4. Reassessing the consumer price index: five years after the Boskin Commission

Dean Baker

INTRODUCTION

Federal Reserve Board Chairman Alan Greenspan started one of the more perverse national political debates in recent memory when he told the Senate Finance Committee in late 1994 that the consumer price index (CPI) substantially overstates the true rate of inflation. At the time, the country was facing large budget deficits. Greenspan pointed out that substantial areas of federal spending, most notably Social Security, and federal income tax brackets were indexed to the CPI. This meant that a reduction in the rate of inflation reported by the CPI could potentially reduce federal spending by hundreds of billions of dollars over the next decade and increase revenue by a comparable amount. According to Greenspan, much of the budget shortfall could be eliminated simply by correcting the CPI.

During the next two and a half years, methodological issues in the construction of the CPI were featured prominently in every major budget story, as changing the CPI was assumed to be the key to eliminating the budget deficit. Briefings by the Bureau of Labor Statistics (BLS) on new aggregation formulas for entry-level items and hedonic quality adjustments became ‘standing-room only’ press events. The Boskin Commission, appointed by the Senate Finance Committee to evaluate the accuracy of the CPI, issued a final report that was, for a brief time, the hottest book in Washington (SFC 1996).

Eventually, the attention to the CPI died down. Part of the reason was political. Cutting Social Security benefits and raising taxes are never popular measures, even if the rationale involves technical issues with the CPI. The interest groups that ordinarily oppose such changes eventually rallied against efforts to tinker with the CPI. Also, the deficit problem largely disappeared by itself, as more rapid economic growth and a surge in tax revenue shifted the deficits to surpluses in short order.
But, there were also serious questions raised about the accuracy of the claims made by the Boskin Commission and others concerning the CPI overstatement of inflation (for example, Moulton and Moses 1997; Madrick 1997; Baker 1997; and Moulton 1996a). Many of the commission’s claims concerning an overstatement of inflation in the CPI were not well supported by research. In some cases, the assertion of a bias rested on no more than introspection and casual observation. It was also possible to identify important areas where the commission appeared to neglect evidence that suggested the CPI could understate inflation. In short, the case against the CPI was far less compelling than the Boskin Commission’s final report implied.

Now that nearly five years have passed and the political heat around this issue has virtually disappeared, it is worth reassessing the evidence. For economists, there is still a great deal at stake. Our whole understanding of economic history and the growth process would be radically altered if the commission’s assessment is largely correct. Annual economic growth in the United States would have been approximately 1.0 percentage point higher than current data indicate. This means that we were much poorer in the recent past than is generally recognized and will be far richer in the near future than current projections suggest is possible. The same would be true for all other countries as well, since few of the criticisms of the CPI would apply to the US measure alone.

A vast amount of economic research on other topics would also be called into question. For example, lifetime saving and consumption patterns would look quite different if the true rate of inflation has been 1.0 per cent less, on average, than what the current data show. Real interest rates and returns on investment would also be far higher than is generally recognized. Macroeconomic studies of the impact of inflation on growth will also have to be reassessed, if inflation has generally been reported as being far higher than its true rate. Industry studies, which examined issues such as the effects of deregulation, would also be called into question, since the reference point for such studies is usually the change in the overall price level. In fact, there would be few areas of economic research which would be left untouched by a large downward adjustment to the consumer price index. For this reason, it is important to come to grips with the Boskin Commission’s claim, even if it is less relevant to the current political situation.

**CATEGORIZING THE ISSUES**

Most of the debate on the CPI has centered on the types of bias identified by the Boskin Commission: formula bias, substitution bias, retail-
outlet substitution bias, quality bias, and new-product bias. While this categorization makes important distinctions between potential sources of bias, it tends to focus the debate too narrowly. In principle, the Boskin Commission was examining the extent to which the CPI served as an accurate measure of the change in the cost of living, the amount of money that is necessary to keep the average consumer at the same level of well-being. While the types of bias identified by the commission would all be relevant to any sort of cost-of-living index, maintaining a level of well-being involves a range of factors that were largely neglected by the Boskin Commission.

One of the key assumptions in the Boskin Commission’s assessment of the CPI as a measure of the cost of living is that it is possible to evaluate the cost of living by examining the price and quality of goods and services apart from the context in which they are consumed. This implies that the quality of publicly provided goods and services and the natural and social environment can be assumed to remain constant. This is a rather heroic assumption, and the failure of the commission to seriously consider this point limited the value of its analysis of the CPI as a measure of the cost of living.

There are numerous examples of situations in which changes in the natural or social environment, or the quality of publicly provided goods and services, could make large differences in the cost of living. The spread of automobiles, and the restructuring of cities to accommodate them, has made individuals who do not have an automobile unambiguously worse off than they were in years before car ownership became common. An accurate cost-of-living index would show some increase reflecting the new need for an automobile to keep one’s level of well-being constant.

Similarly, the spread of the Internet as a means of communication has made households without access to the Internet worse off. They are now being excluded from an important source of communication. Information that may have previously been available through other sources now travels largely on the Internet. In addition, they may find it more difficult to maintain social contacts, if their friends come to rely on the Internet for communication, while they lack access. Again, a true cost-of-living index would show an increase in the cost of living based on the need of individuals to pay for Internet access to keep their well-being constant.

There are numerous other examples where the physical or social environment and provision of public goods can make a large impact on the cost of living, if it is defined as maintaining a level of well-being. The quality and accessibility of public schools, parks and recreational facilities, clean drinking water, and crime-free neighborhoods are all factors that would have a large impact on people’s well-being. An accurate cost-of-living index would have to take account of all of these factors.
However difficult it might be to try to quantify some of the effects noted above, these effects, at least in principle, are amenable to inclusion in the construction of a cost-of-living index. They can be perceived of as individualistic, in the sense that the well-being or harm derived from these factors can in principle be assessed by examining their impact in isolation. There is a much greater problem in constructing a cost-of-living index if well-being is inherently a relative concept, in other words, if the well-being that individuals derive from consuming a bundle of goods is inherently linked to the consumption patterns of others as argued by Thorstein Veblen or Krugman (1996). In this case, there is no basis for constructing an index, since well-being for individuals would depend on their relative position in society, not an absolute level of consumption.

It is important to recognize that there is no theoretical justification for excluding relative factors, or Veblen effects, from a cost-of-living index. The only basis for such an index is some implicit concept of utility. If individuals derive utility from their relative consumption, rather than their absolute consumption, then any cost-of-living index that excludes the effects of relative consumption is not measuring the cost of living. Such an index would lack any theoretical foundation. Whether the effects of relative consumption patterns on well-being are sufficiently small that they can be ignored is an empirical matter, not something that can be determined a priori on theoretical grounds.

The questions raised concerning the physical and social environment in which consumption takes place are quite complex. The issues raised by the importance of relative patterns of consumption are even more complex. In both cases, it is questionable whether these topics can be well addressed by standard economic analysis. Clearly they have not been to date either in the construction of the CPI by the BLS, or by the critics of the CPI. This chapter will examine these issues only to establish their importance to individuals’ well-being. The fact that these factors are important, and do not lend themselves easily, or at all, to standard modes of economic analysis, argues against efforts to treat the CPI as a cost-of-living index. The CPI was designed by the BLS as a price index, and this is how it is best regarded, since economists lack the expertise and possibly the theoretical basis for constructing a genuine cost-of-living index.

The remainder of this chapter is divided into four sections. First, we examine how the social and physical structure in which consumption takes place affects individuals’ well-being. Second, we assess the importance of relative consumption in individuals’ well-being. The next and largest section examines the types of biases noted in the Boskin Commission’s report, and assesses the evidence for its claims. Finally, some concluding comments on the accuracy of the CPI as a price index are presented.
CONSUMPTION: THE PHYSICAL AND SOCIAL INFRASTRUCTURE

In order for the CPI to provide a useful measure of the cost of living, it must be the case that the changes in the physical and social infrastructure, including the changes in publicly provided goods, are sufficiently small that their impact on well-being can be ignored. There is little reason to assume that this is the case. This section briefly examines some of the ways in which the changes in infrastructure in recent years can be viewed as important. The factors discussed will exclude issues of relative consumption (although these cannot be completely separated), which will be dealt with in the next section.

Table 4.1 shows each of the major categories within the CPI, along with their relative importance in the index, and lists some of the ways in which changes in the physical and social infrastructure over the last half-century could be an important factor in the well-being derived from the consumption of the goods and services in each category. As can be seen, there is no area of consumption in which it can be assumed that these changes are not at least potentially quite large.

The discussion below is necessarily largely anecdotal. It would be difficult, if not impossible, to rigorously quantify the impact of many of these effects. The purpose of the discussion is not to argue that there has been a general pattern to these effects that would lead the CPI to either understate or overstate the true increase in the cost of living. Rather, the purpose is to establish that these effects are likely to be large, and that changes in the physical and social infrastructure cannot be ignored in a serious analysis of the cost of living.

