

Editor's Overview

THIS TWELFTH ISSUE OF THE *International Productivity Monitor*, published by the Centre for the Study of Living Standards, differs from past issues. Five of the six articles address one topic, namely the impact of the Boskin Commission after one decade on price measurement. A final article discusses the role of information technology in the US growth resurgence.

In December 1996, the Boskin Commission released its final report, *Toward a More Accurate Measure of the Cost of Living*, prepared for the US Senate Finance Committee. The Commission investigated possible sources of bias in the US Consumer Price Index (CPI) and concluded that the CPI in 1995-96 was upward biased by 1.1 percentage points per year. This startling finding had important ramifications for price measurement in both the United States and throughout the world. The articles in this symposium, by leading researchers in the price measurement field, examine from different perspectives the impact of the Boskin Commission after one decade.

In his brief introduction to the symposium, **Jack E. Triplett** of Brookings Institution, and organizer of the session at the January 2006 annual meeting of the American Economic Association where these papers were originally presented, highlights the importance of price measurement for reliable productivity estimates. For Triplett, accurate price indexes are essential for reliable productivity measurement. Indeed, he points out that a one percentage point upward bias in price changes results in a one percentage point downward bias in real output growth and by consequent productivity growth. An upward bias in price indexes implies that productivity growth is being underestimated.

In the first article in the symposium, **Robert J. Gordon** from Northwestern University, one of the five members of the Boskin Commission,

summarizes the report's methods, findings, and recommendations, and then reviews the comments and criticisms that appeared soon after the report was issued. Changes in CPI methodology are also summarized and assessed, as is recent research on related issues. Gordon sharply distinguishes two questions. First, with what we know now, what should the Commission have concluded about CPI bias in 1995-96? Second, what is the bias now after the many improvements introduced into the CPI since the Commission's report?

On the first question, Gordon notes that his own recent research on apparel and rental housing indicates a substantial downward bias in the CPI over much of the twentieth century, diminishing in size after 1985. Incorporating these findings into the Boskin matrix would reduce its 0.6 percentage point annual upward bias due to quality change and new products to a smaller 0.4 point bias. However, this is more than offset by the stunning discrepancy over 2000-06 in the chain-weighted C-CPI-U compared to the traditional CPI-U, indicating that the Commission greatly understated the magnitude of upper level substitution bias, that is the substitution between broad consumer expenditure categories. This retrospective evaluation suggests that the bias estimate for 1995-96 should have been 1.2 to 1.3 percentage points, not 1.1 points.

Gordon estimates that the upward bias in the CPI has declined from the revised 1.2-1.3 percentage points in the Boskin era to about 0.8 points today. Yet he notes that the Boskin report,

like most contemporary studies of quality change, failed to accord sufficient importance to the value of new products and increased longevity. Allowing for these, he concludes that the current upward bias in the US CPI is at least 1.0 percentage points per year.

In the second article in the symposium **John S. Greenlees**, Associate Commissioner of Prices and Living Conditions at the US Bureau of Labor Statistics (BLS), provides a BLS response to the Boskin Commission from the perspective of ten years following the release of the report. Greenlees documents the research on price indexes done at the BLS in the first half of the 1990s that pointed to upward CPI bias, and discusses how these results attracted the attention of the US Senate, leading to the appointment of the Boskin Commission in 1995.

Greenlees provides a detailed discussion of the methodological changes to the CPI made by the BLS between 1996 and 2002 in three areas corresponding to the categories of bias identified by the Commission: upper and lower level substitution bias, quality change and new products, and outlet bias. A key change in the first area was the introduction of a chained CPI (C-CPI-U) that captured consumer substitution as much as possible. This was the first official superlative CPI produced by a statistical agency in the world. In the second area, the BLS has introduced more hedonic models to capture quality change, but the overall quantitative impact has been small. The BLS has also recognized the need to use a product and outlet sample that was as representative as possible of current consumer spending patterns. Viagra was quickly included in the CPI.

Greenlees concludes that the Boskin Commission, by forcing the BLS to scrutinize the strengths and limitations of its CPI procedures and by highlighting and publicizing the budgetary impacts of the CPI, paved the way for various CPI improvements.

In the third article in the symposium, **Jack E. Triplett** of the Brookings Institution begins by highlighting the extremely salutatory effect the Boskin Commission has had on international price statistics, promoting open discussion of price measurement issues, engendering dialogue between statistical agencies and users, and encouraging research. Less positive in Triplett's view has been the Boskin Commission's popularization of "guestimates," through its widely cited 1.1 percentage point CPI bias figure. Triplett characterizes a guestimate as a number produced when one does not have research results, but he does acknowledge that without its guestimate the report would have likely had minimal impact.

Triplett argues that the Commission ignored the possibility that quality improvements could actually produce a net downward bias to CPI components because the implicit quality adjustments inherent in the BLS procedures may over-adjust. Triplett points out that the motivation for the appointment of the Boskin Commission was highly political, namely a desire to reduce Social Security expenditures by indexing benefits to a lower rate of increase than the CPI. He feels that a mix of politics and statistics seldom produces an output that is favourable to economic statistics. For Triplett, it would have been preferable to separately address the distinct issues of CPI measurement and principles for allocation of resources to the dependent population.

In the fourth article in the symposium, **Ernst R. Berndt** from MIT provides a political economy interpretation of the rise and fall of public interest in price measurement, placing these developments in the context of the attempt by Congress and the White House to deal with growing deficits in the early to mid-1990s. He provides a detailed discussion of initiatives since the Boskin Commission, such as the National Academy of Sciences panel, to improve CPI measurement.

Berndt examines the thorny issue of the CPI for health care, with particular reference to the Boskin Commission recommendation that BLS move from pricing health care inputs to pricing health care outcomes. Because of the formidable measurement challenges in adjusting medical care expenditures for changes in outcome quality, little progress has been made in this area. Berndt concludes that the BLS has responded constructively to the recommendations from the various price measurement initiatives. By implementing many of the methodological changes suggested, the BLS has reduced net CPI inaccuracy and increased professional confidence in the reliability of the CPI.

In the fifth and final article in the symposium, **Martin Neil Baily** of the Institute for International Economics, and a former Chair of the US Council of Economic Advisors, discusses the policy implications of the Boskin Commission. He begins by offering support for the type of the back-of-the-envelope calculations of CPI bias that the Commission used so effectively to attract public attention to its report. In the area of quality adjustment, however, Baily criticizes the Boskin Commission for what he calls “premature extrapolation,” that is moving too quickly from a limited number of examples to a broad conclusion.

Baily stresses the importance of high-quality data for policy decisions. He observes that a better allocation of existing resources can improve economic statistics, suggesting that the creation of a unified statistical agency in the United States, like Statistics Canada, would streamline data collection and analysis. In terms of the issue of Social Security solvency, Baily argues that use of the CPI to adjust social security benefits downward is not a preferred option. Echoing

Triplet, Baily concludes that the Commission should have advised Congress that it did not have an adequate scientific basis to recommend a specific quantitative adjustment to the CPI index used to adjust federal programs.

The US economy has enjoyed a remarkable rebound in productivity and output growth in the last decade. In the sixth and final article in the issue, **Daniel E. Sichel** of the Federal Reserve Board reviews the book *Information Technology and the American Growth Resurgence* by Dale Jorgenson, Mun Ho, and Kevin Stiroh, which provides a detailed analysis of this rebound. Sichel begins by noting that the book can be considered a “Users’ Guide” to growth accounting and is highly recommended in this regard. The basic story as told by Jorgenson et al. and to which Sichel is sympathetic is as follows. In the mid-1990s the constant-quality prices of semiconductors fell substantially, leading to rapid declines in the price of Information Technology (IT) capital goods. Firms responded by substituting capital purchases toward IT capital, resulting in a surge in IT capital deepening and labour productivity growth.

Sichel reviews in an even-handed manner the critiques that have been put forward of the growth accounting methodology from which Jorgenson et al. derive their results. His bottom line is that while many of the critiques make valuable points, there is currently no alternative methodology to growth accounting that offers such a comprehensive framework for assessing the sources of economic growth. Sichel notes that one limitation of the book is that it provides no analysis of the post-2000 US productivity growth acceleration, which has taken place in a period when rapid IT capital deepening was not occurring.

The Boskin Commission Report After a Decade: Introduction to the Symposium and Implications for Productivity

Jack E. Triplett
Brookings Institution

ACCURATE PRICE INDEXES ARE essential for reliable productivity measurement. Most measures of output, the numerator of the productivity ratio, are produced through deflation of an expenditure total by a price index — direct measurement of output is rare. Though the labour input is formed from measures of hours (so it is a quantity index, where different labour types are typically weighted by compensation), the other input measures used in measuring multifactor productivity (intermediate inputs and capital services) are, like the output measure, estimated through deflation.

With the deflation procedure, any error in the price index translates into an equal error of opposite sign in the associated deflated quantity measure. Thus, if the output price index rises, say, one percentage point per year too rapidly, the growth in output is understated by the same amount. The output growth error would in turn translate into a downward error in the estimate of productivity growth, provided the inputs are measured accurately (but they might of course have their own biases).

In late 1995, the Commission to Study the Accuracy of the Consumer Price Index, known

as the Boskin Commission after the Chair Michael Boskin of Stanford University,¹ issued its “interim report.” This was followed at the end of 1996 by a “final report.” The Commission estimated that the U.S. CPI was upward biased by 1.1 percentage points per year. This is a huge error when the *measured* annual rate of consumer inflation since 1991 has only averaged 2.6 per cent per year.

Although the main focus of the Commission’s report was possible over-statement in the escalation of income payments to U.S. Social Security recipients, CPI components are used for deflation of components of the national accounts, not only in the United States, but in every country. The Commission’s estimate implied that output growth was understated in the United States, though the precise amount is difficult to determine, since national accounts deflation does not use the overall CPI, but rather its components, and other price indexes are also used (Producer Price Index components, for example). Nevertheless, the Commission’s estimate, if accurate, had strong implications for productivity measurement, whether or not the price indexes used for deflating capital and other inputs had their own biases.

1 The other members of the Commission were Ellen Dulberger from IBM, Robert J. Gordon from Northwestern University, and Zvi Griliches and Dale Jorgenson from Harvard University.

The Commission's report had a tremendous impact, not only in the United States, but also internationally because economists in other countries rightly saw that CPI measurement error was not specific to one country. A decade has passed since the Commission's interim report set off the international discussions, repercussions, and controversy, time enough to form an assessment of the report's analysis, its findings and its influence. The five articles in this symposium, based on papers originally presented at a session entitled "The Boskin Commission After a Decade: Is the CPI Still Biased?"

held at the annual meeting of the American Economic Association in Boston, January 6-8, 2006, provide such an assessment.

The American Economic Association session where the papers were presented was well attended, which is testimony that the issues the Commission raised are still live ones a decade later. The papers from the session are being published in this issue of the *International Productivity Monitor* because the accuracy of price indexes is of vital concern to those who produce and to those who use productivity measures.

The Boskin Commission Report: A Retrospective One Decade Later

Robert J. Gordon¹
*Northwestern University and
National Bureau of Economic Research*

MORE THAN TEN YEARS HAVE now elapsed since the formation of the Boskin Commission, more formally the “Advisory Commission to Study the Consumer Price Index,” which was appointed by the Senate Finance Committee in June, 1995. The Commission's final report (Boskin et al., 1996) was issued on December 4, 1996 and concluded that the U. S. Consumer Price Index (CPI) overstated inflation by 1.1 percentage points per year for the time period of the Commission's deliberations, 1995-96. The report suggested that the CPI bias might have been larger before 1995 and predicted that it would be lower after 1997.

A sharp distinction must be drawn between the technical and scientific issues which led the Commission to its bias estimate of 1.1 percentage points, and the political and redistributive implications of that conclusion. The Commission report contained not only the technical background for its overall bias estimate but also its estimates of the vast amount by which that bias had increased the federal budget deficit looking back into the past and would increase

the deficit looking forward into the future.² The suggestion that the bias had caused excessive growth in Social Security and other benefits evoked a sharp and damning political reaction, as the AARP (American Association of Retired Persons) sent its lobbyists scurrying through the corridors of Congress to throw cold water on those senators and representatives who had initially been sympathetic to reducing the budget deficit by adjusting the indexation formula by some fraction of the Commission's bias estimate, the so-called “CPI minus X” approach to indexation. This paper ignores the implications of the bias estimate for fiscal policy and instead concentrates entirely on the issues of interest to the worldwide measurement community.

The Boskin Commission's Approach and Method

The Boskin Commission represented the first extensive external evaluation of the nation's price statistics in more than three decades; the previous report was the famous Stigler (1961) Commission Report. There were several

1 The author served as one of the five members of the Boskin Commission. Portions of the summary of the Boskin Commission findings and its reactions to its critics are adapted from Boskin et al. (1998) and Gordon (2000). The evaluations and opinions contained herein are those of the author only and should not be taken to represent the view of any other member of the Boskin Commission. I am grateful to Jerry Hausman for helpful discussions. Email: rjg@northwestern.edu

2 The Commission calculated that the 1.1 percentage points bias would contribute \$1.07 trillion to the U.S. national debt over the period 1997-2008 as compared to an alternative scenario in which Social Security and other programs were indexed by a formula which subtracted the bias from the published CPI.

Table 1
Boskin Commission Estimates of Bias in
the Consumer Price Index in 1995-96
 (percentage points per annum)

| Source of Bias | Estimate |
|-------------------------------|---------------|
| Upper Level Substitution | 0.15 |
| Lower Level Substitution | 0.25 |
| Outlet Substitution | 0.10 |
| New Products / Quality Change | 0.60 |
| Total | 1.10 |
| Plausible Range | (0.80 — 1.60) |

important differences between the two Commissions. The Boskin report concerned only the CPI, while the Stigler report also covered the Producer Price Index (PPI) and agricultural price indexes. The Stigler report did not produce any numerical bias estimates, whereas the mandate of the Boskin Commission included the provision for a point estimate of the overall bias in the CPI. The Stigler Commission had a substantial budget to commission new research studies, whereas the Boskin Commission had no research budget at all. Thus the Boskin Commission report of necessity was a survey article based in part on ongoing research, not a producer of new research.

A novel aspect of the Commission's method was to divide up the CPI into 27 categories and develop a separate estimate of quality change bias for each. This required some extrapolation from categories for which research existed (e.g. personal computers and TV sets) to related categories where no research existed (e.g. electronic toys). While some commentators have complained of the subjective nature of the Commission's assessment of quality change bias, the Commission felt that to assume that the bias was exactly zero in categories which had not been the subject of previous research was just as subjective as to extrapolate the results of related research.

Summary of the Commission's Findings

Findings Related to Substitution

The Commission's findings are summarized in Table 1. In this section we discuss the findings related to substitution, comprising the first three lines in the table. Next we discuss the category of new products and quality change.

The CPI relied on fixed weight Laspeyres indexes which did not account for consumer substitution among commodities. These Laspeyres measures of inflation were inherently an upper bound, and empirical studies led the Commission to conclude that this source of bias amounted to about 0.4 percentage points per year. Of this 0.4 percentage points, 0.15 came from estimates of the effect of substitution among the upper-level "strata" (apples vs. bananas) and 0.25 from the effect of substitution among the lower-level categories (Red Delicious apples vs. Jonathan apples).

In addition to substitution bias among commodities, there is an outlet substitution bias, the third line in Table 1, which refers to the practice of the BLS in ignoring differences in prices for the same item across outlets. Since price data are collected *within* outlets, the shift of consumers to purchases from discounters does not show up in the CPI as a price decline even though consumers reveal their preferences for these outlets by their purchases, measured by the steady shift in market share. We estimated this adds another 0.1 percentage point of upward bias.

Findings and Recommendations Related to New Products and Quality Change

The three types of substitution bias listed in Table 1 account for just a little under half (0.5 percentage points) of the 1.1 percentage point bias identified by the Boskin Commission. Slightly over half (0.6 percentage points) results from the difficulty of adjusting fully for quality

change and the introduction of new products. Economists have known since Hicks (1940) that the introduction of a new product should be dealt with in a COL index by estimating its reservation price and including the consumer surplus attributable to the introduction of the product. While accepting this framework, the Boskin Commission did not make any of its own estimates of the consumer-surplus value of new products. Instead, it took the more cautious view of primarily including estimates of explicit dimensions of quality change and the late introduction of major new products into the index. VCRs, cellular phones, and other products were included in the CPI a decade or more *after* they had penetrated the market and *after* their price had fallen by 80 per cent or more. Previous research allowed an estimate to be made of the bias in the CPI that occurred as a result of the late introduction of these new products.

The Commission attempted to identify commodity categories in which the CPI was biased for failing to take account of improvements in quality. To carry out this task, it examined separately 27 subcomponents of the CPI to determine the extent of quality bias, if any, and established an estimate for quality bias within each of the 27 categories. In eight of the categories the Commission discovered no research evidence or other factor that would indicate a bias other than zero. In the other 19 categories bias estimates were assigned which ranged as high as 3.0 percentage points per year for medical and hospital services

and 5.6 percentage points for appliances and electronic goods.³ When these 27 bias estimates were weighted by the relative importance of each category in 1995 (based on 1982-84 expenditure weights), the overall quality and new product bias was determined to be 0.6 percentage points per year.⁴

While the Commission formulated a point estimate of CPI bias related to quality change and new products, it did not attempt any quantitative estimate of changes in the quality of life. Nevertheless, it did present an informal discussion of changes in the quality of life and concluded that the “good” outweighed the “bad.” It cited a reduction in air and water pollution, a decline in crime by various measures, a decline in the suicide and infant mortality rates, and an increase in life expectancy.⁵ It also cited a wide range of improvements in goods related to quality and new products for which the Commission's estimates made no allowance, namely “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.”⁶ As other improvements for which no allowance was made, it cited the spread of cable and satellite TV and the new availability of the World Wide Web to owners of personal

3 Table 2 of the Boskin Commission report attempted to discriminate between different periods in assigning the estimates for quality change and the late introduction of new products. Thus, for instance, the bias estimate for prescription drugs was 3.0 percentage points for 1970-95 and 2.0 percentage points for 1995-96, reflecting the change in CPI methodology in 1995 which recognized the introduction of a generic version of a given drug as a price decline.

4 The only case in which alternative weights were used was in the case of consumer appliances and electronic goods, where the 1982-84 CPI weights were deemed to be too low, and weights were taken instead from the National Income and Product Accounts. See Boskin et al. (1996), Table 2, note a.

5 The one serious negative which was cited was the increase in births to unmarried women. Other “intangible negatives” cited were “increased job insecurity, possible increased inequality, and decreased job opportunities for workers with only a high school education” (Boskin et al., 1996:76).

6 Although not identified as such in the Boskin report, this quote was copied directly from Gordon (1990:560).

computers (another product which was introduced into the CPI long after it began to become common in homes). The Commission's section on quality-of-life issues concluded that "...the major increase in longevity... perhaps swamps everything else. Accordingly, our estimate of the current bias in the CPI is, if anything, probably understated" (Boskin et al., 1996:77).

The Commission's Recommendations

The Commission's first and overarching recommendation was that the BLS should establish the COL index as its objective in measuring consumer prices. All of the other, more specific recommendations were aimed at achieving this goal. The emphasis on this first recommendation may seem strange to economists, but the BLS in its publications for decades had explicitly stated that the CPI is not a COL index. The Commission stated that a fundamentally sound COL index could and should be developed.⁷

As part of the short-run set of recommendations, the Commission recommended that the single current CPI, which can never be revised (due to its use in legal contracts and for indexation) be supplemented by a second "research-based" index which is published annually rather than monthly and is continuously revised to incorporate new research results. The timely monthly index would continue to be called the CPI and would move toward a COL concept by adopting a "superlative" index formula to account for changing market baskets, abandoning the Laspeyres formula at both the upper level and lower level of aggregation.

At the upper level the Commission recommended that the BLS should move away from the assumption that consumers do not respond at all to price changes in close substitutes, moving instead to a "trailing Tornqvist" index (weighted

geometric mean of price relatives) or another approximation to a superlative index, and also, concurrently, to geometric means of price relatives at the elementary aggregation level. These changes would eliminate the problem of the growing irrelevancy of market baskets based on decade-old consumption patterns, reduce significantly the substitution and (any remaining) formula bias, and facilitate the speedier introduction of new goods and services into the index.

The distinction between the "timely" CPI and the new annual research-based index rests on the fundamental proposition that the basic monthly CPI can never be revised. The Commission recommended that as subsequent data became available, the weights were updated, and new goods were introduced and the history of their price changes was extended backward, the information incorporated in the published CPI could undergo retroactive revision of a new annual cost-of-living index using a compatible "superlative-index" formula that would no longer be affected by the lag in the availability of the required expenditure weight data. This alternative 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.

Continuing with recommendations that would require a somewhat longer term for implementation, the Commission felt that the BLS should revise its approach to sampling. The Commission was astonished at the number of price quotations that the BLS collected on ordinary products like bananas that were not subject to quality-change or new-product bias, relative to the effort devoted to collecting price data on new products like cellular phones, personal computers, and computer peripherals. In addition to reducing the data collection effort devoted to apples and bananas, the

7 The BLS has stated (U.S. Bureau of Labor Statistics, 1997) that it has embraced the Commission's overarching recommendation concerning the objective of the CPI.

Commission believed that the data-collection effort should be divided up between national and local goods. In this revised scheme there would no longer be any collection of data on appliances, other consumer durables, and imported produce (like tomatoes and bananas) in each separate city — data for these national goods could be collected in a much smaller sample applied to the entire country. This would free up resources to do a better job of collecting prices for local goods with components that might vary across cities, e.g., fuel costs, rent, household services, and non-imported produce.

Criticism of the Commission's Findings and Its Reaction

The report of the CPI Commission received much attention. Most of the findings and recommendations have generally been accepted by the economics profession, if we are to judge by various symposia in which prominent academic economists have been invited to comment on the Commission's findings. External corroboration comes from several sources, including the Federal Reserve's study of productivity (Slifman and Corrado, 1996), Nordhaus' (1998) analysis of survey data on changes in economic well-being, and Diewert's (1998) thoughtful complementary analyses of bias by type.

It is noteworthy that few if any criticisms addressed the Commission's findings related to substitution. Instead, most of the criticisms of the Commission's findings centered on its treatment of quality change and new products.

- The Commission did not give adequate attention to quality deterioration (Abraham, 1997, U.S. BLS, 1997, Abraham et al., 1998);

- The BLS already makes many quality adjustments, which the Commission did not adequately credit (Abraham, 1997, U.S. BLS, 1997, Moulton, 1996, Moulton and Moses, 1997);
- The Commission made too many back-of-the-envelope calculations and was too willing to generalize from research on one item to research on related items, both with respect to quality change and with respect to the desirability of geometric means to deal with lower level substitution bias (Abraham et al., 1998, Nordhaus 1998);
- The Commission's estimates, whether right or wrong, are of limited use because they cannot be generated from a mechanical procedure implementable in a month-to-month price program (U. S. BLS, 1997, Abraham, 1997);
- We should have been more aggressive in our estimates of the value of new products (Hausman, 1997b, Nordhaus, 1998).

The Commission's Response to the Critiques⁸

Because there was so little criticism of the Commission's estimates of substitution bias, it is perhaps worth noting that research by Shapiro and Wilcox (1997) suggests that the bias could have been a tenth of a percentage point or so higher than the estimate in our report, which was based heavily on BLS research.⁹ The BLS response to our recommendations about lower-level substitution bias pointed out that, while the degree of substitution between some goods like white shirts and blue shirts may be very high, that between other goods like particular types of pharma-

⁸ This section combines sections of the jointly authored paper by commission members (Boskin et al., 1998) with my own further observations.

⁹ Many of the best studies of substitution bias have been done by BLS researchers (see, e.g., Aizcorbe and Jackman, 1993). The studies from systems of demand equations at high levels of aggregation generally also come to an estimate of 0.2 to 0.25 on both U.S. data and data from other countries. See Greenlees (1997) for further elaboration of the BLS research.

ceuticals may be very low (Abraham et al., 1998:32). However, pharmaceuticals are the exception, not the rule, and within the vast majority of the 207 strata there is ample latitude for substitution, whether we are talking about types of food, shelter, apparel, cars, tires, appliances, entertainment, or personal care items. In fact, pharmaceuticals comprise only 1.2 per cent of the upper-level weight in the CPI. If anything, the unitary elasticity of substitution embodied in the Commission's recommendation of geometric weights at the lower level probably understates the extent of substitutability within the majority of the strata.

The Debate over Quality Change and New Product Bias

Most of the criticism focused on our extensive analysis of quality change and new product bias and our willingness to extrapolate a bias estimate from one category to another, and this critique of “guesstimates” is a major theme of Triplett's (2006) paper in this symposium. The notion that assuming zero bias is scientific, whereas attempting to generalize cautiously from related goods or practical reasoning is not precise enough, throws out available information. For instance, even though we will never precisely measure the value of the invention of the jet airplane, as economists we *know* that consumer surplus triangles have an area that is positive rather than zero.

Most of the Commission's estimates of quality change were based on the collection of price data from independent sources and the careful quality adjustment of those independent data. Independent sources of price data were employed in our bias estimates for shelter, appliances, radio-TV, personal comput-

ers, apparel, public transportation, prescription drugs, and medical care. Estimates derived from these categories were extrapolated, sometimes partially rather than fully, to other house furnishings, nonprescription drugs, entertainment, commodities, and personal care. This left only a few remaining categories where we added a bias estimate to the CPI category in which there were already quality adjustments, rather than computing the bias estimate indirectly by subtracting an independent estimate from the CPI estimate for the same category. These categories were food and beverages, other utilities, new and used cars, motor fuel, and personal expenses. The BLS did not object to our “down in the trenches” approach to the problem. Indeed, Moulton and Moses (1997:308) state, “This is the first time that a systematic analysis of quality bias has been done category by category, which we consider to be a noteworthy accomplishment of the Commission ...[the] overall approach seems to us to be a sensible and useful way to approach the problem of coming up with an overall assessment of bias, and we expect this type of structure will prove to be useful in the future.”

Some outside critics of the Commission argued that the BLS already does a great deal of quality adjustment, and that the Commission report is flawed for ignoring the extent of the BLS adjustments.¹⁰ However, for most categories, the extent of current BLS quality adjustments is irrelevant to an assessment of the Commission's treatment of quality change. We were comparing our own evidence to the corresponding CPI indexes — however they are quality adjusted, in a major or minor way — and thus our estimates of quality change bias are a *residual* that remains after the BLS has completed its efforts.

10 In fact most of the Moulton-Moses paper (1997) is devoted not to a direct critique of the Commission's estimates, but to an explanation of how the CPI is adjusted for quality change and to an attempt to estimate the quantitative significance of those adjustments (see pp. 322-48).

However, it is still instructive to discuss what the BLS calls quality adjustment, since it illustrates the substantive and communication difficulties in this field.¹¹ There is very little explicit adjustment for quality change (Nordhaus, 1998). Most of the reported “quality adjustment” by the BLS, 1.65 out of the 1.76 percentage points in Moulton and Moses (1997) comes from “linking” procedures, where a missing item is replaced by another (excluding outliers, commodity pairs where the implicit price-quality differential exceeds 100 per cent, the quality adjustment number shrinks to 0.3 per cent). No judgment at all is made about the quality differential between the new and old item. Roughly one out of three items disappear sometime during the year and have to be replaced by a different item in the same general class, such as a larger versus a smaller package of yogurt or a blue raincoat versus black. But this churning is not what we had in mind by “quality change,” which rather involves the appearance of new and improved goods, greater speed, durability, variety, convenience, safety, and energy efficiency.

The area in which the Commission's bias estimates are surely likely to be too low, rather than too high, is in the area of new products. We did not make explicit allowance for the late introduction into the CPI of numerous new products. We indicated that the appropriate way to deal with new products is to value the consumer surplus from their introduction, as first demonstrated by Hicks (1940), and then elaborated and applied by Hausman (1997a, 1999) to the case of cellular telephones. We chose to deal with this by being deliberately cautious, but indicating that there was an asymmetrical bias with more potential bias on the upside than the downside. We believed

that our overall estimates were conservative, both by ignoring the many intangible aspects of quality change, such as the improved safety of home power tools or the improved quality of stereo sound and TV pictures, and by omitting any explicit valuation of truly new products.

Problems of Implementation

The BLS response to some of the Commission's recommendations has been that they are very difficult to implement in real time in a monthly CPI program using mechanical rules that are straightforward to implement. These difficulties explain why we proposed that the BLS calculate a second index, published annually, that would constantly be updated and continuously revised, and which need never be final. This second index, if it had been implemented, would have addressed many — not all — of the BLS concerns about practicality.¹²

Changes in the CPI Since the Report

Despite its initial set of critical comments, the BLS moved with surprising speed to implement some of the Commission's recommendations. At the same time the BLS introduced other changes which were planned previously to the release of the Commission report and for which the Commission can take no credit. Subsequently, we will discuss the new experimental BLS index based on chain weights at the upper level; this is not included in the following list of CPI improvements because it has not been incorporated into the basic CPI-U. Here we list the most important changes that the CPI has introduced since the December, 1996, release of the Commission report.¹³ Additional perspective on

11 Some of the material in this section comes from my published discussion of Moulton and Moses (1997).

12 The Stewart and Reed (1999) paper introduces a research-based and retrospective CPI as a one-time project, not as an ongoing annual activity of the BLS.

13 The list of changes comes from background material supplied to Commission members with a 1999 General Accounting Office (GAO) survey (U.S. General Accounting Office, 2000) regarding post-Boskin CPI measurement changes. This list has been checked against the list in Greenlees (2006).

these changes is provided by Greenlees (2006) in this symposium.

1. Lower-level Geometric Weighting.

Effective with data for January, 1999, the BLS introduced the geometric mean estimator for index categories that comprise approximately 61 per cent of total consumer spending in the CPI. This was expected to reduce the rate of increase of the CPI by about 0.2 percentage points (Dalton et al., 1998), and this estimate was later confirmed by Greenlees (2006).

2. More Rapid Change in Upper-Level Weights.

In another major change that appears to be a response to the Commission's recommendations, the BLS now changes upper-level weights much more rapidly than in the past (U.S. BLS, 1999). Eleven years elapsed between the initial use of 1982-84 weights in 1987 and the switch to 1993-95 weights in 1998. More recently the lag has dropped to a mere three years: 1999-2000 weights were introduced in January, 2002 and apply to CPI calculations for the years 2002 and 2003, implying an average three-year lag between the middle of the 1999-2000 period used to calculate the weights and the 2002-2003 period over which the CPI is calculated using those weights. With the same three-year lag structure, new weights were introduced in January, 2004, January, 2006, and so on in the future.

3. Change from Area- to Item-Based Sample Rotation Procedures.

In a change planned before release of the Commission report, the CPI switched its sample rotation methodology in 1998 (Cage, 1996). It shifted its point-of-purchase survey from time-consuming in-person visits to computer-assisted telephone surveys, which allow for an increase in sample size and focus on specific item categories where products turn over rapidly and where new products are frequently introduced.

4. Changes in the Methods for Pricing of Hospital Services.

In another change planned before the release of the Commission's report and implemented in January, 1997, the BLS has improved its procedures for pricing hospital services (Cardenas, 1996). Instead of the old approach, which was a straightforward input cost index that did not reflect shifts in the use of inputs (e.g. shorter hospital stays for a given procedure or a shift from inpatient to outpatient treatment), the new methodology obtains prices for a sample of specified treatments for particular diseases, rather than for a day in the hospital. This approach had been introduced into the PPI in 1992.¹⁴ Even though the PPI made no explicit allowance for improvements in medical technology of the types incorporated into the research reviewed by the Boskin Commission, during the 1992-96 period the PPI for hospital services increased at an annual rate 2.0 to 2.5 percentage points less than the equivalent CPI index, not far from the Commission's estimate of a 3.0 percentage point upward bias in the CPI for medical care services. Allowing for the value of technological advances might imply a CPI bias above 3.0 percentage points for the pre-1997 period.

5. Treating Mandated Pollution Control Measures as Price Increases.

As discussed in Fixler (1998), the BLS switched in January, 1999 to treat changes in vehicle or motor fuel characteristics arising from air pollution mandates as a change in price rather than quality. This reverses a BLS policy in effect since 1971 and follows from a recommendation in the Commission report. Since most of the changes in automobile technology introduced to reduce air pollution occurred during the 1970s and 1980s, this change will have little future effect on the CPI. This change provides another example of

¹⁴ This shift in methodology is discussed in Triplett (1999:3-4).

the need for a second, annual research-based index that can be revised into the past, and in this case the revisions would raise the growth rate of the CPI.¹⁵

6. Hedonic Price Indexes for Electronic Products. Again following in the footsteps of the PPI, which had used the hedonic regression technique to adjust personal computer prices as long ago as 1991, the CPI adopted the hedonic regression approach for personal computers in 1998. Similarly, the CPI introduced a hedonic regression approach to adjusting television set prices, beginning in January, 1999. This technique was introduced following the research of Moulton, LaFleur, and Moses (1998). Greenlees (2006) shows that so far this additional research on quality change has had a very small impact on the overall CPI, because the weights of the products involved are small, and that small sample sizes often preclude developing hedonic indexes from the regular CPI sample.

The Big Surprise: Upper-level Substitution Bias is More Important than We Thought

The above list of changes refers to the basic, most-often quoted version of the CPI known as the CPI-U. The list does not include the introduction of chain weights at the upper level, because the BLS decided not to take this step in the CPI-U. Instead, in a history described in detail in Cage et al. (2003) and in Greenlees (2006), the CPI decided to use chain weights not

in the basic CPI-U but in a new index called the C-CPI-U that is intended as an official supplemental index rather than an experimental research index.

