Policy Implications of the Boskin Commission Report

Martin Neil Baily¹
Institute for International Economics

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

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

The Boskin Report suggested that further work be done to provide more robust estimates of the biases in the CPI and, subsequent to the report, it was suggested that a new National Academy of Sciences/National Research Council panel should examine in greater depth the extent of CPI biases. Both wishes came to pass. Several studies have appeared evaluating the sources and magnitude of biases in the CPI (see for example, Shapiro and Wilcox (1997), Hausman and Leibtag (2004) and Nordhaus (1997)). And an NAS/NRC panel chaired by Charles Schultze released its report in 2002 (Schultze and Mackie (2002) and Schultze (2003) for a summary). The Schultze panel was more cautious in its findings than had been the Boskin Commission, concluding that the basis for estimating biases in the CPI remained rather tenuous and that the BLS should not change its collection and estimation procedures substantially (beyond what had already been done) until more research had been completed. In particular, the panel recommended against a major increase in the use of hedonic estimates in the short run. The conservatism of the Schultze panel may have been overstated, however, ¹ The author is Senior Fellow at the Institute for International Economics. From 1998 to 2000 he served as Chair of the U.S. Council of Economic Advisors. This article is based on a presentation to the panel session “The Boskin Commission Report After a Decade: Is the CPI Still Biased?,” held at the annual meeting of the American Economic Association, January 6-8, 2006, Boston, Mass. Email: mbaily@iie.com
because it offered support for the approach to
hedonics formulated by Ariel Pakes (2002). A
main objection to hedonics as currently used is
that the coefficients in the hedonic regressions
are unstable and may be incorrect for the pur-
pose. Pakes is able to explain why these coeffi-
cients are likely to be unstable and could be
wrong, and offers a methodology for making
better hedonic estimates. The reluctance of the
Schultze panel to expand the use of hedonics
reflects their agreement with Pakes’ findings,
while at the same time believing that BLS lacks
the resources to apply the Pakes approach
widely.

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

He was being too kind to physicists and too
hard on economists. I took undergraduate
classes on quantum mechanics and general rela-
tivity many years ago and I do not remember
back-of-the-envelope estimates being used to
make the results of these theories more intuitive.
To understand modern physics one has to slog
through heavy-duty mathematics because the
material makes no sense at all based on everyday
experience. And on the other side, economics is
not entirely devoid of intuitive estimates. But,
broadly speaking, I agree with his criticism of
economics. Too often economists develop elab-
orate models without exploring whether the
results really make sense or are consistent with a
range of different data. Good back-of-the-enve-
lope estimates are a valuable tool for economics
as well as physics.

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

In its assessment of the quality bias, the
Boskin commission suffered from premature
extrapolation — moving too quickly from a lim-
ited number of specific examples to a broad con-
clusion. I do not think it would be necessary to
estimate quality change for every good and service in the CPI to get a good estimate of overall quality change. But it is necessary to draw a broad sample that is representative of the overall composition of the index.

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

There was a particular problem in the area of quality change and new goods. Jack Triplett has emphasized over the years that the CPI already accounts for quality change because when a new model is introduced at a higher price, 100 per cent of this price increase is attributed to quality improvement (see his paper in this symposium). Given the evidence currently available, it is not possible to say for sure if there is a significant positive or negative bias in the CPI as a result of quality change. The Boskin commission should have done more to estimate not only how large is the change in quality over time, but also how large is the quality change that is already being captured in the CPI. Michael Boskin et al. (1998) have argued that even though there is considerable uncertainty around estimates of the amount of quality change, it is better to make an estimate of the amount of such change than simply to assume a zero effect. But that is not a valid argument. The origin is not at zero and the real question is whether we know that current indexes over or underestimate quality change. My own intuition is that the CPI probably does understimate quality improvement, but intuition is no substitute for more facts and analysis.