The first category in the table is food and beverages. The two items listed here are the increased need to consume food away from home, as a result of an increase in working hours per adult, and changes in access to clean drinking water. The first effect is the result of the increasing number of two-earner households, or families headed by a single adult. The percentage of families in which the woman was not in the paid labor force fell from 67.1 per cent in 1951 to 29.5 per cent in 1998 (Mishel et al. 2001, Table 1.6). The fact that it is far less likely that a family has an adult at home who is not working for a wage means that the typical family has far less time available to prepare food. In order to maintain the same quality of food, it would be necessary to either buy more prepared foods, which are more expensive, or to eat out more frequently. The relative importance of food away from home in the CPI rose from 4.6 per cent in 1953 to 5.7 per cent in 2001, a period in which the relative importance of the rest of the food component dropped from 25.0 per cent to 9.6 per cent of the index (Bureau of Labor Statistics 1966, 2001).
<table>
<thead>
<tr>
<th>CPI category</th>
<th>Relative importance</th>
<th>Infrastructure changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food and beverages</td>
<td>16.20</td>
<td>Increased family work hours – less time for food preparation – more food consumed away from home</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Changed availability of clean drinking water</td>
</tr>
<tr>
<td>Housing</td>
<td></td>
<td><strong>Shelter</strong> 30.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Increased distances both within and between metropolitan areas – more time spent commuting, on shopping trips, or to visit friends and family – increased needs for communication (for example, long-distance phone calls or Internet) and short- and long-distance travel, including increased need for hotels/motels</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Increased crowding – less availability of open space/recreational areas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Increased or decreased pollution of air/water/soil</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crime – greater precautions, including restricting behavior, and added expenses (for example, car and house alarms)</td>
</tr>
<tr>
<td>Insurance</td>
<td>0.37</td>
<td>Liability laws, social norms on lawsuits</td>
</tr>
<tr>
<td>Apparel</td>
<td>4.45</td>
<td>Changing standards for work attire</td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
<td><strong>New and used vehicles</strong> 7.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congestion, size and weight of other vehicles, safety regulations and enforcement (for example, drunk-driving laws)</td>
</tr>
<tr>
<td>Public transportation</td>
<td></td>
<td><strong>Insurance</strong> 2.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of routes covered and frequency, congestion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Congestion, size and weight of other vehicles, safety regulations and enforcement (for example, drunk-driving laws), liability laws</td>
</tr>
<tr>
<td>Motor fuel</td>
<td>3.48</td>
<td>Availability of gas stations</td>
</tr>
<tr>
<td>Medical care</td>
<td>5.81</td>
<td>Improved knowledge of factors affecting health (for example, diet, exercise, smoking, and drinking)</td>
</tr>
</tbody>
</table>
An analysis that measured the increased wages associated with more workers per family, but did not also make some allowance for the increased expenses associated with paying more for prepared food or eating meals out, would overstate the increase in real family income. In principle, an accurate cost-of-living index would measure the increased expenditures on buying prepared food or eating out which would leave a family just as well off as they had been previously, when they had more time available to prepare their own meals. As a practical matter, this would not be easy, since families are likely to consume more prepared foods or eat out more frequently, not just because they have less time, but as a result of the fact that they have more income.

The second item under this category, the availability of clean drinking

Table 4.1 (continued)

<table>
<thead>
<tr>
<th>CPI category</th>
<th>Relative importance</th>
<th>Infrastructure changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recreation</td>
<td>5.91</td>
<td>Changed exposure to hazards in the physical environment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Changed exposure to hazards in the social environment (for example, stress)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exposure to new diseases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Quality of performance (for example, sports teams affected by expansion),</td>
</tr>
<tr>
<td></td>
<td></td>
<td>quality of facilities (for example, replacement of old stadiums by modern facilities)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>General patterns of recreation (for example, availability of partners for sports)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Changed television content due to factors such as the spread of cable</td>
</tr>
<tr>
<td></td>
<td></td>
<td>New broadcast formats</td>
</tr>
<tr>
<td>Education and</td>
<td></td>
<td>Change in the quality of public schools</td>
</tr>
<tr>
<td>communication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>2.80</td>
<td></td>
</tr>
<tr>
<td>Computers</td>
<td>0.80</td>
<td>Change in software standards</td>
</tr>
<tr>
<td>Other goods and</td>
<td></td>
<td></td>
</tr>
<tr>
<td>services</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal care goods</td>
<td>0.73</td>
<td>Changing standards of work appearance</td>
</tr>
<tr>
<td>Personal care services</td>
<td>0.99</td>
<td>Changing standards of work appearance</td>
</tr>
<tr>
<td>Tobacco</td>
<td>1.31</td>
<td>Restrictions on smoking</td>
</tr>
</tbody>
</table>

water, may be more a question of increased concern over health than an actual change in the quality of the water provided through public water systems. However, there has been a large increase in consumption of bottled water and in purchases of water filtration devices over the last three decades. Insofar as these expenditures have been necessary simply to offset an actual deterioration in the quality of public drinking water, it would constitute a large increase in the cost of living.\(^2\) (Two bottles of water a day, would cost approximately $700 per year, or roughly 2 per cent of the annual expenditure of a typical family.)

There are several issues with respect to the items in the housing category, which could be very important in assessing the cost of living. The most obvious is simply the question of location itself. The index never makes direct comparisons between the cost of housing in different locations, it only measures changes in the cost of specific housing units. This means that, insofar as there are gains or losses associated with the new units added to the housing stock, these will be left out of the index. To some extent, this raises quality issues of the housing itself, which will be discussed later; however, at least as important to the value of the housing unit are factors determined by its location, which is distinct from the physical quality of the house itself.

The first set of issues concerns the distances that it may be necessary to travel as a result of changes in living patterns.\(^3\) If housing becomes increasingly dispersed, then it may be necessary for people to own cars, who may not otherwise have needed to own a car. In addition, they may have to spend more time commuting to work, traveling to and from stores, and to visit friends. The availability of open land to allow construction of new housing units is a factor that limits the increase in housing costs compared to a situation where an increased population has to be housed on a fixed amount of land. In this sense, the CPI as constructed, should pick up reasonably well the gains associated with the increased dispersion of housing. However, by excluding all the additional costs implied by greater dispersion, it would be understating the increase in the cost of living. (Of course the increased ability to work at home as a result of the Internet is an important factor going in the opposite direction.)

A comparable set of factors would apply to increasing dispersion of family members across the nation or world. Economic theory implies that an increasingly mobile labor force should be a source of greater efficiency, since workers can find jobs in locations where their skills can be most fully utilized. The gains from this greater mobility should be picked up in more rapid real wage growth than would otherwise have been possible. However, costs that could result from such dispersion, such as the increased travel and communication expenses that are necessary for families to stay in contact, do not get included in the CPI.
There is a range of other factors that can significantly affect the utility that individuals derive from a particular housing arrangement. The population density and the availability of open spaces for recreation could have a very large impact on people’s well-being. People often travel long distances and/or spend large amounts of money to have access to open space. Similarly, the cleanliness of the air, water, and soil will affect people’s enjoyment from living in a particular location, as well as their health. Crime and the fear of crime also can be tremendously important factors in determining individuals’ well-being.\(^4\)

The well-being that individuals derive from clothing will depend to an overwhelming extent on the social norms governing fashion. The expense of purchasing sufficient material for body warmth is trivial compared with the expense of buying clothes that conform to accepted standards within society. Clearly, the CPI makes no effort to factor in the impact of changing norms on clothing patterns, nor is it obvious how it possibly could. In addition, workers generally have to maintain certain standards of attire on their jobs. This can lead to increases or decreases in the cost of living as these standards change. For example, if more casual clothing becomes common, there could be less of a need for spending on business suits and dry cleaning. Since work clothes are generally paid for by workers (except for uniforms), changes in expectations concerning work clothes can significantly affect the cost of clothing to workers.

There are numerous infrastructure factors that would affect the utility that an individual can derive from owning a car. At the top of the list would be traffic congestion and the quality of roads. The regulations governing driving, such as speed limits or restrictions on off-road driving, and the extent of their enforcement, could also have an important impact on the utility derived from car ownership. In addition, the risks associated with driving will depend to a large extent on the type and number of other vehicles on the road. For example, the proliferation of sports utility vehicles in recent years has increased the dangers faced by drivers of other types of cars.

With public transportation, the frequency of train, bus, or plane trips would be an important factor in determining their usefulness, as would the number of routes covered. The latter is a difficult issue to assess, since it may be necessary for the number of routes to expand through time in order to maintain the same quality of service, if residential patterns are increasingly dispersed. The CPI makes no effort to account for any of these factors.

Crowding of public transportation is also a significant issue affecting well-being. In an airplane that is one-third empty, every passenger can have an empty seat next to them. Passengers pay a large premium for the extra space that goes with a business- or first-class seat. If higher capacity utilization significantly reduces the probability that passengers will have an
empty seat next to them, then this is a major deterioration in service quality. Travel delays also have a large impact on quality.

The social influences on medical care are likely to be very large. In general, health is probably far more dependent on diet, exercise, and environmental factors than the medical care provided by the health-care system. Yet these factors, whether positive or negative, are not in any way picked up in the CPI. In fact, the CPI measures of inflation in this sector could easily go in the opposite direction of an accurate measure of health-care costs. For example, if anti-smoking campaigns led to a large enough reduction in some kinds of diseases, it is possible that the treatment for these diseases would rise, as the cost of drug research or expensive equipment would be spread over a smaller number of patients. Conversely, a decline in the price of the drugs developed to treat a new disease, such as AIDS, would be recorded only as fall in the price of medical care, even though the cost of medical care would obviously be still lower if the disease had never appeared in the first place. It is worth noting that the CPI measures the cost of specific procedures, such as treating knee injuries, or specific drugs. If the cost of health insurance rises because individuals are on average receiving more tests or more drugs, this would not appear as an increase in the cost of medical care in the CPI. For this reason, the price changes measured by the CPI in the medical-care sector may have relatively little relationship to how most people experience rising health-care costs.

In the area of recreation, social norms play a very large role in determining costs. It becomes impossible to engage in many types of recreation, if it is not possible to find partners. If people increasingly choose individualistic types of entertainment, such as watching television or playing video games, then it would become increasingly difficult for someone who preferred to play baseball or basketball, or to just sit on a porch talking to neighbors, to pursue their chosen form of recreation. Such a loss would not be taken into account in the CPI. Similarly, the CPI obviously does not attempt to assess the quality of sports or events (for example, movies, concerts, and plays) that people pay to see. If there is a systematic improvement or deterioration in quality through time, this will not be picked up in the CPI.

