price Measurements And Their Uses,* Studies in Income and Wealth, Vol. 57, Chicago, University of Chicago Press for NBER.

Repetto, Robert, *et al.* 1990. *Has Environmental Protection Really Reduced Productivity Growth?* World Resources Institute, 1990.

Schultz, B. J. 1994. "Choice of Price Index Formula at the Micro-Aggregation Level: The Canadian Empirical Evidence," paper presented at the International Conference on Price Indices (October 31–November 2, 1994), Ottawa, Statistics Canada.

Shapiro, M. D. and D. W. Wilcox. 1996a. "Alternative Strategies for Aggregating Prices in the CPI," paper prepared for the Conference on Measuring Inflation and Real Growth, Federal Reserve Bank of St. Louis, October 1996.

Shapiro, Matthew D., and David W. Wilcox. 1996c. "Mismeasurement in the Consumer Price Index: An Evaluation." *NBER Macroeconomics Annual 1996.*

Slifman, L., and C. Corrado. 1996. "Decomposition of Productivity and Unit Costs," mimeo, Federal Reserve Board of Governors, November 1996.

Statistical Abstract of the United States. 1978.

———. *1990.*

———. *1995.*

The U.S. Consumer Electronics Industry in Review 1995, Electronic Industries Association, p. 13.

Tornqvist, Leo. 1936. "The Bank of Finland's Consumption Price Index," *Bank Of Finland Monthly Bulletin,* Vol. 10, pp. 1–8 (1936).

Trajtenberg, Manuel. 1990. "Product Innovations, Price Indices, and the (Mis)Measurement of Economic Performance," NBER Working Paper No. 3261, February 1990.

Triplett, Jack. 1990. "The Theory of Industrial and Occupational Classifications and Related Phenomena," *Proceedings of the Bureau of the Census 1990 Annual Research Conference,* August 1990.

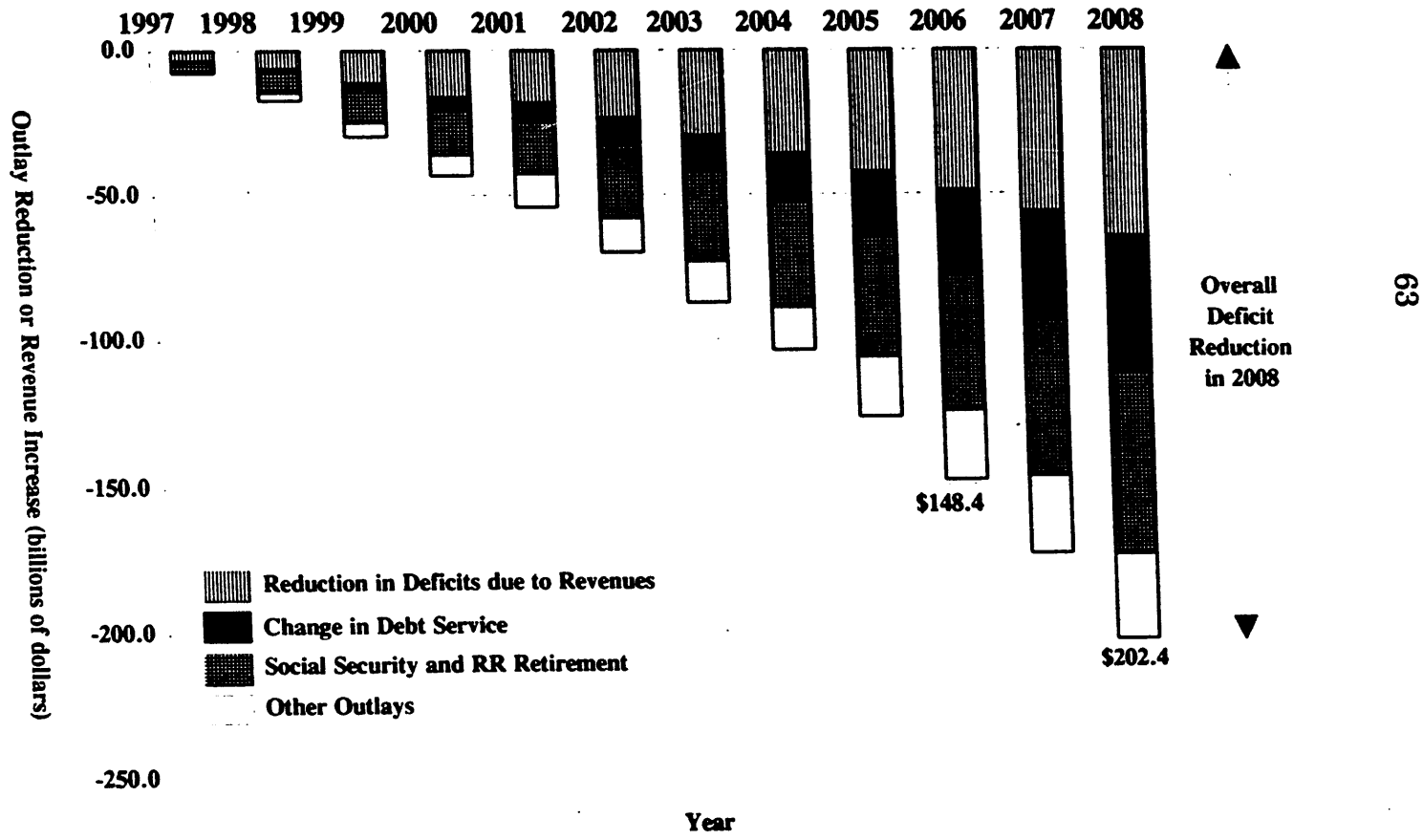
U.S. Census Bureau. 1993. *American Housing Survey, 1993.*

U.S. Census of Housing, 1970.