The remarkable surprise after six years of experience with the C-CPI-U, which the BLS currently publishes from January 2000 to the most recent month, is that the bias between upper-level chain weights and Laspeyres weights is much larger than anyone would have guessed, including BLS staff and the Boskin Commission. Despite the more frequent updating of weights in the Laspeyres CPI-U now than in 1996, the difference between the C-CPI-U and the CPI-U is very large, 0.38 percentage points per year over the six years between January, 2000, and January, 2006.¹⁶ In an amazing coincidence, the difference for essentially the same period, 1999:Q4 to 2005:Q4, between the chain-weighted PCE deflator and the CPI-U is exactly the same, 0.38 percentage points.¹⁷

These facts lead to a reassessment of the quantitative importance of the Boskin evaluation and the changes in the CPI over the past decade. The Boskin Commission found substitution bias at the upper level of 0.15 percentage points and at the lower level of 0.25 points. One would have thought that more rapid updating of weights at the upper level and the movement to geometric weights at the lower level would have eliminated all but perhaps 0.1 points. But we now are faced with six years of evidence of upper-level bias of 0.38 points, suggesting that the Boskin Commission, relying primarily on previous BLS

15 Gordon (1990:351) Table 8.10 shows the time series of BLS quality adjustments for new automobiles divided among safety, environmental mandates, and other factors. The environmental adjustments had the effect of reducing the rate of inflation of new auto prices by 1.22 percentage points per year over the period 1967-85. The Boskin Commission report found no net bias in the CPI for autos, taking a 0.94 percentage points per year downward bias due to the treatment of environmental adjustments as a quality change, and cancelling that out by a 0.95 percentage points upward bias due to the CPI's neglect of the increased durability of autos. The Commission's treatment of increased auto longevity is validated by White (2006), who contrasts a median auto lifetime of 10.5 years and 107,000 miles in 1977 with 13 years and 152,000 miles in 2001.

16 The annual growth rate of the CPI-U over this six-year period was 2.68 per cent and of the C-CPI-U was 2.31 per cent (the 0.38 percentage point difference quoted in the text allows for rounding error). It should be noted that data after 2004 are preliminary.

17 The annual growth rate of the CPI-U over this period was 2.75 per cent and of the PCE deflator was 2.37 per cent.

research, substantially underestimated the amount of upper-level substitution bias.¹⁸

Implications of Research Since the Boskin Commission Report

The final section of this paper examines the implications of research carried out since the 1996 release of the Commission Report. Our topics are recent research on outlet substitution bias, long-term historical evidence on CPI bias, and further research on quality change and new products.

Outlet Substitution Bias and the Wal-Mart Effect

The Commission had estimated outlet substitution bias to contribute 0.1 percentage point per year to the overall upward bias in the CPI, which ignores changes in the level of prices between full-priced and discount stores, assuming that the price differential is fully offset by a service differential. But shifts in market share contradict that assumption; when shoppers shift from full-price to discount outlets, they are “voting with their feet” that the price differential is worth more to them than any service differential. In fact, much of the shift in market share over the past two decades has been from higher priced self-service stores like Sears and K-Mart to lower-priced and more efficient self-service stores like Wal-Mart and Target. The level of service is often not an issue, as shoppers move from one type of self-service outlet to another.

Important new evidence on the Wal-Mart effect for food at home is provided by Hausman and Leibtag (2005), who study both bar-code data on prices charged by each outlet as well as household panel data that can track household

shopping patterns over time. Their discount sellers of food include supercenters, warehouse clubs, and mass merchants, and they cite sources estimating that these outlets began selling food in the late 1980s and by 2003 had achieved a 25 per cent market share of total food expenditures. They find an average benefit of the introduction of discount outlets to be 25 per cent of food expenditures, consisting of 20.2 per cent for the direct effect of lower prices at the supercenters, and an additional 4.8 per cent coming from the competitive responses of lower prices at traditional outlets. They also find that the average rate of price change at the supercenters and the traditional outlets is the same; that is, the benefit of the supercenters comes from their much lower prices when each supercenter opens for business, a consumer benefit that is linked out by the CPI.

If we assume that the 4.8 per cent price decline at traditional outlets is accurately measured by the CPI, then we can calculate the impact of the supercenters as a 25 per cent market share, times a 20 per cent price differential, or 5 per cent, spread over roughly 15 years between 1988 and 2003, or an outlet substitution bias of 0.33 percentage points per year for food, which has a 12 per cent weight in the CPI. Thus food alone would contribute 0.04 percentage points to the outlet substitution bias estimated by the Boskin Commission to be 0.1 percentage points. Presumably other durable and nondurable goods would contribute the rest. It is doubtful that the total of outlet substitution bias could be appreciably above the Boskin estimate of 0.1 percentage points, because housing, medical care, and numerous other types of consumer services are not sold by discount stores.¹⁹

18 The matter is more complicated than this, because there are numerous differences between the CPI-U and the PCE deflator other than weighting schemes, whereas the difference between the CPI-U and C-CPI-U results only from upper-level weighting differences. Thus the identical differences cited above over the past six years may be in part a coincidence.

19 We note that discount chains now sell haircuts, and that big box retailers like Pet Smart now offer both dog grooming and veterinarian services.

The Hulten-Brueghel Paradox and its Implications for Quality Change Bias

The Boskin Commission's CPI bias estimate of 1.1 percentage points per year was explicitly applied to the period 1995-96, and the Commission suggested that the bias was 0.25 points higher prior to 1995 and extending back to 1978, due to so-called "formula" bias in the arithmetic mean formula used to combine individual item prices.²⁰ What do we know about the bias prior to 1978? Conjectures by Nordhaus (1997) and responses by his discussant Hulten (1997) led to the recognition that an upward bias in the CPI in the order of magnitude of 1.5 percentage points a year cannot be extended back for a century or two without the implication that the standard of living in 1800 (Hulten, 1997) or in 1569 (Gordon, 2005) was implausibly low. By Gordon's calculations, which extended Hulten's analysis back several more centuries, an annual price index bias of 1.5 percentage points implies for 1569 a median *annual income in today's prices* of only \$5.60, enough to buy 0.8 ounces of potatoes per day, with nothing left over for food or shelter. Gordon had chosen the year 1569 as the year of death of Pieter Brueghel the Elder, who had painted happy burghers "often shown as overfed, content, well-clothed, and with solid-looking houses in the background" (Gordon, 2005:4).²¹

To resolve the paradox Gordon suggested that at some point in the past the CPI bias must have been zero or even negative, and to address this possibility he carried out research on two of the three major necessities, clothing and shelter. For apparel his major finding is based on applying

two methodologies to the same data for women's dresses from the Sears Roebuck catalog over the period 1914-88. Hedonic regressions are compared with a matched-model index that duplicates the CPI pre-1988 methodology by comparing dresses from one year to the next that are absolutely identical in every quality attribute. It had long been suggested that the matched-model methodology would miss price increases that occur with model changes, and indeed this is what occurred in the Sears data. The annual growth rate of the hedonic index was fully 2.9 percentage points faster than the matched model index based on the same data and 1.3 percentage points faster than the CPI for women's dresses.²² Gordon's results include a close comparison of quality in the 1914 and 1988 dresses, something that is not possible with the CPI, with the conclusion that quality in the Sears sample deteriorated over the full period. He suggests that the CPI for apparel is roughly accurate for 1914-47 but is downward biased by roughly 1.5-2.0 percentage points per year for 1947-88. He also suggests that methodological improvements in the CPI after 1988 may have largely or completely eliminated the downward bias, and in any case he has no evidence after that date based on hedonic regressions.

Rental housing is the most important single component of the CPI, because the price of rental housing is used as a proxy for owner-occupied housing. Gordon and vanGoethem (2005) examine a wide variety of evidence, including a large biennial set of panel data on rental apartments from the American Housing Survey, and conclude that the CPI for rental housing is downward biased for most of the period from

20 The formula bias was partly eliminated by the January, 1995 introduction of a procedure called "seasoning" that is explained in Greenlees (2006).

21 Brueghel dropped the "h" from his name in the last ten years of his life, but his sons retained the "h".

22 These results are taken from Gordon (2005:Table 13). The fact that the Sears matched-model index grew 1.62 percentage points slower than the CPI could reflect differing sample sizes that causes the Sears index to miss more of the price changes, and also a kind of outlet substitution bias reflecting Sears' relatively low prices compared to other merchants in the first two-thirds of the sample period.

1914 to 2003 at an average rate of roughly one percentage point per year. They find that the period of most rapid downward bias was in the first half of the postwar era, the same time interval that yielded the most rapid downward bias in the Gordon apparel results, and that methodological improvements in the CPI had reduced the downward bias to roughly one-third point per year from 1995 to 2003.

These results on apparel and rental shelter require a retrospective downward revision in the Boskin Commission estimates of quality change bias. The Boskin bias estimate for apparel was +1.0 percentage points per year, and this should be reduced to zero.²³ The Boskin bias estimate for rental housing was +0.25 percentage points per year, and this should be reduced to the Gordon-vanGoethem regression-based estimate of -0.46 percentage points per year.²⁴ Taking these together and using the weights in the Boskin report reduces the Boskin estimate of quality change bias from 0.612 to 0.429 points per year for the 1995-96 period.

Post-Boskin Research on Quality Change: Medical Care and Pharmaceuticals

In looking back to the Commission's estimate of a 3.0 percentage points per year upward bias in the CPI for medical care services, fully 2.0 percentage points are accounted for by the change from input costs to a treatment basis, introduced into the PPI in 1992 and the CPI in 1997. This leaves only 1.0 percentage point for all remaining improvements in medical care

technology, and if anything, recent studies suggest that medical care technology improvements may be reducing the true price of medical care by more than 1.0 percentage point per year.²⁵

In a newer version of research that the Commission had reviewed, Cutler, McClellan, and Newhouse (1999) use information from hospital records to price heart attack treatments using the CPI's traditional input-cost approach and find an upward bias of 2.0 percentage points. Adding in the value of improved life expectancy from better treatment procedures, the bias rises to the range of 3.1-3.5 percentage points, and further research raises the bias estimate to 5.0 percentage points per year. A study that implies a much larger bias was carried out by Frank, Berndt, and Busch (1999) on treatments for mental depression. There is no component of the CPI that is directly comparable, but this research implies a possible bias of more than 10 percentage points per year.²⁶ Ellison and Hellestein (1999) analyze a large data set on the prices of the cephalosporin class of antibiotics and find a price increase for 1988-96 of 0.76 per cent per year compared to the PPI cephalosporin component which rose at 4.54 per cent a year, for a bias estimate of 3.78 percentage points per year.²⁷ Newhouse (2001, Table 5) surveys available evidence at that time and considers a 3 percentage points upward CPI bias to be conservative.

The number of aspects of medical care subject to research has been growing, and it seems reasonable to extrapolate from areas where studies have occurred to those which may have similar characteristics. Shapiro, Shapiro, and Wilcox

23 The Boskin estimate had been based on an earlier version of the Gordon paper that compared the CPI with a matched-model index for all apparel (not just womens dresses) based on Sears catalogue data. The hedonic study of womens' dresses had not yet been performed at the time of the Commission deliberations.

24 This is the average of their estimate for 1985-95 of -0.58 and for 1995-2003 of -0.33.

25 Some of this research appears in Triplett (1999), a book with a comprehensive and thoughtful introduction that provides the best available introduction into the issues, problems, techniques, and results in this area.

26 A follow-up study by Berndt et al. (2000:15) supports the earlier paper and "does not materially change findings from previous research on treatment of depression."

27 Related research available to the Commission was Griliches-Cockburn (1994) and Berndt et al. (1996).

(1999) point to the numerous dimensions of cataract surgery. An operation that once required a week in the hospital is now a brief outpatient procedure. Subsequent recovery is much faster; complication rates have declined; intraocular lenses have replaced cumbersome cataract spectacles. As a result of improved technology and a drastic decline in prices, the rate of cataract surgery among individuals in the United States aged 65 years has increased by almost a factor of four. Because the operation now occurs earlier in the disease, the period of steadily obscured vision has been eliminated, with enormous benefits in welfare. The authors point to similar benefits in angioplasty, joint replacement, and laparoscopic removal of gall bladders.

Post-Boskin Research on Quality

Change: Other Products

Scattered pieces of new research have emerged for other products. Moulton et al. (1998) concluded that there was an upward bias in the CPI for television sets of three to five percentage points over the period 1993-97. Ohashi (1999) developed hedonic price indexes for VCRs during the first decade of their introduction (1978-87) prior to their introduction into the CPI and found an average rate of price decline of 12 per cent per year over that period. Hausman (1999) criticizes the Boskin Commission for understating the CPI bias related to cellular phones and estimates that the CPI for telephone service (taking account of the gradually increasing weight on cell phones) is biased upward for 1985-97 by between 0.8 and 1.9 percentage points per year.

Abel, Berndt, and White (2003) find that the prices of Microsoft software declined over 1993-2001 at an annual rate of 4.26 per cent, even without taking into account any improvements in quality. A much more rapid rate of decline of 15 to 18 per cent is recorded by White et al. (2004) for personal computer operating systems

and of 13 to 16 per cent for software productivity suites. An even faster annual rate of decline of 21 to 26 per cent for PDAs (Personal Digital Assistants) is found by Chwelos et al. (2004).

Other than Hausman's studies of Cheerios, cellular phones, and Wal-Mart, there has been relatively little new research on the impact of new products and new outlets, and yet speculative estimates of the value of these improvements were at the core of the Boskin Commission estimates of CPI bias for quality change and new products. The traditional supermarket industry is in a state of upheaval as consumers shift market share both toward the cheaper supercenters studied by Hausman and Leibtag (2005) and at the same time toward the greater variety offered by such upscale markets as Whole Foods. Both low-priced supercenters and high-priced markets that emphasize organic and in-house prepared foods are viewed as attractive new products, as shown by the revealed preference of consumer behavior.

Conclusion

In evaluating the Boskin Commission report ten years later, two issues must be sharply distinguished. First, did the Commission overstate the CPI bias for the period to which it referred, 1995-96? Second, how much have improvements in the CPI reduced that bias?

On the first question, recent research on apparel and housing suggests that the Commission's estimate of quality change and new product bias may have been roughly 0.2 percentage points per year too high. However, the striking 0.38 percentage point annual difference over 2000-2006 between the CPI-U and the chain-weighted C-CPI-U suggests that the Commission's estimate of 0.15 points greatly understated the significance of upper-level substitution bias. This is especially true, given the much more rapid updating of upper-level weights in the CPI for the 2000-06 period from which the 0.38 percentage

points number is calculated. It is possible that, with this new information, the Commission's estimate of upper-level substitution bias for the 1995-96 interval should have been 0.45 to 0.50 points, instead of 0.15 points, more than offsetting the Boskin overstatement of quality change and new product bias. Thus my own retrospective view is that the upward bias in the CPI in 1995-96 was if anything higher than the Boskin estimate of 1.1 percentage points and was perhaps 1.2 or 1.3 percentage points.

An important piece of evidence on the nature of CPI bias comes from the longstanding excess of the annual growth rate of the CPI over the PCE deflator. Any such excess is notable, because the PCE deflator uses the same underlying micro price indexes as the CPI but weights them differently. As discussed above, the CPI-U grew 0.38 percentage points faster than the PCE deflator during 1999:Q4 to 2005:Q4, a period when most of the improvements in the CPI were already in effect. Over the period more relevant to the Boskin Commission bias estimates, 1992-98, the CPI-U grew 0.63 percentage points faster than the PCE deflator. This 1992-98 difference, which can be explained only by item and category substitution effects (because the PCE has the same exact treatment of outlet substitution, quality change, and new product effects), provides *prima facie* evidence that the Boskin bias estimate may have been understated.

What about the second question, reductions in the bias due to CPI improvements since 1996? The greatest amount of progress has been made in reducing substitution bias, both at the upper level and lower level, although the thorny issue of outlet substitution bias remains untouched. However, in light of the continuing large difference in growth rates of the C-CPI-U and CPI-U, these improvements have reduced substitution bias from a higher base than the Commission recognized. Several specific improvements address part of the Commission's estimated

upward bias involving quality change and new products, including new hedonic indexes for television sets and personal computers as well as an improved treatment-based methodology for measuring medical care prices.

What is my own estimate of the current CPI bias? The new C-CPI-U evidence suggests that category and item substitution bias appears to remain at about 0.4 percentage points per year, outlet substitution bias remains at about 0.1 percentage points per year, and bias attributable to quality change and new products has been reduced from the Boskin-era 0.4 percentage points (revised downward from 0.6 percentage points as explained above) to perhaps 0.3 percentage points primarily as a result of the switch from input prices to treatment prices for medical care. This sums to 0.8 percentage points per year.

Concluding as did the Boskin Commission on issues of unmeasured improvements and deterioration, I think that these point estimates substantially understate the value of inventions, new products, and increased longevity. A century ago, our forefathers had to shovel coal, carry water into their dwellings, and heat it manually before bathing or the tedious scrubbing of clothes could take place. The value of running water, water heaters, forced air heating fueled by natural gas and the cleaner air that has resulted, is enormous, even if converted to an annual growth rate over 100 years. Recent research on the value of increased life expectancy creates growth rates in welfare that swamp the Boskin debates about 10 basis points here or there. In particular, Nordhaus (2002) and related research concludes that the value of increased longevity over the past century is as large as the value of measured growth in all non-health goods and services.

Let me conclude with a small personal example. Despite global warming, it still snows occasionally in Chicago, but I don't have to touch a snow shovel. My trusty Toro snow blower is a

new product from the perspective of 50 years ago. The value of its invention is not included in the CPI. But, not only does it greatly ease the job of removing the snow compared to the old-fashioned snow shovel, but it indirectly has contributed to my life expectancy.

References

- Abel, Jason R., Ernst R. Berndt and Alan G. White (2003) "Price Indexes for Microsoft's Personal Computer Software Products," NBER Working Paper 9968, September.
- Abraham, Katharine G. (1997) "The CPI Commission: Discussion," *American Economic Review*, 87 (2), May, pp. 94-98.
- Abraham, K., Greenlees, John S., and Moulton, Brent R. (1998) "Working to Improve the Consumer Price Index," *Journal of Economic Perspectives*, 12 (1), Winter, pp. 27-36.
- Aizcorbe, Ana M. and Patrick C. Jackman (1993) "The Commodity Substitution Effect in CPI Data, 1982-1991," *Monthly Labor Review* 116 (12), December, pp. 25-33.
- Berndt, Ernst R., Anupa Bir, Susan H. Busch, Richard G. Frank and Sharon-Lise T. Normand (2000) "The Medical Treatment of Depression, 1991-1996: Productive Inefficiency, Expected Outcome Variations, and Price Indexes," NBER Working Paper 7816, July.
- Berndt, Ernst R., I. Cockburn, and Z. Griliches (1996) "Pharmaceutical Innovations and Market Dynamics: Tracking Effects on Price Indexes for Anti-Depressant Drugs," *Brookings Papers on Economic Activity*, Microeconomics Issue, pp. 133-188.
- Boskin, Michael J., E. Dulberger, R. Gordon, Z. Griliches, and D. Jorgenson (1996) "Toward a More Accurate Measure of the Cost of Living," Final Report to the Senate Finance Committee, December 4.
- Boskin, Michael J., E. Dulberger, R. Gordon, Z. Griliches, and D. Jorgenson (1998) "Consumer Prices, the Consumer Price Index, and the Cost of Living," *Journal of Economic Perspectives*, 12 (1), Winter, pp. 3-26.
- Cage, Robert (1996) "New Methodology for Selecting CPI Outlet Samples," *Monthly Labor Review* 119 (12), December, pp. 49-61.
- Cage, Robert, Greenlees, John S., and Jackman, Patrick (2003) "Introducing the Chained Consumer Price Indices," in Thierry Lacroix (ed.), *International Working Group on Price Indices (Ottawa Group): Proceedings of the Seventh Meeting* (Paris: INSEE) pp. 213-246.
- Cardenas, Elaine M. (1996) "Revision of the CPI Hospital Services Component," *Monthly Labor Review* 119 (12), December, pp. 40-48.
- Chwelos, Paul D., Berndt, Ernst R., and Cockburn, Iain M. (2004) "Faster, Smaller Cheaper: An Hedonic Price Analysis of PDAs," NBER Working Paper 10746, September.
- Cutler, David, McClellan, Mark, and Newhouse, Joseph P. (1999) "The Costs and Benefits of Intensive Treatment for Cardiovascular Disease," in Triplett (ed.), 1999, pp. 34-71.
- Dalton, Kenneth V., Greenlees, John S., and Stewart, Kenneth J. (1998) "Incorporating a Geometric Mean Formula into the CPI," *Monthly Labor Review* 121 (10), October, pp. 3-7.
- Diewert, W. Erwin (1998) "Index Number Issues in the Consumer Price Index," *Journal of Economic Perspectives*, 12 (1), Winter, pp. 47-58.
- Ellison, Sara Fisher, and Hellerstein, Judith K. (1999) "The Economics of Antibiotics: An Exploratory Study," in Triplett (ed.), 1999, pp. 118-143.
- Fixler, Dennis (1998) "Treatment of Mandated Pollution Control Measures in the CPI," *CPI Detailed Report*, September, pp. 4-7.
- Frank, Richard G., Berndt, Ernst R., and Busch, Susan H. (1999) "Price Indexes for the Treatment of Depression," in Triplett (ed.), pp. 72-117.
- Gordon, Robert J. (1990) *The Measurement of Durable Goods Prices* (Chicago: University of Chicago Press for NBER).
- Gordon, Robert J. (2000) "The Boskin Commission Report and its Aftermath," in Mick Silver and David Fenwick (eds.), *Proceedings of the Measurement of Inflation Conference* (Cardiff UK: Cardiff University), pp. 258-82. Also available as NBER Working Paper 7759, June.
- Gordon, Robert J. (2005) "Apparel Prices 1914-93 and the Hulten/Brueghel Paradox," in C. Corrado, E. Diewert, and C. Hulten (eds.), *Price Index Concepts and Measurement*, Conference on Research in Income and Wealth (Chicago: University of Chicago Press for NBER, forthcoming) Also NBER Working Paper 11548, August.
- Gordon, Robert J. and vanGoethem, Todd (2005) "A Century of Downward Bias in the Biggest CPI Component: The Case of Rental Shelter, 1914-2003," in E. Berndt and C. Hulten (eds.), *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, Conference on Research in Income and Wealth (Chicago: University of Chicago Press for NBER), forthcoming, also NBER Working Paper 11776, November.
- Greenlees, John S. (1997) "Expenditure Weight Updates and Measured Inflation," unpublished.

- Greenlees, John S. (2006) "The BLS Response to the Boskin Commission Report," *International Productivity Monitor*, Number 12, Spring, pp. 23-41.
- Griliches, Z. and I. Cockburn (1994) "Generics and New Goods in Pharmaceutical Price Indexes," *American Economic Review*, December, pp.1213-1232.
- Hausman, Jerry (1997a) "Valuation of New Goods Under Perfect and Imperfect Competition," in Bresnahan, T, and Gordon, Robert J. (eds.), *The Economics of New Goods* (Chicago: University of Chicago Press for NBER) pp. 209-37.
- Hausman, Jerry (1997b) "The CPI Commission: Discussion," *American Economic Review*, 87 (2), May, pp. 94-98.
- Hausman, Jerry (1999) "Cellular Telephone, New Products and the CPI," *Journal of Business and Economic Statistics*.
- Hausman, Jerry and Leibtag, Ephraim (2005) "Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Walmart," NBER Working Paper 11809, December.
- Hicks, John R. (1940) "The Valuation of the Social Income," *Economica*, 7 (26), May, pp. 105-24.
- Hulten, Charles R. (1997) "Discussion of Nordhaus," in Bresnahan T., and R.J. Gordon (eds.), *The Economics of New Goods* (Chicago: University of Chicago Press for NBER) pp. 66-70.
- Moulton, Brent R. (1996) "Bias in the Consumer Price Index: What is the Evidence?" *Journal of Economic Perspectives*, 10 (4), Fall, pp. 159-177.
- Moulton, Brent R. and Moses, Karin E. (1997) "Addressing the Quality Change Issue in the Consumer Price Index," *Brookings Papers on Economic Activity*, 28 (1), pp. 305-49.
- Moulton, Brent R., LaFleur, Timothy J., and Moses, Karin E. (1998) "Research on Improved Quality Adjustment in the CPI: The Case of Televisions," presented to the Conference of the Ottawa Group, April.
- Newhouse, Joseph P. (2001) "Medical Care Price Indices: Problems and Opportunities, the Chung-Hua Lectures," NBER Working Paper 8168, March.
- Nordhaus, William D. (1997) "Do Real-Output and Real-Wage Measures Capture Reality? The History of Light Suggests Not," in Bresnahan T., and R.J. Gordon (eds.), *The Economics of New Goods* (Chicago: University of Chicago Press for NBER) pp. 29-66.
- Nordhaus, William D. (1998) "Quality Change in Price Indexes," *Journal of Economic Perspectives*, 12 (1), Winter, pp. 59-68.
- Nordhaus, William D. (2002) "The Health of Nations: The Contribution of Improved Health to Living Standards," NBER Working Paper 8818, March.
- Ohashi, Hiroshi (1999) "Quality-Adjusted Price Indexes for Home Video Cassette Recorders in the US, 1978-87," presented at NBER Summer Institute, July 21.
- Shapiro, Matthew D., and David W. Wilcox (1997) "Alternative Strategies for Aggregating Prices in the CPI," NBER working paper 5980, March.
- Shapiro, Irving, Matthew D. Shapiro, and David W. Wilcox (1999) "Quality Improvement in Health Care: A Framework for Price and Output Measurement," NBER working paper 6971, February.
- Slifman, L., and C. Corrado (1996) "Decomposition of Productivity and Limit Costs," *Occasional Staff Studies*, Federal Reserve Board, November 18.
- Stewart, Kenneth J. and S. B. Reed (1999) "CPI Research Series Using Current Methods, 1978-98," *Monthly Labor Review* 122 (6), June, pp. 29-38.
- Stigler, George, ed. (1961) *The Price Statistics of the Federal Government. Report to the Office of Statistical Standards*, Bureau of the Budget. New York, National Bureau of Economic Research, 1961.
- Triplett, Jack (ed.) (1999) *Measuring the Prices of Medical Treatments*, Washington, Brookings.
- Triplett, Jack (2006) "The Boskin Commission Report After a Decade," *International Productivity Monitor*, Number 12, Spring, pp. 42-60.
- U.S. Bureau of Labor Statistics (1997) "Measurement Issues in the Consumer Price Index," *June 1997, response to the U.S. Congress*, Joint Economic Committee.
- U. S. Bureau of Labor Statistics (1999) "Scheduled Updates for Expenditure Weights in the Consumer Price Index," *CPI Detailed Report*, February, pp. 5-6.
- U. S. General Accounting Office (2000), *Consumer Price Index: Update of Boskin Commission's Estimate of Bias*. Washington, draft report, January.
- White, Alan G., Abel, Jason R., Berndt, Ernst R., and Monroe, Cory W. (2004) "Hedonic Price Indexes for Personal Computer Operating Systems and Productivity Suites," NBER Working Paper 10427, April.
- White, Joseph B. (2006) "Cars Last Longer, Driving a Shift in All Aspects of the Auto Sector," *Wall Street Journal*, February 28, pp. D3.

The BLS Response to the Boskin Commission Report

John S. Greenlees¹
U.S. Bureau of Labor Statistics

I HAVE BEEN ASKED TO PROVIDE the Bureau of Labor Statistics (BLS) perspective on the Boskin Commission Report, and a description of its influence on the Consumer Price Index (CPI) program. I am a natural choice for this assignment because I was the chief of the CPI program during the period of the Commission's deliberations. In fact, I became Assistant Commissioner for Consumer Prices and Price Indexes on July 10, 1995, only two weeks after the Commission came into existence.

In this paper my focus will be on the arguments presented and the decisions made by the BLS during the Boskin Commission period. Although the paper makes many references to BLS documents and publications, it contains no review of the economic literature that formed the basis for the Commission's criticisms and recommendations. Thus, the large volume of theoretical and quantitative work on the CPI by Erwin Diewert, Jack Triplett, the individual Commission members themselves, and many others will go mostly unrecognized here. Moreover, I will make no attempt to characterize the present state of knowledge with respect to the issues raised by the Commission, nor will I present any official BLS position on the existence of upward bias in the CPI.

The context of the CPI in 1995

It is difficult to exaggerate the impact of the Boskin Commission on CPI day-to-day activities during and immediately following the Commission's deliberations. Although only a handful of formal meetings were held between the BLS and the Commission, there was extensive communication through email and telephone correspondence, and CPI staff responded to numerous information requests, especially from Zvi Griliches and Robert Gordon. The indirect effects on BLS activity were probably even greater than the direct interaction. Commissioner Katharine Abraham testified before several Congressional committees, and BLS conducted briefings for the press and public, for Congressional staff, and for Executive Branch officials. Meanwhile, numerous papers and reports were issued, either reporting the results of BLS research on "Boskin" issues² or presenting BLS views and action plans.³ Completely aside from any programmatic changes that resulted from this period, the Boskin Commission years called forth a great deal of analysis and clarification of the Bureau's goals and methods in producing the CPI.

-
- 1 The author is Associate Commissioner for Prices and Living Conditions, U.S. Bureau of Labor Statistics. This paper has benefited from comments by Dennis Fixler, Patrick Jackman, Rob McClelland, Brent Moulton, Philip Rones, Ken Stewart and Jack Triplett. All responsibility for errors remains with the author. Email: Greenlees.John@bls.gov
 - 2 See, for example, Moulton (1996a, 1996b), Aizcorbe et al. (1996), Reinsdorf and Moulton (1997), McClelland (1996, 1998), Moulton and Moses (1997), Greenlees (1998).
 - 3 For example, Abraham (1995, 1996, 1997b), Abraham et al. (1998), Bureau of Labor Statistics (1997), Greenlees (1996, 1997).

It is important to note, however, that from the BLS perspective at least, the period of heightened outside concern about CPI bias began several years before the formation of the Boskin panel. One major trigger was the paper presented in 1990 by BLS economist Marshall Reinsdorf⁴ in which he attributed a significant upward bias to the CPI from its failure to account for the lower prices offered in discount outlets.

In the wake of the Reinsdorf paper and the considerable interest it generated, the BLS devoted the December 1993 issue of the *Monthly Labor Review* to four articles on consumer price measurement issues written by economists in the Office of Prices and Living Conditions. Notable among these was the paper by Ana Aizcorbe and Patrick Jackman, who compared indexes based on the CPI's Laspeyres formula to indexes employing the "superlative" Fisher Ideal and Tornqvist formulas for the period 1982 to 1991. That paper was the first to construct superlative indexes using detailed CPI series defined by geographic area and item category, and it provided several estimates of the upward substitution bias resulting from use of the fixed-weight Laspeyres formula to aggregate those series. These Aizcorbe-Jackman results became the basis for many subsequent estimates, including those by the Boskin Commission, of what came to be called "upper level" substitution bias.⁵

Another article in the December 1993 *MLR* was an analysis authored by Brent Moulton of a complex but important problem in the way

the roughly 9,000 elementary area-item indexes in the CPI were calculated. This analysis built on the empirical anomaly, noted earlier by Reinsdorf, that many CPI indexes for food and energy items rose faster than the averages of the prices on which those index series were built. Although Reinsdorf had first attributed this difference to the influence of new discount outlets, in subsequent research he demonstrated that the BLS's operational implementation of the arithmetic mean formula to combine individual item prices led to a systematic upward bias in CPI indexes *relative to the Laspeyres objective*.⁶ This discovery, which I will refer to as "formula bias" in this paper,⁷ and Moulton's empirical demonstration that this bias was especially severe in the volatile food indexes, led the BLS to implement a technical, and costly, procedural change called "seasoning" in the CPI food-at-home components effective in January 1995. It was further discovered that a related bias existed in the CPI shelter (Rent and Owners' Equivalent Rent) indexes, and that bias also was corrected in January 1995. The BLS estimated that these changes together would reduce annual CPI growth by about 0.14 percentage point.⁸

Moulton (1993) and the subsequent paper by Reinsdorf and Moulton (1997) demonstrated that a geometric mean, unlike the arithmetic mean, is not vulnerable to the bias described in the previous paragraph. Moulton therefore suggested that the BLS consider moving to a geometric mean formula in the CPI. Such a change, he argued, would have the additional value of

4 This conference paper was subsequently published as Reinsdorf (1993).

5 The Boskin Commission, having no budget for research, necessarily relied on work by BLS and others for many of its quantitative bias estimates.

6 That research was first presented in 1993 and subsequently published as Reinsdorf (1998). The analysis of this problem by Reinsdorf and Moulton is also reported in their 1997 article.

7 There has been some semantic ambiguity in the term "formula bias." As noted later in this section, the Boskin Commission used the term in a broader sense than I do here.

8 Including other changes in the CPI made at the same time, the estimated effect was a negative 0.12 percentage point. See Armknecht et al. (1995).

reflecting consumer responses to relative price change; as is well known from cost-of-living index theory, use of the geometric mean yields a true cost-of-living index when consumers have Cobb-Douglas preferences characterized by unitary elasticities of substitution among items. The arithmetic mean or Laspeyres form, by contrast, is consistent with Leontief preferences or zero substitution elasticities. Under certain assumptions, movements in a Laspeyres index will tend to overstate movements in a cost-of-living index if consumers do, in fact, change their purchasing patterns in response to changes in relative prices.

Under Paul Armknecht, then the CPI chief, the CPI program did begin building the capability to shift to a geometric mean formula. By mid-1995 a “requirements” document had been completed, detailing the computer system processing changes that would be necessary to implement the shift. Certainly no decision had been made to adopt the geometric mean, but significant resources had been committed to a planning effort.

Public interest in the CPI accelerated greatly after testimony by Federal Reserve Board Chairman Alan Greenspan to the Senate Finance Committee in January 1995. In that testimony he estimated that the CPI measure of consumer price inflation was overstated by about one percentage point annually. This estimate was based on a paper by Board economists David Lebow, John Roberts and David Stockton (1994) that was an update of a 1992 paper by the same authors and the first of several published analyses of upward CPI bias arising from different sources.⁹ Greenspan’s testimony led to Congressional hearings in which subsequent testimony on CPI issues was heard from 13 witnesses.¹⁰ The outgrowth of these hearings was, of course,

the formation of the Advisory Commission to Study the CPI, the “Boskin Commission.”

Another BLS activity in the period immediately prior to the Boskin Commission’s formation was the preparation of a report on CPI issues in response to a request from the House Budget Committee. This report, completed in April 1995, broke little new ground. It did, however, include a discussion of previous research on potential CPI biases. It also indicated that within the next year the Bureau hoped to publish an evaluation of the geometric mean and other formulas for aggregating item prices.