Public schools are a very important publicly provided good. If there is a deterioration in the quality of these schools, so that more people feel a need to send their children to private schools, then this could be a large increase in the cost of living, which would not be picked up in the CPI. Similarly, insofar as parents need to pay for daycare for their children, because they do not have a nonworking parent at home, this would be an important cost not picked up in the CPI. Child care now accounts for an average of nearly 0.4 per cent of consumption expenditures (Bureau of Labor Statistics 2001).
The CPI has shown an extraordinary rate of price decline for computers in recent years. While this may reflect real improvements in computer technology, to some extent these improvements are needed to operate ever more complex software. New software may offer real advantages to many consumers, but others may feel the need to buy it simply to stay in communication with individuals and websites that use the more complex software. The cost of upgrading software, simply to be able to stay in contact with others, could be a significant unmeasured cost associated with computer ownership.

The other goods and services category includes personal-care items, such as hair cuts and styling. The need for this would be dependent to a very large extent on social norms, including grooming requirements at workplaces. This category also includes legal and accounting services. The need for legal services would depend on social norms and the complexity of the legal system. For example, if laws are written so that it is very simple to write a will or buy a house, then the demand for lawyers would be significantly reduced. Expenditures for legal and financial services now account for 0.7 per cent of consumer spending, with auto and home insurance accounting for another 2.8 per cent (Bureau of Labor Statistics 2001). Insofar as an increase in spending on these items is attributable simply to the greater complexity of society, this could be a source of substantial increase in the cost of living. Tobacco also appears in the other goods and services category. While restrictions on smoking may have increased the well-being of nonsmokers, and those who are trying to quit, they would reduce the utility that individuals derive from smoking. The CPI does not attempt to pick up either of these effects.

To summarize, the list of changes in the social and physical environment that would have affected the utility that individuals derive from consuming a particular bundle of goods over the last 50 years is quite extensive. The ones noted above are at best a very small subset of the factors that would be relevant to individuals’ utility. It is likely that the cumulative effect of these changes is large, although it is impossible to determine without more careful investigation whether the net impact is positive or negative on individual well-being. A true measure of the cost of living would have to assess the quantitative importance of these changes along with the changes in the prices of market goods and services.

RELATIVE CONSUMPTION

The standard assumption in nearly all economic work is that individuals only derive utility from the goods they directly consume, not what they
consume relative to what others consume. In many situations this may be a fine working assumption, even if it is not entirely accurate. Measuring changes in the cost of living is not one of those situations.

The *only* reference point for a cost-of-living index is individual utility. It involves comparing different bundles of goods and services of varying quality through time. There is no common metric, except individual utility, against which these bundles can be measured. If individuals view their relative consumption as an important factor in their well-being, economists have no justification whatsoever not to take this into account. If an identical bundle of goods and services leaves a person feeling worse off in a situation in which their relative position has deteriorated, then a cost-of-living index that failed to measure this impact is inaccurate. If relative patterns of consumption are actually quite important to individual well-being, a cost-of-living index that ignored relative consumption patterns would not really be measuring anything at all.

The actual importance of relative consumption to individual well-being is, of course, an empirical question, not a theoretical one. It is not clear how the importance of relative consumption to well-being can be determined or if economists even possess tools that readily lend themselves to this sort of assessment. In any case, it is clear that few economists have attempted to take issues of relative consumption patterns seriously or tried in any serious way to quantify such effects. This section will examine the major categories of the CPI to note areas in which relative consumption patterns are likely to be important to the utility derived. The discussion is intended to be suggestive, since a comprehensive analysis of the impact of relative consumption patterns is beyond the scope of this chapter.

Before directly considering specific categories of goods and services, it is worth again noting the distinction between the nature of the problem posed by relative consumption affecting individuals’ utility and the type of problems noted in the previous section. That section noted changes in the social and physical infrastructure that would affect individuals’ well-being. A genuine cost-of-living index would have to take these changes into account. The CPI makes no effort to include these sorts of factors, nor are there any serious proposals to incorporate such effects in the future, but in principle, it should be possible to measure their impact on well-being. On the other hand, it is not clear how an index could even theoretically include the impact of changes in relative consumption, even if this could somehow be measured. In other words, if relative consumption patterns are actually important to individuals’ well-being, it is not clear that a cost-of-living index is a meaningful concept.

Table 4.2 lists the major categories of the CPI, along with their relative weights and the ways in which relative consumption patterns may prove to
Table 4.2  Importance of relative consumption to utility

<table>
<thead>
<tr>
<th>CPI category</th>
<th>Relative importance</th>
<th>Nature of potential impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food and beverages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foods</td>
<td>9.56</td>
<td>Utility derived from various types of food may be affected by the status associated with</td>
</tr>
<tr>
<td></td>
<td></td>
<td>them</td>
</tr>
<tr>
<td>Alcoholic beverages</td>
<td>0.62</td>
<td>Utility derived from various types of beer, wine, or hard liquor may be affected by the</td>
</tr>
<tr>
<td></td>
<td></td>
<td>status associated with them</td>
</tr>
<tr>
<td>Food away from home</td>
<td>6.02</td>
<td>Utility associated with specific restaurants depends on status and location</td>
</tr>
<tr>
<td>Housing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rent</td>
<td>7.08</td>
<td>The value attached to a particular location depends on the status attached to living or</td>
</tr>
<tr>
<td></td>
<td></td>
<td>visiting a specific area</td>
</tr>
<tr>
<td>Owners’ equivalent rent</td>
<td>20.46</td>
<td></td>
</tr>
<tr>
<td>Hotels and motels</td>
<td>2.10</td>
<td></td>
</tr>
<tr>
<td>Household furniture (including</td>
<td>4.61</td>
<td>Utility depends in part on relative quality</td>
</tr>
<tr>
<td>appliances)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apparel</td>
<td>4.45</td>
<td>Utility depends on conforming to style changes</td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New and used vehicles</td>
<td>7.48</td>
<td>Utility is a function of relative quality and newness of vehicle</td>
</tr>
<tr>
<td>Public transportation</td>
<td>1.41</td>
<td>Segregating travel by class provides utility apart from the qualitative features associated</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with the differences</td>
</tr>
<tr>
<td>Medical care</td>
<td>5.81</td>
<td>Standards of health and appearance could be relative</td>
</tr>
<tr>
<td>Recreation</td>
<td>5.91</td>
<td>Desirability of specific types of recreation could be affected by the social status</td>
</tr>
<tr>
<td></td>
<td></td>
<td>associated with them</td>
</tr>
<tr>
<td>Education and communication</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education tuition and fees</td>
<td>2.80</td>
<td>Quality of schools is a relative measure</td>
</tr>
<tr>
<td>Total</td>
<td>78.31</td>
<td></td>
</tr>
</tbody>
</table>

be important to individual's utility. Most of the major items in the CPI do appear on the table, although it is not easy to determine the extent to which relative factors are important to individuals' utility.

For example, in the case of food, it is not clear to what extent a preference for lobster over mackerel is attributable to intrinsic human characteristics, as opposed to the relative expense and status associated with lobster. The same sort of issue arises with a wide variety of foods, in addition to quality upgrading within a food type. It is at least plausible that status is a very important factor in the utility that individuals derive from the food they consume. The view that status plays an important role in the utility from consuming various types of beer, wines, and spirits seems even more plausible.

There are huge differences in prices charged by restaurants that may outwardly appear to be comparable in the quality of service and food. These differences may in part be explained by the status associated with the particular restaurant and the status associated with the location. (A restaurant overlooking a river in rural Pennsylvania will not command the same premium for its river view as a restaurant overlooking the Hudson River in New York City.)

It is reasonable to believe that status considerations play an enormous role in the utility derived from housing. There are enormous price differentials for comparable housing in different parts of the country or different parts of the same metropolitan area (Moulton 1995). A small apartment near Central Park in Manhattan can easily be rented for a higher price than a large house in a rural part of the country. Since housing is such a large part of consumption spending (the shelter component has a 30.3 per cent relative importance in the CPI (Bureau of Labor Statistics 2001)), this issue can pose very large problems in assessing living standards. For example, rents can rise rapidly in an area either because of pure inflation, or because the area becomes relatively more desirable. (As noted earlier, the CPI only measures price changes for the same units of housing, it does not make comparisons across units.)

There is no obvious way to distinguish between these two phenomena, but it does raise a very important issue. Everyone cannot live in Manhattan. If most people do not necessarily want to live in Manhattan, then this is not a problem. However, if through time it becomes increasingly important to more people to live in Manhattan, there is greater sense of deprivation caused by the limited housing space there. A cost-of-living index has to assume that the latter sort of shifts in taste do not occur, or at least that they are relatively unimportant.

The claim that the utility derived from clothes is in part relative should not be too contentious. It would be difficult to imagine that the desire to
change clothing styles with fashion changes is intrinsic to the human psyche. Similarly, the point at which clothes are considered too old to wear is probably less dependent on their physical usefulness than clothing standards within different social strata. The situation with cars and trucks is similar. The claim that people select a luxury car (or almost any car) at least in part for the status it conveys is probably not too controversial.

Some types of public transportation, such as air travel and now trains, are explicitly segregated by class of travel. It is at least plausible that part of the premium for a first-class ticket on an airplane is attributable to the status it conveys, and the desire to be among a higher class of traveler. Informal types of segregation in travel are probably at least as important as the formal types, for example, choosing trains rather than buses, or cabs rather than shuttles.

Relative factors could be quite important to the types of medical care that people receive. If social norms change so that higher standards of health become common, individuals may expect more from their medical care. For example, if it becomes common to engage in strenuous physical activity later in life, there will be an increased demand for doctors that specialize in sports medicine. Similarly, if certain types of largely cosmetic medical procedures become common (for example, braces for children with ill-formed teeth), then those who do not get these procedures will suffer as a result.

The utility derived from many types of recreational events is probably to a large extent a function of relative factors. For example, it is certainly plausible that the value that individuals will attach to a ticket to a rock star’s concert is in part attributable to the fact that he or she is popular. Seeing the same star would provide less utility, if he or she were less popular. Similarly, for some people, part of the attraction of a sport like racquetball, instead of sports like bowling, probably stems from the relative status associated with the two sports.