**Outlet Substitution Bias**

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

The price collection techniques used by BLS do not capture the benefit to consumers of this consumer-driven evolution of the retailing sector. An impressive study by Hausman and Leibtag (2004) finds that outlet substitution in grocery retailing could account for a significant overstatement of inflation in the food-at-home category of the CPI. And since the evolution of retailing applies much more broadly than just to groceries, this study could potentially be extrapolated to a significant fraction of the CPI. In
addition, the productivity studies of retailing by
the McKinsey Global Institute (see Baily and
Zitzewitz (2001) for a discussion) suggest that
this sector is contributing to overall productivity
growth and it seems that this is not being cap-
tured in the current CPI measures.

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

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

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

Policy Implications

Better Allocation of Spending
on Statistics

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

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

In my judgment, it would also be helpful for
the United States to create a unified economics
statistical agency, like Statistics Canada. This
would allow a streamlining of the data collection
and analysis process. One counter-argument is
that having several different agencies creates
competition, increasing performance. That
argument looks weak. There is no competitive
market among the different agencies. It would
be much better to have a single agency with a
system of outside review of the agency’s performance. Again the problems here are political. Each cabinet office guards its own turf and protects its statistical arm. We need someone with the clout to reorganize US statistical operations and reallocate the budget in a way that more closely matches the actual economy.

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

**CPI and the Budget**

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

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

Now that we are back in an era of endless budget deficits, there is a search on to find ways to cut spending. So far, I am not aware of serious efforts to change the way federal pensions or tax brackets are indexed by moving to CPI minus X per cent. But there are proposals to change Social Security, and so modifying the indexing procedure may be put on the table.

**Using the CPI to Adjust Social Security Payments**

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

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

---

2 The figure of 22 per cent comes from the website of the Social Security Administration. T. Lynn Fisher (2005) has pointed out, however, that some other sources of income are excluded in this calculation, most notably drawings from defined contribution pension plans unless they are based on an annuity. This is a potentially serious problem.
that average Social Security benefits have been driven up by over-indexing beyond the level that would be seen as a minimum income needed to avoid poverty.

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

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

Despite the fact that CPI indexing seems to be doing a pretty good job of keeping the benefits of the very elderly at a level comparable to the level of those just retiring, there has been a lot of interest in indexing among policymakers in Washington DC, driven by a desire to reduce the level of Social Security payments in the years ahead by reducing the rate of growth of the CPI. This is driven in turn by the fact that the aging of the population will increase the costs of the system, leaving it insolvent at some future time unless taxes are increased or benefits cut. I am not as averse to tax increases as are most Americans, although I am quite averse to increases in payroll taxes, because they distort work incentives and fall heavily on low-wage workers. But even if tax increases were ruled out, it would be possible to find other ways to make social security solvent, for example by giving less to those with strong private pensions or other assets, while giving the same, or even a bit more, to those that rely solely on Social Security.

The Impact of Innovation on the Level of Social Security Benefits

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

It is worth picking up on this health care issue both because it is such a large fraction of GDP and because it illustrates a more general point. An important issue for indexing Social Security benefits is to determine the extent to which new and improved goods and services represent, on the one hand, a true reduction in the amount of
money needed to maintain a minimum living standard and, on the other hand, the extent to which they represent new opportunities for consumption that are realized only with a higher level of income. In the former case, the minimum level of income is lowered by the innovation. In the latter it is not, and may well be increased.

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

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

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

In practice, I do not think our political system is capable of making good objective decisions about Social Security payments on a regular basis. Congress would be unwilling to delegate to a technical group the power to raise benefit payments, given that there are budget implications of that decision. And it would be a nightmare to have to make adjustments every year following Congressional debate and decision. That means we are very much in a second or third-best world. Some form of wage indexing has considerable appeal, as it would keep the minimum living standard in relation to the incomes of those still working. However, this would generally involve larger benefit increases
than under CPI indexing and, given the solvency problems of the Social Security system, such a change looks politically infeasible. Faute de mieux I would stick with the current CPI adjustment rather than moving to CPI minus X per cent. The very old are getting about as much as recent retirees, which seems about right. So, like policymakers in the 1990s, I would not try to balance the budget “on the backs of the very old.”

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

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

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

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

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

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

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

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

**Conclusion**

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

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

**References**