Finally, any discussion of the CPI program and the Boskin Report would be incomplete without noting that 1995 was also the year in which the BLS embarked on a multi-year CPI revision effort. This was the sixth such major revision in the program’s history. Like the previous revisions, it included an updating of the CPI’s underlying expenditure weights and its geographic area and housing unit samples based on the latest Decennial Census data. Also like its predecessors, this revision included other CPI improvement efforts: a thoroughgoing redesign of the Point of Purchase Survey (POPS), which is used to generate the sampling frame for CPI retail outlets; an overhaul of data processing systems for the Consumer Expenditure Survey, which generates CPI weights; a reworking of the item classification structure used in the CPI; a modernization of the housing index estimation system, including an improved method for recovering homeowner shelter costs from rental market data; and the implementation of computer-assisted data collection for CPI prices, replacing paper survey forms with pen-pad computers and electronic survey instruments.¹¹

9 See also Wynne and Sigalla (1994), Shapiro and Wilcox (1996) and the later study by Lebow and Rudd (2003).

10 The hearings were held in March, April and June. See U.S. Senate (1995).

11 See Greenlees and Mason (1996) for a summary of the revision effort.

The BLS Response to the Boskin Report Recommendations

The Boskin Report and the recommendations therein were important drivers of CPI activity during the late 1990s and beyond. Interestingly, the aspect of the report that received the most attention was not its 16 recommendations for action by BLS and others. Instead, the Commission's estimates of CPI upward bias were the most widely cited, and they also were the elements of the report that were most closely associated with subsequent BLS methodological changes. For example, the Boskin recommendations contained little about dealing with product quality change or new goods; but their estimates of CPI bias arising from inadequate handling of those phenomena were widely cited and debated, and intense interest surrounded the subsequent expansion of hedonic quality adjustment for quality change in the CPI.

One recommendation that was widely cited was the Commission's first, overarching recommendation that the BLS should establish the economic concept of a cost-of-living index (COLI) as the measurement objective for the CPI. Viewed from the context of statistical agencies around the world, this recommendation was relatively controversial. As Triplett (2001) has noted, the COLI is rejected as a measurement objective in many countries, including the United Kingdom and Australia. The recommendation was accepted rather readily by the BLS, however. In Congressional testimony in January 1997, Commissioner Abraham stated simply that "this seems basically right to me."¹² For many years, the *BLS Handbook of Methods* had contained language indicating that although for several reasons the CPI could not be considered a COLI,

cost of living theory was used as a guide to operational problems in constructing the index. Moreover, the 1995 BLS report to the House Budget Committee had characterized the CPI as a subindex of the general COLI concept (Greenlees, 2001). Thus, formal acceptance of the Boskin recommendation did not represent a major shift in concept or practice. By the same token, however, it may have made it easier for the BLS subsequently to take steps to make the CPI a closer approximation to a COLI.

The remainder of this section presents the methodological changes made by BLS between 1996 and 2002. It is convenient to divide these changes into three areas, roughly corresponding to the categories of bias identified by the Commission: upper and lower level substitution bias; quality change and new products; and new outlets.

Substitution Bias

Constructing the CPI involves two stages of aggregation. At the "lower level" of aggregation, price changes for individual items are averaged together to produce basic item-area CPI indexes such as the index for apples in Chicago. The item weights used in this lower-level aggregation are primarily a function of the probability-sampling rates used to select individual outlets and items for pricing. In "upper-level" aggregation, the item-area indexes are combined using spending data taken from the Consumer Expenditure Survey to produce the overall U.S. All Items CPI along with subaggregates such as the CPI for Chicago or the CPI for Food and Beverages. The issue of potential substitution bias, and related issues of formula and weighting, arise at both of these two levels of aggregation, although in somewhat different forms due to the different types of information available.

¹² Abraham (1997a). See also Abraham et al. (1998:27).

Lower-level formula

In their September 1995 Interim Report (Advisory Commission to Study the CPI, 1995), the Commission members attributed an upward bias of 0.5 percentage point to “formula bias,” which was their term for problems arising from the arithmetic mean formula then used to compute basic item-area CPI indexes. Their estimate of bias was based on comparisons of simulated CPI series to geometric mean equivalents presented in Moulton (1993), and their criticism — and their definition of formula bias, as opposed to the one used elsewhere in this paper — thus included both aspects of the geometric mean: its immunity to the item-weighting bias discovered and explained by Moulton and Reinsdorf, and its reflection of consumer substitution behavior.

A fundamental principle accepted by the Commission was that the CPI should as closely as possible approximate a cost-of-living index. The closeness of the approximation resulting from any index formula will depend on the accuracy of the embedded assumption about consumer demand elasticities, and no CPI data exist with which to estimate elasticities within item-area cells. The Commission argued that the Cobb-Douglas preference assumption incorporated in the geometric mean formula was much closer to the truth than an assumption of zero substitution elasticities, and that therefore the geometric mean was clearly preferable to the formula used in the CPI. Under the assumption that the BLS was about to introduce a geometric mean formula, the Commission believed that formula bias was a thing of the past.

During the months following the Interim Report, the BLS devoted much additional attention and research to this issue of “lower-level” formula. As viewed by Bureau management, disentangling the two geometric mean effects — formula bias and consumer substitution — was

crucial. It was clear that using the geometric mean would systematically reduce the rate of growth in the CPI. This was clearly justified to the extent that the reduced rate of growth arose from the complete elimination of formula bias, which had thus far only been addressed in the food and shelter indexes. On the other hand, incorporating consumer substitution in the CPI had never been done before; as noted above, the BLS had always been comfortable portraying the index as an upper bound to a COLI due to its use of the Laspeyres formulation. Commissioner Abraham agreed that taking such a step would require careful analysis to ensure that the Cobb-Douglas assumption did, in fact, provide a reasonable characterization of consumer preferences within CPI cells. BLS practice also would require that CPI users and stakeholders be given notice of such a major methodological change.

These considerations would have been moot if it could be established that the formula bias accounted for nearly all of the difference between the CPI and simulated geometric mean indexes. BLS research on CPI microdata, however, demonstrated that this was not the case: arithmetic-mean indexes that were purged of formula bias remained systematically higher than geometric mean indexes.¹³ Therefore, it was decided to expand the “seasoning” approach from the food indexes to the entire CPI in 1996, which, combined with another simultaneous operational change, could be expected to align the CPI with its Laspeyres measurement objective. The BLS estimated that the changes would reduce the CPI’s aggregate annual growth rate by about 0.1 percentage point (Stewart, 1996). Meanwhile, the BLS would continue its research on an experimental geometric mean index.

Recognizing the 1996 changes, the December 1996 Boskin Commission Final Report gave an estimate of 0.25 upward bias from lower-level substitution bias, this arising from the consumer

13 See McClelland (1996), Erickson (1995, 1996), and the discussion in Moulton (1996a).

substitution issue. Adoption of the geometric mean formula was one of the Commission's formal recommendations, and it is one with which the BLS ultimately did agree. In April 1997, the BLS unveiled and began releasing monthly the CPI-U-XG, an experimental index covering the historical period from December 1990 forward.¹⁴ A literature review was conducted by BLS to evaluate the Cobb-Douglas assumption, and several outside experts were retained to provide recommendations on whether the geometric mean should be adopted. In April 1998, based on all these considerations and the performance of the CPI-U-XG, the BLS announced that effective with data for January 1999 it would employ the geometric mean in the overwhelming majority of the 211 CPI item categories, the exceptions being 15 in which it was unreasonable to expect consumers to respond quickly or smoothly to relative price changes.¹⁵ The change was justified explicitly on the basis that it would reflect consumer response to relative price changes. The estimated impact on CPI growth was 0.2 percentage point annually.

Upper-level formula

Two Commission recommendations were concerned with the formula for upper-level aggregation. The Boskin members argued that the preferred formula for a cost of living index was a superlative index formula like the Fisher Ideal formula used in the National Income and Product Accounts chain indexes or the Tornqvist formula used in the BLS Multifactor Productivity series. Based on BLS updates of the Aizcorbe and Jackman (1993) results (Aizcorbe et al., 1996), along with estimates by Matthew Shapiro and David Wilcox (1996), they estimated that the substitution bias resulting from the CPI's Laspeyres formula was about 0.15 percentage point annually.

They recognized, however, that a superlative index cannot be computed in "real time" due to the inherent delays in receipt of the consumer expenditure data required for weighting the monthly price changes. They also recognized that a primary value of the CPI is as a timely measure of consumer inflation. Consequently, their Recommendation 3 proposed that the official, "timely, monthly" CPI be constructed using a geometric mean formula with annually-updated but not contemporaneous expenditure weight information. (The Boskin Report referred to this formula as a "Trailing Tornqvist" index and called it "superlative," notwithstanding the fact that the use of simultaneous weight data is an inherent feature of superlative indexes.) Meanwhile, Recommendation 3 also proposed an annual index that would employ a superlative formula and would also incorporate retroactive revisions to reflect new information on the historical introduction of new goods and improved methodologies.

By the time that the Commission's Final Report was released in December 1996, the BLS had already committed internally to dealing with the upper-level substitution bias issue. Earlier that year, as part of the submission process for the Fiscal Year 1998 budget, the BLS proposed a funding package called the Timeliness and Accuracy, or CPI Improvement (CPII), Initiative. Among the components of that initiative was the development of a superlative CPI.

The standard practice is for budget requests to be closely held until they are formally included in the President's Budget, but the BLS received permission to announce its CPII initiative slightly earlier, in Commissioner Abraham's testimony to the Senate Finance Committee in January 1997. The details of the proposed index were not given until much later, and in fact the BLS took years to ana-

14 See Moulton and Stewart (1999) for a description of the CPI-U-XG and several other experimental BLS indexes.

15 This BLS analysis is summarized in Dalton et al. (1998). Two additional item categories were shifted to the geometric mean in 2004, and the proportion of CPI weight now indexed using a geometric mean is about 61 per cent.

lyze whether the index should be annual or monthly, whether it would be published in a current or only retrospective form, and what superlative formula it would employ. The BLS did make clear that the new index would be inaugurated in 2002; that it was intended to provide a closer approximation to a COLI than the existing CPI-U and CPI-W; and that it would be a complement to those series, rather than a replacement for either.

The logic underlying the BLS superlative plan was that its superlative index, ultimately named the Chained CPI for all Urban Consumers or the C-CPI-U, would be an official CPI product rather than an experimental series. Users, whether in Congress, the Administration, or the private sector, would be given the opportunity to use whichever CPI series best suited their needs. The C-CPI-U would be attractive to users who wanted an index that reflected consumer substitution as much as possible. Conversely, the CPI-U or CPI-W might be more attractive to users who valued that those series are in final form when first published, unlike a superlative index, which must be published either substantially after the fact or subject to revision.

This BLS approach fell short of satisfying Recommendation 3 of the Boskin Report. Most importantly, the upper-level Laspeyres formula in the existing CPI series was left unchanged, so there was no effect on indexed federal programs. Any impact on future tax receipts and benefit payments would require Congressional action to change the basis for indexation from the CPI-U or CPI-W to the C-CPI-U.

Moreover, the BLS superlative plan contained no provision for incorporating the sort of historical revisions that the Boskin Report recommended. The C-CPI-U is subject to two annual revisions when more recent Consumer Expenditure (CE) data become available, but *ex post* infor-

mation such as the results from historical studies of hedonically-adjusted price indexes are not incorporated retrospectively in any of the three official CPI indexes.¹⁶

The details of how the BLS superlative index would be constructed were reported in 2002.¹⁷ The C-CPI-U was the first official superlative CPI produced by a government statistical agency anywhere in the world. Its development required that the BLS resolve numerous issues that had not been addressed in the mainly theoretical literature on superlative indexes. The most difficult operational question arose from the fact that the CE survey can provide expenditure share data for about 8,000 item-area cells on a monthly basis, but those expenditure shares, as well as the associated monthly price indexes, are subject to considerable (and uncorrelated, since they come from different surveys) sampling error. Of course, superlative theory assumes that the observed expenditure shares reflect consumer decisions based on the observed price index movements. If both those data series are affected by sampling error, use of the superlative formula could yield inappropriate inferences. A decision had to be made, therefore, about the extent to which those prices and expenditures should be averaged across either time or geography, or both, so as to retain the benefits of the superlative formula. A discussion of the ultimate BLS decisions is provided in Robert Cage et al. (2003).

Empirically, the movements of the C-CPI-U relative to the Laspeyres CPI-U have been somewhat surprising. BLS simulations using data from the early 1990s suggested that a superlative index would rise by about 0.15 percentage point per year less than an otherwise-comparable Laspeyres index. Later simulations reported in Cage et al. (2003) showed a larger gap in the

16 As discussed below, this Boskin recommendation was addressed to some degree by the development of the experimental CPI-U-RS series.

17 See Cage et al. (2003). These details are available on the BLS website at www.bls.gov/cpi/superlink.htm.

later 1990s, however, with the average annual difference for the 1990-99 period being approximately 0.3 percentage point. As of this writing, official C-CPI-U official data in final form are available for 2000 through 2004. The gap between the C-CPI-U and the CPI-U has continued at about 0.3 percentage point in the last four of those years, but was nearly 0.8 percentage point in 2000, apparently because of some extreme movements in energy prices in that year.¹⁸

Expenditure weight updates

Closely related to the choice of index formula is the choice of the frequency of expenditure weight updates. In fact, the two issues were sometimes confused in popular discussions of upper-level substitution bias.

In terms of its upper-level construction, the CPI in 1995 could be characterized by two features. First, it used a Laspeyres-type formula.¹⁹ Second, its expenditure weights were updated once per decade; the last update had been in January 1987, when CE expenditure data from the three-year period 1982-84 became the basis for weighting basic component CPI indexes. Another update was scheduled for January 1998, when expenditure weights from 1993-95 would be incorporated.

The fact that CPI weights were updated only once a decade was among the most easily and often criticized aspects of the index. The statistical agencies in most other developed countries updated their CPI weights more frequently. At one time the BLS could have justified its policy by the lack of timely expenditure data. However, the Consumer Expenditure Survey, the source of CPI

weights, had been put on a continuous basis during the early 1980s, so that beginning in 1986 the CPI had available expenditure data that would have supported more frequent, even annual, weight updates.

From the perspective of economic theory, all Laspeyres indexes have the same characteristic as an upper bound to a COLI, and there is no theoretical argument to be made for one base period over another. On the other hand, it seems natural to view an index with a more recent base period as being more representative of price movements in what consumers are purchasing currently, and this argues for more frequent updating. The strongest argument *against* frequent updating was the potential for “chain drift”: an upward bias that could result from oscillations in prices and consumer spending. That is, one feature usually considered attractive of a fixed-weight index, relative to a chained index, is that when prices change and then return to their original position, the index itself will return to its original level.

Aside from chain drift, however, it was assumed by most analysts that a more frequently updated index, even one with a Laspeyres formula, would rise less rapidly than a less frequently updated one. This would occur if there were persistent secular trends in relative prices, and consumers tended over time to purchase more of the goods and services with falling relative prices. Somewhat surprisingly, the BLS found little evidence of this in its simulations of how the CPI would have moved differently had it employed more frequent updating.²⁰

Nevertheless, in the FY 1998 CPII budget initiative referred to above, the BLS asked for and

18 Recent CPI-U and C-CPI-U index movements are compared in Johnson et al. (2005).

19 More precisely, it used what is now often termed a Lowe index formula, since the expenditure weight base period (1982-84) was “linked in” at a subsequent date (January 1987). For the same reason the BLS usually employed the term “modified Laspeyres” to describe the CPI formula. On the properties of a Lowe index see International Labour Office (2004).

20 Greenlees (1998) presents an analysis of this question. It should also be noted that it had always been BLS policy to publish “overlap” indexes for six months following an expenditure weight update, comparing the movements of the CPI under the two weighting structures.

received funding to expand the CE sample size by 50 per cent and to accelerate the editing and processing system for CE data. This would have three advantages. First, it would enable CE weights to be used in the CPI after one year instead of two (had the new system been in effect, the 1993-95 base period could have been employed in January 1997 instead of 1998). Second, a two-year base period would be sufficient to provide the same accuracy of weights previously provided by three (so a 1994-95 base period could have been implemented in January 1997). Both these improvements would make CPI weights more current when first used. Third, the BLS argued that the CE enhancements would make it more feasible to update weights in the future on a more frequent basis.

In December 1998, the BLS did announce that starting in January 2002 the expenditure weights for both the CPI-U and CPI-W would be updated biennially using two-year base periods (BLS, 1999). A 1999-2000 base period was introduced in 2002, a 2001-2002 base period in 2004, and so on. More frequent updating was rejected on the basis that biennial updating would make the weights sufficiently representative, and for fear of chain drift or other unforeseen consequences from annual weighting with necessarily overlapping base periods. At the time of the 1998 announcement, based on its simulation evidence, the BLS declined to make any estimate of the effect on CPI growth from moving to biennial updating.

In retrospect, it seems likely that the introduction of more frequent expenditure weight updating is a change that was accelerated by the increased public attention to CPI bias issues. Prior to the Boskin period, the BLS had been reasonably comfortable with decennial updating, on the basis that more frequent updating would be costly and that the available evidence was insufficient to show it would have any major effect. With significant outside criticism, how-

ever, it became difficult to defend the policy and more feasible to obtain funding to change it.

Quality Change and New Products

The Boskin Commission devoted approximately one-third of its report to a component-by-component review of the potential biases attributable to the CPI's allegedly insufficient recognition of the benefits from the introduction of new products and the improved quality of existing products. This part of the report received considerable attention both within and outside BLS, for two major reasons. First, the Commission's aggregate estimate of bias from new products and quality change was 0.60 percentage point per annum, more than half of their total point estimate of 1.1 percentage points from all sources of bias. Second, their 0.60 point estimate was built up from a large number of sometimes rough component estimates based on little firm evidence, making it vulnerable to criticisms and counter-arguments. By contrast, most economists readily accepted the concepts of upper- and lower-level substitution bias, and the Commission estimates of bias in those areas were based largely on BLS research.

To quote Shapiro and Wilcox (1996), "quality change is the house-to-house combat of price measurement." Each product area presents different measurement problems, and there is no single formula or approach that can solve those problems. For the same reason, it is impossible here to go through the various component bias estimates in the Boskin Report. The interested reader can consult Moulton and Moses (1997) or Bureau of Labor Statistics (1997) for a relatively detailed BLS rejoinder to those estimates.

It is probably worth noting that the terms "quality adjustment" and "quality change" themselves led to some semantic confusion in the context of the Boskin bias estimates. The BLS, and indeed all statistical agencies, must adjust each period for quality differences

between items leaving and entering their product samples. In the great majority of cases, however, this is not done by making an explicit comparison of the relative qualities of the two items being compared. Rather, the quality adjustment is made implicitly, through a “law-of-one-price” assumption. Typically, through a procedure called “linking,” the rate of “true” price change between the new item in period t and the old item from period $t-1$ is assumed to be equal to the mean change observed for the sampled items that appeared in both periods; all the rest of the price difference, positive or negative, is assumed to be the value of quality difference. Armknecht and Weyback (1989) had demonstrated using 1984 data that, in aggregate, BLS quality adjustments had a significant impact in reducing CPI price change relative to the estimate that BLS would have obtained had it not adjusted for quality differences at all. Moulton and Moses (1997) updated that analysis, and showed that about half of the growth of the CPI in 1995 was attributable to the treatment of replacement items. Those results were sometimes misinterpreted, though, either by drawing an exaggerated conclusion about the extent to which BLS explicitly adjusted for quality change (most of the quality adjustment was by the implicit linking method), or by using the Moulton-Moses results to derive a measure of aggregate quality improvement in the consumer product sector (many of the most significant quality improvements occur in ways that would not show up in BLS item sample comparisons).²¹

The extent and detail of the Boskin Report’s focus on quality bias was not mirrored in their recommendations. Only Recommendation 8, which advocated a flow-of-services approach for all durable goods and a revised treatment of

insurance, referred to the handling of specific item categories, and those issues were not linked closely to discussions in the body of the report. Somewhat remarkably, the report did not recommend an expanded use of hedonic modeling. Nevertheless, the report may have been a major stimulus to the subsequent incorporation of several new CPI hedonic models into the CPI, beginning with personal computers.

Until 1998, the weight of computers in the CPI was extremely small, not surprisingly given that CPI weights were based on consumer expenditures during the 1982-84 period. Even when 1993-95 weights and a new category for Personal Computers and Peripheral Equipment were introduced in 1998, the relative importance of that component was only 0.234 percent. In October 1997, the BLS announced that in January 1998 it would extend to the CPI the use of the hedonic model for PCs that had been regularly estimated and employed in the Producer Price Index (PPI) since 1990.²²

A second hedonic model, for televisions, was adopted for use beginning in January 1999. That model was estimated using sample CPI data, and had been developed and reported by Moulton et al. (1999). The insufficiency of CPI item samples acted (and continues to act) as a roadblock to broader use of hedonics, however. Often, the CPI sample sizes that are adequate for the usual matched-model CPI methodology are too small to support estimation of a hedonic regression. Moreover, a matched model index may not require a large amount of accurate information on the characteristics of the products being priced, whereas such information is crucial to estimation of hedonic coefficients.

To address this problem, in 1996 the BLS included, as part of its 1998 CPI budget initia-

21 Triplett (2000) provides a precise analysis of this point.

22 That hedonic model for PCs had been developed originally at the Bureau of Economic Analysis. As noted in Johnson et al. (2005), the growth of readily available component cost data via the Internet has led the BLS to discontinue the use of the PC hedonic model in the CPI.

tive, regular funding to collect special samples of prices and characteristics of products for the purpose of hedonic regression estimation. This activity bore fruit in 1999 and 2000, as several models were estimated and employed for a variety of products. Most of these were consumer durables, such as refrigerators, washing machines, and VCRs (Fixler et al., 2000).

The recent paper by David Johnson et al. (2005) presents estimated impacts on overall CPI growth of the use of hedonic models. Echoing a point made earlier by Charles Schultze, their estimates show that the quantitative effect of the models introduced since 1998 has been small.²³ The product categories to which these models have been applied have a small overall weight, and in the aggregate the hedonic quality adjustments have not been dramatically different from the implicit quality adjustments that would have been made using the default CPI methodologies. Further expansion of hedonic methods in the CPI has been hindered by the problems in obtaining satisfactory data sets for regression estimation, and by the inherent difficulty of applying hedonic methodology in the heavily weighted services areas. This continues to be a very important research area at the BLS, despite those obstacles.

Medical care is the other major product area in which the BLS has attempted to find improved methods of decomposing price and quality change. During the tenure of the Boskin Commission, the CPI was taking a set of costly and difficult operational steps to reduce its reliance on list, or “Chargemaster,” rates for hospital services (Cardenas, 1996). This improvement, parallel to a change made in the PPI several years earlier, involved moving from pricing individual compo-

nent services, such as room charges, to pricing patterns of services associated with individual patient hospital stays. This moved the CPI unit of pricing closer to the service for which insurers typically reimbursed hospitals. Combined with other changes in data collection procedures, this made it easier for the BLS to collect third-party reimbursement rates for hospital care rather than having to fall back on the more readily available Chargemaster fees.

Even recognizing this improvement, the Boskin Report estimated an annual upward bias of 3 percentage points for hospital and related services in the CPI, due to the failure to capture the benefits of new and improved technologies and treatment methods. Since the report, the BLS has investigated two different approaches to addressing this well-recognized but difficult problem. The first was the approach of directly pricing health insurance: measuring the cost of hospital services, and other medical care covered by insurance, through the collection (and quality-adjustment) of health insurance premiums. Unfortunately, the BLS decided that the problems of adjusting premiums for utilization and other quality changes were prohibitive.²⁴

A second approach to solving the medical care pricing problem is through the use of insurance claims files, as proposed in National Research Council (2002). In accord with that proposal from a National Academy panel, the BLS contracted with MEDSTAT, Inc., to develop experimental medical care price indexes using claims data and comparing those to simulated series replicating current CPI methodology. The paper by Song et al. (2004) reports on that work, and the BLS continues to be interested in this approach.

23 It should be noted that the BLS had employed hedonic models in some key areas of the CPI prior to 1998, notably in apparel beginning in 1991 and for age-bias adjustments in the shelter indexes beginning in 1987. The weight of these components and the quantitative impact of these earlier hedonic models have been significant, as shown for example in Stewart and Reed (1999).

24 See Bureau of Labor Statistics (2000) for a discussion of this proposal. The approach is discussed in National Research Council (2002:186-188). Health insurance had been priced directly in the CPI prior to 1964, as described by Armknecht and Ginsburg (1992).

New Outlets

The Boskin Commission linked the problem of new goods with that of quality change, but one could also link new goods with new outlets as a type of conceptual issue arising in price index construction. As others have noted, to the extent that a particular good is replaced in a CPI sample by a successor good, perhaps a new model produced by the same firm, this would usually be handled by explicit or implicit methods of quality adjustment. By contrast, when classes of goods are replaced by wholly new products, such as VCRs by DVD players, the usual statistical agency procedures typically do not allow for direct comparisons of quality or estimation of consumer welfare effects. This same situation applies with innovations in outlet type, such as online booksellers or large discount retailers. The CPI does not now reflect the benefits (or losses) to consumers of changes in the distribution of outlet types. The Commission attributed a 0.1 point annual upward bias to the new outlet problem, but I will discuss the BLS reaction in the context of new goods as well, since conceptually the issues are the same.

For many years BLS has employed a complete and sophisticated process of probability sampling for outlets and items. This has set the CPI apart from its counterparts in other countries, which have relied heavily on judgmental sampling. Although probability sampling is a major strength of the CPI, its goal of maintaining a representative sample can sometimes be in conflict with the goal of timely incorporation of new item and outlet categories. In the case of new outlets, inclusion in the CPI sample would require that they be reported as shopping destinations by households in the POPS survey. The data from the POPS, after delivery to BLS by the Census Bureau, which administers the survey, are used to construct the CPI's outlet sampling frame. Once sampled, these outlets

would be visited by CPI data collectors, who would select items for subsequent pricing. At the time of the Boskin Commission, the POPS survey was being extensively improved and reworked from an annual personal-visit survey to a telephone-based quarterly survey called TPOPS (Cage, 1996). It continued to follow a five-year rotation process, however, meaning that in addition to the lags between appearance of an outlet type and the first use of its prices in the CPI, there would be several additional years before that outlet type was fully represented in the CPI sample.

Similar lags in inclusion would apply for those new goods that are associated with unique new outlet types (an example might be pad thai, introduced into wide availability through the expansion of Thai restaurants into more communities). New goods that are sold in traditional outlets could be included in the CPI much more rapidly if they are chosen as replacements when other products disappear from shelves, but otherwise they would tend to be included only through the outlet rotation process.

Finally, the longest lags would arise for a uniquely new product that did not fall clearly into any CPI category. The famous example of this scenario, featured prominently in the Boskin Report and by other writers during that period, was cellular phones. These were not brought into the CPI through outlet rotation because they were not considered in scope for any of the CPI's communications categories; in particular, the category Telephones was part of the Housing major group, and cellular phones were not judged to be part of the cost of Housing. As a consequence, despite the growing consumer use of cellular phones, they were not scheduled for inclusion in the CPI until the 1998 decennial revision, when a new Education and Communication major group would be defined with compo-

ment categories for both land-line and cellular telephone services.²⁵

The interest in new outlet bias in the CPI was based on the previously-noted work by Marshall Reinsdorf of BLS. Reinsdorf (1993) attributed an upward bias to the failure to incorporate the impact of discount stores. Meanwhile, Jerry Hausman (1997, 1999) argued that the CPI accrued an important upward bias by failing to reflect the consumer surplus gains from new cereal products and cellular telephones. These and similar analyses and critiques, repeated in the Boskin Report, led to an increased recognition on the part of the BLS that the CPI should improve its efforts to include new products quickly. At a fundamental level, a decision was made within the CPI program that the item categorization implemented in January 1998 would be viewed — operationally as well as conceptually — as encompassing the universe of consumer goods and services; that is, no new products would be treated as out of scope for CPI collection in the way that cellular telephones had been. The program also accepted the idea that one of its goals would be to achieve a product and outlet sample that was as representative as possible of *current* consumer spending patterns.

In addition, the 1998 CPII initiative included funding for special data collection activities to sample new products. In its early years this money was used for special efforts to include Viagra (a new product with decreasing price and rapidly increasing market share) quickly. New procedures also were instituted to accelerate the resampling of prescription drugs, an area where product innovations were especially prevalent (Lane, 2000).

Somewhat later, using funding from another CPI budget initiative for Fiscal Year 2002, the

BLS made two significant improvements aimed at making its item and outlet samples more current. First, it reduced the TPOPS outlet sample rotation period from five years to four. It also instituted a broad program of faster item rotation. In many product categories, the item sample is rotated midway between outlet rotations; that is, within a given outlet the item sample is rotated after two years.

The BLS has not, as yet, attempted to implement Hausman's recommendations for reflecting consumer welfare gains from the introduction of new goods or outlets. Although the measurement objective is accepted, the technique has been considered too untested for use by a statistical agency.²⁶ The BLS also has not devised satisfactory procedures for comparing the quality and price at disappearing and appearing outlets, although this has remained an issue of continuing interest, as discussed, for example, in Walter Lane (2000).

Other recommendations

At the request of Chairman Jim Saxton of the Joint Economic Committee (JEC), the BLS submitted a report in June 1997 that addressed the issues raised by the Boskin Commission. That report included a response to each of the Commission's thirteen recommendations to BLS. It is outside the scope of this paper to discuss all the recommendations and responses, but a few points are worth noting here.

In two cases, the BLS expressed misgivings about the advisability of the changes that the Commission proposed. Recommendation 5 suggested that CPI sampling and data collection activities should focus on providing "information on the future longer-term movements in the index." This seemed to run counter to what the

25 It is also true that analysts tended to exaggerate this problem, forgetting that much use of cellular phones during the early 1990s was business use out of scope for the CPI.

26 The National Academy panel was divided on the conceptual desirability of including consumer surplus, while agreeing that implementing the Hausman approach was premature. See National Research Council (2002:157-161).

BLS has considered the fundamental goal of the CPI, which is to measure current price change. BLS sampling procedures are designed to minimize the variance of estimated overall price change subject to the program's constraint on data collection cost, not to support development of an inflation forecast or estimate the degree of inflationary pressure. Therefore, the BLS essentially rejected this Commission recommendation.

Another recommendation urged that CPI sampling should de-emphasize geography by sampling commodities at a national level. In its report to the JEC, the BLS argued that "the practical meaning of this recommendation is somewhat unclear," and discussed the practical difficulties of obtaining national frames for commodity samples. Underlying the Commission's recommendation may have been the belief that the markets for many products are national rather than local, so that it is unnecessary to follow the CPI process of pricing those products in numerous different geographic areas. Subsequent BLS econometric research by Dennis Fixeler and Robert McClelland (1999), however, casts doubt on that assumption by rejecting the hypothesis of a common national price trend for most of the item categories studied.

Several recommendations were accepted by the BLS as identifying reasonable directions of research or effort. One of these was the recommendation that the BLS have "a permanent mechanism for bringing outside information, expertise, and research results to it." The report to the JEC indicated that the Bureau was studying the possibility of creating an academic advisory committee, and in 2000 this came to pass with the formation of the Federal Economic Statistics Advisory Committee (FESAC), which advises the BLS as well as the Census Bureau and Bureau of Economic Analysis. The FESAC model does not satisfy the Boskin Commission's recommendation that the advisory group be

organized by an independent professional entity, but it does constitute a formal and permanent source of input on methodology from the economic and statistical academic community.

The Commission's Recommendation 8 stated that the price of consumer durables such as cars should be measured using a flow-of-services approach such as the rental equivalence method used in the CPI to measure the price of owner-occupied housing. In the JEC report the BLS noted that it was introducing an automobile leasing index in the CPI, reflecting the growth of that market. Data collected for that index could at some point provide the basis for a flow-of-services (i.e., leasing equivalence) approach to vehicle pricing. It is interesting to note, however, that in recent years the rapid increase in housing prices has led many outside analysts to question the appropriateness and accuracy of the rental equivalence approach. Notwithstanding those criticisms, the BLS remains fully committed to the flow of services concept for housing.²⁷

The three longer-run Boskin recommendations involved fundamental research efforts, to look beyond the CPI's "market basket" framework, to investigate the CPI's embedded assumption of price equilibrium, and to undertake data collection initiatives such as on time use and quality of life. The BLS did institute the American Time Use Survey in 2003, although that program is not directly associated with the CPI. Meanwhile, the concerns underlying these longer-run recommendations have much in common with the issues addressed by the National Academy panel organized and partially funded by the BLS in 2000. That panel, under the chairmanship of Charles Schultze, was asked to explore the implications of COLI theory for index construction and to address, for example, the design of indexes for particular purposes and the role of public goods and the environment.

27 BLS discussions of this issue are found in Poole et al. (2005) and Verbrugge (2005).

The Schultze panel's report (National Research Council, 2002) has helped to clarify and extend the debate on many of the same issues dealt with by the Boskin Commission.²⁸

Finally, in the body of the Boskin Report, not in its formal recommendations, the Commission argued that the CPI should treat changes in automobile prices due to additional anti-pollution devices as price rather than quality increases. This was a viewpoint that had been advocated by the BLS decades earlier, but which had failed to gain support from the overall federal statistical community. In 1999 the BLS did change its handling of mandated clean air improvements in both the vehicle and motor fuel components of the CPI, treating those improvements as equivalent to indirect taxes (Fixler, 1998).

Overall Impacts of the Boskin Report

I cannot comment on how the Boskin Commission's work may have changed the way economists viewed published price index movements, or how the report was used in the consideration of ways to reduce the federal budget deficit during the mid-1990s. Certainly the Commission debates had important repercussions on statistical agencies around the world. Many of those agencies differed from the United States by not accepting the COLI as a measurement objective for their CPIs. Nevertheless, they had to discuss how the Commission's arguments applied to their index construction procedures, such as formulas they employed and their techniques for quality adjustment.

As for the US CPI, as discussed earlier, the impact of the Boskin Commission is difficult to separate from the impacts of Federal Reserve criticism or the attention from Congressional committees. The Commission Report did, how-

ever, prolong and greatly accentuate public attention on CPI methods. It is likely that by forcing the BLS to scrutinize the strengths and limitations of all its CPI procedures, and by highlighting and publicizing the budgetary impact of the CPI, the Commission paved the way for the various index improvements made by the BLS in the wake of the report.

Some, although not all, of those CPI improvements had predictable impacts on CPI growth. At the request of the General Accounting Office, the Boskin Commission members themselves provided updated estimates of bias in 2000; their median updated point estimate of bias was 0.8 percentage points annually, indicating that they estimated a 0.3 percentage point impact of the BLS changes to that point. The 1999 *Economic Report of the President* contained a table indicating that the CPI changes had slowed measured growth by 0.68 percentage point annually (Council of Economic Advisers, 1999:93-94). That table, and a similar table presented by Maurine Haver (1999), drew largely on formal estimates provided by the BLS. Finally, in 1999 the BLS introduced an experimental index, the CPI-U-RS, showing estimated CPI movements from 1978 to the present under the assumption that the current CPI-U methodology had been used throughout that period. Comparison of the CPI-U-RS to the official CPI-U provides a measure of the impact of subsequent CPI improvements, including those in the late 1990s such as the elimination of formula bias, use of the geometric mean formula, and application of hedonic models. The CPI-U-RS, described in Kenneth Stewart and Stephen Reed (1999), can be thought of as a partial response to the Boskin Commission's call for a historically revised CPI.