The component of education that goes to tuition and fees can be viewed as heavily affected by relative factors. It is very unlikely that a parent or student would be satisfied paying the same tuition for a school that stood near the bottom in its rankings as one near the top. (Education could arguably be treated as an investment good, and therefore excluded from the CPI altogether, since much of the purpose is to increase future income.)

The items listed in Table 4.2 account for 78.3 per cent of the weight of the CPI, and even this figure could be low. For example, it is certainly plausible that status considerations affect the well-being that individuals derive from different types of tobacco, or even personal-care items, although these are not included in the table. Of course, being included in the table does not establish that relative factors are important to the utility derived from the
consumption of these items. However, it is at least plausible that relative factors are important, and it would be arbitrary to assume the opposite in the absence of any supporting evidence.

ANALYZING THE CPI AS A PRICE INDEX

The purpose of the preceding discussion was to place the analysis of the accuracy of the CPI in its proper context. The CPI was not designed as a cost-of-living index, and could not easily be transformed into one. As noted above, a true cost-of-living index would have to evaluate the impact of all the changes in the social and physical infrastructure as well as the provision of publicly provided goods and services. The BLS does not even attempt to make this sort of assessment, nor have its critics. At the very least, maintaining this assessment on an ongoing basis would require an enormous expansion of the index into new realms, where it is not clear that economists have the necessary expertise.

The problem of constructing a cost-of-living index goes from difficult to impossible, if it is the case that many of the sources of utility are inherently relative, as noted in the previous section. Here also, there is no research to support the contention that relative factors are unimportant to well-being, simply a hope that reality corresponds to convenience.7

The focus of the CPI must be narrower than an all-encompassing measure of the cost of living. The CPI is defined as a price index for consumption goods and services, and this is the only purpose that it can reasonably be expected to serve. An index that attempts to quantify the additional convenience associated with 24-hour food stores or ATMs (automated teller machines), as some have proposed, but does not pick up the benefits lost, which are related to these developments, from changing norms under which it is less acceptable to borrow items from neighbors, or less common for stores to accept personal checks, is not more accurate than an index that ignores both sets of changes in society.

To evaluate social progress or well-being it will be necessary to look at many other factors that go beyond the measures of real wages or income obtained with the CPI. These factors could include various health-care statistics, education measures, incarceration rates, and numerous other measures that may reflect the overall progress of society. The CPI is not a comprehensive measure that can be used for this purpose, and there is no plausible way in which it can be transformed into such a measure.

This section will evaluate the evidence on the accuracy of the CPI as a price index. It will examine the types of bias that have appeared in the CPI, and the extent to which their size has varied over the last half-century.
While the main topic in this discussion is the CPI-U, the standard CPI reported in news accounts, this section will also make reference to the CPI-U-RS, an index that the BLS recently constructed which applies the current CPI methodology to the years between 1978 and 1998 (Stewart and Reed 1999). 8

**Formula Bias**

The simplest form of bias that has appeared in the CPI in recent decades is formula bias. The issues here are rather straightforward. Because of the way that the BLS substitutes new items into its survey, it overweighted items that were selling at temporarily low prices and therefore likely to experience larger than average price increases, and underweighted goods that were temporarily selling at high prices and therefore likely to experience price declines. This caused the index to overstate the true rate of inflation for a wide variety of goods in the period from when it began its rotating sample method in 1978, to 1995 and 1996 when corrections to the flawed pricing procedure were put in place. BLS research estimates that this formula bias caused the CPI-U to overstate the rate of inflation by approximately 0.25 percentage points annually, with the largest effects on food and clothing (Bureau of Labor Statistics 1996; Armknecht et al. 1995).

In addition to the problem with sampling rotation, there were two other notable areas in which CPI formulas led to bias in the index. An incorrect aggregation formula used for the owner-equivalent rent component led it to rise at a rate of approximately 0.6 percentage points annually compared to a proper index, over the 1987–95 period. This would have led to an upward bias of approximately 0.12 per cent in the index as a whole over this period. On the other side, failure to impute rent increases for vacant units led to an understatement of inflation of approximately 0.1 percentage points annually in both the rent and owner-equivalent rent component in the years 1978–95, which would have biased downward the reported inflation rate in the CPI as a whole by approximately 0.03 percentage points a year over this period (Henderson and Smedley 1994). These specific sources of formula bias are no longer present in the CPI, and the CPI-U-RS has been adjusted to correct for their impact in previous years, as has the reconstructed CPI in Baker (1996).

**Substitution Bias**

The issue of substitution bias raises a basic methodological question about what the CPI should be trying to measure. Historically, the CPI has been constructed as a Laspeyres index, which held the consumption bundle fixed.
at base-level quantities. Assuming that the social and physical infrastructures remain unchanged, this is supposed to guarantee that the index would place an upper bound on the true increase in the cost of living. If it is accepted that the CPI is not and cannot be a true cost-of-living index, then the question to be asked is which set of consumption items should it track? From this perspective, there is no obvious justification for maintaining a fixed basket, especially if the rate of inflation for this basket diverges significantly from the rate of inflation on the goods that households are actually consuming.

It is also worth noting that the difference between the inflation rate shown by a Laspeyres index and a superlative index will vary through time, depending on the overall rate of inflation and its evenness across sectors. In most of the last 15 years, the substitution incorporated into a superlative index has lowered the measured rate of inflation by less than 0.15 percentage points compared to the CPI (Aizcorbe and Jackman 1993, 1997). However, in some years – notably 1990, when the price of oil rose sharply – the difference has been considerably larger (0.6 percentage points in 1990). Bryan and Cecchetti (1993), using a somewhat different methodology, estimated the size of the substitution effect at close to 0.8 percentage points annually in the 1960s and 1970s and close to zero in subsequent years. (The Aizcorbe and Jackman series only goes back as far as 1983.) If there are large divergences between the rates of inflation shown by a Laspeyres index and a superlative index, the inflation rate shown by a superlative index would better reflect the rate of inflation that consumers are actually experiencing at the time. For this reason, as far as possible, it would be desirable to use the inflation rate shown by superlative indices, when examining historical inflation rates.

The above discussion refers to ‘upper-level’ substitution, where individuals substitute between categories of goods. There is also an issue concerning ‘lower-level’ substitution, where individuals substitute between items within a category. (For example, a switch from apples to oranges would be an example of a switch at the upper level. Switching between types of apples would be a switch at the lower level.) From the standpoint of constructing the index, the main distinction is that the BLS actually has data (albeit with a significant lag) on shifts in consumption patterns at the upper level, from the Consumer Expenditure Survey (CES). It never actually collects data on shifts in consumption at the lower level. Prior to 1998, it had used a Laspeyres index to aggregate the inflation rates shown at the lower level. With the 1998 revision to the CPI, the BLS began using geometric means to aggregate rates of inflation across items at the lower level in most categories of the index. This had the effect of lowering the reported rate of inflation by approximately 0.2 percentage points annually (Moulton and
Smedley 1995, Reinsdorf 1996; Reinsdorf and Moulton 1995). The impact of this change has been incorporated into the CPI-U-RS for years after 1978. It is important to note that this has led to a serious inconsistency with the data for years prior to 1978. The BEA has used deflators constructed with geometric means to measure output in years after 1978. These revised output numbers have been, in turn, used by the BLS to measure productivity growth. Since the deflators used for years prior to 1978 continue to use the Laspeyres aggregation, the reported level of output and productivity growth is still biased downward for these years. Any serious effort to compare growth rates for the pre- and post-1978 periods should adjust for the impact of this methodological change.

**Retail Outlet Substitution Bias**

The third major source of bias identified by the Boskin Commission was retail outlet substitution bias: the failure of the index to pick up potential gains to consumers that result from buying goods at discount stores. This bias could arise because the index never makes direct comparisons of prices across outlets. It is assumed that the quality-adjusted price is identical in different outlets, so that price differences – for example between discount stores and traditional retail stores – reflect reduced service quality. The Boskin Commission and others have argued that this approach may cause the index to miss real gains associated with the switch to discount outlets. They pointed out that their share of the retail market has risen significantly in recent years which could suggest significant benefits to consumers (Senate Finance Committee (SFC) 1995, 1996; Oi 1992; Reinsdorf 1993, 1996).

While the Boskin Commission originally estimated this source of bias as adding 0.2, and possibly as much as 0.4 percentage points to the reported rate of inflation, more careful analysis indicates that its impact is almost certainly far smaller, and could even go in the opposite direction (SFC 1995). First, it is important to recognize that the bias only applies to a subset of the index that could in principle be sold in discount stores. Large areas of the index – shelter, medical care, utilities – are excluded altogether. The areas of the index that can properly be included – primarily food at home, clothes, appliances, and household items – only comprise about 30 per cent of the index. Second, the bias could only apply to the percentage of sales in these categories that switch in any given year. According to a study in the early 1990s (MacDonald and Nelson 1991, cited in Moulton 1996), approximately 0.7 per cent of the sales in these categories switched each year from traditional retail outlets to discount stores, which means that the bias could apply to approximately 0.2 per cent of the whole index.
Assuming a 10 per cent pure (quality-adjusted) price difference between discount stores and traditional stores, this source of bias would add 0.02 percentage points to the reported rate of inflation.\textsuperscript{11}

It is also important to recognize that the bias could go in the other direction. If traditional retail stores, which are valued by consumers, close because the owners choose to abandon the business, then there could be a loss associated with the shift. In addition, traditional stores may reduce their service in order to more effectively compete in price with discount stores, a deterioration in service quality that would not be picked up in the CPI.\textsuperscript{12} In short, there is no clear evidence concerning either the size or the direction of this bias, although it seems implausible that it is very large. Based on very little evidence, the Boskin Commission’s final report estimated the size of retail outlet substitution bias at just 0.1 percentage point annually.