Woolford, K. 1994. "A Pragmatic Approach to the Selection of Appropriate Index Formulae," paper presented at the International Conference on Price Indices (October 31–November 2, 1994), Ottawa, Statistics Canada.

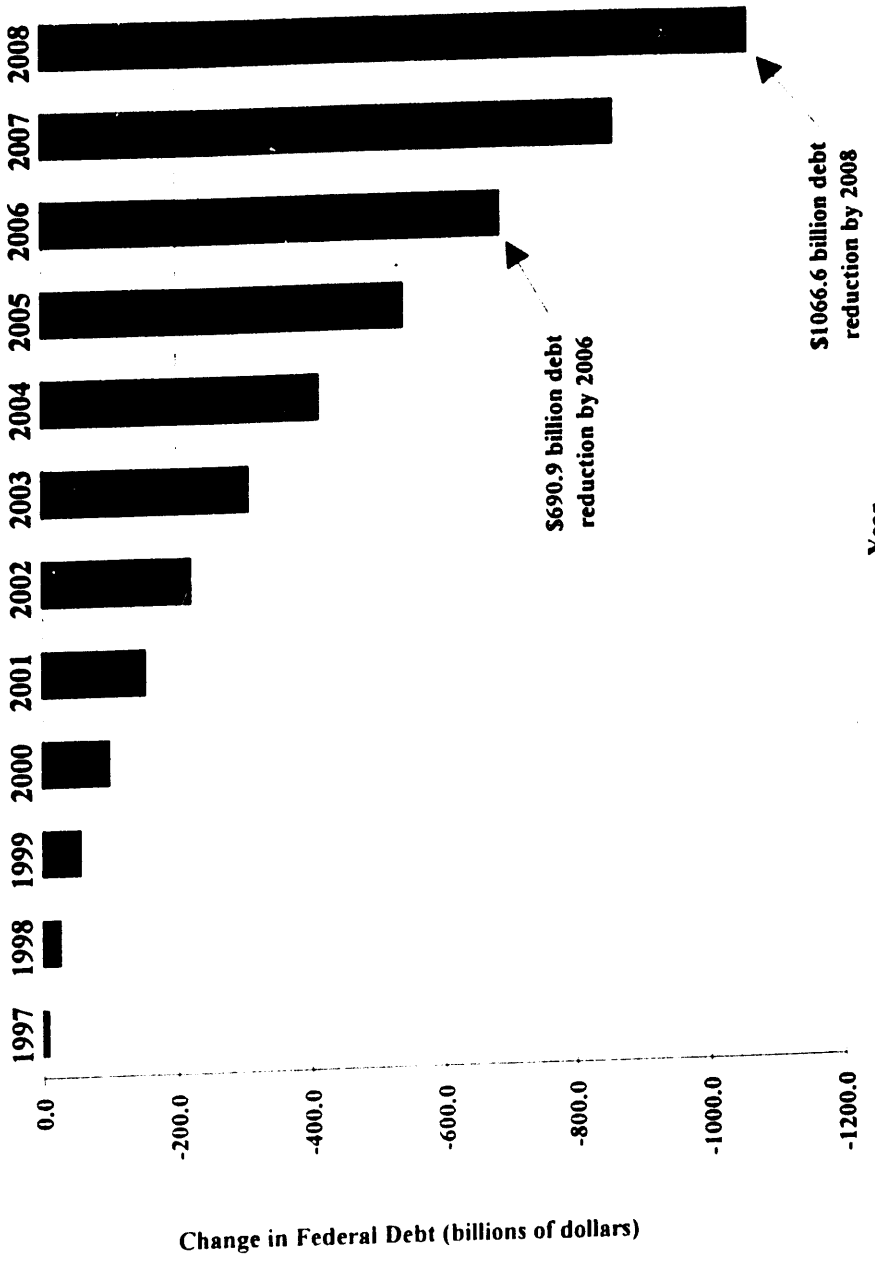
APPENDIX

Figure A-1
Effect of Correcting a 1.1 Percentage Point Overstatement in the CPI on Annual Federal Deficits



Source: 1997-2006 CBO; 2007-2008 Commission estimate.

Figure A-2
Effect of Correcting a 1.1 Percentage Point Overstatement in the CPI on Future Federal Debt



Source: 1997-2006 CBO; 2007-2008 Commission estimate.

MEMBERS OF THE ADVISORY COMMISSION TO STUDY THE CONSUMER PRICE INDEX

Michael J. Boskin, Ph.D., Chairman
Tully M. Friedman Professor of Economics
& Senior Fellow, Hoover Institution
Stanford University
Stanford, California

Ellen R. Dulberger, Ph.D.
Director of Marketing Strategy
IBM Personal Computer Company
Somers, New York

Robert J. Gordon, Ph.D.
Chairman, Department of Economics
& Stanley G. Harris Professor in the Social Sciences
Northwestern University
Evanston, Illinois

Zvi Griliches, Ph.D.
Paul M. Warburg Professor of Economics
Harvard University
Cambridge, Massachusetts

Dale Jorgenson, Ph.D.
Chairman, Department of Economics
& Frederic Eaton Abbe Professor of Economics
Harvard University
Cambridge, Massachusetts