As suggested in the previous section, to some extent the Boskin Report led to the subsequent

28 One of the Boskin Commission members, Zvi Griliches, was a member of the National Academy panel until his untimely death.

decision by the BLS to support formation of the Schultze panel. Many of the underlying issues that the Boskin Commission did not have time to address thoroughly, such as the treatment of environmental changes, or that the Commission took as given, including the COLI framework, were the subject of detailed analysis by the Schultze panel. Some of the positions taken in National Research Council (2002) were controversial, such as the panel's reluctance to fully endorse the COLI objective, its caution about rapid expansion of hedonic adjustment in the CPI, and its recommendation not to employ estimates of consumer surplus from new goods. In other areas, such as medical care, the Schultze panel advocated lines of research that are being followed by the BLS.

In characterizing the BLS reaction to the Boskin Commission, it is important to note the support given to the CPI program by Commissioner Abraham and the rest of the BLS leadership, including Kenneth Dalton and William Barron. The Commissioner played a crucial role in explaining and defending CPI methods to Congress and other audiences, and she successfully led the effort to secure funding for the 1998 CPI budget initiative (as well as the subsequent 2002 initiative). More crucial, perhaps, was her energy in questioning the CPI program about its methods, and her firm support for methodological changes whenever she was convinced of their validity.

Another factor that cannot be overestimated is the role of BLS's Division of Price and Index Number Research (DPINR) as a driving force in many of the methodological debates during the Boskin Commission period. This is not to downgrade the value of other CPI program staff, who not only implemented many methodological improvements but also wrote many of the BLS papers and reports cited here. Under the leadership of Brent Moulton and then Dennis Fixler, however, DPINR provided the conceptual guidance for the BLS positions on numer-

ous difficult measurement issues; devised specific procedures for handling technical problems such as solving formula bias or implementing a superlative formula; completed a prodigious number of quantitative research papers that provided much of the knowledge base for both the BLS and its critics; participated closely in the preparation of BLS testimony, official announcements, and other materials; and often acted as the link between the BLS and the academic community, explaining CPI methodology and demonstrating the Bureau's technical competence. Formation of a price research division was one of the recommendations of the 1961 Stigler Committee. The value of that recommendation was highlighted during the Boskin years, and DPINR continues to play a key role today.

Finally, in comparing the CPI program before and after the Boskin Report, there appears to be a closer relationship now with the economics research community. The issues raised in the report undoubtedly increased the interest in price index methodology among academic economists. Much credit, however, must go to Ernst Berndt, Charles Hulten, Jack Triplett, and others who have worked assiduously through such activities as the Conference on Research in Income and Wealth, the Brookings program on economic measurement (Triplett and Bosworth, 2004), and the NBER Summer Productivity Conference to strengthen the ties between BLS and outside researchers.

References

- Abraham, Katharine G. (1995) "The Consumer Price Index: Its Uses and Limitations as a Cost-of-Living Proxy," *CPI Detailed Report*, July, pp. 3-4.
- Abraham, Katharine G. (1996) "Statistics Under the Spotlight: Improving the Consumer Price Index: Statement," *Proceedings of the Section on Government Statistics*, Papers Presented at the Meeting of the American Statistical Association, August 4-8, pp. 54-56.

- Abraham, Katharine G. (1997a) "Testimony before the Senate Budget Committee," January 30.
- Abraham, Katharine G. (1997b) Comment in "The CPI Commission: Discussion," *American Economic Review* 87 (2), pp. 94-98.
- Abraham, Katharine G., John S. Greenlees and Brent R. Moulton (1998) "Working to Improve the Consumer Price Index," *Journal of Economic Perspectives* 12 (1), pp. 27-36.
- Advisory Commission to Study the Consumer Price Index (1995) "Toward a More Accurate Measure of the Cost of Living, Interim Report to the Senate Finance Committee," September 15.
- Aizcorbe, Ana M. and Patrick C. Jackman (1993) "The commodity substitution effect in CPI data, 1982-91," *Monthly Labor Review* 116, December, pp. 25-33.
- Aizcorbe, Ana M., Robert A. Cage and Patrick C. Jackman (1996) "Commodity Substitution Bias in Laspeyres Indexes: Analysis Using CPI Source Data for 1982-1994," paper presented at the Western Economic Association International Conference in San Francisco, July.
- Armknrecht, Paul A. and Daniel H. Ginsburg (1992) "Measuring Price Changes in Consumer Services," in Griliches (ed.), *Output Measurement in the Service Sectors*. NBER Studies in Income and Wealth 56, University of Chicago Press, Chicago, pp. 109-156.
- Armknrecht, Paul A. and Donald Weyback (1989) "Adjustments for Quality Change in the U.S. Consumer Price Index," *Journal of Official Statistics* 5 (2), pp. 107-123.
- Armknrecht, Paul A., Brent R. Moulton and Kenneth J. Stewart (1995) "Improvements to the Food and Home, Shelter, and Prescription Drug Indexes in the U.S. Consumer Price Index," Bureau of Labor Statistics Working Paper No. 263.
- Bureau of Labor Statistics (1997) "Measurement Issues in the Consumer Price Index," U.S. Department of Labor, June, reprinted in *Statistical Journal of the United Nations ECE*, 15 (1998), pp. 1-36.
- Bureau of Labor Statistics (1999) "Scheduled Updates for Expenditure Weights in the Consumer Price Index," *CPI Detailed Report*, February, pp. 5-6.
- Bureau of Labor Statistics (2000) "Direct Pricing of Health Insurance in the Consumer Price Index," paper presented at the meeting of the Federal Economic Statistics Advisory Committee. Bureau of Labor Statistics, December.
- Cage, Robert (1996) "New methodology for selecting outlet samples," *Monthly Labor Review* 119, December, pp. 49-61.
- Cage, Robert, John S. Greenlees and Patrick Jackman (2003) "Introducing the Chained Consumer Price Index," in *International Working Group on Price Indices (Ottawa Group) Proceedings of the Seventh Meeting*, Thierry Lacroix, (ed.), Paris: INSEE, pp. 213-246.
- Cardenas, Elaine M. (1996) "New Revision of the CPI hospital services component," *Monthly Labor Review* 119, December, pp. 40-48.
- Council of Economic Advisers (1999) Economic Report of the President, (Washington, D.C.: Government Printing Office) February.
- Dalton, Kenneth V., John S. Greenlees and Kenneth J. Stewart (1998) "Incorporating a geometric mean formula into the CPI," *Monthly Labor Review* 121, October, pp. 3-7.
- Erickson, Timothy (1995) "Formula Bias: A Simple Description in an Idealized Setting," Bureau of Labor Statistics manuscript.
- Erickson, Timothy (1996) "Effects of Mismeasuring Base Period Prices When Estimating the Laspeyres Index: Some Idealized Cases," Bureau of Labor Statistics Working Paper No. 284.
- Fixler, Dennis (1998) "Treatment of Mandated Pollution Control Measures in the CPI," *CPI Detailed Report*, September, 4-7.
- Fixler, Dennis, Charles Fortuna, John Greenlees and Walter Lane (2000) "The Use of Hedonic Regressions to Handle Quality Change: The Experience in the U.S. CPI," in *Proceedings of the Ottawa Group Fifth Meeting* (Reykjavik: Statistics Iceland).
- Fixler, Dennis and McClelland, Robert (1999) "The Use of a National Price Index to Overcome Small Sample Bias," presented at the NBER Summer Institute, July.
- Greenlees, John S. (1996) "Comment on Matthew D. Shapiro and David W. Wilcox: 'Mismeasurement in the Consumer Price Index: An Evaluation,'" *NBER Macroeconomics Annual 1996*, pp. 143-147.
- Greenlees, John S. (1997) "A Bureau of Labor Statistics Perspective on Bias in the Consumer Price Index," Federal Reserve Bank of St. Louis *Review*, May/June, pp. 175-178.
- Greenlees, John S. (1998) "Expenditure Weight Updates and Measured Inflation," in *Proceedings of the Third Meeting of the International Working Group on Price Indices*, B. M. Balk (ed.), Statistics Netherlands, pp. 57-64.
- Greenlees, John S. (2001) "The U.S. CPI and the Cost-of-Living Objective," paper presented at the Joint ECE/ILO Meeting on Consumer Price Indices in Geneva, Switzerland, November 2.
- Greenlees, John S. and Charles C. Mason (1996) "Overview of the 1998 revision of the Consumer Price Index," *Monthly Labor Review* 119, December, pp. 3-9.
- Hausman, Jerry A. (1997) "Valuation of New Goods Under Perfect and Imperfect Competition," in

- Bresnahan and Gordon (eds.), *The Economics of New Goods*, NBER Studies in Income and Wealth 58 (University of Chicago Press, Chicago) pp. 209-237.
- Hausman, Jerry A. (1999) "Cellular Telephone, New Products, and the CPI," *Journal of Business and Economic Statistics* 17 (2), pp. 188-94.
- Haver, Maurine A. (1999) "The Statistics Corner: recent and planned improvements to the CPI," *Business Economics*, January.
- International Labour Office (2004) *Consumer Price Index Manual: Theory and Practice* (Geneva: International Labour Office).
- Johnson, David S., Stephen B. Reed and Kenneth J. Stewart (2005) "What Has Happened to Price Measurement Since the Boskin Report? The U.S. Experience," paper presented at the OECD Conference on Inflation Measures, Paris, June.
- Lane, Walter (2000) "Addressing the New Goods Problem in the Consumer Price Index," paper presented at the Conference on Issues in Measuring Price Change and Consumption, Bureau of Labor Statistics, June.
- Lebow, David E., John M. Roberts and David J. Stockton (1992) "Understanding the Goal of Price Stability," Board of Governors of the Federal Reserve System, October.
- Lebow, David E., John M. Roberts and David J. Stockton (1994) "Monetary Policy and 'The Price Level,'" Board of Governors of the Federal Reserve System, July.
- Lebow, David E. and Jeremy B. Rudd (2003) "Measurement Error in the Consumer Price Index: Where Do We Stand?" *Journal of Economic Literature* XLI, March, pp. 159-201.
- McClelland, Robert (1996) "Evaluating Formula Bias in Various Indexes Using Simulations," Bureau of Labor Statistics Working Paper No. 289.
- McClelland, Robert (1998) "On the Use of the Geometric Mean and Laspeyres Formulas in the CPI," Bureau of Labor Statistics manuscript.
- Moulton, Brent R. (1993) "Basic components of the CPI: estimation of price changes," *Monthly Labor Review* 116, December, pp. 13-24.
- Moulton, Brent R. (1996a) "Bias in the Consumer Price Index: What is the Evidence?" *Journal of Economic Perspectives*, Fall, pp. 159-77.
- Moulton, Brent R. (1996b) "Estimation of Elementary Indexes of the Consumer Price Index," *Proceedings of the Business and Economic Statistics Section*, American Statistical Association.
- Moulton, Brent R. and Karin E. Moses (1997) "Addressing the Quality Change Issue in the Consumer Price Index," *Brookings Papers on Economic Activity* (1), pp. 305-349.
- Moulton, Brent, Karin Moses and Timothy LaFleur (1999) "Research on improved quality adjustment in the CPI: The case of televisions," *Proceedings of the Fourth Meeting of the International Working Group on Price Indices* (Washington, D.C., Bureau of Labor Statistics) January.
- Moulton, Brent R. and Kenneth J. Stewart (1999) "An Overview of Experimental U.S. Consumer Price Indexes," *Journal of Business and Economic Statistics* 17 (2), April, pp. 141-151.
- National Research Council (2002) *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes. Panel on Conceptual, Measurement, and Other Statistical Issues in Developing Cost-of-Living Indexes*. Charles L. Schultze and Christopher Mackie (eds.), Committee on National Statistics, Division of Behavioral and Social Sciences and Education (Washington, DC: National Academy Press).
- Poole, Robert, Frank Ptacek and Randal Verbrugge (2005) "Treatment of Owner-Occupied Housing in the CPI," paper prepared for presentation to the Federal Economic Statistics Advisory Committee, December 9.
- Price Statistics Review Committee (1961) *The Price Statistics of the Federal Government*, National Bureau of Economic Research, New York.
- Reinsdorf, Marshall (1993) "The Effect of Outlet Price Differentials on the U.S. Consumer Price Index," in Foss, Manser and Young (eds.), *Price Measurements and Their Uses*, NBER Studies in Income and Wealth 57 (Chicago: University of Chicago Press) pp. 227-254.
- Reinsdorf, Marshall (1998) "Formula Bias and Within-Stratum Substitution Bias in the US CPI," *Review of Economics and Statistics* 58 (2), pp. 175-187.
- Reinsdorf, Marshall and Brent R. Moulton (1997) "The Construction of basic components of cost-of-living indexes," in Bresnahan and Gordon (eds.), *The Economics of New Goods*, NBER Studies in Income and Wealth 58 (Chicago: University of Chicago Press) pp. 397-436.
- Shapiro, Matthew D. and David W. Wilcox (1996) "Mismeasurement in the Consumer Price Index: An Evaluation," *NBER Macroeconomics Annual 1996*, pp. 93-142.
- Song, Xue, William D. Marder, Onur Baser, Robert Houchens, Jonathan E. Conklin and Ralph Bradley (2004) "Can Health Care Claims Data Improve the Estimation of the Medical CPI?" paper presented at the CRIW Conference on Price Index Concepts and Measurement in Vancouver, Canada, version of October 14.
- Stewart, Ken (1996) "Extending the Improvement in CPI Sample Rotation Procedures," *CPI Detailed Report*, June, pp. 9-10.

- Stewart, Kenneth J. and Stephen B. Reed (1999) "Consumer Price Index Research Series Using Current Methods, 1978-1998," *Monthly Labor Review*, June, pp. 29-38.
- Triplett, Jack E. (2000) *Handbook on Quality Adjustment of Price Indexes for Information and Communication Technology Projects*, OECD Report OECD/DSTI/EAS/IND/SWP/(2000)2.
- Triplett, Jack E. (2001) "Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index?" *Economic Journal* III (472), pp. F311-F334.
- Triplett, Jack E. and Barry P. Bosworth (2004) *Productivity in the U.S. Services Sector* (Washington, D.C.: The Brookings Institution).
- U.S. Senate, (1995) Hearings before the Committee on Finance, March 13, April 6, and June 6.
- Verbrugge, Randal (2005) "The Puzzling Divergence of Rents and User Costs, 1980-2004," mimeo, Bureau of Labor Statistics.
- Wynne, M.A. and F.D. Sigalla (1994) "The Consumer Price Index," *Economic Review*, Federal Reserve Bank of Dallas, Second Quarter, pp. 1-22.

The Boskin Commission Report After a Decade

Jack E. Triplett¹
Brookings Institution

THE ADVISORY COMMISSION TO Study the Consumer Price Index, known as the Boskin Commission after its chairman, Michael Boskin, published its final report at the end of 1996 (Boskin et al., 1996, hereafter Commission Report). The Commission Report has been exceptionally influential, not only in the United States but worldwide. Its chairman noted recently that “statistical agencies around the world have been using the Commission report as a major input to their own agendas for improvement” (Boskin, 2005:2).

To review the Commission Report and its impact, I have selected six topics. Other papers in this symposium elaborate the report’s influence on U.S. data, so I specialize by discussing the report’s world-wide influence (Section one). Section two concerns one of its less salutatory legacies, the increased popularity of “guestimates.”

Two parts of the Commission’s report have not stood up to the test of time. In its analysis of quality change bias (the largest part of its famous 1.1 percentage point CPI upward bias estimate), the Commission ignored the possibility that quality improvements could nevertheless produce a net downward bias to CPI components because the implicit quality adjustments inherent in BLS procedures might over-adjust (Section three). The Commission’s analysis of CPI basic components likewise has been overtaken

by more recent analysis, which has focused on consumer search behavior; the research agenda for measuring CPI basic components does not now seem to be solely Konüs-type commodity substitution, one level down, as the Commission supposed (Section four).

In Section five, I consider the Commission’s neglected recommendation on CPI classifications, which I contend ought to be implemented by BLS. The concluding section of the paper includes some remarks on the politics of the Commission Report, and the lost opportunity for distinguishing more clearly the difference between improving the CPI and articulating improved bases for escalating payments to Social Security participants.

I neglect the Commission’s perhaps major recommendation — that the CPI should be an approximation to a cost of living index (COLI) — because I thoroughly agree with it. My position is presented in Triplett (2001) and in Reinsdorf and Triplett (2004).

The Report’s Influence World-Wide

I begin with an anecdote. The “Ottawa Group” is a recurring international conference on price indexes that is attended by many statistical agencies of the countries of the Organization for Economic Cooperation and

¹ The author is Visiting Fellow, Brookings Institution. This paper was originally presented at the session “The Boskin Commission Report After a Decade: Is the CPI Still Biased?,” held at the annual meeting of the American Economic Association, January 6-8, 2006, Boston, Mass. B.K. Atrostic’s perceptive comments improved this paper substantially, as did suggestions from the editor. Email: jtriplett@brookings.edu

Development (OECD). At the 1994 meeting of the Ottawa Group (Statistics Canada, 1994), I introduced a discussion of the survey paper by Wynne and Sigalla (1993), which was one of the first attempts to pull together research findings in order to quantify bias in the CPI. After the discussion, representatives of several countries said: It is very interesting that in the U.S. you have these public criticisms of your CPI. That could never happen in our country.

At the time (1994), they were more or less right. Most countries had no tradition of outside reviews and criticisms by economists and statisticians of the statistical outputs produced by government agencies, and no organizations (such as the Conference on Research in Income and Wealth) to facilitate reviews and exchanges of views between analytic users and statistical agency staffs.

The Boskin Commission changed all that. After publication of the Commission Report — indeed after the Commission’s “Interim Report” was published in 1995 — many countries’ statistical agencies found, to their surprise, that the questions that the Commission asked were for the first time being asked of their CPIs as well. Though they had not been very vocal in the past, economists in other countries shared the Commission’s conception of the measurement issues that surround the CPI, and were persuaded by the force of the Commission’s analysis that CPI bias was a topic to be taken seriously.

The Commission Report, and other major U.S. studies such as Wynne and Sigalla (1993), Lebow, Roberts, and Stockton (1994), and Shapiro and Wilcox (1996), were followed by studies similar in intent, such as Oulton (1995) for the U.K., Hoffmann (1998) for Germany, and Shiratsuka (1999) for Japan. But these published or formal studies are just the visible tip of the ferment, for more commonly central banks and

treasury or finance ministries posed questions to their statistical agencies that were conditioned by the Commission Report, or by the controversy in the U.S. that originally spawned the Commission. Some of these queries led statistical agencies to publish their own self-evaluations, Lequillier (1997) for France being an example (Wynne and Rodriguez-Palenzuela, 2001, mention some others).

Many statistical agencies outside the U.S. initially responded to the Boskin Commission’s bias estimates by claiming that they did not apply to their indexes. Eurostat, the statistical agency of the European community, summarized this point of view in its response to an OECD country survey in 1997 (OECD, 1997:27): “Most member states feel that Boskin biases apply to the United States’s CPI and that the United States aims at a COLI while that is not the case in the EU. There is no bias that could arise due to the use of the HICP as a basis for measuring inflation instead of any sort of COLI.” [The HICP are the harmonized indexes of consumer prices that Eurostat designed, with input from member states, for all of the European Union countries.] That is, because the HICP were not intended to be COLIs, Eurostat was contending that they were not biased measures of inflation.²

International response to the Commission’s detailed bias estimates typically took the following forms.

Upper Level Substitution Bias

Upper level substitution bias is the classic Konüs (1925) type substitution in response to relative price change among the components of the CPI. The Commission’s point estimate of bias from upper level substitution was 0.15 index points per year.

As noted above, many countries claimed that their indexes were not intended to be cost of

2 Eurostat and some other countries’ statistical agencies insisted that inflation was not measured by a COLI. A paper by Peter Hill conveys the joint position (Hill, 1997). I did not find it convincing, for the reasons in Triplett (2001).

living indexes (COLI). In any case, they also pointed out that they updated their weights more frequently than did the BLS. At the time the Commission wrote, weights in the U.S. CPI were about 13 years old, which was indeed older than CPI weights for any other major OECD country. Eurostat's standard for HICP indexes was one to two years, ideally, and no more than five years maximum. Thus, their assertions had some plausibility, though no other country outside the U.S. had an actual estimate of substitution bias at that time.³

Lower Level Substitution Bias

Substitution within a CPI detailed component (examples of detailed components are bananas in the fruit category and musical instruments in the entertainment category) the Commission called lower level substitution bias. The Commission set lower level substitution bias at 0.25 points, or more than upper-level substitution bias, and suggested that the BLS move to the unweighted geometric mean formula, in place of the arithmetic average of price relatives (APR) that had been in place since 1978.⁴

European countries were quick to point out that Eurostat had already endorsed the geometric mean formula for basic components in HICPs, and indeed had banned the BLS APR method. Accordingly, they believed their indexes were immune to lower level substitution bias (even if they were considered as COLIs, which they were not). Again, this European contention had merit, though some countries were subsequently discovered still using arithmetic means.

Outlet Substitution Bias

This bias arises from consumers shifting to lower cost outlets. Outlet substitution behavior is not recorded in the CPI because the CPI is

constructed from a matched outlet sample. The Commission put outlet substitution bias at 0.10 index points annually, but it did not recommend any CPI improvements that would reduce the bias, other than perhaps bringing new outlets more quickly into the sample.

Many countries claimed that outlet substitution was no problem in their CPIs because their retail sectors were less dynamic than United States retailing. It is true that strong multi-factor productivity (MFP) growth in the U.S. retail sector in the 1990s (Triplett and Bosworth, 2004) seems without a parallel in Europe at the same time (Timmer and Inklaar, 2005), which would be evidence in favor of the European position.

On the other hand, at the time of the Commission Report the only existing European study relating to outlet effects in the CPI (Saglio, 1994, on chocolate bars in France) found substantial shifts from traditional small French retail shops to hyper-markets and so forth. Covas and Silva (1999) found outlet bias for Portugal that varied by year and by methodology; it averaged a little under 0.1 point per year in one method (which the authors believed understated substitution) and 0.49 in the other, probably an overstatement. These estimates, then, are not lower than the Commission's estimate for the U.S.

Another factor bearing on this whole matter is sampling. Few countries outside the U.S. employ much of a probability sample for items and outlets in their CPIs. The U.K. and Sweden choose outlets by probability sample, but they are exceptions (an international comparison of CPI methods is OECD, 1997). Outlet bias might well be larger in a non-probability sample than in a probability sample (the Stigler Committee thought so — see Price Statistics Review Committee, 1961), though this might not be outlet *substitution* bias.

3 This has now changed. For a sophisticated substitution bias estimate for the U.K., see Blow and Crawford (2001), who implemented the procedures used for the U.S. by Manser and McDonald (1988).

4 In Section four, I suggest that the Commission and the BLS were too quick to judgment on the lower level substitution bias question.

On balance, it is not clear that other countries' protestations — that their CPIs were not subject to outlet substitution bias — were correct.

Quality Change Error

The Commission put error arising from inadequate adjustments for quality change at +0.60 index points per year. As suggested in Section three, this estimate is very shaky. Surprisingly, the Commission did not recommend that BLS use hedonic methods to produce better quality adjusted indexes.

Some countries contended that their quality change error was lower than in the U.S. CPI. Some of their comments revealed lack of knowledge about U.S. quality change procedures, and also, more surprisingly, lack of understanding of the biases in the quality adjustment procedures that many of them were using.⁵ The Eurostat statement mentioned earlier maintained that in the CPI the “main bias is agreed to come from quality changes. Since there is no agreement on what could or should be done, there can be no agreement on what biases exist” (quoted in OECD, 1997:27).

Quality change is a major problem in all countries' CPIs, and little reason exists for believing that European countries have devised better methods than those used in North America. Dalén (2002) and Ribe (2002) provide revealing analyses of European country practices on quality change. Hoffman (1998) showed that differences in treating quality change in the German CPI created enormous differences in the price indexes for household appliances,⁶ and Silver and his colleagues found substantial quality change error in U.K. price indexes (for example, Silver and Heravi, 2001). No OECD country, with the exceptions of France and Sweden, had any hedonic price indexes for computers before

the Commission wrote (of course, neither did the U.S. CPI), and in most of them the environment was very hostile toward hedonic indexes. By roughly the turn of the century, however, countries outside North America became more receptive to new methods, perhaps for more reasons than the ferment caused by the Commission Report, but the report certainly contributed to their changed viewpoints.

Overall Impact

The Commission Report had an extremely salutary effect on international price statistics. Countries differ greatly in the quality of their statistics, in their agencies' receptiveness to new methods, and in their lines of communications with analytic users. Accordingly, generalizations do not fit individual country experiences. Yet, several generalizations apply across many countries.

- The report brought into open discussion a set of price measurement issues that are properly concerns in every country, not just in the U.S.
- The report, and the controversy that surrounded it, engendered a more extensive dialogue between statistical agencies and their analytic users than had been the norm before.
- The report, and the world-wide discussion of it, jarred some agencies out of their complacency about the quality of the numbers they were producing.
- The report also encouraged price index research among economists and statisticians in countries outside of North America, where it had been difficult to find before. For example, a 2001 Workshop in the Measurement of Price Indices sponsored by the European Central Bank and the Centre for Economic Policy Research contained

5 This is a more complex matter than is sometimes appreciated. An analysis of all methods for handling quality change in price indexes, and the probable errors in each of them, is found in Chapter 2 of Triplett (2004).

6 Production of the German CPI is not centralized, unlike the situations in the U.S., Canada, and most other OECD countries.

contributions on all the price measurement topics addressed by the Commission and estimates of price index bias that applied to a variety of components and countries (Camba-Mendez, Gaspar and Wynne, 2001). Wynne and Rodriguez-Palenzuela (2001) pointed to much recent research but even so, found insufficient research in Europe to permit an evaluation of measurement bias in HICP indexes.

Guestimates

Ironically, what I liked least about the Commission Report was exactly what made it so influential — its guestimate of 1.1 percentage points of bias.

A guestimate is a number produced when one does not have research results. The Commission (and others who have followed its lead) used ad hoc reasoning to come up with a number. I did not think that we economists knew enough to estimate the overall CPI bias when the Commission wrote its report, and I still think we do not know enough.

But this seemingly so precise 1.1 number caught the eyes not only of the press and the politicians, but also of economists in the U.S. and in other countries. Jacob Ryten wrote just after the release of the Commission's interim report: "...for the first time ever, a blue ribbon commission dared give a number for the estimate of total bias and detailed each of the contributing factors" (Ryten, in Ducharme, 1997:3). Without the guestimates, the Commission Report was just another dry, academic study to be perused by professionals (and perhaps only those professionals who were previously interested in price index research). Without the guestimates, the report would likely have had minimal impact.

Conversations with Commission members suggest that some, at least, were ill at ease themselves with guestimates, but I take it that they felt that their mandate from the Senate Finance Committee compelled them to brew one. The report notes the shortage of research findings on which to base a better and more scientific estimate. Moreover, the Commission presented a "plausible range" (also a guestimate, of course) of 0.80 to 1.60 index points, so they themselves were not asserting the degree of precision that some of their readers took away from the report.

The Commission's 1.1 point guestimate, made to carry out their Senate Finance Committee mandate, is one thing. My personal preference is to resist the seductive blandishments of politics and politicians, but not everyone feels that way. But the Commission's 1.1 point guestimate seems to have legitimized the making of guestimates by professionals who are not under the sway of the Senate Finance Committee, and who seem eager to charge into what apparently the Commission members entered into with some reluctance.

I should specify more precisely what I mean by guestimate, since they have now been so much used that the distinction I am drawing may no longer be clear to some economists. By guestimate, I do not mean a compilation of research results. The Commission's estimate of substitution bias (0.15 points) was not a guestimate. It was a professional distillation of research results that encompassed a number of different studies (most of them conducted within the BLS) that used a variety of methods and reached closely similar quantitative conclusions on the size of the bias.⁷

Little controversy exists about the Commission's substitution bias estimate because the Commission could cite the research studies on which its estimate was based. Of course,

7 The major studies are Christensen and Manser (1976), Manser (1975), Braithwait (1980), Manser and McDonald (1988), and Aizcorbe and Jackman (1993). BLS research on substitution bias is summarized in Reinsdorf and Triplett (2004). Recent BLS publications suggest a somewhat larger substitution bias estimate, but those BLS numbers are suspect because they commingle commodity substitution with area effects (substitution of cabbage in New York for carrots in Los Angeles).

judgment is involved in integrating these studies. I believe that any competent group of professionals, viewing the same evidence, would come up with a similar judgment.

However, for its measure of the impact of quality change (at 0.6 points, much larger than its substitution bias estimate), the Commission had no comparable body of research on which to rely. Instead, the commission entered into a variety of more or less ad hoc reasoning exercises.

In some cases the Commission's indirect logic seems persuasive. But persuasive or not, its conclusions were not backed up by firm empirical estimates of the size of the bias. There is no contradiction in saying at the same time that the Commission's guesstimates were better than some previous ones (partly because they derived them on a component-by-component basis, which imposes a certain consistency and check on the total) and in saying that they were nevertheless inadequate. They are inadequate in terms of what one should demand in a scientific estimate — that it is evidenced-based and reproducible. Evidence-based, in this case, means a comparison of actual CPI indexes with others computed on a different basis or with different data, such as comparisons involving conventional methods for handling quality change and hedonic indexes, or else an empirical analysis of the impacts of applying alternative methods.

The guesstimates made parts of the report's 1.1 point bias estimate unreliable, especially its 0.6 point estimate for quality change and to a lesser

degree its 0.1 point estimate for what the Commission called lower level substitution bias. More regrettably, the Commission's guesstimates have spawned a flurry of other guesstimates in the U.S. and in other countries. Recent ones are Lebow and Rudd (2003) for the U.S. Federal Reserve Board and Rossiter (2005) for the Bank of Canada. Especially in their estimates of quality change error, these new guesstimates are flawed in the same way that the Commission's guesstimates were flawed, and for less reason. In the next section, I give some concrete reasons why guesstimates fail.

In my view, the economics profession would be better served if the resources put into producing new guesstimates were instead devoted to price index research. New research would not only improve our information on the accuracy of the CPI, but perhaps as well improve the index.⁸

The Boskin Commission on Quality Change

Quality change is the biggest problem in obtaining accurate price indexes. The Commission was surely right that estimating the bias caused by quality change — in fact, doing any meaningful research on quality change — must proceed on a component by component basis. To illustrate the principle, it is now well established that the rate of quality-corrected price decline differs substantially across different categories of computer and communications equipment, and across different categories of semiconductors (see the review of studies by Mark Doms, Ana Aizcorbe, and collaborators, in Triplett and

8 In an interesting parallel, the Mitchell Committee that investigated alleged bias in the CPI during World War II guesstimated that the downward bias (not upward bias) in the index from forced trading up and quality deterioration was 1-3 per cent for food, 4-5 per cent for clothing, and 8-11 per cent for house furnishings. These numbers are all several multiples of the comparable component estimates in Boskin et al. (1996, Table 2), and of course are of opposite sign because of wartime conditions. George Stigler, head of the 1961 Stigler Committee on price indexes, declined to make an estimate when asked by a Congressional committee to do so, saying that the committee did not know enough. However, one of the Committee's members, Richard Ruggles, inserted a footnote suggesting 3 percentage points of bias, a guesstimate nearly 3 times that of the Boskin Commission. The Mitchell and Stigler reports are reviewed in Reinsdorf and Triplett (2004). I suspect that the clear downward trend in the magnitudes of professional guesstimates about CPI bias does not reflect entirely a consensus that the BLS is doing a better job. Rather, more recent professional judgments draw on more empirical research. Speculative judgments that are not informed by research seem invariably to produce larger numbers than are documented in research findings.

Bosworth, 2004, chapter 10). Each component needs its own study. One cannot just presume that “electronics” price indexes fall equally. The same is true of components in other product groupings. There is no substitute for what Shapiro and Wilcox (1996) termed the “house to house combat” approach to quality change.

The Commission examined research estimates for those cases where research existed. Some research missed their net, and in one or two cases, regrettably, they substituted their own judgments for research findings. However, as the Commission also observed, in many categories “There is little if any published evidence that allows us to reach a determination” (Boskin et al., 1996:41). For those cases, they used their intuitions and a variety of indirect data and hypotheses — guesstimates.

What is logically wrong with guesstimates of quality change error?

Most economists, I suspect, believe that our dynamic and competitive economy generates an overall improvement in the quality and variety of goods and services that are available for consumption. I share that presumption. Counter examples can be cited — the most recent versions of some software that I use may not be better than older versions, quality deterioration in some services is often alleged, and some traditional manufactured products (desk telephone sets and toasters) seem less reliable than their ancestors. But overall, the direction of quality change is undoubtedly positive, not negative.

From the presumption that quality is on balance improved, economists have often inferred that quality change must bias price indexes upward, that $\Delta QI > 0$ implies $\beta > 0$ (where I use ΔQI to represent quality change, or its value, and β to indicate the price index bias). The converse expectation is thought for those cases when $\Delta QI < 0$. Accordingly, an additional infer-

ence goes, if one can form some estimate of the direction and size of ΔQI , that information can be used to estimate the price index bias. Even if the price index agency makes some adjustments, these economists infer that information on ΔQI will at least provide a bound on β .

All three of these inferences are wrong. Information on ΔQI , even if it were reliable, does not determine the magnitude of β , or even its sign.

The quality error that is incorporated into the index must be a product of:

- a) the type and extent of quality change that takes place and
- b) what the price index agency did about it.

If the price index agency did nothing, of course, either ignored the quality change or did not notice it, then the full amount of the quality change passes through into a bias in the price index. In this case *only*, quality improvements cause upward bias to price indexes — that is, $\Delta QI > 0$ implies $\beta > 0$ — and quality deteriorations results in downward biased price indexes.

Obviously, if the price index agency does something about the quality change that it observes, then the full amount of the quality error does not pass through untouched into index number bias. That is probably well understood.