**Quality Bias and New-goods Bias**

The most controversial aspect of the Boskin Commission’s report was its estimate of the size of quality and new-goods bias. The commission estimated that the combined impact of these two sources of bias led to an overstatement of inflation by 0.6 percentage points annually. While these are distinct sources of bias, the commission chose to treat them together in its final report. Following the commission, this discussion will also treat the two types of bias together. However, before examining the issues that arise in each category of the CPI, it is worth distinguishing conceptually between the basis of the two types of bias.

Quality bias could appear in the index because the BLS has failed to accurately measure changes in the quality of goods and services through time. In principle, the CPI is measuring the change in the price of goods and services of constant quality. If the price of items rises because their quality has improved through time, then this should not be counted as inflation in the CPI. It is important to recognize that the BLS has always made quality adjustments for most items. According to BLS data, quality adjustments reduced the reported rate of inflation by 1.77 percentage points in 1995, an increase of nearly 0.6 percentage points in the size of annual quality adjustments from the mid-1980s level (Smedley and Moulton 1997). The claim that the CPI has a quality bias is therefore not a claim that the quality of consumption items is improving through time. Rather, it is a claim that the quality of these items is rising more rapidly than the BLS is finding in its measures.

There are two issues that arise with new-goods bias. The first stems from the fact that new goods are not immediately incorporated into the CPI
survey. The sample rotation method in place from 1978 to 1998 would gradually have brought new items into the index on average over a five-year period. The new rotation system put in place in 1998 should typically get new goods into the index within a year of their appearing on the market. This delay could be important because new items tend to fall rapidly in price during the period just after they are introduced into the market. For example, products like hand calculators and cellular phones both had very rapid price declines in the first years after they appeared on the market. The failure to include this price decline in the index leads to an overstatement of the true rate of inflation.

While this can be a genuine source of bias in the index, there are two important points about the size of this bias that should be noted. First, for most items, the size of this bias is likely to be very small. Smedley and Moulton (1997) estimate the size of the bias that resulted from excluding cellular phones as 0.02 percentage points annually for the years from 1986 to 1994. While this bias is only from a single good, cellular phones were both an extraordinarily important innovation, and also, by chance, fell outside of the CPI’s normal rotation system. It is extremely unlikely that the new-product bias from the late entry of any other individual items will have an impact on the index of the same order of magnitude.

The other important point about this bias is that it is likely to be relevant primarily to the consumption patterns of wealthier households. When new goods are first introduced and are relatively expensive, most of the buyers are likely to be higher-income households. The CPI is an expenditure-weighted index, so that the consumption patterns of higher-income households have disproportionate weight in the index. Arguably, an index that was democratically weighted, with the consumption patterns of each household counting equally, would be a more appropriate price index for most of the uses of the CPI. Recent research has found relatively little difference between the inflation rate measured by an expenditure-weighted index and a democratically-weighted index (Garner et al. 1996; Kokoski 1987). However, if new-products bias causes the CPI to significantly overstate inflation, then a correct index would probably show a substantial difference between the inflation rate shown with expenditure weights compared with democratic weights. In this case, eliminating a bias that exists in the CPI as it is constructed, could actually make it less appropriate as a price index for the items purchased by typical consumers.

It is also important to note that the size of this bias has almost certainly been reduced through time. Prior to 1978, items could only enter the index at major revisions to the survey. The previous two revisions had taken place in 1964 and 1951. It is likely that this earlier system, where the BLS tracked the prices of an unchanging sample of goods and services, meant that there
would be several instances comparable to that of the cellular phone, where important consumption items did not appear in the CPI until long after they had been in common usage. For example, home air conditioners did not get included in the CPI until 1964, the same year that air travel was first part of the index. The fact that new goods entered the CPI so much more slowly in the years prior to 1978 probably means that the new-goods bias would have been considerably larger in those years than it was in the 1980s and 1990s, and far larger than it is with the methodology put in place in 1998.

There is one other aspect to the type of new goods bias cited by the commission that is worth noting. Following Hausman (1994), the commission cited increased variety itself, apart from any specific goods, as an unmeasured benefit to consumers that leads to an upward bias in the CPI. While increased variety may be a real source of gain, this is the sort of benefit that does not properly belong in a price index like the CPI. It is not measuring either the price or attributes of goods themselves. The benefits of increased variety would best be left to broader measures of the quality of life.

The following discussion examines the evidence produced by the Boskin Commission for quality and new-goods bias in each of the major categories of the CPI. Following the commission, this discussion uses the CPI categories in place in 1997. It summarizes a larger discussion in Baker (1997). Table 4.3 presents the amount of bias the commission estimated for each CPI component and the basis for the estimate.

1. **Food and beverages** The commission estimated that the CPI understates the quality improvements in this component by an average of 0.3 percentage points annually. This estimate was based entirely on introspection, the commission’s belief that the rate for food and beverages, including food away from home, was on average about 10 per cent better in 1996 than it had been in 1976.14

   How much would a consumer pay to have the privilege of choosing from the variety of items available in today’s supermarket instead of being constrained to the much more limited variety available 30 years ago? A conservative estimate of the extra variety and convenience might be 10 per cent for food consumed at home other than produce, 20 per cent for produce where the increased variety in winter (as well as summer farmers’ markets) has been so notable, and 5 per cent for alcoholic beverages where imported beer, microbreweries, and greatly improved distribution of imported wines from all over the world have improved the standard of living. (SFC 1996, pp.41–2)

   The commission’s discussion of quality bias in this component gave no evidence that it recognized the size of the quality adjustments already included in the CPI for food and beverages. For the year 1995,
### Table 4.3 Estimates of quality bias in the CPI

<table>
<thead>
<tr>
<th>Major and selected minor components</th>
<th>Relative importance</th>
<th>Current (1995) quality imputations</th>
<th>Commission’s estimate of bias</th>
<th>Source of estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Food and beverages</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food at home other than produce</td>
<td>8.54</td>
<td>0.3</td>
<td>Introspection</td>
<td></td>
</tr>
<tr>
<td>Fresh fruits and vegetables</td>
<td>1.34</td>
<td>0.6</td>
<td>Introspection</td>
<td></td>
</tr>
<tr>
<td>Food away from home</td>
<td>5.89</td>
<td>0.3</td>
<td>Introspection</td>
<td></td>
</tr>
<tr>
<td>Alcoholic beverages</td>
<td>1.57</td>
<td>0.15</td>
<td>Introspection</td>
<td></td>
</tr>
<tr>
<td>2. Housing</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shelter</td>
<td>28.29</td>
<td>0.25</td>
<td>Introspection</td>
<td></td>
</tr>
<tr>
<td>Other utilities, including telephone</td>
<td>3.22</td>
<td>1</td>
<td>Introspection</td>
<td></td>
</tr>
<tr>
<td>Appliances, including electronic</td>
<td>0.81</td>
<td>5.6</td>
<td>Gordon (1990)</td>
<td></td>
</tr>
<tr>
<td>Other house furnishings</td>
<td>2.64</td>
<td>0.3</td>
<td>Introspection</td>
<td></td>
</tr>
<tr>
<td>3. Apparel and upkeep</td>
<td>5.52</td>
<td>3.37</td>
<td>1</td>
<td>Gordon (1996)</td>
</tr>
<tr>
<td>4. Transportation</td>
<td>16.95</td>
<td>1.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New vehicles</td>
<td>5.03</td>
<td>0.59</td>
<td>Gordon (1990)</td>
<td></td>
</tr>
<tr>
<td>Used vehicles</td>
<td>1.34</td>
<td>0.59</td>
<td>Gordon (1990)</td>
<td></td>
</tr>
<tr>
<td>Motor fuel</td>
<td>2.91</td>
<td>0.25</td>
<td>Introspection</td>
<td></td>
</tr>
<tr>
<td>5. Entertainment</td>
<td>4.37</td>
<td>4.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commodities</td>
<td>1.98</td>
<td>2</td>
<td>Gordon (1990)</td>
<td></td>
</tr>
<tr>
<td>6. Other goods and services</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal care</td>
<td>1.17</td>
<td>0.9</td>
<td>Gordon (1990)</td>
<td></td>
</tr>
<tr>
<td>Personal and education expenses</td>
<td>4.34</td>
<td>0.2</td>
<td>Introspection</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.3 (continued)

<table>
<thead>
<tr>
<th>Major and selected minor components</th>
<th>Relative importance</th>
<th>Current (1995) quality imputations</th>
<th>Commission's estimate of bias</th>
<th>Source of estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Medical care</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prescription drugs</td>
<td>0.89</td>
<td></td>
<td>2</td>
<td>Berndt et al. (1993, 1997)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Griliches and Cockburn (1994)</td>
</tr>
<tr>
<td>Nonprescription drugs and medical supplies</td>
<td>0.39</td>
<td></td>
<td>1</td>
<td>none</td>
</tr>
<tr>
<td>Professional medical services</td>
<td>3.47</td>
<td></td>
<td>3</td>
<td>Cutler et al. (1996) and</td>
</tr>
<tr>
<td>Hospital and related services</td>
<td>2.26</td>
<td></td>
<td>3</td>
<td>Shapiro and Wilcox (1996)</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>1.76</td>
<td>0.612</td>
<td></td>
</tr>
</tbody>
</table>

these adjustments lowered the rate of inflation reported for this component by 1.39 percentage points (Smedley and Moulton 1997). If this year was typical, then the commission’s introspective assessment of the rate of quality improvement in this category would imply that the CPI was substantially overstating the rate of quality improvement and understating the rate of inflation. At the very least, the commission has provided no evidence that the actual rate of quality improvement for food and beverages was more rapid than indicated by the CPI.

2. Housing The commission estimated that the shelter component of the CPI, the largest component of the housing category, understated the rate of quality improvement by 0.25 percentage points annually. It based this estimate on the fact that the average unit rent in the Census Bureau’s housing survey had risen by 1.0 percentage point a year more than the shelter index in the CPI, over the previous 20 years. It noted that the average size of a unit had risen by 1.0 per cent annually during this period. Based on this calculation it concluded that the index had accurately measured the increase in housing size, but had missed improvements in average housing quality, which it estimated at 0.25 percentage points annually.

As with the estimate of bias in the food component, it is difficult to evaluate a claim based on introspection. It is worth noting that much of the cost of housing depends on location. There was a nationwide shift of population during this period from the northeast, a relatively expensive region, to the south, the least expensive region. There was also a large shift of population from inner cities and inner suburbs to relatively inexpensive outer suburbs. This means that an index that held location constant probably would have risen far more rapidly than the Census Bureau’s index. The Boskin Commission’s estimate also depends on the counterintuitive assumption that square footage is priced at a constant marginal cost.

Among other items in the housing category, the commission assumed that there was a 1.0 percentage point bias in the reporting of telephone prices based on its belief that there were unmeasured improvements in the quality of telephone service over the last two decades. It is worth noting that the CPI would pick up the price declines associated with the introduction of competition to the industry, but none of the search costs that consumers face as a result of the fact that there is no longer a monopoly in the industry.

The commission also estimated a substantial amount of quality bias in the pricing of household appliances. It relied on Gordon (1990) for this estimate. In fact, the evidence in this study is far more ambiguous. Compared to one of the indices constructed, the CPI does show a
substantial bias, however, when compared to an index constructed from Sears catalogue price quotes, the CPI actually understated the true rate of inflation for much of the period examined. This study cannot be viewed as providing convincing evidence of a quality bias in the CPI in this area.

3. Apparel The commission estimated a quality bias equal to 3.37 percentage points a year in the apparel component based on a study that compared the CPI index with an index based on Sears catalogue prices (Gordon 1996). The CPI apparel index had serious problems in this period due to formula bias. However, it is not clear that there were also problems with quality bias. In the period after 1989, when the formula-bias problem in apparel was corrected, the CPI index actually showed a lower measured rate of inflation than the Sears catalogue index. Therefore, the Gordon study provides little reason to believe that the CPI measure of apparel prices is biased due to unmeasured quality improvements.

4. Transportation The commission estimated that the CPI understated quality improvements in new cars by 0.59 percentage points and used cars by 1.59 percentage points, based on its belief that it was not picking up improvements in durability. The commission noted the findings of Gordon (1990), who found that the CPI and his own hedonic index tracked each other reasonably well. However, it noted that the Gordon study did not include durability as a characteristic, and therefore concludes that it would have understated the true rate of quality improvement in cars. In fact, in its methodology, the CPI would factor in additional costs associated with improvements in durability. The claim that these quality improvements are somehow being missed by the BLS really lacks any clear evidence.

The commission also concluded that the CPI overstated the rate of inflation for gasoline by 0.25 percentage points annually since it did not factor in the benefits of pay-at-the-pump credit-card readers. It estimated the benefit from these readers at 5 per cent of the purchase price, and obtained the 0.25 per cent annual inflation bias estimate by assuming a 20-year phase-in period. The commission had no evidence of the actual value that consumers place on this convenience, nor did it factor in the decreased availability of gas stations, a fact noted in its report, when assessing the net change in convenience to consumers.

5. Entertainment The commission attributed an annual quality bias of 2.0 percentage points to the items, such as VCRs and televisions, that were included in the CPI’s entertainment component. This estimate was also based on estimates of bias in Gordon (1990). It is worth noting that the BLS has recently introduced hedonic pricing for most
major appliances. Contrary to the commission’s expectation, this has had almost no effect of the measure of inflation in these items, and in some cases it may actually have led to a slightly lower measured rate of inflation than the quality adjustment system previously in place (see Liegey and Shepler 1999; Liegey 2000a,b,c; Shepler 2000; Kokoski et al. 2000; and Thompson 2000).

6. Other goods and services

The commission attributed a 0.9 percentage point annual bias in the personal-care component of the index, based on its estimate of the CPI’s large undercounting of the rate of quality improvement in personal-care items, such as hairdryers. This estimate was also based on Gordon (1990). As noted above, recent research from the BLS indicates that there never was a large amount of quality bias in the BLS measure.

7. Medical care

The commission attributed an annual quality bias of 3.0 percentage points to the medical-care component of the index. This estimate relied largely on two studies that examined distinct sets of issues. The first study examined the cost of treating cataracts through time and compared it to a hypothetical CPI index (Shapiro and Wilcox 1996). It found that their index measuring the actual cost of treatment would have fallen by more than 3.0 percentage points annually measured against the hypothetical CPI index. The main reason for the difference is that the CPI index is focused narrowly on specific procedures, such as the cost of the surgery, and specific services, such as the price of a hospital room for a night. The actual cost of treating cataracts declined sharply during the period examined because it went from a type of surgery that generally required one to two weeks of hospitalization to a procedure that could be performed on an outpatient basis. The gains from this change would have been completely missed by the BLS procedures in place at the time.

While this study did expose a serious flaw in BLS methodology in pricing medical procedures, it is not clear that its impact was anywhere near as large as the findings of this study suggest. The BLS actually adopted a procedure similar to the one used in the Wilcox and Shapiro study in January 1997. There was no noticeable break in the medical services index as a result of this change. (The medical services component increased by 3.05 per cent in the year prior to January of 1997 and by 2.80 per cent in the subsequent year.) In other words, the Boskin Commission’s assessment, that the different methodology would lead to a far lower measured rate of inflation, was not borne out by the evidence. Any bias that may have been introduced by a flawed procedure was clearly far smaller than the 3.0 percentage points annually assumed by the commission.
The second study cited by the commission in its estimate of bias in the CPI’s medical-care component raises a qualitatively different set of issues. The study, Cutler et al. (1996), constructed a price index for treating heart attacks based on the years of expected life following a heart attack. This sort of quality adjustment brings in considerations that take it beyond a narrowly defined price index. For example, if other factors change that lead to a reduction in life expectancy following a heart attack (for example, more traffic deaths or a rise in the homicide rate), this methodology would imply that heart attack treatment should show a quality decline and therefore an increase in the price of treatment.17

As noted earlier, the pricing of medical care raises some very fundamental problems. In addition to the problems mentioned in the previous sections, there are also some attributable to fact that most private medical expenses are paid through insurance, and not directly by individuals. Since the CPI prices specific treatments, and not insurance premiums, it is possible that individuals’ payments could increase or decrease in ways not picked up by the CPI. For example, if the cost of a typical premium rises because the average person receives more care, this would not be picked up by the CPI as an increase in medical-care costs.

This methodology can lead to some serious anomalies. The nation has a large uninsured population, many of whom are in poor health and therefore cannot afford standard insurance policies. If the government were to insure this population by directly paying for their insurance, this cost would be directly visible in the form of the necessary government expenditures and the taxes to pay for this spending. However, as an alternative, the government could require that all insurers use a system of community rating, under which all individuals pay the same premium, regardless of their health. In this situation, the average cost of insurance would almost certainly rise, since the average health of beneficiaries would have deteriorated, but there would be no price increase reported by the CPI’s medical-care component as a result of this change. Clearly, this asymmetric treatment of these two policies is inappropriate.

There is a simple alternative approach that gets around this problem. Until 1964, the CPI simply priced the cost of medical insurance premiums, rather than the cost of individual procedures. This method has the advantage of simply measuring what individuals actually pay for their medical care.18 It does not evaluate the quality of the care they are receiving. In effect, this methodology would measure the rate at which individuals’ spending on health care increases independent of any
assessment of what they are getting for their expenditure. The latter can better be measured by examining data on health. Presumably what people value is their health – the ability to live a long and healthy life. The quality of the medical services they receive is a secondary consideration.

By coincidence, an index that tracked the increase in out-of-pocket health-care expenditures and insurance premiums would have shown roughly the same rate of inflation as the CPI’s medical-care component over the 1970–91 period: 8.12 per cent annual growth in per capita out-of-pocket health spending and 8.17 per cent annual growth in the medical-care index in the CPI (Cowen and McDonnell, 1993, p. 228). This is due to the fact that a shift to employer and government-paid health care largely offsets the more rapid increase in insurance premiums over this period. However, there is no reason to expect that in general this will be the case, so in the future there could be substantial divergences between these two measures. The CPI will be a more meaningful price index if it measures what people pay for their health care – a reasonably straightforward task – as opposed to trying to assess the quality of the health care they receive.

In sum, there is remarkably little evidence to support the claim of a substantial quality bias in the CPI in the period examined by the Boskin Commission. More than a quarter of the quality bias they attributed to the index was based on nothing more than causal reflection. There was no research to support their assessment. Most of the other estimates of bias involved dubious extrapolations from a very limited body of research. In several of these cases, such as the hedonic pricing of appliances and the new procedures for pricing medical care, subsequent BLS research and changes in the index itself have shown that the commission’s estimates of bias were grossly exaggerated.

There is one last point worth noting on the issue of quality adjustment. All quality adjustments, in trying to price the gains (or losses) from the changes in quality, assume that consumers are homogeneous and therefore would all place the same value on these changes (for example, Griliches, 1971; Lancaster, 1977). This can be an important source of overstatement of the size of quality improvements, since many consumers may value improvements little, or possibly not at all. But consumers generally do not have the option to buy the ‘unimproved’ version of the product, so they end up paying for the quality improvement regardless of whether they want it or not.

For example, the hedonic methodology now in place for pricing computers has consistently shown quality improvements in computers of more
than 30 per cent annually. As a result, there has been a reported rate of price decline in some years of close to 40 per cent. The rate of price decline reported for computers over the last six years would have reduced their cost by a cumulative total of more than 90 per cent since 1995. This would imply that a 1995 computer, equipped with fully web capable Windows 95 software, could be purchased today for $150. While some consumers may prefer this option to paying three or four times this price for a much better computer, it is unlikely that they could purchase a new computer for this price.