Less well understood is that most of the procedures used by price index agencies create *implicit* quality adjustments when quality change is encountered. Those implicit adjustments can be either too small or too large, and a fair amount of evidence has accumulated that the implicit adjustments are frequently too large. *When the implicit adjustments are too large, quality change can create an error of the opposite sign from the sign of the quality change itself.*

Why is the sign of the error not determined by the sign of the quality change?

I first examine the Canadian CPI because its method is somewhat easier to analyze. Statistics

Canada frequently uses the “link-to-show-no-change” method when quality changes are encountered (see Statistics Canada, 1995). Under this method, if a new model has higher quality than the model that it replaces and its price is also higher, the price difference between the new model and the old one is taken as the value of the quality change, so: $\Delta P \approx \text{est } \Delta QI$ (the estimated quality adjustment). Note that the new and old typically are not observed in the same month, so $\Delta P = P_{nt} / P_{o,t-1}$, where n and o designate new and old models. This implies that when the new model first enters the sample, *it can generate no price change in the index*, by construction.

Suppose that the seller takes the opportunity of the new model to introduce a price increase, above the value of the quality change the new model represents. In this case, $\Delta P > \Delta QI$. The implicit quality adjustment is too high, relative to the true quality change, because $\text{est } \Delta QI (\equiv \Delta P) > \Delta QI$; the price index bias is negative, $\beta < 0$, even though the quality change is positive.

Conversely, suppose that $\Delta P < \Delta QI$ — either the price increase is less than the value of the quality change or some price decline accompanies the improved new model (frequently the case with electronics, for example). Then, the implicit quality adjustment given by ΔP is too low; any price decline is missed or adjusted out of the index, and $\beta > 0$, the bias is positive.

Even when quality is improving, therefore, *the quality change bias could go either way*. The bias depends on the ratio of $\Delta P / \Delta QI$, and whether that ratio is greater or lesser than unity; it does not depend on whether $\Delta QI > 0$ or $0 < \Delta QI$, that is, whether quality is improving or deteriorating.

Note that the quality change bias is more nearly a function of the direction of price change than of the direction of quality change. When prices are rising, one can infer that the index will be biased downward, it will show too little infla-

tion because the method tends to remove price increases from the index. The proposition is completely symmetric: If prices are falling (true of electronic goods), the bias is upward because the method picks up too little of the price decline, which is one reason that hedonic indexes for electronic goods so frequently fall more than indexes constructed with the usual statistical agency methods (a review of studies is in Triplett, 2004, chapter 4). It has been shown repeatedly that in many CPI components price changes occur most frequently when quality changes occur (Moulton and Moses (1997), Armknecht and Weyback (1989), and the discussion of Armknecht’s findings in Shapiro and Wilcox (1996)).

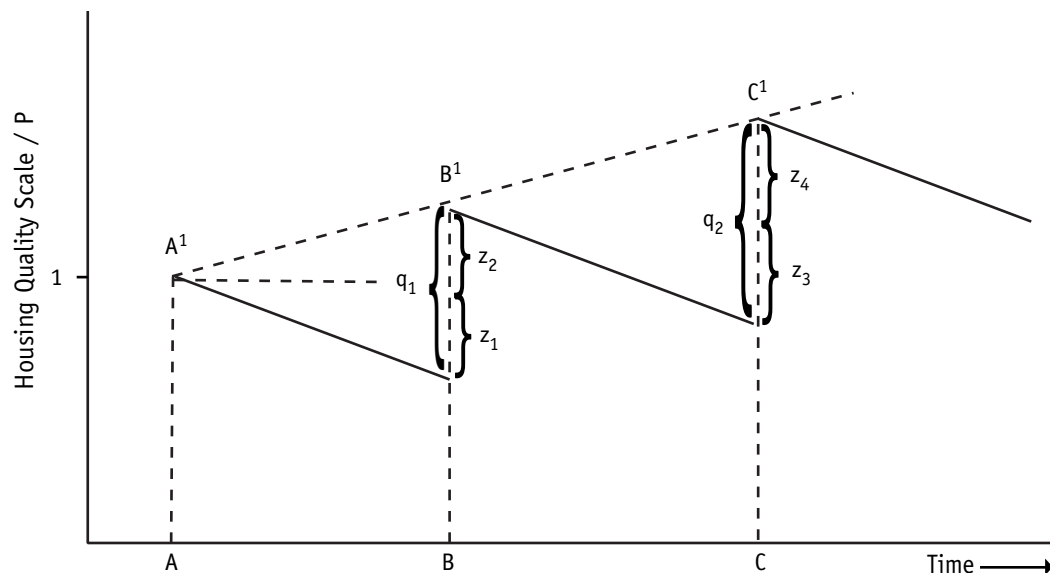
In cases where quality is deteriorating, a parallel analysis applies, which need not be detailed here.

The U.S. does not use Statistics Canada’s link-to-show-no-change method in the CPI, but rather another linking method often termed “deletion.” In the deletion linking method, the prices of the item whose quality changed (both the old item and its replacement) are deleted from the index for the adjacent periods in question. The *price change* for the missing item (not its quality change, and not its price level, as sometimes said) is imputed from price movements of other items that have not changed in quality. Again, this method creates an implicit quality adjustment.⁹

Although the U.S. method is different from the Statistics Canada method, the analysis is similar: The quality bias in the index depends on the size of the implicit quality adjustment and whether the implicit quality adjustment is too high or too low. The BLS found a number of years ago that the deletion method is biased (though not so much as the Canadian method), so it now uses a modification of the deletion method developed by Armknecht (Armknecht and Weyback, 1989), in an attempt to reduce the

9 The deletion method, its probable bias, and the algebra that provides its analysis are presented in Triplett (2004, Chapter two).

Chart 1



bias. They call the modification the “class mean” method. The best empirical evaluation of the implications of the deletion method is Moulton and Moses (1997), whose results suggest that it is on balance biased toward overadjusting for quality improvements.

No guestimate of quality change bias has paid much attention to where the bias comes from and how its origin influences one’s expectations about it. The guestimating methods that have been used to estimate CPI quality change bias (B) are predominantly guestimates about ΔQI . But the sign of the quality error in the index is not determined by whether quality is improving or deteriorating, and the amount of the error or bias is not determined by the amount of quality change. Accordingly, the guestimates are flawed, possibly even with respect to their sign. The following example illustrates.

An Example: Housing

The Commission made a guestimate for how much it thought the quality of housing had improved and used that guestimate as the basis for its CPI housing bias estimate. Studies have shown (the major one is Randolph, 1988) that

even though housing quality is steadily improving, the quality bias is negative in the CPI rent and homeowners equivalent rent housing components, unless a correction is made.

In Chart 1, the vertical axis measures simultaneously a housing quality scale and the price of rental housing (I assume for simplicity in drawing the diagram that competitive markets operate to bring about this equality instantaneously). At point A, a new rental housing unit comes into the sample. I normalize its price and quality at 1.0, and assume that no housing inflation is taking place (again for simplicity in drawing the diagram).

In use, the rental unit begins a slow, probably imperceptible, process of deterioration depicted in Chart 1 as a decline in the unit’s quality (and price) over time. The monthly deterioration would not normally be detected when BLS collects the price for this rental unit for the CPI. The price index is accordingly biased downward (the term in the literature is “aging bias”) over the interval from A to B because the price/quality decline shown from A to B should have been adjusted out of the index.

At point “B” on Chart 1, the rental unit is renovated. It may not only be painted, repaired, and

so forth, but it may also receive new and better appliances, perhaps it is centrally air conditioned where it was not before, and so forth. Thus, the renovation may create a rental unit of higher quality than it was originally, by the amount z_2 in Chart 1. The amount z_2 is the correct measure of improvement in housing quality, and if the Commission (or other guestimators) arrived at the right number, they would have estimated improved housing quality in the amount of z_2 .

However, the price index agency would never compare the deteriorated unit with the renovated unit, they differ too much to make a valid price index comparison. If the agency used Statistics Canada's method (the simplest to show on the diagram), it would apply the quality adjustment q_1 , that is, the sum of z_1 and z_2 , as the quality adjustment. This quality adjustment is too large because z_1 is just the amount that restores the rental unit's quality to its as-new level. Put another way, when quality deterioration occurred with aging, no adjustment was made because the monthly changes were too small to be observed, so the index was biased downward. Then when the quality was restored, there is so much difference in the unit that it does not meet the standards for a the month-to-month "match," so an implicit quality adjustment would be made — the change in rent associated with the renovation is linked out. That means, however, not only that the negative quality error from aging bias remains in the index, but also that the value of its corresponding renovation is linked out of the index (the adjustment for z_2 is not at issue). Overall, the index is biased downward, even if on average housing quality is improving, as shown in Chart 1.

The U.S. CPI contains an aging correction, based on Randolph (1988). Aging bias has become very well known in the housing literature — see, for recent examples, McCarthy and Peach (2002) and Krone, Nakamura, and Voith (2002). Aging bias was rejected by the Commission, (but see Gor-

don's contribution to this symposium), and it is apparently not known at all by those who have followed the Commission by trying to estimate, in effect, z_2 and adjusting it out of the index.

Wrap Up

The implications of the linking and deletion methods have been in the price index literature for a long time (at least since Triplett, 1971). The Commission ignored the matter, and though I think that was a flaw in their report, perhaps at the time the Commission wrote, it was still regarded as controversial. But since then, Moulton and Moses (1997) produced for the first time estimates of the aggregate impact of the method, and the Committee on National Statistics Panel report (Schultze and Mackie, 2002) spelled out clearly and plainly the price index implications of this form of handling quality change. Gordon's (2003, 2004) studies on clothing and housing imply that the same negative bias from linking in the matched model method infests historical price measures (that is, he confirms with historical data the earlier BLS methodological studies by Liegey (1993) on clothing and Randolph (1988) on housing). Yet, the implications of deletion, and other methods that the agencies use for handling quality change, are still ignored in the most recent U.S. guestimate (Lebow and Rudd, 2003) and the implications of Statistics Canada's somewhat different methods are similarly overlooked by Rossiter (2005).

Boskin et al. (1996:32) contend that generating a positive number as the quality change bias estimate is better than accepting zero, because zero "represents an extreme one-sided answer to the question as to whether the components of the CPI subject to relatively little research are biased." The analysis in this section, however, shows that when quality is improving, the lower bound on the quality error that is incorporated into the index is not zero. Conversely, when

_quality is deteriorating, the upper bound of the quality change error is not zero. When one has no empirical information on the magnitude of the quality change error, one also does not know its sign. That is, $\Delta QI > 0$ may imply $\beta > 0$ or $\beta < 0$, and similarly $\Delta QI < 0$ may imply $\beta > 0$ or $\beta < 0$. Accordingly, zero may be the best point estimate.

Extrapolations are not always invalid. In some circumstances, one has information on related products or related sampling conditions that may facilitate estimating a bound. For example, studies of electronics have overwhelmingly shown that prices of those products are falling. One might well extrapolate falling price indexes to electronic products that have not been studied, on the grounds that existing studies have also shown the reasons why the conventional linking techniques have missed price decline. Or, one might be willing to extrapolate results from one study to a class of related products that are thought to exhibit similar behavior, or similar pricing and measurement problems.

The point is: One needs a study to extrapolate from. Without some empirical work to base them on, guestimates have little value.¹⁰

What the price index agencies do to adjust for the quality changes they encounter (“inside the sample” quality change) is but one aspect of the more general problem of quality change. For example, bias may occur because the sampling methodology systematically misses the varieties of products on which quality change occurs (“outside the sample” price and quality change, such as in Berndt, Griliches and Rosset (1993), and other research questions exist. But guestimates do not address them adequately, either.

The Commission on CPI Basic Components¹¹

Examples of CPI basic components include musical instruments in the entertainment category and bananas in the food indexes. CPI basic components are formed from microdata collected from matched retail outlets. Two measurement problems have dominated recent discussion in the U.S. — what the Commission called “lower level substitution bias” and the related question that the BLS called “formula bias.”

Formula bias was caused by procedures adopted in the 1978 revision of the CPI to implement probability sampling for varieties and outlets, as recommended by the Stigler committee.¹² Before 1978, the (unweighted) formula for a CPI basic component was: $(\sum p_{i,t+1} / n) / (\sum p_{i,t} / n)$, where t is the last price observation (month for monthly pricing), $t+1$ the current month, and the calculation was done separately for each city. In words, then, the basic component index for, say, refrigerators in the San Francisco area was the change in the average price of a matched sample of refrigerators in this city in the two months. In the index number literature, this is known as the ratio of average prices method, or RAP. Note that it is a form of arithmetic mean.

After 1978, BLS estimated outlet sampling probabilities from a new “Point of Purchase survey” and item selection was done by probability methods within each retail outlet. However, random sampling of varieties and quality levels precluded continued use of averages of prices. If a piano and a guitar pick are probability selec-

10 Lebow and Rudd (2003) present a table that lists CPI components, their estimated biases, and the sources for the authors’ bias estimates. Components for which they believe they have good research estimates amount to only 7 per cent of the CPI and contribute only 0.02 points to their total bias estimate. Components for which they have no research whatever account for over half the CPI, and they themselves mark these estimates as “almost entirely subjective.” They remark that this table is “sobering;” it ought to have been discouraging.

11 Portions of this section are drawn from Section II.C in Reinsdorf and Triplett (2004), which also reviews related topics in reports of the Mitchell, Stigler and Committee on National Statistics panels, in 1945, 1961, and 2002, respectively.

12 Probability sampling of outlets, but not varieties, began with the 1964 revision of the CPI.

tions for musical instruments, an example presented by Moulton and Moses (1997), the average price movement would be dominated by the piano. This suggested changing the computation for basic component indexes from ratios of average prices (or RAP), used before 1978, to averages of price ratios (APR), used thereafter. The simple APR discussed by McCarthy (1961) was an unbiased sample estimator of a Laspeyres price index **if** the specific items priced were selected with probabilities proportional to expenditures in the base period and price collection began in the base period.

However, neither McCarthy (1961), nor Adelman (1958), nor Westat (the statistical consulting firm brought in to design the BLS move to probability sampling) considered problems that arose in practical application. The BLS produced some long and complicated analyses that focused on the expectation of the APR sample estimator with respect to a Laspeyres population parameter, and how the particular implementation of that estimator was biased because the perfect measurement of base period expenditures and prices assumed by McCarthy is far from achievable in practice (Reinsdorf, 1998; Reinsdorf and Moulton, 1997). BLS called the resulting bias “formula bias,” and proposed a solution (they called it “seasoning”) that attenuated the bias.

Additionally, Reinsdorf (1993) showed that use of a geometric mean as the basic component aggregator in the CPI gave lower rates of price change than the ARP that the BLS had been using. As soon as “geometric-arithmetic” appeared in the discussion, some economists seized on the apparent parallel between a Laspeyres index (weighted arithmetic mean) and a superlative index (weighted geometric mean or a geometric combination of arithmetic means) to raise the well-known Konüs substitution bias as an explanation for the geo-arithmetic mean dif-

ference at the basic component level. Before the dust had settled, the Commission appeared on the scene and took up this “lower level substitution bias” explanation for the difference between APR and geometric mean indexes. The Commission proposed that the BLS use the geometric mean to bring basic components more closely in line with COLI theory, a recommendation that the BLS accepted for most components of the CPI.

What Economic Behavior Applies?

Economists have sometimes interpreted the difference between arithmetic mean and geometric mean aggregators for basic components as just the classic substitution bias paradigm drawn from Konüs (1925), only applied one level down. In this context, much (probably too much) has been made of the fact that Cobb-Douglas behavior justifies a (weighted) geometric mean price index — that is, a weighted geometric index is a COLI if the elasticity of substitution is everywhere equal to unity. The question concerns unweighted means, not weighted indexes.

It is a mere mechanical fact that an unweighted arithmetic mean of positive quantities will be greater than an unweighted geometric mean. The difference between the two is not evidence of substitution bias. No inference about lower level substitution can be drawn from the fact that the geometric mean basic component gives a lower estimate of price change than the arithmetic mean, since that will always be the case, substitution or no. The fixed-weight Laspeyres index, on the other hand, only exceeds the COLI when commodity substitution takes place and will equal the COLI when substitution is zero.

Empirically, geo-arithmetic mean divergences have usually been associated with “price bouncing,” the periodic if not monthly sales that dominate some sections of retailing. Schultz (1994),¹³ though not the first study of

13 An earlier paper by Carruthers, Sellwood and Ward (1980) influenced Europeans more than North Americans, one example of numerous trans-Atlantic divides in price statistics.

its kind, brought the formula for basic components to attention in North America. He reported huge variances from different formulas applied to microdata from Statistics Canada's CPI. His subject was a single brand and size of soft drink in a single Canadian city — clearly not a *commodity* substitution story of the classic Konüs kind. Prices for this soft drink fluctuated, often dramatically, from month to month in a single retail outlet. When prices returned to normal from the sales prices, the increases dominated the movements in an arithmetic mean but less so in a geometric mean, even though sales weights were not available,.

This example and others suggest that the arithmetic mean-geometric mean differences that dominated the discussion at the time the Commission was preparing its report were not, as the Commission supposed, evidence of commodity substitution at the lower level (Granny Smith for American Delicious apples was a favorite example). To be sure, substitution must take place there. But the ubiquity of the difference even when substitution was unlikely to provide the full explanation suggests looking beyond lower level substitution.

The Commission was evidently unimpressed with the BLS analysis of formula bias, the bias of the estimator with respect to the population Laspeyres index. How highly does one rank exact computation of a Laspeyres index if the objective is to produce a COLI? I agree with the Commission that the COLI is the way to think about the problem of CPI basic components.

Shopping behavior vs. substitution behavior in a COLI

Commodity substitution behavior is undoubtedly one relevant concern for basic components since many components are made up from samples of substitutable commodities. When con-

sumers substitute in response to relative price changes within a basic component, then the geometric index is a better approximation to the COLI, and that is all that needs to be said.

A theory of basic components, however, must be applicable to all basic components, not just some of them. It must explain differences between arithmetic and geometric means for components such as the CPI banana price index, or the Schultz (1994) study of a single brand of soft drink, cited above. Surely there is no room for commodity substitution within a single size and brand of one product.

When soft drinks go on sale, consumers do not necessarily consume more of them (as the theory of commodity substitution has it), they stock up and store the soft drinks. The standard model of commodity substitution — though clearly relevant to the construction of basic components — must be supplemented by a model of consumer search with costly and imperfect information, combined with inventory and storage behavior, as Pollak (1998), Feenstra and Shapiro (2003), and Triplett (2003) contend.¹⁴ The Committee on National Statistics Panel (Schultze and Mackie, 2002:24) called for more research on CPI basic components because: “Consumer responses to price differences may reflect something other than substitution behavior: for example, a consumer stocks up on particular items when sales occur but does not change the amount of those items purchased per month or per year.”

Search, storage and so forth are not necessarily modeled by simply switching to a superlative index or a geometric index, since the theory that lies behind those index number formulas is not the theory that explains consumer search and storage behavior. To apply to basic component index numbers, a theory of consumer behavior must model consumers' choices across *sellers* of a homogeneous commodity, not just consumers'

14 See Baye (1985), Anglin and Baye (1987), and Reinsdorf (1993).

choices across different (substitutable) commodities. A geometric mean of matched sellers prices is not in any way related to household search and inventory behavior, nor does it effectively incorporate it into the CPI.

The CPI collection strategy implies that the price frontiers faced by individual consumer units can be calculated from *prices collected from matched retail outlets*. This is demonstrably wrong. Indeed, Triplett (2003) presents a simple numerical example to show that with an imputation for search costs, no standard price index formula applied to prices collected from matched retail outlets will measure the COLI of households who shop. Hendel and Nevo (2002) show that neglect of consumer storage and shopping behavior results in an over-estimate of ordinary demand elasticities, surely a fatal problem if one proposes to model index number substitution bias at the basic component level with a simple Konüs system.

In some cases characterized by shopping behavior, a unit value index may perhaps be justified. If the average price paid drops because information has become easier to obtain, so that consumers are better able to find the lowest prices, the COLI should decline, even though no price has changed. On the other hand, Bradley (2005) shows deficiencies of unit value indexes for almost every purpose, so more analysis of the shopping/search behavior question will be required.

BLS Change to Geometric Means

Implementing the recommendation of the Commission, the BLS changed most (but not all) components of the CPI to the geometric mean formula in 1999. Their unpublished study in support of the change analyzed commodity substitution at a detailed level, thus showing

BLS acceptance of the commodity substitution paradigm at the basic component level, as proposed by the Commission. The unpublished study was exemplary. But the difference between geometric and arithmetic mean indexes for commodities such as bananas cannot, as we have noted above, be explained by commodity substitution. The study was incomplete.¹⁵

A crucial issue is the source of the price data. Laspeyres and Paasche indexes computed from matched sellers prices are not necessarily bounds on the index needed for a COLI in the presence of costly information, because consumers switch between sellers. Studies of commodity search and shopping behavior and also consumer storage behavior (stocking up at sales price time) will need data collected from households — that is, purchasers' prices — not just data obtained from matched sellers.

Attempts to fit the basic component problem into the standard Konüs commodity substitution model lack insight into the nature of the problem and have yielded misleading conclusions. Pollak (1998:73) put it well: "I argue against the view of the Boskin Commission and Diewert (1995) that the 'elementary aggregate' problem, which the Commission calls 'lower level substitution bias,' is primarily a problem of choosing an appropriate formula for combining the prices of items." Both the Commission and the BLS rushed to judgment on the geometric mean. Likely, the geometric mean will prove better than the arithmetic mean of price relatives, for statistical reasons, if no other. However, COLI theory supports neither the Commission's recommendation nor the BLS action: More research is required to determine why the geometric mean is a better measure, if it is one.

15 BLS found no studies that could be used to analyze commodity substitution within the individual basic components, so the analysis proceeded at a higher level, e.g., fruit (a first-level aggregation), not bananas and citrus fruits, which are basic components in this case. BLS accepted the geometric mean for a basic component when substitution was non-negligible at the first-level aggregate, but retained the arithmetic mean when less substitution was found (e.g., medical items). Strictly, the BLS findings supported abandoning the Laspeyres index for the first-level aggregate in the CPI — combining apples and oranges, say, into an index for fruit. It did not support the decision to change the formula for the basic components to the geometric mean.

A Neglected Recommendation: Classification of CPI Commodities.

The Commission's eighth recommendation to the BLS reads: "The BLS should investigate the impact of classification, that is item group definition and structure..." So far as I can determine, this seems to have brought no comment elsewhere and no response of any kind from BLS. The Commission's recommendation is a good one, and BLS ought to take it seriously.

Classifications are the kernels of economic statistics, they group economic data into the units that economists use for their analysis. Classifications — that is to say, groupings — are vital to economic analysis because the way that data are grouped limits the analyses that can be done with them. It is accordingly surprising how little attention economists pay to their classifications — unlike biologists, who understand that classifications need to be done according to a theory.

The theory of economic classifications is the economic theory of aggregation. For the CPI, the relevant research is small, the main items being Pollak (1975), Blackorby and Russell (1978), and Triplett (1990). The theory says that CPI groupings ("fruit," or the higher level aggregate "food") depend on separability of the direct or indirect utility function. Obviously, the theory is hard to implement empirically, but other theoretical abstractions (substitution bias) are now tractable, and so too might this one become, with sufficient work.

As the Commission suggested, better groupings might facilitate analysis of lower level substitution. Regrettably, the BLS has recently implemented a new CPI classification that makes no reference to economic theory. As well, the international agencies seem committed to atheoretical approaches, as demonstrated in the new international price

index manuals. More work needs to be done on the theory, and BLS (and others) should try to work out implementations that are more consistent with the theory, which will be a challenging task.

Conclusion (and Politics)

The Commission Report has been very influential and many of its conclusions stand up to research in the intervening period. Sections three and four of this paper suggest that its reviews of quality change bias and of methods for estimating CPI basic components would need rethinking if the Commission were writing its report now. But I suspect that members of the Commission did not believe they were writing a timeless document. Moreover, they are not to blame if their imitators at making guestimates have not always followed best practice (Section two).

Greg Mankiw, in commenting on the paper by Shapiro and Wilcox (1996:154), "expressed the view that the current debate about the CPI was really a political debate about how, and by how much, to cut real entitlements." Quite so.

The Commission wrote a technical document about CPI measurement. All of the members, no doubt, were aware why Congress was so interested in arcane details of the CPI: Senator Moynihan's complaints about deficits that stretched "as far as the eye can see" were not made in a private setting. Being aware of the political setting is not the same thing as being a captive of it, though the Commission Report did repeat much of the rhetoric of the time, and indeed succumbed to the lure of political statements in its choice of language to describe the effect of CPI measurement errors on Social Security expenditures.

In the debate after release of the Commission Report, political charges were made. It should have been possible to have a debate on the merits of the Commission's report on the CPI that was separate from a debate on whether the depen-

dent part of the population was or was not getting a fair share (whatever that means, which is part of the debate) of the social product. Professionals, at any rate, should understand that improving the accuracy of the CPI is not the same thing as improving the basis for the allocation to the dependent population, and not even the same thing as improving the basis for escalation of their Social Security payments. Zvi Griliches made that point in his testimony before the Senate Finance committee, and it is perhaps instructive that his wisdom on this won few adherents (but see the well stated passages in the Berndt and Baily papers in this symposium).

Many criticisms of the Commission's report were motivated by the belief that benefits to Social Security recipients should not be cut, rather than balanced judgments about the CPI or about the Commission Report. To be sure, parts of the report could be criticized by non-professionals, and that contributed to the politics of the debate: One real estate agent was quoted in the press at the time to the effect that (put in economists' language, not the language the agent used) hedonic functions for housing did not go through the origin.¹⁶

It was inevitable. Congress was not so much interested in the CPI as such, but rather (as one Congressman candidly put it) in finding a way to get BLS to cut the deficit so they did not have to vote on it. I was reminded of an earlier (1981) controversy on CPI escalation of Social Security benefits: In the midst of that one, Senator Goldwater introduced a bill that would have prevented any change in CPI methods if the effect was to lower Social Security benefits. A mix of politics and statistics seldom produces an outcome that is favorable to economic statistics. I have seen many political debates about economic statistics: They are always debates about

something, other than statistics, that is under the table.

We still need discussion of CPI methods. But we need even more a debate on principles for allocation of resources to the dependent population, not only Social Security payments but also Medicare. The two debates need to be conducted separately. The debate ten years ago on the CPI effectively served to thwart the debate on the more important issue.

References

- Adelman, Irma (1958) "A New Approach to the Construction of Index Numbers," *Review of Economics and Statistics* 40, pp. 240-9.
- Advisory Commission to Study the Consumer Price Index (1996) See Boskin et al., 1996.
- Aizcorbe, Ana M. and Patrick C. Jackman (1993) "The Commodity Substitution Effect in CPI Data, 1982-91: Anatomy of Price Change," *Monthly Labor Review* 116:12, December, pp. 25-33.
- Anglin, Paul M., and Michael R. Baye (1987) "Information, Multiperiod Search, and Cost-of-Living Index Theory," *Journal of Political Economy*, 95(6), December, pp. 1179-95.
- Armknacht, Paul and Donald Weyback (1989) "Adjustments for Quality Change in the U.S. Consumer Price Index," *Journal of Official Statistics* 5:2, pp. 107-23.
- Baye, Michael R. (1985) "Price Dispersion and Functional Price Indices," *Econometrica* 53:1, January, pp. 213-223.
- 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*, 11(3), July, pp. 251-64.
- Blackorby, Charles and Robert R. Russell (1978) "Indices and Subindices of the Cost-of-Living and the Standard of Living," *International Economic Review* 19:1, February, pp. 229-240.
- Blow, Laura and Ian Crawford (2001) "The Cost of Living with the RPI: Substitution Bias in the UK Retail Prices Index," *Economic Journal*, 111(472), June, pp. F357-82.
- Boskin, Michael J., Ellen R. Dulberger, Robert J. Gordon, Zvi Griliches, and Dale Jorgenson (1996) "Toward a More Accurate Measure of the Cost of Living," Final Report to the Senate

16 The agent was criticising the Commission's price per square foot measure in its discussion of housing. "Price per square foot goes down as square footage increases." Bette Gorman, as quoted by John M. Berry, Washington Post, December 19, 1996, page E10.

- Finance Committee, Advisory Commission to Study the Consumer Price Index, December 4, <http://www.ssa.gov/history/reports/boskin-rpt.html>.
- Boskin, Michael (2005) "Causes and Consequences of Bias in the Consumer Price Index as a Measure of the Cost of Living," *Atlantic Economic Journal* 33, pp. 1-13.
- Bradley, Ralph (2005) "Pitfalls of Using Unit Values as a Price Measure or a Price Index," *Journal of Economic and Social Measurement* 30 (1), pp. 39-61.
- Braithwait, Steven D (1980) "The Substitution Bias of the Laspeyres Price Index: An Analysis Using Estimated Cost-of-Living Indexes," *American Economic Review* 70(1), March, pp. 64-77.
- Camba-Mendez, Gonzalo, Vitor Gaspar, and Mark Wynne (2001) *Measurement Issues in European Consumer Price Indices and the Conceptual Framework of the HICP* (Frankfort and London: European Central Bank and Centre for Economic Policy Research).
- Carruthers, A.G., D.J. Sellwood, and P.W. Ward (1980) "Recent Developments in the Retail Prices Index," *The Statistician* Vol 29, No. 1, pp. 1-32.
- Christensen, Laurits R. and Marilyn E. Manser (1976) "Cost-of-Living Indexes and Price Indexes for U.S. Meat and Produce, 1947-1971," In Nestor E. Terleckyj, ed. *Household Production and Consumption*, Conference on Research in Income and Wealth, Studies in Income and Wealth 40: 399-446 (New York: Columbia University Press for the National Bureau of Economic Research).
- Covas, Francisco, and Joao Santos Silva (1999) "Outlet Substitution Bias," *Bank of Portugal Economic Bulletin*, September, pp. 77-85.
- Dalén, Jorgen (2002) "Personal Computers in Different HICPs," paper presented at the Brookings Workshop, "Hedonic Price Indexes: Too Fast? Too Slow? Or Just Right?" Washington, D.C., February 1, available on the Brookings Institution website.
- Diewert, W. Erwin (1995) "Axiomatic and Economic Approaches to Elementary Price Indexes," NBER Working Paper No. 5104, May (Cambridge, MA: National Bureau of Economic Research).
- Ducharme, Louis Marc, ed. (1997) *Bias in the CPI: Experiences from Five OECD Countries*. Statistics Canada, Analytical Series, Prices Division No. 10, Ottawa: Statistics Canada, Catalogue no. 62F0014MPB, No. 10.
- Feenstra, Robert C. and Matthew D. Shapiro (2003) "High-Frequency Substitution and the Measurement of Price Indexes," in Robert C. Feenstra and Matthew D. Shapiro (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research Studies in Income and Wealth, Vol. 64. (Chicago and London: The University of Chicago Press) pp. 123-46.
- Gordon, Robert J. (2003) "A Century of Downward Bias in the Most Important Component of the CPI: The Case of Rental Shelter, 1914-2003," in Ernst R. Berndt and Charles R. Hulten (eds.), *Hard to Measure Goods and Services: Essays in Honor of Zvi Griliches*, forthcoming CRIW-NBER conference volume.
- Gordon, Robert J. (2004) "Apparel Prices and the Hulten-Bruegel Paradox," in Carol Corrado, W. Erwin Diewert, and Charles R. Hulten, (eds.), *Price Index Concepts and Measurement*, forthcoming CRIW-NBER conference volume.
- Hendel, Igal and Aviv Nevo (2002) "Sales and Consumer Inventory Behavior," National Bureau of Economic Research Working paper 9048.
- Hill, Peter (1997) "The Measurement of Inflation and Changes in the Cost of Living," prepared for the Conference of European Statisticians, Joint ECE/ILO Meeting on Consumer Price Indices, Geneva, November 24-27.
- Hoffmann, Johannes (1998) "Problems of Inflation Measurement in Germany," Discussion Paper 1/98, Economic Research Group of the Deutsche Bundesbank (Frankfurt, Germany: Deutsche Bundesbank).
- Konüs, A.A. (1925) Translated into English and published as "The Problem of the True Index of the Cost of Living," *Econometrica* 7, January 1939, pp. 10-29.
- Krone, Theodore M., Leonard I. Nakamura, and Richard Voith (2002) "The CPI for Rents: A History," Brookings Program on Economic Measurement, paper available on the Brookings Institution Website.
- Lebow, D. E., Roberts, J. M., and Stockton, D. J. (1994) "Monetary Policy and 'The Price Level'," Washington D.C., Board of Governors of the Federal Reserve System.
- Lebow, David E. and Jeremy B. Rudd (2003) "Measurement Error in the Consumer Price Index: Where Do We Stand?" *Journal of Economic Literature* (43), pp. 157-201.
- Liegey, Paul R. Jr (1993) "Adjusting Apparel Indexes in the Consumer Price Index for Quality Differences," in Murray F. Foss, Marilyn E. Manser, and Allen Young (eds.), *Price Measurements and Their Uses*, conference on Research in Income and Wealth, Studies in Income and Wealth 57: 209-26 (Chicago: University of Chicago Press for the National Bureau of Economic Research).
- Lequillier, Francois (1997) "Does the French Consumer Price Index Overstate Inflation?" Chapter 3 in Ducharme (ed.).