The same situation would apply with most of the quality improvements in goods and services. The treatment of quality improvements can be seen as exactly analogous to a situation in which distinct items were sold in indivisible bundles. For example, if bacon and eggs were always sold together, consumers would have to pay for both items even if they only wanted one. Since the characteristics that are falling rapidly in price (for example, speed and memory in computers) are likely to be the ones that get added in most to products through time, the standard methods of quality adjustment are likely to lead to somewhat of a downward bias in the index. If these quality adjustments increase through time, as has almost certainly been the case, the size of the downward bias would increase.

**Evaluating CPI Bias through Aggregate Measures**

It is worth noting another strand in the literature on CPI bias, which was not mentioned in the Boskin Commission’s report. Several researchers have sought to argue that the index contains an overall bias by examining measures of aggregate living standards, rather than by finding specific areas where a bias exists. For example, Nordhaus (1998) compared changes in real median family income with positive economic assessments given by respondents in the Michigan Survey of Consumer Confidence. He showed that a CPI bias of approximately 1.0 percentage points would support the conclusion that most people will give a positive assessment of the economy in years in which the real median family income is rising. Kreuger (1998) uses a more complete analysis of the income distribution to show that an assumption of zero bias in the CPI is consistent with most families providing a positive assessment when their real income is rising.

Nakamura (1995) uses very broad consumption categories in an Engels curve analysis to argue that the shift in consumption shares between 1978 and 1994 is consistent with an upward bias in the CPI of more than 1.5 percentage points annually. This analysis ignores changes in relative prices and also the fact that most medical care is not paid for directly by consumers, with approximately half being paid by the government. Most of the shift
in shares found by Nakamura was attributable to the growth in medical-care expenditures as a share of total consumption expenditures during this period.

Costa (2000) makes a similar argument based on the decline in food and entertainment shares in total consumption. However, Bils and Klenow (2001a) show that this argument depends on using only these two broad categories in the analysis. When they replicated the methodology using 106 more narrowly defined consumption categories, their results showed a significantly slower rate of share shifts than would be predicted by the measured rise in income, thereby implying that the CPI understates inflation. They therefore reject this sort of Engels curve analysis.

This study instead notes that the largest increase in spending shares took place in the CPI categories with the greatest rate of noncomparable product substitution in the CPI survey. The study considers the rate of noncomparable product substitution a measure of a sector’s dynamism. Therefore, they take their result to imply that consumers are switching their spending to the sectors which have the greatest dynamism in terms of new product offerings. They note that these are also the categories of the CPI that show the largest rate of price increase, and infer that the CPI is missing much of the quality improvement in goods in its substitution process.

This analysis is carried further in Bils and Klenow (2001b). This study constructs measures of quality income elasticity for 66 categories of durable goods based on an analysis of the price of the goods that households reported purchasing in the CES over the years 1980–96. The study found that these quality elasticities were good predictors of the rate of price change in goods over this period, with the goods with the highest-quality elasticities showing the most rapid rate of price increase. The study then regressed the predicted rate of price increase on the rate of price increase reported for these categories in the CPI to get an estimate of unmeasured quality improvement and therefore CPI overstatement of inflation. This exercise produced an estimate of an average overstatement of inflation in the CPI of 2.2 percentage points annually in the goods examined.

While this analysis is a rather ingenious use of the data in the CES, there are some obvious problems with the methodology. First, the analysis would pick up all forms of bias in the CPI’s measure of inflation in these goods, not just quality bias. For example, any bias attributable to the CPI’s sample rotation procedure during this period would be picked up as quality bias. Similarly, if consumers were achieving savings on goods as a result of buying at discount outlets, this would be picked up as quality bias in this analysis, even if the savings were attributable to actual declines in service quality.

A more serious problem is the possibility that the measures of quality
income elasticity may be biased by the actual price increases in the goods in the study. The CES has a relatively small cross-section sample of 5000 households a year. This means that much of the income variation in the study would be taking place through time as reported incomes rise, rather than simply across time. This would be especially true of goods that are infrequently purchased, such as playground sets, since the cross-sections of the households purchasing these items in any given year are small. It is worth noting that most of the goods categories with a quality elasticity that it is at least one standard deviation from the mean calculated for the sample, have sample sizes that are smaller than the median (15 of 22), exactly the relationship that would be predicted if the measures of quality elasticity were biased by actual changes in price during the period.

It is also worth noting that there is a simple alternative explanation to the correlation between categories with an increase in expenditure shares and the number of noncomparable substitutions, found in Bils and Klenow (2001a). If firms tend to alter products in ways that make price increase less apparent to consumers, then it would be expected that categories of goods that had the largest increases in prices, would also have the most instances of product substitutions in the CPI. If demand for goods generally is inelastic, then these categories would also experience the largest increase in expenditure shares.

In conclusion, the evidence in these studies that the CPI has overstated the true rate of inflation in consumer goods is at best questionable. Furthermore, it would be difficult to know how to assess the meaning of evidence of an overall overstatement of inflation, apart from evidence of bias in the measure of inflation in specific categories of goods and services. If there are gains to well-being not directly embodied in the consumption basket, for example as a result of increased variety, this would be an issue that gets beyond a measure of prices. As argued earlier, making this sort of larger assessment of well-being requires an evaluation of a broad set of issues that are far beyond the scope of the CPI as it is presently constructed. In short, these aggregate studies can be suggestive of the existence of a bias and the possible mechanisms responsible. However, before this evidence can be taken as conclusive, it is necessary to actually show where the bias is affecting the measurement of inflation in specific areas of the CPI.

MEASURING CONSUMER PRICES: IS THE CPI THE RIGHT YARDSTICK?

The CPI has had a number of serious flaws through time. These include problems of formula bias, an inaccurate treatment of owner-occupied
housing, and varying degrees of attentiveness to quality change through time. While these problems make it less than an ideal measure of the rate of inflation, as the previous discussion demonstrates, there is no compelling evidence for the claim that it has a large upward bias in its measure of the rate of inflation. It is worth noting that almost none of the economists or policy experts who have advanced this argument has been willing to embrace the implications of this view for other areas of economics or for policy debates – other than indexing benefits and tax brackets.

The CPI-U-RS series constructed by the BLS corrects many of the problems in the CPI in the period from 1978 to 1998. It provides a relatively consistent measure of inflation over this period. Unfortunately, this series has not been applied to years prior to 1978. The largest and most consistent difference in the inflation rate reported with the methodology used in the CPI-U-RS and the CPI-UX1 is a gap of approximately 0.2 percentage points annually due to the use of geometric means to aggregate inflation rates at lower levels in the CPI. To maintain as consistent series as possible, it would be appropriate to subtract 0.2 percentage points from the rate of inflation reported in the CPI-UX1 in years prior to 1978.

It is likely that the other changes that result from using the CPI-U-RS methodology would be of less consequence, and would be harder to backcast through time. The CPI-U-RS does not treat pollution abatement costs as quality improvements, as the BLS did in the CPI prior to 1998. Extending the RS series backward before 1978 would have led to some increase in the measured rate of inflation in some items, notably autos and gasoline, since additional costs attributable to catalytic converters and lead-free gas were treated as quality improvements in the CPI. However, in general, less attentiveness to quality improvements during this earlier period and the fact that new products were never brought into the index, except during revisions to the CPI, should have caused the CPI to show a somewhat higher measured rate of inflation. In other words, for periods prior to 1978, it is likely that the CPI was higher relative to the true rate of inflation than it has been during the more recent period, although there is little basis for an accurate measure of this difference.

While there is a widely held view that greater product variety, increased innovation, and an increased focus on quality by producers has had a greater impact in the last two decades than in prior periods, there is little real evidence for this position. The prior three decades also saw the spread of extremely important innovations such as home air conditioning, television, and air travel. There were also large improvements in product quality, as shown in an earlier generation of studies of the CPI (Griliches 1961; Triplet and McDonald 1977; Gordon 1990). In addition, there was also a significant increase in the variety of goods and services available in the
years prior to 1980. It is important to recognize that the issue here, even in the context of a cost-of-living index, is the value to consumers of increased variety, not any absolute measure of the rate of growth of variety. For example, there is no a priori reason to assume that an increase in the number of breakfast cereals from 40 to 400 would provide more benefit than an increase from 20 to 40. It is entirely possible that relatively small increases in the number of items available when the selection is very small, provide more benefits to consumers than large increases in a period in which there is already significant variety.

It is also worth noting that the size of the CPI's substitution bias has varied over time, depending on the rate and evenness of inflation. The substitution bias would almost certainly have been significantly larger in the late 1960s and 1970s than it was through most of the 1950s, 1980s, and 1990s. Ideally, a measure of the rate of inflation should take this into account. Unfortunately, there is not a consistent series measuring substitution bias in the CPI for years prior to 1982. It would be a valuable project to attempt to piece such a series together, with the best data presently available.

The Boskin Commission rushed out its report in the heat of a political battle over the budget deficit. As the members of the commission acknowledged, their estimate of a 1.1 percentage point bias involved a considerable degree of speculation, since there was relatively little hard evidence available. In the last five years, new studies, in addition to changes in the index itself, have provided a firmer basis for assessing bias in the CPI. The surviving members of the Boskin Commission recently estimated that the changes implemented in the last five years have had the effect of lowering the measured rate of inflation by approximately 0.3 percentage points annually (General Accounting Office 2000). The BLS had a sound basis for implementing most of the changes that led to this lowering in the measured rate of inflation. There is very little economic research that can be cited to support the rest of the 0.8 percentage points of bias identified by the commission.

In short, researchers can feel reasonably comfortable in using the CPI-U-RS series as a measure of the rate of inflation faced by consumers in the years since 1978. The CPI-U-X1 series for the years prior to 1978 does not use the same methodology and therefore may show a rate of inflation that is slightly higher (approximately 0.2 percentage points annually) relative to the true rate of inflation. However, it is important that economists be humble and recognize that these are measures of price inflation, not the cost of living. The construction of a true cost-of-living index would require procedures that are well beyond what economists have developed to date. Those who are interested in the question of whether there has been an
increase or a decrease in well-being over time will want to use measures of real wages and incomes as starting points, but they will also have to turn to other indices – life expectancy, the availability of leisure time, crime rates, and so on – to get an accurate view of social progress. Economists do not have a simple number that will provide an answer to this question.

NOTES

1. Whether the CPI should be thought of as a cost-of-living index is itself a matter for debate. The official documentation for the CPI describes it as a price index for consumption goods, as opposed to a cost-of-living index (BLS 1995; Fixler 1993). The former would not imply any claims about a direct relationship between the measure shown by the index and consumers’ well-being, whereas the latter is defined exclusively in reference to a measure of consumers’ welfare.

2. It is possible that people switch to consuming other fluids as well, because they view the drinking water as being unsafe.

3. Two researchers at the Texas Transportation Institute estimated that average commuting cost per family had increased by $85 a year (1993 dollars) between 1986 and 1993 (Schrank and Lomax 1996). This is equal to approximately 0.3 per cent of the average family’s annual expenditure.

4. In its assessment of the accuracy of the CPI as a measure of the cost of living, the Boskin Commission briefly examined trends in crime rates over the previous three decades. This at best only gives part of the picture. Insofar as crime has gone down, or remained stable in part as a result of additional spending on security measures, such as alarm systems or security guards, the CPI would be missing a rise in the cost of living. More importantly, if people have been forced to change their behavior as a result of crime, for example not taking evening walks, then the CPI would be missing an important source of deterioration in living standards, and by implication, a large increase in the cost of living.

5. It is difficult to know exactly how much of this spending is an increase from prior years. The personal business expenditure category rose as a share of total consumption expenditures in the NIPA from 3.6 per cent in 1953 to 9.4 per cent in 1999 (NIPA Table 2.4). But this includes spending that is explicitly investment related, such as brokerage fees, which does not belong in a consumption index.

6. This issue has been examined to some extent by sociologists, perhaps most notably by Hirsch (1976).

7. It is interesting to note that Krugman (1996) asserted that status considerations are the primary determinants of well-being in the context of the debate over the accuracy of the CPI as a measure of the cost of living. This assertion was an effort to reconcile the fact that a substantial overstatement would imply widespread poverty (by the current measure) in the United States in the very recent past. Krugman argued that although people in the 1950s and 1960s might have been very poor by current standards, most people did not perceive themselves to be living in poverty because they were keeping pace with the prevailing living standards of the time.

8. Baker (1996) constructs a historical index dating back to 1953, although it relies on the methodology that was in use in the CPI at the time.

9. Actually a Laspeyres index does not necessarily show a higher rate of inflation than a superlative index for any specific period. Over time, it should be expected that an index that allows for substitution will rise less rapidly than an index that holds the consumption bundle fixed at the base-year levels, but in any specific year, it is entirely possible that the prices of the goods and services consumed in the base year will rise less rapidly than the goods and services actually consumed that year. The BLS was recently given funding so that it could update its sample every second year, instead of at 10–15-year intervals.
as had been the case in the past. While this more frequent updating is desirable from the standpoint of producing a more accurate index, it appears as though Congress may have approved the funding under the mistaken view that a newer basket will generally report a lower rate of inflation, and therefore reduce spending on indexed programs (see SFC 1995; General Accounting Office 1997).

10. In CPI strata where the sort of substitution implied by geometric means seemed implausible, such as medical procedures, the Laspeyres methodology continues to be used.

11. The assumption of a 10 per cent pure price difference is quite large, since it means that after adjusting for any quality differences, discount store prices average 10 per cent less than prices in traditional stores, yet enough consumers still shop at the traditional stores to keep them in business.


13. Cellular phones were not rotated over a five-year cycle as ordinarily would have happened, apparently because the CPI specialists in electrical appliances had assumed that cellular phones would be covered under phones, whereas the specialists in the phone division expected them to be covered as an electrical appliance.

14. The commission did cite Hausman (1994) for evidence of the benefit of increased variety in breakfast cereals – a benefit that it was argued does not properly belong in a price index.

15. Gordon (1990) is by far the most important source for the commission’s estimates of quality bias in various categories of the CPI. This study contrasted inflation rates measured by the CPI with hedonic indices constructed for a variety of durable goods using price data and descriptions from Sears catalogues and Consumer Reports. Depending on the good, the series go as far back as the late 1940s, and they end in 1983. The most striking feature in most of the estimates is the reduction in the gap between the Sears catalogue index and the CPI over the period. In most cases, the CPI rose substantially against the Sears index in the 1950s and early 1960s, whereas the two indices tend to move at close to the same rate after the late 1960s. In fact, in the years after 1970, the Sears index shows a slightly higher rate of inflation than the CPI for the goods examined in the study. For this reason, it provides a rather dubious basis for the commission’s estimates of quality bias.

16. The CPI methodology for evaluating the quality improvements in new cars relies on cost estimates of improvements provided by the manufacturers. It seems unlikely that this method would lead to an understatement of quality improvement.

17. The Boskin Commission cited several papers to support its assessment that the CPI overstated the rate of inflation in prescription drugs by 2.0 percentage points annually in the years after a change in procedure in 1995, and by 3.0 percentage points in prior years (Griliches and Cockburn 1994, 1996; Berndt et al. 1996). The main argument of these papers was that new classes of drugs being developed were qualitatively better than the existing drugs in ways that were not being picked up by the CPI. The new generation of anti-depressant drugs was the prime focus of two of these studies. Subsequent BLS research found that the impact of the change in BLS procedures instituted in 1995 on the rate of inflation in the prescription drug component, was only 0.4 percentage points annually, less than half the size estimated by the commission (Stewart and Reed 1999). More important for the commission’s claim of unmeasured quality improvement in drugs has been new medical research which found that the new generation of antidepressants was no more effective than the earlier generation, and possibly no more effective than placebos (Kahn et al. 2000).

18. The CPI only measures the cash expenditures by individuals for medical care, it does not attempt to measure the payments by employers to insurers or from the government. The Boskin Commission argued that all medical expenditures should be included in the CPI, which would approximately double their weight in the index. This change would make the index less accurate as a deflator of cash expenditures by households, which is its main purpose.
19. This figure combines the price decline in the GDP deflator for computers in 1995–97 with the decline in the CPI computer component since 1998, because prior to that date, the CPI did not use a hedonic pricing method for computers.

20. The BLS has not kept records on the extent to which measured quality improvements have lowered the reported rate of inflation each year, but it did have data that showed that quality improvements lowered the reported rate of inflation by an average of 1.17 percentage points in the years 1983 and 1984, as compared to 1.77 percentage points in 1995 (Armknecht and Weyback 1992; Smedley and Moulton 1997).

21. Noncomparable substitutions arise in the CPI when a price collector cannot find the identical product, and no replacement product is considered close enough to be treated as identical for purposes of the survey. It is therefore necessary to make an implicit or explicit quality adjustment to assess the quality difference between the two items.

22. Quality elasticity is defined as the percentage increase in the average price of the goods purchased within a category (conditional on a purchase being made) relative to a 1 percentage-point increase in income.

23. If the price of a good rises more than average, this would lead to a higher measured rate of quality income elasticity, since the higher reported nominal incomes in later years would be associated with larger price increases. By contrast, in the case of goods that had relative price declines, the higher reported incomes in later years would be associated with the purchases of less expensive items, thereby biasing the estimate of quality elasticity toward zero. This bias would be limited if the study properly controlled for inflation in its measure of household income over the period, but it is questionable whether this is the case.

24. The convergence of the rates of inflation shown by Gordon’s (1990) Sears catalogue indices and the CPI suggests a diminishing degree of quality bias during the period from the 1950s to the late 1970s.

25. Zvi Griliches, one of the members of the commission, died in 2000.

REFERENCES


Berndt, E., I. Cockburn and Z. Griliches (1997), ‘Pharmaceutical innovations and...


Moulton, B.R. (1996a), ‘Bias in the consumer price index: what is the evidence?’, 
Moulton, B.R. (1996b), ‘Estimation of elementary indexes of the consumer price 
aggregates of the CPI’, unpublished paper, Washington, DC: Bureau of 
Labor Statistics.
Moulton, B.R., and K.E. Smedley (1997), ‘Addressing the quality change issue in 
Perspectives*, 12, 1 (winter), 59–68.
Oi, W.Y. (1992), ‘Productivity in the distributive trades: the shopper and econom-
ies of massed resources’, in Zvi Griliches (ed.), *Output Measurement in the Service 
Sector*, National Bureau of Economic Research Conference on Research in 
price index’, in M.F. Foss, M.E. Manser and A.H. Young (eds), *Price 
Measurements and Their Uses*, National Bureau of Economic Research 
Conference on Research in Income and Wealth, vol. 57, Chicago: University of 
cost of living indexes’, Bureau of Labor Statistics working paper no. 261, 
Volume 2: Methodology and Urbanized Area Data*, College Station, TX: Texas 
Transportation Institute.
Senate Finance Committee (SFC) (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: US 
Senate.
Senate Finance Committee (SFC) (1996), *Toward a More Accurate Measure of the 
Cost of Living*, final report to the Senate Finance Committee from the Advisory 
Shapiro, M., and D. Wilcox (1996), ‘Mismeasurement in the consumer price index: an 
Shepler, N. (2000), ‘Developing a hedonic regression model for refrigerators in the 
Smedley, K., and B. Moulton (1997), ‘Replication of Armknecht–Weyback analy-
Labor Statistics.
Stewart, K., and S. Reed (1999), ‘CPI research series using current methods, 
Triplet, J., and R. McDonald (1977), ‘Assessing the quality error in output mea-