- Manser, Marilyn E (1975) "A Note on Cost-of-Living Indexes and Price Indexes for U.S. Food Consumption," BLS Working Paper 57 (Washington, D.C.: U.S. Bureau of Labor Statistics).
- Manser, Marilyn E. and Richard J. McDonald (1988) "An Analysis of Substitution Bias in Measuring Inflation, 1959-85," *Econometrica*, July, pp. 909-30.
- McCarthy, Jonathan and Richard Peach (2002) "Housing Trends in the 1990s: Measurement in the CPI," Brookings Program on Economic Measurement, paper available on the Brookings Institution website.
- McCarthy, Phillip J. (1961) "Sampling Considerations in the Construction of Price Indexes with Particular Reference to the United States Consumer Price Index," Staff Paper number 4 in Price Statistics Review Committee.
- Moulton, Brent R. and Karin E. Moses (1997) "Addressing the Quality Change Issue in the Consumer Price Index," *Brookings Papers on Economic Activity* (1), pp. 305-49.
- Organisation for Economic Co-operation and Development (1997) "Synthesis Paper on Shortcomings of the Consumer Price Index Measure of Inflation for Economic Policy Purposes," document number ECO/CPE/WP1(97)12, (Paris: OECD, Economics Department) September.
- Oulton, Nicholas (1995) "Do UK Price Indexes Overstate Inflation?" *National Institute Economic Review* (152), pp. 60-75.
- Pollak, Robert A. (1975) "Subindexes of the Cost of living," *International Economic Review*, 16(1) February, pp. 135-50. Reprinted in Robert A. Pollak, *The Theory of the Cost-of-Living Index*, Chapter 2, pp. 53-70, 1989 (Oxford, U.K.: Oxford University Press).
- Pollak, Robert A (1998) "The Consumer Price Index: A Research Agenda and Three Proposals," *Journal of Economic Perspectives* 12(1), pp. 69-78, Winter.
- Price Statistics Review Committee (1961) *The Price Statistics of the Federal Government*, U.S. Congress, Joint Economic Committee, Government Price Statistics, Hearings, Part 1. 87th Congress, 1st Session, 1961; also published as: National Bureau of Economic Research, General Series, Number 73.
- Randolph, William C (1988) "Housing Depreciation and Aging Bias in the Consumer Price Index," *Journal of Business and Economic Statistics*, July, pp. 359-71.
- Randolph, William C. (1988) "Housing Depreciation and Aging Bias in the U.S. Consumer Price Index," *Journal of Economic and Business Statistics*. 6-3, pp. 241-246, May.
- Reinsdorf, Marshall (1993) "The Effect of Outlet Price Differentials on the U.S. Consumer Price Index," In Murray F. Foss, Marilyn E. Manser, and Allen Young (eds.), *Price Measurements and Their Uses*. Conference on Research in Income and Wealth, Studies in Income and Wealth 57: 227-60 (University of Chicago Press for the National Bureau of Economic Research).
- Reinsdorf, Marshall B. (1998) "Formula Bias and Within-Stratum Substitution Bias in the US CPI," *Review of Economics and Statistics* 80, no. 2., May, pp. 175-87.
- Reinsdorf, Marshall and Brent R. Moulton (1997) "The Construction of Basic Components of Cost-of-Living Indexes," in *The Economics of New Goods*, Timothy F. Bresnahan and Robert J. Gordon (eds.) (University of Chicago Press) pp. 397-423.
- Reinsdorf, Marshall and Jack E. Triplett (2004) "A Review of Reviews: Eighty Years of Professional Thinking About the CPI," presented at the NBER-CRIW conference on Price Measurement, Vancouver, June.
- Ribe, Martin (2002) "Quality Adjustment (QA) for New Cars in Austria and Sweden," paper presented at the Brookings Workshop, "Hedonic Price Indexes: Too Fast? Too Slow? Or Just Right?" Washington, D.C., February 1, available on the Brookings Institution website.
- Rossiter, James (2005) "Measurement Bias in the Canadian Consumer Price Index," Ottawa: Bank of Canada, Working Paper 2005-39, December.
- Saglio, A. (1994) "Comparative Changes in Average Price and a Price Index: Two Case Studies," in Statistics Canada (1994).
- Schultz, Bohdan J. (1994) "Choice of Price Index Formulae at the Micro-Aggregation Level: The Canadian Empirical Evidence," in Statistics Canada (1994).
- Schultze, Charles and Christopher Mackie (eds.) (2002) *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes*, panel on Conceptual, Measurement, and Other Statistical Issues in Developing Cost-of-Living Indexes, Committee on National Statistics, National Research Council (Washington, D.C.: National Academy Press).
- Shapiro, Matthew D. and David W. Wilcox (1996) "Mismeasurement in the Consumer Price Index: An Evaluation," in Ben S. Bernanke and Julio Rotemberg (eds.), *NBER Macroeconomics Annual* (Cambridge, MA: MIT Press) pp. 93-142.
- Shiratsuka, Shigenori (1999) "Measurement Errors in Japanese Consumer Price Index," Federal Reserve Bank of Chicago, Working Paper Series (WP-99-2), February (this paper is a summary of a longer work published in Japanese).
- Silver, Mick and Saeed Heravi (2001) "Scanner Data and the Measurement of Inflation," *Economic Journal*, 111(472), June, pp. 383-404.

- Statistics Canada (1994) *International Conference on Price Indices: Papers and Final Report*, First Meeting of the International Working Group on Price Indices, Ottawa, Canada, November.
- Statistics Canada (1995) *The Consumer Price Index Reference Paper*, Ottawa, Canada: Minister of Industry, catalog No. 62-553, July.
- Timmer, Marcel, and Robert Inklaar (2005) "Productivity Differentials in the U.S. and EU Distributive Trade Sector: Statistical Myth or Reality?" Research Memorandum GD-76, Gronigen Growth and Development Centre, University of Gronigen, April.
- Triplett, Jack E. (1971) "Quality Bias in Price Indexes and New Methods of Quality Measurement," in Zvi Griliches, ed., *Price Indexes and Quality Change: Studies in New Methods of Measurement* (Cambridge, MA: Harvard University Press) pp. 190-94.
- Triplett, Jack E. (1990) "The Theory of Industrial and Occupational Classifications and Related Phenomena," 1990 *Annual Research Conference Proceedings*: 9-25. U.S. Bureau of the Census August.
- Triplett, Jack E. (2001) "Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index?" *The Economic Journal* 111, no. 472, June, pp. F311-F334.
- Triplett, Jack E. (2003) "Using Scanner Data in the Consumer Price Indexes: Some Neglected Conceptual Considerations," In Robert C. Feenstra and Matthew D. Shapiro (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research, Studies in Income and Wealth, Vol. 64 (Chicago and London: The University of Chicago Press) pp. 151-162.
- Triplett, Jack E. (2004) *Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes*, Directorate for Science, Technology and Industry, Paris: Organization for Economic Co-operation and Development, STI Working Paper 2004/9, Document DST/DOC2004/9.
- Triplett, Jack E. and Barry P. Bosworth (2004) *Services Productivity in the United States: New Sources of Economic Growth* (Washington, D.C.: Brookings Institution Press).
- Wynne, Mark A. and Fiona Sigalla (1993) "A Survey of Measurement Biases in Price Indexes," *Federal Reserve Bank of Dallas Research Paper* 9340, October.
- Wynne, Mark A. and Diego Rodriguez-Palenzuela (2001) "Measurement Bias in the HICP: What Do We Know and What Do We Need to Know?" paper presented at the CEPR/ECB Workshop on Issues in the Measurement of Price Indices, Frankfurt, November.

The Boskin Commission Report After a Decade: After-life or Requiem?

Ernst R. Berndt¹

*Massachusetts Institute of Technology and
National Bureau of Economic Research*

ISSUES OF PRICE INFLATION measurement and the appropriate uses of Consumer Price Indexes in the U.S. have long been of concern to a relatively small cadre of economists and statisticians, particularly those affiliated with the Conference on Research in Income and Wealth, and the Productivity Program at the National Bureau of Economic Research.² In the early 1990s, however, issues of price measurement again began percolating more widely. The late Zvi Griliches (1997:169) explained this phenomenon as follows:

“Why has the measurement of changes in consumption (and output) prices suddenly become a popular topic? We were a small band, wandering in the wilderness — nobody was listening to us. And now, the measurement issue is attracting attention... The big difference recently has been that the Chairman of the Board of Governors and many politicians have become interested in price measurement, and suddenly it has become a hot political issue. If you can do something ‘to the CPI,’ then you can do something to the growth in entitlements and growth in taxes.”

It is clear that the rise and fall of public interest in price measurement issues, including the Boskin Committee report and its legacy, needs

to be interpreted in the political economy context of Congress and the White House attempting to deal with growing budget deficits in the early to mid-1990s, particularly following the 1994 elections.

It is now a little more than ten years since the Boskin Committee issued its interim report on September 15, 1995, and not quite a decade since its final report was released on December 4, 1996 (Boskin et al., 1995 and 1996). Budget deficits are again looming as a significant political economy issue, and the venerable Chairman Alan Greenspan of the Board of Governors of the Federal Reserve Board has recently been replaced by CEA Chairman Ben Bernanke. But one does not hear much these days about biases in the Consumer Price Index, or about adjusting downward mandated growth in entitlement program benefits due to an upward biased CPI.

There is but one exception of which I am aware, and that is one that suggests a somewhat more solemn interpretation, albeit an ambiguity on the nature of any bias. Specifically, on April 13, 2005, that esteemed and legendary news outlet, *The Onion* (2005), disclosed that:

“A report released Monday by the Federal Consumer Quality-of-Life Control Board indi-

1 The author is Professor, Sloan School of Management at MIT and Director of the Program on Productivity at the National Bureau of Economic Research. This article is based on a presentation at the 2006 annual meeting of the American Economic Association session entitled “The Boskin Commission After A Decade: Is the CPI Still Biased?” January 6-8. The views and opinions expressed herein are those of the author, and are not necessarily those of any of the institutions with whom he is affiliated. Email: eberndt@mit.edu

2 For an historical overview, see Reinsdorf and Triplett (2005), and the references cited therein; also see Stigler et al. (1961).

cates that the cost of living now outstrips life's benefits for many Americans. 'This is sobering news', said study director Jack Farness. 'For the first time, we have statistical evidence of what we've suspected for the past 40 years: Life really isn't worth living.'"

Putting aside the idiosyncratic Onion pronouncement, why is it that inflation and price index measurement issues are no longer in the news? Did the Boskin Committee lay these issues to rest? Has subsequent research effectively addressed and resolved the issues identified by the Boskin Committee? Has the BLS implemented all the principal Boskin Committee recommendations so that the Boskin Committee is now but an interesting footnote in the political economy history of measurement? Or are Congress, the public and the economic policy community still in stubborn denial concerning the continued importance of an upward bias of about 1.1 percentage points annually in the CPI — perhaps the most memorable conclusion reached by the Boskin Committee?

What has transpired since the Boskin Committee descended from the mountain and tabled their conclusions and recommendations? While much could be and perhaps some day will be written about the legacy of the Boskin Committee, given the constraints of time and space, here I will offer only several observations.

The First Five Years after the Boskin Committee Report: Updates and More Commissions

In the opening paragraph of the Conclusion section in both its interim and final report, the Boskin Committee chose to highlight the budgetary implications of their findings:

"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, overcompensate (on average) for changes in the true cost of living."³

Thus, technical issues regarding the net inaccuracy of the CPI were immediately blended in with issues of bias, redistribution and equity by the Boskin Committee.

Initially, some in Congress responded by seeking legislation that would explicitly trim the rise in benefits and reduce the tax-bracket indexing by some given amount off CPI growth, while others argued that the BLS should get it right in the first place and not accommodate arbitrary judgments about CPI accuracy from outside the agency (see, for example, Black, 1997, and Calmes, 1997). The Boskin Committee itself recommended establishing a rotating expert advisory committee that would periodically recommend the "Inflation Adjustment Factor" that in its judgment would best represent the modification of the CPI rise needed to offset the actual change in the cost of living — a recommendation endorsed by Chairman Greenspan and others (Feldstein, 1997).

While the Boskin Committee report was offered at a time when such a "big fix" was

3 Boskin et al. (1995:26-27); Boskin et al. (1996:87-88). In the 1996 final report, the words "for the foreseeable future" and "tax brackets" replaced the words "for the next few years" and "tax rules" that appeared in the interim report. In both the interim and final report, the same selected words appeared in bold.

politically attractive to some, Committee member Zvi Griliches was particularly vocal in arguing that redistribution and CPI measurement issues ought to be separated. Regarding indexing for the retired elderly on pensions, for example, Griliches advocated that the elderly should share the burden with workers receiving flexible wages, stating that “Compensation arrangements should be based on a price index of the domestic value-added components of the various consumption goods — or perhaps on something like the median wage” (Griliches, 1997:172). Griliches then went on to say:

“It seems kind of backwards to say we are going to reduce the rate of growth of transfers to, for example, a person with paraplegia by 1 per cent per year without asking whether the actual support level is the correct level for him. There is a cowardice in our political system that is very depressing. The real question is: What is the right level of transfers? The rate of escalation may have a first-order impact on the budget, but it is second-order relative to the real issues involved.”⁴

After the dust finally settled on budgetary negotiations in 1997 and Congress passed the Omnibus Balanced Budget Act of 1997 that effectively sidestepped CPI indexing issues, what was the next response to the Boskin Commission? Not surprisingly, the solution was quite straightforward: Form some more committees, who will then issue new reports and updates.

First there was The Conference Board, who formed a Study Group on the CPI which, like that of the Boskin Commission, consisted solely of economists and had a clear bipartisan flavor: Paul W. McCracken and James Tobin served as co-chairs, and Charles R. Hulten,

Marvin Kosters and Robert D. Reischauer as study group members.⁵ The Conference Board 1999 report has received remarkably little attention, although some of its recommendations overlapped with those of the Boskin Committee and the subsequent National Academy of Sciences Panel. Perhaps the relative obscurity of this report reflects in part its inaccessibility — even today it is priced at \$295 US by The Conference Board.⁶

Responding to a request from Senator Moynihan, in June 1999 the U.S. General Accounting Office (GAO) initiated a study identifying methodological changes the BLS made to the CPI since the Boskin Committee issued its final report in December 1996. It was also asked to obtain the opinions of the five former Boskin Committee members on how much of the bias in the CPI remained after changes were implemented by the BLS. In its February 2000 report, the GAO identified seven changes that had been implemented by the BLS, and three that had been announced but not yet implemented (U.S. Government Accounting Office, 2000). The four remaining members from the former Boskin Committee estimated that these BLS changes in CPI measurement had reduced the annual upward CPI bias from 1.1 percentage points to between 0.73 and 0.90 points.⁷

Next, later in 1999, with sponsor funding from the Bureau of Labor Statistics, and after a series of exchanges among the BLS and others regarding the extent to which the CPI should be conceptualized within the framework of the economic theory of the cost-of-living,⁸ the Committee on National Statistics in the Division of Behavioral and Social Sciences and Education at the National Academy of Sciences (NAS) formed

4 Griliches (1997:173). Also see Kuttner (1997) on this point.

5 The Conference Board coordinators of the study were Edgar Fiedler and Gail Fosler (1999).

6 <http://www.conference-board.org/publications/describe.cfm?id=437>, last accessed 1 January 2006.

7 Boskin Committee member Zvi Griliches passed away in late 1999.

8 For one observer’s views of these exchanges on the role of the cost-of-living framework, see Triplett (2001).

a Study Panel of experts “to investigate conceptual, measurement, and other statistical issues in the development of cost-of-living indexes.” (Schultze and Mackie, 2002:2). The study panel was chaired by Charles L. Schultze, with economist Christopher Mackie serving as its Study Director.

The NAS study panel differed from that of the Boskin Committee and Conference Board Study Group in several important respects. First, it was much larger — initially consisting of thirteen experts, whereas both the Boskin Committee and the Conference Board each had only five members. Second, the composition of the NAS panel was considerably more diverse than both the Boskin Committee and the Conference Board, whose members were all economists; the NAS panel included not only economists,⁹ but also sociologist Christopher Jencks, psychologist Norbert Schwarz, and statisticians such as Kirk Wolter and Albert Madansky. Thus the NAS panel was considerably more interdisciplinary and heterogeneous. Even among its economists, the NAS panel had a broader mix of empirical and theoretical members than did the Boskin Committee and The Conference Board Study Group.

Third and related, while the Boskin Committee members were seemingly for the most part in consensus regarding the appropriateness of the cost-of-living framework for evaluating the CPI, and in believing that on net there was an upward CPI bias of around 1.1 percentage points per year over the true cost-of-living, by contrast the NAS panel took nothing for granted, and started from scratch, vigorously arguing at considerable lengths amongst themselves on these and many other issues (e.g., the usefulness of the representa-

tive consumer paradigm, Engel curves and the implications of non-homotheticity for index numbers, and the validity of utility maximization in the context of experimental research findings from psychology, behavioral economics, and behavioral finance). Managing such a diverse group of social scientists and statisticians and coordinating the authorship of a very detailed and thorough report was most challenging. Chairman Charles Schultze merits accolades for his able and professional leadership as well as his persistence, good humor and thoughtfulness.

Fourth and finally, issues of budget deficits and overstated rates of inflation were not very visible in the booming and exuberant U.S. “new” economy of the late 1990s and early 2000s. Hence the absence of a strong political economy “big fix” mandate enabled the NAS panel to operate at a more leisurely and academic pace, deliberating issues in greater depth than did the Boskin Committee.

The NAS panel met numerous times for about two years, and then issued its report in late 2001 (Schultze and Mackie, 2002). Other papers in this symposium may comment in detail on its recommendations regarding the competing frameworks of the Cost of Living Index and the Cost of Goods Index, the appropriate role for hedonic regressions, and how these and other recommendations differed from those of The Conference Board and the Boskin Committee. Here I simply note that the NAS report received considerable publicity and discussion, in spite of the National Academy Press inexcusably making the report available online at that time to potential readers only by downloading it one page at a time, instead of as a single easily readable or downloadable PDF.¹⁰

9 Economists on the NAS panel included Charles Schultze, Ernst R. Berndt, Angus Deaton, W. Erwin Diewert, Claudia Goldin, the late Zvi Griliches, Van doorn Ooms, Robert Pollak and Richard Schmalensee.

10 That policy has since changed somewhat since 2001. Currently the hardback version of *At What Price* is available for \$44.96, the hardback plus a pdf version sells for \$54, a 6.3 mb version of the PDF book sells for \$34, and chapters in pdf format sell for \$5.20 per chapter. Online readers can gain free access to the book, but the format is page-specific, and readers must engage in clumsy commands to read successive pages. See <http://www.nap.edu/catalog/10131.html>, last accessed January 2, 2006.

The Second Five Years After the Boskin Report: Revisionism?

Subsequent to the Boskin Committee report, the BLS has reaffirmed its commitment to operating within the cost-of-living framework, even though the NAS panel had noted that “for many (perhaps even most) purposes, the distinctions (between the cost-of-living and cost-of-goods approaches) are less important than they might seem.”¹¹ For example, in its most recent Handbook of Methods, the BLS states:

“Although the CPI cannot be said to equal a cost-of-living index, the concept of the COLI provides the CPI’s measurement objective and the standard by which we define any bias in the CPI. BLS long has said that it operates within a cost-of-living framework in producing the CPI. That framework has guided, and will continue to guide, operational decisions about the construction of the index.”¹²

Even as the BLS began implementing a number of changes recommended by the Boskin Committee, academic research documented that in some categories, the CPI is likely to have been downward biased, rather than upward biased. This research initially was in response to a well-known paper by Nordhaus (1997b) on the price of light, whose price increase he argued would have been overstated using CPI methods by around 1.4 percentage points per year since about 1800. Hulten (1997) argued that if this bias were true for the overall CPI and constant over time, then the implied standard of living for U.S. households in 1800 would have been implausibly low. Gordon (2004) calculated that had the bias in the overall CPI been 1.4 percentage points annually since 1800, then in 1800 the

median U.S. household would have been able to purchase only 1.3 pounds of potatoes per day, with nothing left over for clothing, shelter or other goods. Hulten’s and Gordon’s argument echoes a point made at least back to Triplett (1971), who pointed out that the CPI bias was likely to be negative for some product categories and positive for others, suggesting that the sign of the overall bias was indeterminate.¹³

Noting that in his earlier research he had documented that for durable goods the CPI had been upward biased between 1948 and 1983, former Boskin Committee member Gordon argued that if one makes the plausible assumption that the CPI for durable goods was upward biased over the entire 20th century, then in order for real income levels to have been plausible in the early 1900s, some other major component of the CPI must have been downward biased (Gordon and vanGoethem, 2005). Gordon (2004) presents persuasive evidence that for apparel, there has indeed been a downward bias in the CPI, due primarily to the inability to link style changes reliably; moreover, for related reasons hedonic pricing methods are unlikely to mitigate this bias. This downward bias may well persist into the present context.

Gordon and vanGoethem (2005) document CPI downward bias for shelter (tenant rent), reflecting in part non-response by tenants who moved just as rents were being raised.¹⁴ Using a variety of data sources, Gordon and vanGoethem estimate that since 1914 and up through the mid-1980s when the BLS implemented a number of changes, the CPI bias was surprisingly consistent at about -1 percentage points per year. Since the mid-1980s, however, the shelter CPI bias is likely to be negligible.

11 Johnson, Reed and Stewart (2005:3). I have spelled out their COLI and COGI abbreviations.

12 U.S. Bureau of Labor Statistics (2005), as quoted in Johnson, Reed and Stewart (2005). An earlier statement is that by Greenlees (2001).

13 This suggests as well that use of the term “bias” may be unnecessarily pejorative, and that a better description of the measurement goal is to assess the “net inaccuracy” of the CPI.

14 This rationale was apparently first developed by Crone, Nakamura and Voith (2003).

Table 1**BLS Estimates of Lower and Upper Level Substitution Methodological Variations**

(Annualized Per cent Changes, December 1999 through December 2004)

| BLS Index/Difference: | CPI-U-XL | CPI-U | C-CPI-U | Lower | Upper | Total |
|--------------------------------|-----------------|--------------|----------------|--------------|--------------|--------------|
| All Items | 2.77 | 2.49 | 2.09 | 0.28 | 0.40 | 0.68 |
| CPI Major Groups: | | | | | | |
| Food and Beverages | 2.9 | 2.6 | 2.3 | 0.3 | 0.3 | 0.6 |
| Housing | 3.0 | 3.0 | 2.8 | 0.0 | 0.2 | 0.2 |
| Apparel | -0.3 | -1.8 | -2.2 | 1.5 | 0.4 | 1.9 |
| Transportation | 2.4 | 2.1 | 1.8 | 0.3 | 0.3 | 0.6 |
| Medical Care | 4.5 | 4.4 | 4.3 | 0.1 | 0.1 | 0.2 |
| Recreation | 1.8 | 1.2 | 0.7 | 0.6 | 0.5 | 1.1 |
| Education/Communication | 2.5 | 1.9 | 0.0 | 0.6 | 1.9 | 2.5 |
| Education | 6.5 | 6.3 | 6.5 | 0.2 | -0.2 | 0.0 |
| Communication | -1.4 | -2.3 | -4.8 | 0.9 | 2.5 | 3.4 |
| Other Goods & Services | 3.5 | 3.2 | 2.8 | 0.3 | 0.4 | 0.7 |
| Special Aggregates: | | | | | | |
| Food | 2.9 | 2.6 | 2.3 | 0.3 | 0.3 | 0.6 |
| Energy | 6.8 | 6.5 | 6.1 | 0.3 | 0.4 | 0.7 |
| All Items Less Food and Energy | 2.4 | 2.1 | 1.7 | 0.3 | 0.4 | 0.7 |
| Commodities & Services | | | | | | |
| Commodities | 1.8 | 1.3 | 0.6 | 0.5 | 0.7 | 1.2 |
| Services | 3.5 | 3.3 | 3.2 | 0.2 | 0.1 | 0.3 |

Source: Johnson, Reed and Stewart (2005:5), Table 1.

In terms of substitution bias, the Boskin Committee estimated the upper level bias to be 0.15 per cent per year, while that at the lower level was 0.25 per cent, yielding a total substitution bias of 0.40 per cent annually. In 1999, the BLS converted the CPI from its traditional Laspeyres-type computation to one using geometric means at the item strata level, thereby altering lower level price computations for about 61 per cent of the index (Johnson, Reed and Stewart, 2005:4). Consistent with recommendations from the Boskin Committee, The Conference Board Study Group, and the NAS Panel, in 2002 the BLS addressed upper level substitution bias by using a Tornqvist formula and expenditure data from both the base and current period in

the upper level aggregation, and publishing the resulting chained CPI index (“C-CPI-U”) as a separate and distinct index from the CPI-U.¹⁵

BLS officials have recently published estimates of the effects of these upper and lower level substitution methodological variations on the growth of the CPI, annualized over the December 1999 — December 2004 time period. The BLS results are reproduced in Table 1, where CPI-U-XL is the (now experimental) traditional Laspeyres index, CPI-U employs geometric mean aggregation at the lower level, C-CPI-U is the experimental chained CPI, Lower is the difference between CPI-U-XL and CPI-U, Upper is the difference between CPI-U and C-CPI-U, and Total is the sum of Lower and Upper.

15 Since expenditure data are only available with a time lag, a geometric means formula is used to estimate the indexes initially, and then the figures are revised when the final expenditure data become available. Another index, based on the old Laspeyres CPI methodology, is now published experimentally, and is dubbed the “CPI-U-XL” index. See Johnson, Reed and Stewart (2005:4-5).

As expected, with easier substitution apparel has a large lower effect, whereas housing and medical care, with little lower level substitution, only have small lower level effects. Overall, the lower level impact is estimated at 0.28 per cent annually, while that for the upper level is larger at 0.40 per cent, yielding a total substitution effect of 0.68 per cent per year. However, BLS notes that while this upper level impact is on average 0.4 per cent annually, between 2000 and 2004 it varied considerably, being much larger in 2000 than in later years at 0.80, 0.28, 0.36, 0.17 and 0.41, respectively, between 2000 and 2004.¹⁶

In summary, regarding substitution bias, while the lower level substitution bias has been mitigated by implementing a geometric means procedure, at the upper level the difference between the experimental chained index and the flagship CPI-U remains considerable at about 0.4 per cent per year, although there is some evidence suggesting a decline in more recent years.

A set of issues receiving considerable attention in the Boskin Committee and especially the NAS Panel report was that of quality adjustment and hedonics. Hedonic methods have been used since February 1988 in the rental housing component of the CPI to make minor adjustments based on the aging of the housing units sampled, and have also been used in the apparel component of the CPI since January 1991. These two adjustments have tended to increase rather than decrease the rate of CPI growth — by 0.39 percentage points per year for apparel compared to previous methods, and by 0.31 points for the affected housing indexes (Johnson, Reed and Stewart, 2005:11).

Between January 1998 and October 2000 the BLS introduced hedonic regression adjustments for eleven product classes, including computers (January 1998), televisions (January 1999), audio

equipment and video cameras (January 2000), VCRs and DVD players (April 2000), refrigerators/freezers, microwave ovens and college textbooks (July 2000), and washers and dryers (October 2000). Excluding housing, the combined weight of the item categories undergoing hedonic price adjustment is about 3.01 per cent, and if one excludes in addition apparel, the weight is but 0.85 per cent. Hence, hedonic adjustment is of relatively minor importance. In some cases the hedonic adjustments increased the CPI (washers), whereas in other cases they decreased it (dryers); excluding personal computers the net effect of hedonic price adjustment on the growth of the all-item CPI has been less than one hundredth of one per cent per year, i.e. 0.005 per cent (Johnson, Reed and Stewart: 9-11).

Personal computers have long been the focus of hedonic price adjustment research.¹⁷ Even as the NAS panel was deliberating, the BLS began considering the impacts of employing hedonic methods in real time for personal computers, such as those developed by Ariel Pakes (2003). However, since September 2003 the BLS has replaced hedonic-based adjustments for personal computers with attribute pricing using web-based specific manufacturer's component cost information to estimate values for model features. According to the BLS,

“The attribute cost adjustment process has a database of 250 to 300 variables/items which are updated monthly. This alternative method for quality adjustments allows for more adjustments to be calculated, as many of the items that change in a PC are not specifically covered in a hedonic model.”¹⁸

The BLS reports that compared to the previous hedonic method, between April 2004 and September 2004 the new attribute method resulted in a slightly higher decline in the PC

16 Johnson, Reed and Stewart (2005:7), Table 2. All items lower level effects were less volatile, though falling over the same time period, from 0.30 per cent in 2000 to 0.23 per cent in 2004.

17 See, for example, Berndt and Rappaport (2001) and the references cited therein.

18 Johnson, Reed and Stewart (2005), p. 12 and footnote 24, p. 16.

index, an annualized rate of -9.78 per cent compared to -8.58 per cent for the hedonic method (Johnson, Reed and Stewart, 2005: 13).

In summary, while a modest amount of hedonic pricing was implemented in the BLS' CPI program since the Boskin Committee report, since 2000 no new categories have undergone hedonic adjustment, and that for personal computers has been dropped, replaced with an attribute approach. While hedonic price adjustments to the CPI have received considerable attention and been the source of much controversy, in fact their contribution to the CPI is currently miniscule.

Finally, regarding updating the market basket, since 2002 the BLS has updated expenditure weights based upon consumer expenditure surveys every two years, considerably more rapidly than the roughly every ten years in the past. The lag time between survey initiation and completion has also been shortened. BLS estimates that for 2004 the increase in the CPI-U was 0.06 percentage points lower than it would have been had the old weights been utilized. Moreover, the outlet sample is now rotated every four rather than five years, resulting in a more up-to-date basket of goods, particularly high tech goods, according to the BLS (Johnson, Reed and Stewart, 2005:13). Just how up-to-date the sampled items are remains somewhat unclear, however. Chwelos, Berndt and Cockburn (2005), for example, have reported that personal digital assistants were not an explicit category within the CPI; moreover, they find that for these PDAs, hedonic adjustments result in substantially higher rates of measured price decline than does the matched model method. During the recent Christmas shopping season, high tech communications equipment integrated with cameras and other personal filing information have become leading-selling electronic items. It would be useful and perhaps reassuring if the BLS periodically announced what new types of

products were being captured in its changing CPI market basket.

Major Unfinished Business: The Medical CPI (and PPI and PCE Deflator?)

The Boskin Committee assigned a medical care CPI bias of 3.0 percentage points annually, noting in particular that in 1995-1996 the PPI (which they interpreted as to some extent taking changing outcomes into account as contrasted with CPI repricing of fixed inputs) had risen about 2.0 percentage points more slowly than the CPI, especially in the physician and hospital categories (Boskin et al., 1996:59).

Three major recommendations of the Boskin Committee were that: (i) the weight assigned medical care should not be based just on consumers' out of pocket medical care and health insurance expenditures, but it should also include employer-financed health insurance contributions, as well as expenses paid by Medicaid and Medicare, thereby approximately doubling the weight of medical care in the CPI, from 7.4 per cent to about 16 per cent (Boskin et al., 1996:58); (ii) that the medical care category should receive a substantial component of the CPI's future research investment; and (iii) "...we strongly endorse a move in the CPI away from the pricing of health care inputs to an attempt to price medical care outcomes." (Boskin et al., 1996:60)

In comparison, the NAS panel recommended:

- that the BLS compile and publish an "expanded scope medical CPI" that included employers' (but not Medicare and Medicaid) health insurance payments;
- the BLS convene a task force in collaboration with the Centers for Medicare and Medicaid Services and others to implement construction and publication of a total medical care expenditure price index, encompassing purchases from all health care

- payers — governments, private third-party insurers, and consumers;
- against the BLS making immediate attempts to adjust medical care expenditures for changes in outcomes quality (mortality, morbidity and quality of life), since the NAS panel members recognized “the formidable measurement challenges and do not know how best to proceed,” as well as the need for “considerably more research, much of it interdisciplinary”; (Schultze and Mackie, 2002:190) and
 - that the BLS consider developing a disease- or diagnosis-based elementary unit for pricing episodes of treatment rather than the current input-based “industry” or medical care strata, initially experimenting with 15 to 40 randomly chosen diagnostic categories drawn from commercially available retrospective medical claims databases (Schultze and Mackie, 2002:188-189).

Regarding this last recommendation, the NAS panel recognized that in 1998 the BLS aggregated inpatient and outpatient hospital services into a single stratum, thereby allowing for substitution between them, but also noted explicitly that hospitals were only one area of potential bias from input substitution (Schultze and Mackie, 2002:188).

How has the BLS responded to these differing, and at times inconsistent, recommendations? First, regarding the use of an expanded scope medical CPI that includes employers’ contributions to employees’ health insurance and medical care costs in weighting the medical care component of the CPI, as far as I can tell this has not received much attention from the BLS. Indeed, a recent presentation by the BLS to the Federal Economic Statistics Advisory Committee (“FESAC”) noted the recommendation, but did not discuss it in any detail, other than pointing out that given the new

2002 weights from the Consumer Expenditure Survey, as of December 2004, the medical care weight was 6.13 per cent, down from the 7.4 per cent cited by the Boskin Committee final report.¹⁹ However, the Bureau of Economic Analysis is examining weight and other medical care price index issues, in part because the Personal Consumption Expenditure component of the implicit Gross Domestic Product price deflator weights medical care by total expenditures, rather than simple out-of-pocket costs (Aizcorbe and Nestoriak, 2005).

Second, regarding inconsistent recommendations from the Boskin Committee and the NAS panel concerning incorporation of changing health care outcomes and other quality adjustment into the medical CPI, as best as I can determine BLS has not done any in-house research in this area, and thus it has implicitly agreed with the NAS panel rather than the Boskin Committee recommendation. Recent research by David Cutler and Rosen (2005), as well as by Daniel Slesnick (2005), is beginning to address these quality adjustment and outcomes issues. Provocative research findings on real output growth and price declines in the mental health sector during the 1990s have also been reported by Berndt, Busch, Frank and Normand (2005).

Third, regarding BLS collaboration with the Centers for Medicare and Medicaid Services, as well as with other agencies, to construct and publish a total medical care expenditure price index, while the opportunities here are considerable, particularly as the Medicare Part D Drug Benefit has now come into effect (on this, see, for example, Platt and Ommaya, 2005), as best I can determine, relatively little has been accomplished to date in this context.

Fourth, concerning experimental episodes-based price indexes based on retrospective claims data, here the BLS has undertaken a

19 Johnson, Velez, and Bradley (2005), slides #16 and 6.

research initiative jointly with Medstat, a commercial retrospective medical claims data base vendor. Using monthly data from January 1998 through December 2002, the researchers utilized an “episode grouper” from Medstat to group claims into disease-treatment episodes, and then randomly sampled 40 conditions using alternatively expenditure vs. population (simple count of number of episodes treated) weights. Medstat claims were drawn from and compared with BLS price quotes from three metropolitan areas — Boston, New York and Philadelphia. Both small samples (the same size as those used by the BLS in collecting data for the CPI) and large samples (about ten times the BLS sample size) were used to create alternative price indexes (Medstat, 2004).

The central finding from this research reported by the authors is that for the most part, while there appear to be very different trends among cities and methods over shorter time periods, after 48 months the cumulative estimated price changes for the various methods are not statistically different. In general and with several exceptions, while point estimates of the disease-based price indexes tend to suggest smaller price index growth after 48 months than does the medical CPI for each of the three cities, the bootstrap-based estimated standard errors are very large (a common characteristic of medical claims data, where a relatively small number of outliers observations can make means relatively volatile), resulting in the inability to reject the null hypothesis of no difference between them at usual p-values (Song et al., 2005).

There are a number of shortcomings to this research, and I have argued elsewhere that the absence of evidence here should not be inter-

preted as evidence of absence (Berndt, 2005). The BLS has acknowledged some of these issues, particularly those involving the large sample variances in both episode prices and utilization across cities and over time;²⁰ much detailed empirical investigative research needs to be done regarding implications for price and quantity measurement of using alternative commercially or publicly available episode groupers.²¹ While these preliminary results are perhaps initially disappointing, it is reassuring to see the BLS commit to carrying out this NAS panel recommendation, and I hope that the BLS, BEA and others will continue to engage fully in this important line of research. Moreover, while the retrospective claims data bases do not in general include the uninsured, it is important that future research also focus on price and quantity measurement of health care received by the uninsured.

One issue not raised by the various committees and panels involves an unforeseen development, and that is adoption of the Medicare Part D Prescription Drug Benefit as of January 1, 2006. For those elderly currently uninsured or receiving only partial prescription drug benefits, the new Medicare Part D benefit will likely result in reduced prices and out-of-pocket payments; for those previously receiving drug benefits from private pension plans, whether out-of-pocket payments and prices are greater or less than pre-January 1, 2006 depends on the nuances of what looks like very heterogeneous benefit designs among various private plans. How the BLS ideally should and will in fact deal with these new out-of-pocket payment regimes, and how they ought and will be linked to previous payment structures, raises very challenging

20 Johnson, Velez and Bradley (2005), especially slides #17-19.

21 On this, see Berndt (2005) and Aizcorbe and Nestoriak (2005), slide #16. I note in passing that while the NAS panel report recommended (Schultze and Mackie (2002:189) that the BLS pay particular attention to the possibility that the retrospective claims based episode treatment price index would “jump” at the linkage point when weights changed (i.e., annually in December — January), this issue was not addressed at all by Song et al. (2005).

new issues, as well as older ones involving non-linear pricing schedules and substitution among insurance plans.²²

A striking feature of the current environment, and a stark reminder of how public discussion of CPI measurement issues emanates from political economy considerations, involves prescription drugs and the elderly. In the 1990s, there was great concern about how uninsured senior citizens were being forced to pay cash prices for prescription drugs, prices that were not benefiting from the growing buying power of managed care organizations, and that as a result prescription drug prices for uninsured senior citizens were rising more rapidly than the prescription drug component of the CPI. Adjusting seniors' entitlement payments downward from growth in the CPI as the Boskin Commission seemed to recommend, it was argued, was therefore truly unfair. Today one does not hear any discussion about the Medicare Part D drug benefit providing a rationale for de-linking entitlement escalation from growth in the CPI, since the prescription drug component of the CPI is now likely to overstate price growth for those now enjoying benefits of Medicare Part D.

Concluding Remarks

Much progress has been made by the BLS over the years in implementing methodological changes that have helped reduce the net inaccuracy of the CPI (and, implicitly, of the PPI as well). While understandably it must be tiring and frustrating for BLS officials to be perpetually accused of biasing upward the CPI, I believe the BLS and its professional staff have generally responded professionally and constructively to recommendations from the various committees and study groups that have released reports

examining the net inaccuracies of the CPI. The academic community, including former commission members, has also demonstrated a willingness to reconsider and perhaps even reverse earlier conclusions regarding CPI bias. Professional confidence in the reliability of the CPI has I believe increased since publication of the Boskin Committee report in 1996.

But much research and hand-to-hand combat with microeconomic data remains to be done, particularly, I believe, with the medical care-related price indexes. While the United States is somewhat unique in the substantial role played by the private sector in paying for medical care, issues of price measurement, cost-effectiveness, and adjusting medical expenditures for changes in quality and health care outcomes pervade all countries — even those that provide health care insurance on a universal basis.²³

References

- Aizcorbe, Ana and Nicole Nestoriak (2005) "PCE Deflators for Medical Care: A Progress Report," Powerpoint presentation to BEA Advisory Committee Meeting, Washington D.C., November 4.
- Berndt, Ernst R. (2005) "Comments on Xue Song, William D. Marder, Robert Houchens, Jonathan E. Conklin and Ralph Bradley, 'Can A Disease-Based Price Index Improve the Estimation of the Medical CPI?'" forthcoming in W. Erwin Diewert, John E. Greenlees and Charles R. Hulten (eds.) *Price Index Concepts and Measurement*, draft manuscript, National Bureau of Economic Research, Conference on Research in Income and Wealth.
- Berndt, Ernst R., Alisa B. Busch, Richard G. Frank and Sharon-Lise Normand (2005) "Real Output in Mental Health Care During the 1990s," Cambridge, MA: National Bureau of Economic Research, Working Paper No. 11557, August.
- Berndt, Ernst R. and Neal J. Rappaport (2001) "Price and Quality of Desktop and Mobile Personal Computers: A Quarter Century Historical Overview," *American Economic Review*, Vol. 91, No. 2, May, pp. 268-273.

22 Some of these issues are discussed in Aizcorbe and Nestoriak (2005), slides #13-15, and in Johnson, Velez and Bradley (2005), slides #12-13 and 23-26.

23 For a discussion linking price index research in health care to cost-effectiveness analyses, see Triplett (1999).

- Black, Chris (1997) "Clinton backs away from CPI panel plan: Congress debates changing inflation index," *Boston Globe*, March 14, pp. D4.
- Boskin, Michael J., Ellen M. Dulberger, Robert J. Gordon, Zvi Griliches and Dale W. Jorgenson (1995) *Toward A More Accurate Measure of the Cost of Living*, Interim Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index (Washington DC: U.S. Government Printing Office) September 15.
- Boskin, Michael J., Ellen M. Dulberger, Robert J. Gordon, Zvi Griliches and Dale W. Jorgenson (1996) *Toward A More Accurate Measure of the Cost of Living*, Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index (Washington DC: U.S. Government Printing Office) December 4.
- Calmes, Jackie (1997) "New Washington Dance: Fix Price Index, Duck Blame," *The Wall Street Journal*, February 25, pp. A24.
- Chwelos, Paul, Ernst R. Berndt and Iain M. Cockburn (2005) "Faster, Smaller, Cheaper: An Hedonic Price Analysis of PDAs," Cambridge, MA, unpublished, June. Revision of "Valuing Mobile Computing: A Preliminary Price Index for Personal Digital Assistants," National Bureau of Economic Research, Working Paper 10746 (Cambridge, MA) September 2004.
- Crone, Theodore M., Leonard I. Nakamura and Richard Voith (2003) "The CPI for Rents: A Revisionist History," paper presented at the Brookings Workshop on Economic Measurement, Washington DC, May 23.
- Cutler, David M. and Allison Rosen (2005) "National Health Accounts: An Overview," Powerpoint presentation (Cambridge, MA: National Bureau of Economic Research) December 9.
- Feldstein, Martin S. (1997) "The Right Way to Adjust for Inflation," *The Wall Street Journal*, Editorial Page, February 26.
- Fiedler, Edgar and Gail D. Foster (1999) *Measuring Prices in a Dynamic Economy: Re-Examining the CPI* (New York: The Conference Board) October.
- Gordon, Robert J. (2004) "Apparel Prices 1914-93 and the Hulten/Bruegel Paradox," paper presented at the CRIW Conference on Price Index Concepts and Measurements, Vancouver, BC, June. Available at <http://faculty-web.at.northwestern/economics/gordon/apparel.pdf>.
- Gordon, Robert J. and Todd vanGoethem (2005) "A Century of Housing Shelter Prices: Is There A Downward Bias in the CPI?," Cambridge, MA: National Bureau of Economic Research, Working Paper No. 11776, November.
- Greenlees, John S. (2001) "The U.S. CPI and the Cost-of-Living Objective," paper prepared for the Conference on European Statisticians Joint ECE/ILO Meeting on Consumer Price Indices, November 2.
- Griliches, Zvi (1997) "Commentary: The Commission Report on the Consumer Price Index," in *Measuring Inflation and Real Growth*, Proceedings of the Twenty-First Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis, *Federal Reserve Bank of St. Louis Review*, Vol. 79, No. 3, May/June, pp. 169-173.
- Hulten, Charles R. (1997) "Comment" on Nordhaus, in Timothy F. Bresnahan and Robert J. Gordon (eds.) *The Economics of New Goods*, Chicago: University of Chicago Press for the National Bureau of Economic Research, pp. 66-70.
- Johnson, David S., Stephen B. Reed and Kenneth J. Stewart (2005) "What Has Happened to Price Measurement Since the Boskin Report?," paper presented at the OECD Conference, "Inflation Measures: Too High – Too Low – Internationally Comparable?" Paris, France, June 21-22.
- Johnson, David S., Frank Velez and Ralph Bradley (2005) "BLS Medical Care Price Indexes: FESAC Presentation," Powerpoint presentation to the Federal Economic Statistics Advisory Committee, Washington D.C., June 10.
- Kuttner, Robert (1997) "The CPI Fix," *Boston Globe*, March 10, pp. A15
- Medstat (2004) *Can Health Care Claims Data Improve the Estimation of the Medical CPI?*, Final Report to the Bureau of Labor Statistics, Santa Barbara, CA, revision dated November 12, 2004.
- Nordhaus, William D. (1997a) "Beyond the CPI: An Augmented Cost of Living Index," *Business Economics*, July, pp. 48-53.
- Nordhaus, William D. (1997b) "Do Real-Output and Real-Wage Measures Capture Reality? The History of Lighting Suggests Not," in Timothy F. Bresnahan and Robert J. Gordon (eds.), *The Economics of New Goods* (Chicago: University of Chicago Press for the National Bureau of Economic Research) pp. 29-66.
- Pakes, Ariel (2003) "A Reconsideration of Hedonic Price Indexes with an Application to PCs," *American Economic Review*, Vol. 93, No. 5, September, pp. 1578-1596.
- Platt, Richard and Alexander Ommaya (2005) "A Beneficial Side Effect of the Medicare Drug Benefit," *New England Journal of Medicine*, Vol. 353, No. 26, December 29, pp. 2742-2743.
- Reinsdorf, Marshall and Jack E. Triplett (2005) "A Review of Reviews: Ninety Years of Professional Thinking About the Consumer Price Index," paper presented at the NBER-CRIW Confer-

- ence on Price Indexes, June 2004, Vancouver, Canada, revised April 2005.
- Schultze, Charles L. and Christopher Mackie (eds.) (2002) *At What Price? Conceptualizing and Measuring Cost-of-Living Indexes*, Panel on Conceptual Measurement, and Other Statistical Issues in Developing Cost-of-Living Indexes, National Research Council, Committee on National Statistics, Division of Behavioral and Social Sciences and Education (Washington DC: National Academy Press).
- Slesnick, Daniel T. (2005) "The Prices of Medical Care Services: Estimates from the MEPS," (Cambridge, MA: Dale Jorgenson Associates) revised August.
- Song, Xue, William D. Marder, Robert Houchens, Jonathan E. Conklin and Ralph Bradley (2005) "Can A Disease-Based Price Index Improve the Estimation of the Medical CPI?," paper presented at the NBER-CRIW Conference on Price Indexes, June 2004, Vancouver, Canada, revised 2005, but undated.
- Stigler, George J., Dorothy S. Brady, Edward Denison, Irving B. Kravis, Philip J. McCarthy, Albert Rees, Richard Ruggles and Boris C. Swerling (1961) *The Price Statistics of the Federal Government*, in U.S. Congress, Joint Economic Committee: Government Price Statistics, Hearings, Part 1, May 1-5, 1961 (Washington DC: U.S. Government Printing Office).
- The Onion (2005) "Cost Of Living Now Outweighs Benefits," Vol. 41, Issue 15, 13 April, available online at <http://www.theonion.com/content/node/30975>, last accessed September 11, 2005.
- Triplett, Jack E. (1971) "Quality Bias in Price Indexes and New Methods of Quality Measurement," in Zvi Griliches (ed.) *Price Indexes and Quality Change: Studies in New Methods of Measurement* (Cambridge, MA: Harvard University Press) pp. 180-214.
- Triplett, Jack E. (1999) "Accounting for Health Care: Integrating Price Index and Cost-Effectiveness Research," in Jack E. Triplett (ed.), *Measuring the Prices of Medical Treatments* (Washington DC: Brookings Institution Press) pp. 220-250.
- Triplett, Jack E. (2001) "Should the Cost-of-Living Index Provide the Conceptual Framework for a Consumer Price Index?," *Economic Journal* 111 (472): June, pp. F311-F334.
- United States Bureau of Labor Statistics (2005) "The Consumer Price Index," ch. 17 in the BLS Handbook of Methods, updated 2005, available online at http://stats.bls.gov/opub/hom/homch17_h.htm.
- United States Government Accounting Office (2000) Update of Boskin Commission's Estimate of Bias, Report to the Ranking Minority Member, Committee on Finance, U.S. Senate, U.S. Government Printing Office, GAO/GGD-00-50, February.

Policy Implications of the Boskin Commission Report

Martin Neil Baily¹
Institute for International Economics

MANY COMMISSION REPORTS LANGUISH on dusty shelves: but not the Boskin Commission Report (Boskin et al. 1996). It received wide publicity because of the importance of the issue, the courageous (or foolhardy) stance taken to come up with a number for the overall estimated bias of the Consumer Price Index (CPI), and because of the distinguished composition of the commission itself. The report was generally well received because it identified significant potential sources of bias in the CPI — the substitution bias, the outlet bias, the new goods bias and the quality change bias. And it suggested values for the likely magnitude of these biases. It was not without critics, however, who noted that the report used a modest number of examples, rather than a comprehensive analysis, as the basis for its conclusion that the overall bias in the CPI overestimated inflation by 1.1 percentage point a year.

One of the virtues of the Boskin Report is that it provoked the Bureau of Labor Statistics to move more quickly in finding improvements to the index. In the past, important new goods had been left out of the index sample for years or even decades. The weights used to construct the index in 1996 were based on consumption weights from 1982-84, even though more up-to-date information had been collected. Credit should be given to Katherine Abraham and the

staff of the BLS for making substantive changes in the index, notably to lessen the substitution bias and make the expenditure weights more up to date.

The Boskin Report suggested that further work be done to provide more robust estimates of the biases in the CPI and, subsequent to the report, it was suggested that a new National Academy of Sciences/National Research Council panel should examine in greater depth the extent of CPI biases. Both wishes came to pass. Several studies have appeared evaluating the sources and magnitude of biases in the CPI (see for example, Shapiro and Wilcox (1997), Hausman and Leibtag (2004) and Nordhaus (1997)). And an NAS/NRC panel chaired by Charles Schultze released its report in 2002 (Schultze and Mackie (2002) and Schultze (2003) for a summary). The Schultze panel was more cautious in its findings than had been the Boskin Commission, concluding that the basis for estimating biases in the CPI remained rather tenuous and that the BLS should not change its collection and estimation procedures substantially (beyond what had already been done) until more research had been completed. In particular, the panel recommended against a major increase in the use of hedonic estimates in the short run. The conservatism of the Schultze panel may have been overstated, however,

¹ The author is Senior Fellow at the Institute for International Economics. From 1998 to 2000 he served as Chair of the U.S. Council of Economic Advisors. This article is based on a presentation to the panel session "The Boskin Commission Report After a Decade: Is the CPI Still Biased?," held at the annual meeting of the American Economic Association, January 6-8, 2006, Boston, Mass. Email: mbaily@iie.com

because it offered support for the approach to hedonics formulated by Ariel Pakes (2002). A main objection to hedonics as currently used is that the coefficients in the hedonic regressions are unstable and may be incorrect for the purpose. Pakes is able to explain why these coefficients are likely to be unstable and could be wrong, and offers a methodology for making better hedonic estimates. The reluctance of the Schultze panel to expand the use of hedonics reflects their agreement with Pakes' findings, while at the same time believing that BLS lacks the resources to apply the Pakes approach widely.

Good and Bad Uses for Back-of-the-Envelope Estimates

Some years ago I talked about economic methodology with a theoretical physicist. At the time, he was working with the Goddard Space Flight Center developing mathematical models that could explain the stellar observations being collected by the space telescope. He had been exposed to some mathematical economics and econometrics presentations and he observed that neither he nor any of his colleagues would present the results of a modeling exercise unless they could also provide a back-of-the-envelope calculation to show that the magnitudes obtained from the complex model were reasonable. The economists he had listened to, on the other hand, never presented such quick estimates. He judged that this was a weakness of economics.

He was being too kind to physicists and too hard on economists. I took undergraduate classes on quantum mechanics and general relativity many years ago and I do not remember back-of-the-envelope estimates being used to make the results of these theories more intuitive. To understand modern physics one has to slog through heavy-duty mathematics because the

material makes no sense at all based on everyday experience. And on the other side, economics is not entirely devoid of intuitive estimates. But, broadly speaking, I agree with his criticism of economics. Too often economists develop elaborate models without exploring whether the results really make sense or are consistent with a range of different data. Good back-of-the-envelope estimates are a valuable tool for economics as well as physics.

I want to offer this support for back-of-the-envelope estimates because the Boskin commission has been criticized for making such an estimate of the bias in the CPI. For me, the problem is not that such an estimate was made, but that this was the *only* estimation method used for the quality adjustment. Specifically, the quality bias in the CPI was evaluated based on intuitive and convincing evidence collected only from narrow areas of the economy. The Nordhaus (1997) study of light is a wonderful and entirely persuasive exploration of the decline in the cost of lighting. The decline in the quality-adjusted price of computers has been extensively explored and the results are consistent with our own experience as users of computers. Robert Gordon (1990) has worked extensively on capital goods prices. But neither a random sample nor a comprehensive set of case studies was available and still is not available. In fact, there was a suspicion that researchers had found commodities or services where the quality change had been very rapid and had not been captured by existing index numbers. The case studies tended to find large upward biases in standard price estimates. (The study by Robert Gordon (2004) of clothing prices is a very welcome exception to that rule.)

In its assessment of the quality bias, the Boskin commission suffered from premature extrapolation — moving too quickly from a limited number of specific examples to a broad conclusion. I do not think it would be necessary to

estimate quality change for every good and service in the CPI to get a good estimate of overall quality change. But it is necessary to draw a broad sample that is representative of the overall composition of the index.

I am not the first to make this point and I am repeating it here because there is a serious danger in policy analysis. As Tversky and Kahneman (1988) have shown, people are heavily influenced, often over influenced, by specific examples they can identify with. Doctors will often give advice based on their own experience with a few patients rather than based on the results of large controlled studies. People quit smoking when a friend gets cancer. Newspaper and TV reports on economic subjects rarely present systematic analysis, instead they interview an anguished worker adversely affected by economic change. A few striking anecdotes play well in Congressional deliberations. The Boskin commission was charged with coming up with an estimate of the overall bias in the CPI and a few examples supported their conclusion on the impact of quality change. Back-of-the-envelope calculations based on concrete examples should probably be used more in economics, but only if such a calculation can capture the broad range of the phenomenon being considered. And, more importantly, only if they are backed up by deeper analysis. Otherwise, economists will end up reinforcing the natural tendency everyone has to believe specific experience and anecdote more than careful analysis.

There was a particular problem in the area of quality change and new goods. Jack Triplett has emphasized over the years that the CPI already accounts for quality change because when a new model is introduced at a higher price, 100 per cent of this price increase is attributed to quality improvement (see his paper in this symposium). Given the evidence currently available, it is not possible to say for sure if there is a significant positive or negative bias in the CPI as a result of qual-

ity change. The Boskin commission should have done more to estimate not only how large is the change in quality over time, but also how large is the quality change that is already being captured in the CPI. Michael Boskin et al. (1998) have argued that even though there is considerable uncertainty around estimates of the amount of quality change, it is better to make an estimate of the amount of such change than simply to assume a zero effect. But that is not a valid argument. The origin is not at zero and the real question is whether we know that current indexes over or underestimate quality change. My own intuition is that the CPI probably does understate quality improvement, but intuition is no substitute for more facts and analysis.

Outlet Substitution Bias

The retail sector in the United States is very dynamic and has contributed substantially to productivity growth over the past ten years. The sector has evolved in three main directions. Discount and warehouse stores like Wal-Mart and Costco offer low prices. Specialty retailers and high-end department stores like Benetton and Saks offer fashion and/or high service levels. And category killers such as Circuit City and Home Depot generally offer low prices also, but compete with large inventories in their particular category and, in some cases, by offering high service levels (advice in hardware stores, for example).

The price collection techniques used by BLS do not capture the benefit to consumers of this consumer-driven evolution of the retailing sector. An impressive study by Hausman and Leibtag (2004) finds that outlet substitution in grocery retailing could account for a significant overstatement of inflation in the food-at-home category of the CPI. And since the evolution of retailing applies much more broadly than just to groceries, this study could potentially be extrapolated to a significant fraction of the CPI. In

addition, the productivity studies of retailing by the McKinsey Global Institute (see Baily and Zitzewitz (2001) for a discussion) suggest that this sector is contributing to overall productivity growth and it seems that this is not being captured in the current CPI measures.

Two cautions are important, however. First, the reason the BLS assumes that buying a product at a discount store is different from buying at a local supermarket or convenience store is that it *is* different. Discount stores are usually further away from where consumers live or work and the level of customer service is low — deliberately so, in order to offer low prices. The broad problem is that *retailing is not just about prices, it is about the level of retail service being provided*. As noted above, the evolution in retailing in the US has not simply been an expansion of low-price discounters. There has also been a massive expansion of higher priced, higher margin retailers. Retailing has evolved to become bipolar. In the grocery area, for example, stores like the Whole Foods/Fresh Fields chain charge breathtakingly high prices for high quality produce and organic cereals. Applying the Hausman and Leibtag methodology to Whole Foods or Saks would conclude that these stores are increasing the price level — and that is not correct. Since in practice his study was applied to study the impact of discounters, it is providing an upper bound on the actual outlet substitution bias because it neglects the lower level of retail service provided at discount stores.

Second, some of the rapid productivity growth measured in US retailing is not real. The most obvious example is computer retailing, where measured productivity increases because of the decline of the quality-adjusted price of computers. Selling the same box is counted as more retail output when it has a more powerful computer inside it.

In summary, there is more than just intuition behind the view that the CPI contains an upward

bias because it does not capture the ability of consumers to cut their cost of living by shopping at more productive and lower priced retailers. But the magnitude of the bias is not known and it is hard to say what the BLS should do to solve the problem. Scanner data may well be helpful, but scanner data provide no way to measure the value of the retailing service being provided, either in high or low service retail formats.

Policy Implications

Better Allocation of Spending on Statistics

One is preaching to the choir in this audience to say that more money should be available to improve the quality of the CPI, or indeed many other economic statistics. It is just crazy that important policy decisions are being made on the basis of inadequate information. We know how disastrous faulty intelligence can be in the defense and foreign policy arenas. Why are we not making more investment in good economic intelligence?

It is not necessary to spend more dollars on economic statistics. At present way too much is allocated for agricultural statistics and other small sectors of the economy. The problem is political. States where agriculture is important are disproportionately represented in the Senate, and Congress generally has not really recognized that the United States basically has a service economy.

In my judgment, it would also be helpful for the United States to create a unified economics statistical agency, like Statistics Canada. This would allow a streamlining of the data collection and analysis process. One counter-argument is that having several different agencies creates competition, increasing performance. That argument looks weak. There is no competitive market among the different agencies. It would be much better to have a single agency with a

system of outside review of the agency's performance. Again the problems here are political. Each cabinet office guards its own turf and protects its statistical arm. We need someone with the clout to reorganize US statistical operations and reallocate the budget in a way that more closely matches the actual economy.

I want to acknowledge here the contribution that Mike Boskin made to improve statistics while Chairman of Council of Economic Advisors (CEA). He worked very hard to increase data quality. I tried pretty hard too in my time at CEA, but I was not as effective as he.

CPI and the Budget

If there are revisions made to the CPI, this has important budgetary effects. The CPI indexes Social Security, federal retirement payments and the federal income tax brackets. Both pension and tax adjustments mean that a lower rate of increase in the CPI generates a lower budget deficit. About two-thirds of the federal budget impact of any change in the CPI comes from Social Security and retirement, while about one-third comes from taxes.

In the 1990s there were bipartisan efforts to reduce the budget deficit and in the aftermath of the Boskin commission there was widespread support for changing the way that federal programs are indexed — CPI less a half percentage point or something like that. In the end this did not happen because, while both political parties could find support for the change, neither wanted to take the lead. Neither the Clinton Administration nor Newt Gingrich wanted to mess with Social Security, the famous third rail of American politics.

Now that we are back in an era of endless budget deficits, there is a search on to find ways to cut spending. So far, I am not aware of serious

efforts to change the way federal pensions or tax brackets are indexed by moving to CPI minus X per cent. But there are proposals to change Social Security, and so modifying the indexing procedure may be put on the table.

Using the CPI to Adjust Social Security Payments

The Social Security retirement program provides the principal source of income for a large fraction of retirees and elderly in the United States. The initial payments made to newly retiring workers are effectively indexed to wages, while for subsequent years the payments are adjusted by the increase in the CPI. I share the view held by Martin Feldstein (2005) and many others that Social Security should provide a basic minimum standard of living to retirees in order that they not be forced into poverty or forced to work into old age (the Turner Commission in the UK reached the same conclusion). It should not provide an excessive amount because that would unduly discourage saving and work, but it should not provide too little and create a class of elderly poor.

In November 2005 the average Social Security retirement benefit was \$962.10 a month; spouses received \$479.60. A married couple where each received these average amounts would receive \$1,441.70 a month. Overall, Social Security represents 39 per cent of income for recipients. However, for 22 per cent of the elderly it is the only source of income.² That is not much to live on in the most populated urban areas of the United States. Of course, persons 65 and over are also eligible for Medicare, which now provides partial drug coverage. And many elderly own their own homes even if they lack financial assets. Nevertheless, the current level of payments does not make an immediate case

2 The figure of 22 per cent comes from the website of the Social Security Administration. T. Lynn Fisher (2005) has pointed out, however, that some other sources of income are excluded in this calculation, most notably drawings from defined contribution pension plans unless they are based on an annuity. This is a potentially serious problem.

that *average* Social Security benefits have been driven up by over-indexing beyond the level that would be seen as a minimum income needed to avoid poverty.

The main impact of CPI indexing on Social Security benefits is on persons who have been collecting benefits for many years. Is the CPI resulting in over-indexing or are the very elderly being squeezed into poverty? One quick way to check that is to look at how benefit levels vary by age. According to the 2005 Annual Statistical Supplement of the Social Security Administration, the average male retiree aged 65-69 received \$1,125.60 in December 2004, compared to \$1,066.60 for those 70-74, \$1,060.90 for those 75-79, \$1,012.30 for those 80-84 and \$1,097.30 for those 85-89. This is not a perfect test of how indexing is playing out. The rich are healthier and live longer than the poor, so those recipients with higher initial benefits are more likely to live to be in their 70s and 80s than those with low earnings in their work history and low initial benefits — creating a mix effect in averages by age. Nevertheless, as a first cut, these figures suggest that indexing with the CPI is resulting in benefit levels that are roughly constant by age of recipient at a point in time.

A similar pattern over time holds for women, but the level of benefits received is lower. There are clear reports of poverty among elderly widows, but this is more to do with the fact that many do not have their own earnings history and receive only a half of their deceased spouse's benefit than because of indexing *per se*. There is a supplementary income program for those retirees who are below the poverty level.

Despite the fact that CPI indexing seems to be doing a pretty good job of keeping the benefits of the very elderly at a level comparable to the level of those just retiring, there has been a lot of interest in indexing among policymakers in Washington DC, driven by a desire to reduce the level of Social Security payments in the years

ahead by reducing the rate of growth of the CPI. This is driven in turn by the fact that the aging of the population will increase the costs of the system, leaving it insolvent at some future time unless taxes are increased or benefits cut. I am not as averse to tax increases as are most Americans, although I am quite averse to increases in payroll taxes, because they distort work incentives and fall heavily on low-wage workers. But even if tax increases were ruled out, it would be possible to find other ways to make social security solvent, for example by giving less to those with strong private pensions or other assets, while giving the same, or even a bit more, to those that rely solely on Social Security.

The Impact of Innovation on the Level of Social Security Benefits

Innovations that reduce the cost of purchasing a given consumption bundle should clearly lower the price escalator that is applied to retirement or other indexed benefits. An issue that has been raised for Social Security indexation is whether the innovations that occur are benefiting retirees to the same degree as the rest of the population. Michael Boskin et al. (1998) respond to that issue by arguing that the consumption basket purchased by seniors shows about the same rate of increase and is as subject to bias as the basket purchased by wage earners. A key point here is that seniors spend a large fraction of their incomes on health care, despite the provision of Medicare. Health care, argues Boskin, is an area where there are very significant biases in the CPI because of new products and treatments and increased quality.

It is worth picking up on this health care issue both because it is such a large fraction of GDP and because it illustrates a more general point. An important issue for indexing Social Security benefits is to determine the extent to which new and improved goods and services represent, on the one hand, a true reduction in the amount of

money needed to maintain a minimum living standard and, on the other hand, the extent to which they represent new opportunities for consumption that are realized only with a higher level of income. In the former case, the minimum level of income is lowered by the innovation. In the latter it is not, and may well be increased.

Health care illustrates both types of innovation. Gall bladder surgery used to require a major operation and a lengthy hospital stay. Some years ago, laparoscopic surgery was introduced that made the operation much simpler, safer and less costly. For anyone paying part or all of the cost of that operation, this innovation represented a decline in the amount of income needed for a given lifestyle. Some new drugs are the other type of innovation. As people age, their various bodily systems show signs of wear and tear. The pharmaceutical industry has found a bevy of drugs that help to keep people functioning better, such as anti-cholesterol drugs or drugs to fight late onset diabetes. Unambiguously, we are better off as a society because of the availability of these new drugs. People would die sooner or be more distressed if they were not available. Do these innovations allow low-income people to live on a smaller income? No. Unless the drugs are fully covered by insurance, people will have to spend more on health care if they are to take advantage of these new products. And there are many other such innovations, notably mobile phones and cable TV. On the anecdotal level, in my own household I pay about \$200 US a month for bundled cable TV, high-speed Internet and land line phone service and about \$120 US a month for mobile phones for my spouse and myself. I am better off because I can consume services that were not available in the past, but anyone living on a low income or solely on Social Security benefits would not be able to take advantage of these opportunities.

I note also that there is a kicker even to an obvious cost-reducing innovation like laparoscopic gall bladder surgery. Most of us get treatment when our doctors tell us this is a good idea. When the cost of gall bladder surgery fell by a factor of 5, the number of such operations rose by a factor of 5. The overall cost of health care did not fall.

In principle, therefore, I am a strong supporter of the view expressed by Zvi Griliches (1996) that the decision about how much money should be paid to Social Security beneficiaries should be separated from the decision about how to measure the rate of change of consumer prices. The dollar income needed for a minimum standard of living would be set by policymakers and would rise with pure “monetary” inflation; it would be held down by price reducing innovations; and it would be increased when new products or services become available that we think the elderly should have access to. By the same principle, cash benefits would be adjusted up or down depending on the generosity of other programs, notably Medicare. It is not clear, under such a system, that retirees should be fully protected against energy price increases when taxpayers are not so protected.

In practice, I do not think our political system is capable of making good objective decisions about Social Security payments on a regular basis. Congress would be unwilling to delegate to a technical group the power to raise benefit payments, given that there are budget implications of that decision. And it would be a nightmare to have to make adjustments every year following Congressional debate and decision. That means we are very much in a second or third-best world. Some form of wage indexing has considerable appeal, as it would keep the minimum living standard in relation to the incomes of those still working. However, this would generally involve larger benefit increases

than under CPI indexing and, given the solvency problems of the Social Security system, such a change looks politically infeasible. *Faute de mieux* I would stick with the current CPI adjustment rather than moving to CPI minus X per cent. The very old are getting about as much as recent retirees, which seems about right. So, like policymakers in the 1990s, I would not try to balance the budget “on the backs of the very old.”

I note that the conclusions I have drawn in this section reflect value judgments about the adequacy of Social Security benefits. Setting the minimum living standard for the elderly is not something that can be done purely on the basis of economics. It is important, however, to recognize that a statement that CPI indexing is over-compensating the elderly should be assessed by looking at how the very old are doing financially.

Other Policy Issues

There are important additional policy issues tied to indexing that deserve extensive treatment. Unfortunately, I can provide only a rather cursory perspective on two of them, the indexing of tax brackets and monetary policy.

Tax Brackets

US federal tax brackets are indexed using the CPI in order to maintain income cutoffs that are constant in real terms. Many of the same issues that were discussed with respect to Social Security indexation are relevant for tax indexation. However, the distinction made above between innovations that cut the cost of a given consumption basket and innovations that increase consumption opportunities seems less important in the tax context. As taxpayers move into higher tax brackets they are receiving higher levels of income and are able to take advantage of the increased consumption opportunities.

Monetary Policy

Maintaining price stability is one of the goals of central bank monetary policies around the world. In the United States, the Federal Reserve has been unwilling so far to say exactly what they mean by price stability or to give an inflation target. In contrast, the European Central Bank has a target of inflation of less than 2 per cent and many other central banks have inflation targets. At the Federal Reserve, Alan Greenspan has been very aware of the potential biases in the CPI and has spoken about their importance. He looks at a variety of price measures, most notably the core price index for personal consumption expenditure. This index, that removes the volatile energy and food components, is measured differently from the CPI and has generally increased at a slower rate than the corresponding core CPI.

Clearly, it would be helpful to monetary policymakers to have better price indexes as they decide whether or not to vary interest rates. But the problem is not pressing. The key issue for the Federal Reserve is to determine whether or not inflation is moving up or down on a sustained basis. In the absence of clear evidence that biases in the CPI vary significantly in the short run or over the cycle, the existing range of price measures are adequate for monetary policy decisions.

The same conclusion is even true when looking at productivity data, which has been a very important issue for monetary policymakers, particularly in the mid-1990s when there was an acceleration of productivity. That acceleration, together with other data, convinced the Federal Reserve that monetary policy did not need to be tightened even though real GDP growth was faster than pre-existing estimates of the economy’s potential growth rate. Accurate price indexes are vital to accurate productivity measurement of course, but again, the key question was whether or not the growth rate of produc-

tivity had increased. To know that, it was important to have price measures that were consistent over time. A constant bias would not change the answer.

This point has been misunderstood. In the 1990s it was sometimes argued that monetary policy should not tighten because biases in price and hence real output and productivity measures caused an understatement of the economy's potential growth rate. As several economists pointed out, however, that is a fallacious argument. If there are biases in the data, then potential growth is indeed understated, but *so is the actual growth rate*. The key question is whether or not actual growth exceeds potential. Increasing both series by, say, one percentage point a year would not change the answer to that question.

As someone who researches productivity, I would love to know the magnitude of quality change in goods and services production. But I cannot claim that this would make a big difference for monetary policy.

Conclusion

It was very important to draw attention to potential and actual biases in the CPI. It was very important to draw attention to the large policy implications of improving the quality of the CPI and other economic series. It was helpful to spur the BLS and other agencies to move more quickly to eliminate problems in their approach to price measurement and to provide these agencies with a reason for Congress to give them more money to accomplish these improvements. Viewed in this context, the Boskin report was a huge success. The fact that they provided a rough overall estimate of the bias was entirely appropriate. If they did not get it exactly correct, that is fine too, especially since they drew attention to the uncertainty involved in their estimate.

The problems came when the report was used as the basis for policy discussion, notably Social Security indexation. Individual commission

members expressed concern over such use of their estimate, but the report itself invited it and indeed this was the context in which Congress requested the report. There was an established method of indexing federal programs and there had to be a clear basis in economic science to change that approach. It would have been better if the Commission had advised Congress that they did not have an adequate scientific basis to recommend a specific quantitative adjustment to the CPI index used to adjust federal programs.

References

- Baily, Martin Neil and Eric Zitzewitz (2001) "Service Sector Productivity Comparisons: Lessons for Measurement," in Charles R. Hulten, Edwin R. Dean and Michael J. Harper (eds.) *New Developments in Productivity Analysis* (Chicago: University of Chicago Press for the NBER).
- Boskin, Michael J., E. Dulberger, R. Gordon, Z. Griliches, and D. Jorgenson (1996) *Toward a More Accurate Measure of the Cost of Living*, Final Report to the Senate Finance Committee, December 4.
- Boskin, Michael J., E. Dulberger, R. Gordon, Z. Griliches, and D. Jorgenson (1998) "Consumer Prices, the Consumer Price Index, and the Cost of Living," *Journal of Economic Perspectives*, Winter, 12(1), pp. 3-26.
- Feldstein, Martin (2005) "Rethinking Social Security," *American Economic Review*, 95, 1, March, pp. 1-24.
- Fisher, T. Lynn (2005) *Measurement of Reliance on Social Security Benefits*, available from Lynn.Fisher@ssa.gov .
- Gordon, Robert J. (1990) *The Measurement of Durable Goods Prices* (Chicago, University of Chicago Press for NBER).
- Gordon, Robert J. (2004) "Apparel Prices and the Hulten/Bruegel Paradox," presented at the CRIW conference on Price Index Concepts and Measurement, June 28-29, forthcoming.
- Griliches, Zvi (1996) "The CPI and Beyond: Issues of Concept and Measurement," Shirley Kallek Memorial Lecture, available at www.census.gov/prod/2/gen/96arc/vgrilich.pdf.
- Hausman, J. and E. Leibtag (2004) "Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart," paper presented at EC2 conference, Marseille, December.
- Nordaus, William D. (1997) "Do Real-Output and Real-Wage Measure Capture Reality? The His-

- tory of Lighting Suggests Not,” in Timothy F. Bresnahan and Robert J. Gordon (eds.) *The Economics of New Goods*. (Chicago: University of Chicago Press for NBER).
- Pakes, Ariel (2002) “A Reconsideration of Hedonic Price Indices with an Application to PCs,” in Jack E. Triplett ed. *Hedonic Price Indexes: Too Fast, Too Slow or Just Right?* Brookings Workshop on Economic Measurement, February 1, available at www.brookings.edu.
- Schultze, Charles and Christopher Mackie, eds. (2002) *At What Price? Conceptualizing and Measuring Cost-of-Living and Price Indexes* (Washington D.C.: National Academy Press).
- Schultze, Charles L. (2003) “The Consumer Price Index: Conceptual Issues and Practical Suggestions,” *Journal of Economic Perspectives*, 17, No. 1, Winter, pp. 3-22.
- Shapiro, Matthew and David Wilcox (1997) “Alternative Strategies for Aggregating Prices to the CPI,” *Federal Reserve Bank of St. Louis Review*, 79, May, pp.113-25.
- Tversky, A. and Kahneman, D. (1988) “Rational Choice and the Framing of Decisions,” in D.E. Bell, H. Raiffa, and A. Tversky (eds.) *Decision Making*, (New York: Cambridge: University Press) pp. 167-192.

Accounting for Growth from A to Z: Review Article on *Information Technology and the American Growth Resurgence*

Daniel E. Sichel¹
Federal Reserve Board

THE PERFORMANCE OF THE U.S. economy has, in many ways, been remarkable in the past decade. After a quarter century of sluggish growth, labour productivity growth accelerated in the mid-1990s and, to the surprise of many analysts, accelerated further after 2001. The mid-1990s pickup has been widely documented and there is a broad consensus that that speedup in growth was importantly driven by information technology (IT).² This growth resurgence and its sources are the focus of the book by Dale Jorgenson, Mun Ho, and Kevin Stiroh (JHS).³ More than just telling that story, however, this book provides a detailed record of the methodology for analyzing economic growth that Dale Jorgenson has developed and advocated over the past several decades; this book is as much about methodology as it is about patterns of growth in the United States and other countries. The other story that is woven throughout the book is the history of economic growth analysis over the past fifty years, along with a recounting of some of the debates around key issues.

In this essay, I will start by reviewing the methodological contributions described in the book and highlight how the book could be relied on as “Users’ Guide” to growth accounting at the aggregate and industry levels. The methodology described by JHS will be familiar to many readers; it was largely adopted by the Bureau of Labor Statistics (BLS) twenty years ago for its estimates of multifactor productivity and has been adopted as the gold standard by many researchers, including the writer of this review. Nevertheless, I will highlight areas where debate continues. I will then turn to the JHS story for how information technology transformed the economy. Again, much of this material will be familiar to many readers, and I will emphasize parts of the story that might be less well known or where I would tell the story a bit differently. Finally, I will discuss the strengths and weaknesses of the book, and comment on the roadmap for future research proposed by JHS.

1 The author is Assistant Director, Division of Research and Statistics at the Federal Reserve Board. The views expressed are those of the author alone and should not be attributed to the Board of Governors of the Federal Reserve or to other members of its staff. I would like to thank Carol Corrado, Steve Oliner, Andrew Sharpe, Larry Slifman and Kevin Stiroh for very helpful comments and suggestions. Email: Dan.Sichel@frb.gov

2 See Jorgenson and Stiroh (2000), Jorgenson, Ho and Stiroh (2002), Oliner and Sichel (2000 and 2002). In these papers, IT refers to computer hardware, software, and communications equipment. This category often also is referred to as information and communications technology, or ICT.

3 *Information Technology and the American Growth Resurgence*, Productivity, Volume 3, MIT Press, 2005, 400 pp., \$50 US.

Using this Book as a Users' Guide to Growth Accounting

From time to time, I am asked to recommend a source that provides a comprehensive description of how to do growth accounting. This book is just such a source.

Aggregate Growth Accounting

Chapter 2 provides an overview of aggregate growth accounting, while chapters 5 and 6 describe the procedures for constructing capital services and labour input at the aggregate level and at the industry level.

A significant chunk of the methodological discussion in chapter 2 describes the basic measurement choices made by JHS (many of which also have been made by the statistical agencies in the United States). For example, JHS discuss the importance of breaking out IT and of using constant-quality prices for information technology and other capital goods. Breaking out IT is essential to gauge its influence on the economy and its role in the growth resurgence. And, constant-quality price indexes are essential for tracking the remarkable development of IT and also for being able to aggregate across vintages of capital. In chapter 1, JHS note where the National Income and Product Accounts (NIPA) in the United States have incorporated constant-quality price indexes; they also point to areas where they suspect that the price measures used in the NIPAs are inadequate.

In terms of measuring output, JHS focus on an output measure that differs from that used by the Bureau of Economic Analysis in the NIPAs. In particular, JHS advocate using an output measure that includes imputations for the service flows from durable goods used by households; for example, rather than include the purchase of an automobile by a household as consumption in the year of purchase, JHS include an imputation for the service flow from this automobile spread out over the life of the automobile. Conceptually, I

am sympathetic to the JHS argument that the service flow from consumer durables should be included in output rather than the purchase of the durable. Indeed, this is the approach used in the NIPAs for housing. However, that conceptual purity does come at a cost: namely, analyses of the sources of growth by JHS must be adjusted in order to be compared with published output measures or with other analyses that rely on published output measures.

A key theme of the book is the importance of using capital services rather than the capital stock as a measure of capital input for growth analysis. Capital services is an aggregate of individual capital stocks weighted by the user cost of each type of capital. The difference in growth rates between capital services and the capital stock is capital quality, which captures the contribution of the changing composition of capital. Details of how to do these calculations are provided in chapter 5.

Another key theme is the importance of using labour input rather than hours as the measure of labour's contribution to growth. This approach divides workers into categories defined by sex, age, education, and employment class (employed, self-employed, etc.) and then weights up hours in each of these cells by wages for that cell. Assuming that wages provide a rough measure of marginal products, this procedure aggregates different types of labour by marginal products. Details of how to do these calculations are provided in chapter 6.

With regard to the discussion of aggregate growth accounting, I suspect that experienced practitioners of conventional growth accounting will find these parts of the book a bit routine. On the other hand, readers new to the subject will find a thorough description of how to do growth accounting along with a discussion of key conceptual and empirical issues confronted in actually doing it. And, it must be recognized, that the methodology described here has not always been

accepted by researchers. Indeed, JHS provide an entertaining discussion at the end of chapter 2 about the evolution of thought about methodologies for growth analysis. As is clear from that discussion, elements of the methodology described by JHS were controversial when first introduced. And, the use of capital stocks (rather than capital services) still occasionally pops up in empirical work as does the use of hours rather than labour input.

One area where disagreement remains is with JHS' use of the production possibility frontier. This frontier describes efficient combinations of inputs and outputs for the economy as a whole and allows multiple inputs and multiple outputs. In contrast, Greenwood, Hercowitz, and Krussell in a series of papers have challenged the approach in JHS, highlighting the role of investment-specific technical change.⁴ For the reasons described in Ho and Stiroh (2001) and Whelan (2003), I am a bit skeptical of the approach taken by Greenwood, Hercowitz, and Krussell. Having said that, this has been an area of active debate and JHS largely brush it off, rather than providing a serious critique in the book.

Industry-level Growth Accounting

For growth accounting at the industry level, a guide is provided by chapter 7 ("Productivity Growth for U.S. Industries"), along with the industry-specific parts of chapters 4, 5 and 6.

Chapter 4 ("The Changing Structure of Output and Intermediate Inputs") provides a detailed description of the methodology and data used for measuring industry output and intermediate input by industry. As discussed, JHS collapse the 192 industries in the input-output tables down to 44 industries, four of which produce IT — computers and office equipment, electronic components, telecommunications equipment, and computer services (which includes software). The remaining 40 industries

are divided into 13 "IT-using" industries and 27 "non-IT" industries. IT-using industries are defined as those for which IT capital services amounted to 15 per cent of total capital services in 1995 at the beginning of the IT-related growth resurgence.

A couple of methodological points are worth noting. First, the analysis is done on the old SIC basis, rather than the newer NAICS basis. The original papers on which the book is based were done with the SIC data so it is easy to see why that is replicated here. However, there are also significant gaps in the availability of the NAICS industry data needed for this type of analysis, a serious shortcoming in the U.S. data system that should be resolved as quickly as possible.

Second, JHS express a strong preference for doing industry-level analysis using gross output and including intermediate inputs as well as capital and labour, rather than using value added and just accounting for capital and labour as inputs. They cite several reasons for preferring a gross output framework. Gross output provides a fuller description of what an industry actually produces (industries produce gross output, not value added). Also, the use of gross output makes it possible to track the role of intermediate inputs; this approach is particularly useful for IT industries where semiconductors are a very important intermediate input. Finally, although industry productivity figures based on value added can be combined easily to obtain aggregate productivity, strong restrictions on the industry production functions must be satisfied for this aggregation to be appropriate. And, as JHS show in chapter 8, these restrictions appear to be violated, raising questions about the validity of simple value added aggregation. On the other hand, gross output measures of labour productivity will be bounced around by swings in intermediate inputs, while value added measures of labour productivity will not. Thus, if a gross output measure is used, it is

⁴ For examples, see Greenwood, Hercowitz, and Krussell (1997 and 2000).

important to keep track of what is happening to intermediate inputs. On balance, I share JHS's preference for doing industry analysis on a gross output basis, whenever data on intermediate inputs is available.

JHS also discuss the various different sources of industry productivity data for the United States, and they discuss the relative merits and shortcomings of their data (derived from BEA and BLS data) compared with the BEA's industry accounts, the industry productivity figures produced by the BLS, and the industry data from the Federal Reserve's Productivity System, as described in Bartelsman and Beaulieu (2004). As JHS note, "The differences among these estimates can be sizable, but it has proven difficult to provide a comprehensive explanation." Gordon (2001) and others have encountered the same difficulties. These apparently inexplicable differences remain a source of continuing frustration for users of U.S. industry data. Although this probably will not happen any time soon, it would be useful for the statistical agencies to push ahead on the long, hard work of reconciling these differences.

As indicated above, JHS define the IT-producing sector as computers and office equipment, electronic components, telecommunications equipment, and computer services. Although JHS are using the narrowest breakdown for which the U.S. data system provides reasonably complete data, a limitation of this breakdown is that each of these industries covers widely different products with very different markets and price dynamics. This broad coverage is, perhaps, a particular problem for electronic components, an industry that is often used as a stand-in for semiconductors. In addition to including the microprocessor and memory chips that go into computers, this industry also covers much less complex products such as resistors and capacitors.⁵ Thus, the elec-

tronic components industry is a bit broad for tracking the role of integrated circuits, particularly those important for computers.

As mentioned earlier, chapter 5 ("Capital Services and Information Technology") provides a full description of the methodology and necessary data for calculating capital services. The material here is fairly standard (although, again, Jorgenson and coauthors were the key developers and early advocates of this methodology). The one place where I would quibble with JHS is with their generally ready acceptance of geometric depreciation patterns. The BLS assumes non-geometric depreciation in their estimates of capital services. Of course, ultimately the pattern of depreciation is an empirical question, and there is disappointingly little work here since the comprehensive and impressive work of Hulten and Wykoff (1981a and 1981b). However, their work predated the IT revolution. Moreover, the empirical evidence that is available for depreciation patterns for IT products point in the direction of non-geometric depreciation. In particular, Oliner (1993) found a non-geometric pattern of depreciation for mainframe computers, and Doms, Dunn, Oliner, and Sichel (2004) and Antonopoulos and Sakellaris (2005) found a non-geometric pattern of depreciation for personal computers.

Chapter 6 ("Labor Input and the Returns to Education") describes the methodology and necessary data for calculating labour input at the aggregate and industry levels. A very important contribution in this chapter is the calculation of labour input and labour quality by industry. A handful of other studies have done that, but most studies of industry productivity have not tackled this because of the heavy data needs and complexities of the calculations. These results add some very interesting pieces to the story of how IT affected the economy in the past decade.

5 Nominal shipments of integrated circuits made up only about 57 per cent of shipments of Electronic Components in 2004.

Information Technology and the mid-1990s Growth Resurgence

As indicated above, JHS also tell the now well-known story of the mid-1990s growth resurgence and IT's role in that growth pickup. In the book, this story is woven through all of the chapters, and is summarized in chapter 1 ("Understanding the Information Age"). The story begins with the concept of "faster, better, cheaper," which describes the progression of semiconductor technology. In the mid-1990s, constant-quality prices of semiconductors started to fall more rapidly than they had in earlier years, leading to especially rapid declines in prices of IT capital goods. Firms responded to these price declines by increasingly substituting capital purchases toward IT capital, generating a surge in IT capital deepening.

In the JHS framework, the particularly rapid declines in semiconductor prices are taken as signaling especially rapid technological progress at producers of semiconductors, and this technological progress shows up as faster multifactor productivity growth in IT-producing industries. When JHS run the numbers, IT-related capital deepening and the IT-related boost to multifactor productivity growth account for an important part of the U.S. growth resurgence; hence, the conclusion that IT made a critical contribution to that resurgence.

JHS extend this story in a couple of directions. Chapter 3 ("Information Technology and Growth in the G7 Countries") presents aggregate growth accounting results for other large industrialized economies, extending through 2001. This analysis uses internationally harmonized prices developed by Schreyer (2000) and links the data across countries using OECD purchasing power parities for 1999. The basic story in the chapter is that the IT revolution is also evident in G7 countries other than the United

States. Although the chapter is chock-full of tables and charts showing cross-country comparisons, it does not say much about why some countries appear to have benefited more than other countries from the IT revolution in the second half of the 1990s.⁶

As already described above, the other direction in which JHS extend the aggregate story is with industry detail, as summarized in chapter 7. Their basic result is that the 1995 revival of productivity growth in the United States was widespread across many industries and that IT played an important role in many of these industries. In their framework, more than three-fourths of the industries posted stronger growth in labour productivity after 1995 than before and about two-thirds posted an acceleration in multifactor productivity. This result has not gone unchallenged, as discussed in the next section.

JHS also highlight the considerable variation across industries in growth rates of labour productivity and multifactor productivity. Although Corrado and Slifman (1999) and others have suggested that persistent declines in the level of productivity in an industry seem implausible and could reflect mismeasurement, JHS argue that it likely reflects real changes in industries as they respond to industry-specific and macroeconomic shocks.

JHS emphasize that most industries responded to the rapid declines in IT prices and shifted toward IT investment. On the labour side, JHS describe a similar evolution within and across industries, summarizing the story for labour input as one of "... rapid changes and reallocation, particularly toward information age industries that either produce or consume information technology most intensively. This expansion of the IT group involved a disproportionate number of young well-educated workers and pays them relatively well." These parts of the book should be broadly comforting to econ-

⁶ For an interesting discussion of these issues, see van Ark and Inklaar (2005).

omists, as they suggest that markets work and firms respond to price and demand signals.

Critiques of the JHS Story

The broad JHS story about the mid-1990s growth resurgence and the role of IT in that pickup generally sounds right to me, and my work with Stephen Oliner has largely reached the same conclusions. However, challenging questions have been raised about aspects of these results. Of course, these challenges apply to most of the work on aggregate growth accounting, including my work with Stephen Oliner. Critiques fall into a handful of categories. One strand expresses general discomfort with the assumptions underlying the neoclassical framework and with the causal interpretations that have been attached to it. Another strand questions the conclusion that the growth resurgence was widespread across industries. McKinsey Global Institute (2001) makes each of these arguments and Farrell, Baily, and Remes (2005) suggest that the conventional framework overstates the role of IT; instead, they emphasize the role of managerial expertise in key industries.

Another area of concern raised by many researchers about the framework in JHS — such as Gordon (2003) — is the potential importance of learning and possible lags between the installation of IT and the generation of productivity benefits, a phenomenon not captured in the framework used by JHS. Criticism also has focused on the role of adjustment costs, with important contributions by Basu, Fernald, and Shapiro (2001) and Kiley (2001).

Finally, business investment in intangible capital is not fully captured in the NIPAs and therefore is not fully captured by the conventional growth accounting framework using NIPA data. Brynjolfsson and Hitt (2005), Nakamura (1999, 2001, and 2003), and Corrado, Hulten, and Sichel (2005 and 2006) have focused on the role of business investment in intangible capital that

often accompanies investments in IT. Because these intangible investments are largely missed by the official published data and because learning lags likely are important for intangibles, conventional growth accounting — such as that in JHS and in Oliner and Sichel — could be misattributing the sources and timing of contributions from different factors.

For the most part, these critiques have not been fully integrated into the growth accounting framework and, in my judgment, there is no consensus that these critiques overturn the essentials of the story described above for the role of IT in the mid-1990s growth resurgence. But, clearly, more work is needed to nail this down.

One piece of these critiques warrants further discussion. As indicated, McKinsey Global Institute (2001) raised a series of questions about the type of results in JHS. Although by this time, the McKinsey study is a bit out of date, it received a lot of attention at the time, and was widely perceived (and perhaps intended) as a challenge to aggregate and industry growth accounting results that highlighted the role of IT in the growth resurgence. And, it captured a sentiment that appears to have gained currency over time. Thus, even though the book does not focus on these issues, I will discuss the McKinsey report in a little more detail.

As indicated, the McKinsey report questions the JHS conclusion that the growth resurgence was widespread. In particular, McKinsey argued that the mid-1990s acceleration in labour productivity could be accounted for by developments in just a handful of industries: retail trade, wholesale trade, semiconductor manufacturing, computer manufacturing, telecommunications services, and securities. McKinsey also argued that IT was just one of the driving factors behind the productivity pickup in the mid-1990s.

Regarding the point about how widespread was the acceleration in labour productivity, one interpretation of the McKinsey study is that the

disagreement really represents a difference in interpretation, rather than a difference in the underlying numbers. Indeed, McKinsey showed that industries accounting for 70 per cent of employment experienced an acceleration in labour productivity in the second half of the 1990s, similar to the JHS result. Of course, the remaining industries experienced flat or decelerating productivity so that the net pickup was less than that coming from the industries that experienced an acceleration. McKinsey chose to focus on the net result; namely, that six industries could account for the net acceleration. Other industries also saw an acceleration, but in McKinsey's classification scheme, those pickups were netted against the industries with decelerations in productivity.

Whether the McKinsey approach or that in JHS is to be preferred depends on the question being asked. In my view, if one wants to gauge how widespread was the productivity resurgence across industries, it does not seem particularly helpful to focus on the net result, implicitly canceling a number of industries with an acceleration in productivity with a number of other industries that experienced a deceleration. Of course, if one wants to gauge what industries led the productivity resurgence, then McKinsey's approach seems reasonable.

Also, even though the McKinsey study had language raising questions about the aggregate growth accounting results of the type in JHS, McKinsey did not directly challenge these results. The JHS analysis is based on the premise that, over time, firms earn a normal return (net of depreciation) on all investments including IT. Results in that framework can be misleading if capital or the rate of return are mismeasured, but McKinsey does not claim that this is so. Indeed, the report argues that IT behaves much like other capital, which is a maintained assumption of conventional

growth accounting. Thus, the McKinsey study does not really contradict the aggregate implications of studies such as JHS. McKinsey does, importantly, point out a host of other developments that contributed to the late 1990s pickup in productivity growth, including competitive pressures, regulatory changes, and managerial expertise leading to improved efficiency of critical processes. I can not speak for JHS, but for my part, these other sources of the productivity improvement seem complementary to the role of IT. Indeed, it would be difficult to imagine IT boosting productivity growth without improved efficiency of critical processes.

I will now turn to one part of the story told by JHS (and, for that matter, by Oliner and Sichel in earlier work) that, over time, I have become a little less comfortable with. This piece is the linkage between constant-quality semiconductor prices and the pace of technological progress in the semiconductor industry.⁷ As evidence of the speedup in technological progress, JHS cite the mid-1990s shift from a three-year product cycle for semiconductors to a two-year cycle as described in the 2003 Edition of the International Technology Roadmap for Semiconductors. However, the 2005 Edition of the Roadmap has, based on some new information from chip-producing companies, shifted the timing of the switch in product cycles from the mid-1990s to the late 1990s, somewhat muddying the linkage between the shift in product cycles and the observed pattern of semiconductor prices. Moreover, many other factors affect the pace of decline in constant-quality prices for semiconductors. For example, Aizcorbe, Oliner, and Sichel (2006) highlight the role of time-varying markups for semiconductors, a factor that could make price movements alone an inadequate proxy for technical progress. And, Basu, Fernald, Fisher, and Kimball (2005) highlight a

7 For a fuller discussion of these issues, see Aizcorbe, Oliner, and Sichel (2006).

number of reasons why IT price trends could be a poor proxy for technological progress.⁸

As a further example of the rather loose linkage between the trends described in the Roadmaps and constant-quality prices of semiconductors, consider the slowdown in the rate of decline in constant-quality semiconductor prices that occurred in 2001.⁹ The latest Roadmap (along with consultations with people in the industry) suggests that there was not a slowdown in the pace of technological progress in semiconductors around 2001. Thus, it appears that the slower pace of decline that began at that time must be coming from other factors. All told, this new evidence has left me more cautious about how tightly we ought to connect the type of information in the Roadmaps with semiconductor prices, and, conversely, with how tightly we ought to connect swings in constant-quality semiconductor prices to the pace of technical progress in the semiconductor industry.

Strengths and Weaknesses of the Book

The primary strengths of the book include its comprehensiveness, the consistency of methodology, and its care with data. This book is a very good place to look for a careful and thorough description of state-of-the-art growth analysis at the aggregate and industry levels. It also provides the most comprehensive description that I have seen of the myriad ways in which information technology affected almost all aspects of the economy in the past decade. Having said that, the book does have a couple of weaknesses. Because the book is, essentially, a compendium

of work completed by JHS over a number of years, the book does not cover some of the most recent issues and debates in much depth. In particular, the sources of the further pickup in growth of labour and multifactor productivity that began after 2001 is not discussed at all nor is there much discussion about why some countries appear to have benefited more from the IT revolution than did other countries.¹⁰ And, as discussed above, issues surrounding intangible capital are not mentioned in the volume by JHS, and the book provides relatively little discussion about the role of adjustment costs.

Finally, the volume by JHS highlights several areas where further progress by the measurement research community and the statistical agencies would be most welcome. These include development of reliable constant-quality price indexes for a wider range of capital goods where technology is changing rapidly, the need for more detailed industry classifications for the high-tech sector, and the need to extend the complete set of industry accounts on a NAICS basis back to 1947. I would like to add a few more items to the list. For starters, there is a need for additional work on depreciation, particularly for IT capital, but for other types of capital as well. In addition, my Federal Reserve colleagues Carol Corrado and Lawrence Slifman frequently mention some key measurement issues that are worth repeating here. First, for many high-tech categories (like communications equipment), the product lists used by the Census Bureau are woefully out of date. These are the product lists that are used when information on nominal shipments is collected from businesses in a variety of Census programs. For

8 Feenstra, Reinsdorf, Slaughter, and Harper (2005) point to another possible source of slippage between IT price declines and the pace of technological progress; namely, that changes in the terms of trade appear to be an important source of IT price declines and that, therefore, the conventional framework may have overstated the role of technological progress.

9 See Aizcorbe, Oliner, and Sichel (2006) for evidence of the statistical significance of the mid-1990s and 2001 break in the trend of constant-quality semiconductor prices.

10 But see Stirih (2006) for an analysis of the industry composition of the resurgence in labour productivity growth in the 2000s.

example, there are more than a dozen detailed product categories listed under the umbrella category “Broadcast, studio, and related electronic equipment” where nominal shipments totaled a bit less than \$3 billion in 2004. But a host of leading edge products for data communication — including routers, gateways, bridges, and terminal servers — are lumped together in a single category, in which nominal shipments were more than \$10 billion in 2004. This state of affairs does not make sense, and the current classification scheme limits our ability to track the new economy because we do not even know nominal shipments of key IT products at any significant level of detail. Moreover, as JHS point out, we also know very little about prices of these leading edge products.

Also, the BLS and the Census Bureau use different lists of establishments for their various surveys, and they also have taken different approaches to dealing with the conversion from the SIC system to NAICS. This creates some non-comparabilities between the shipments or output data that come from Census' surveys and the hours data that come from BLS surveys. And, these non-comparabilities complicate efforts to match industries at a detailed level and to create productivity statistics at a detailed industry level. Resolving these issues would significantly facilitate analysis of productivity developments.

References

- Aizcorbe, Ana, Stephen Oliner and Daniel Sichel (2006) “Shifting Trends in Semiconductor Prices and the Pace of Technological Progress,” mimeo, January.
- Antonopoulos, Christos and Plutarchos Sakellaris (2005) “Estimating Computer Depreciation Using Online Auction Data,” mimeo, Athens University of Economics and Business.
- Bartelsman, Eric and J. Joseph Beaulieu (2004) “A Consistent Accounting of U.S. Productivity Growth,” Federal Reserve Economics and Finance Discussion paper #2004-55.
- Basu, Susanto, John Fernald, Jonas Fisher and Miles Kimball (2005) “Sector-Specific Technical Change,” paper presented at the NBER Summer Institute, July 21.
- Basu, Susanto, John Fernald and Matthew Shapiro (2001) “Productivity Growth in the 1990s: Technology, Utilization, or Adjustment?” *Carnegie-Rochester Conference Series on Public Policy*, Volume 55, Issue 1, pp. 117-165.
- Brynjolfsson, Erik and Lorin Hitt (2005) “Remarks,” in *Measuring Capital in the New Economy*, Carol Corrado, John Haltiwanger and Daniel Sichel (eds.), Studies in Income and Wealth, Vol. 65 (Chicago: The University of Chicago Press) pp. 557-575.
- Corrado, Carol, Charles Hulten and Daniel Sichel (2005) “Measuring Capital and Technology: An Expanded Framework,” in *Measuring Capital in the New Economy*, Carol Corrado, John Haltiwanger and Daniel Sichel (eds.), Studies in Income and Wealth, Vol. 65 (Chicago: The University of Chicago Press) pp. 11-45.
- Corrado, Corrado, Charles Hulten and Daniel Sichel (2006) “Intangible Capital and Economic Growth,” NBER working paper #11948.
- Corrado, Carol and Larry Slifman (1999) “A Decomposition of Productivity and Costs.” *American Economic Review* 89, No. 2, May, pp. 328-332.
- Doms, Mark, Wendy Dunn, Stephen Oliner and Daniel Sichel (2004) “How Fast Do Personal Computers Depreciate? Concepts and New Estimates,” in *Tax Policy and the Economy*, James M. Poterba (ed.), National Bureau of Economic Research and the MIT Press, pp. 37-79.
- Farrell, Diana, Martin Baily and Jaana Remes (2005) “U.S. Productivity After the Dot.Com Bust,” McKinsey Global Institute.
- Feenstra, Robert, Marshall Reinsdorf, Matthew Slaughter and Michael Harper (2005) “Terms of Trade Gains and U.S. Productivity Growth,” paper presented at the NBER Summer Institute, July 26.
- Gordon, Robert J. (2001) “Did the Productivity Revival Spill Over from Manufacturing to Services? Conflicting Evidence from Three Data Sources,” paper presented at the NBER Summer Institute, July 27.
- Gordon, Robert J. (2003) “Exploding Productivity Growth: Context, Causes, and Implications,” *Brookings Papers on Economic Activity* 2003:2 (Washington, D.C.:The Brookings Institution) pp. 207-298.
- Greenwood, Jeremy, Zvi Hercowitz and Per Krusell (1997) “Long-Run Implications of Investment-Specific Technological Change,” *American Economic Review* 87, no. 3, June, pp. 342-362.

- Greenwood, Jeremy, Zvi Hercowitz and Per Krusell (2000) "The Role of Investment-Specific Technological Change in the Business Cycle," *European Economic Review* 44, no. 1, January, pp. 91-115.
- Ho, Mun S. and Kevin J. Stiroh (2001) "The Embodiment Controversy: You Can't Have Two Prices in a One-Sector Model," mimeo, May.
- Hulten, Charles R. and Frank C. Wykoff (1981a) "The Estimation of Economic Depreciation Using Vintage Asset Prices: An Application of the Box-Cox Power Transformation," *Journal of Econometrics* 15, April, pp. 367-96.
- Hulten, Charles R. and Frank C. Wykoff (1981b) "The Measurement of Economic Depreciation." In *Depreciation, Inflation, and the Taxation of Income from Capital*, Charles R. Hulten (ed.) (Washington, D.C.: The Urban Institute Press) pp. 81-125.
- International Technology Roadmap for Semiconductors (2003) 2003 Edition, Sematech Corporation, Austin, TX, December.
- International Technology Roadmap for Semiconductors (2005) 2005 Edition, Sematech Corporation, Austin, TX, December.
- Jorgenson, Dale W. and Kevin J. Stiroh (2000) "U.S. Economic Growth in the New Millennium," *Brookings Papers on Economic Activity* 2000:1, pp. 125-211.
- Jorgenson, Dale W., Mun S. Ho and Kevin J. Stiroh (2002) "Projecting Productivity Growth: Lessons from the U.S. Growth Resurgence," *Atlanta Fed Economic Review*, Third Quarter, pp. 1-14.
- Kiley, Michael (2001) "Computers and Growth with Frictions: Aggregate and Disaggregate Evidence," *Carnegie-Rochester Conference Series on Public Policy*, Vol. 55, Issue no. 1, pp. 171-215.
- McKinsey Global Institute (2001) *U.S. Productivity Growth, 1995-2000: Understanding the Contribution of Information Technology Relative to Other Factors* (Washington, DC: McKinsey Global Institute) October.
- Nakamura, Leonard (1999) "Intangibles: What Put the *New* in the New Economy?" Federal Reserve Bank of Philadelphia *Business Review*, July/August, pp. 3-16.
- Nakamura, Leonard (2001) "What is the US Gross Investment in Intangibles? (At Least) One Trillion Dollars a Year!" Federal Reserve Bank of Philadelphia Working Paper No. 01-15.
- Nakamura, Leonard (2003) "The Rise in Gross Private Investment in Intangible Assets Since 1978," mimeo, Federal Reserve Bank of Philadelphia.
- Oliner, Stephen D. (1993) "Constant-Quality Price Change, Depreciation, and Retirement of Mainframe Computers," in *Price Measurements and Their Uses*, Murray F. Foss, Marilyn E. Manser and Allan H. Young (eds.) (Chicago: University of Chicago Press) pp. 19-61.
- Oliner, Stephen D. and Daniel E. Sichel (2000) "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?" *Journal of Economic Perspectives* 14, Fall, pp. 3-22.
- Oliner, Stephen D. and Daniel E. Sichel (2002) "Information Technology and Productivity: Where are We Now and Where are We Going?" *Atlanta Fed Economic Review*, Third Quarter, pp. 15-44.
- Schreyer, Paul (2000) "The Contribution of Information and Communication Technology to Output Growth: A Study of the G7 Countries," OECD, Paris, May.
- Stiroh, Kevin J. (2006) "The Industry Origins of the Second Surge of U.S. Productivity Growth," mimeo, January.
- van Ark, Bart and Robert Inklaar (2005) "Catching Up or Getting Stuck? Europe's Troubles to Exploit ICT's Productivity Potential," Research Memorandum GD-79, Groningen Growth and Development Centre, September.
- Whelan, Karl (2003) "A Two-Sector Approach to Modeling U.S. NIPA Data," *Journal of Money, Credit, and Banking*, Volume 35, No. 4, August, pp. 627-656.