CENTRE FOR THE STUDY OF LIVING STANDARDS
CONFERENCE ON SERVICE SECTOR PRODUCTIVITY
AND THE PRODUCTIVITY PARADOX

April 11 and 12, 1997

Chateau Laurier Hotel, Ottawa, Canada

Luncheon - Banquet Room - April 11, 1997 1:00 p.m. - 1:30 p.m.

Luncheon Speaker: **Zvi Griliches,** Harvard University

"The CPI Debate and the Measurement of Productivity"

[Note: The introduction, questions and responses are taken from a transcription of the session, with editing for purposes of clarification.]

Chair: **Jacob Ryten**, Assistant Chief Statistician, Statistics Canada**:**

Bon appetit, although belatedly, to everybody. Some 35 years ago, a former Chief of mine at OECD, in the department where Peter Hill and I worked, called me in and said: "I know that you are working on making comparisons among the CPIs of member countries. I would like you to look at a paper by a chap called Griliches. It deals with something, a Greek word, ... hedonistic ... hedonic - I'm not quite sure - about adjustments to the quality of motor cars. It obviously, given the importance of motor cars, has an influence on the CPI. For all I know, it may be that our collective CPIs are overestimated". My former chief, who had been actively following the work of the European Productivity Agency, said, "Imagine if it turned out that the real product in European countries were consistently under-estimated. Our picture of productivity for the 1950s would be quite different. I know you aren't going to be deterred by the econometrics this chap uses, so let me know what you think."

All this is to make the point that it's really not a necessity to introduce Zvi Griliches, Professor at Harvard University, to you. If you don't know what he is, who he is and you don't know what he's written about real output, hedonic adjustment, productivity, returns to human capital and R&D and so on and so forth, you shouldn't be here at a seminar on productivity in the first place.

So it is really a matter of protocol and a courtesy that prompts me to introduce him; it's also a matter of great personal satisfaction to me that he is with us here today. The subject, strangely enough, is CPI and productivity. Zvi.

Zvi Griliches: Zvi Griliches

I actually do not have much new to say on this topic. It is not news to most of us that if prices are off, so are also the implied measures of productivity. But having been immersed in the CPI Commission's work and the debate that followed, I have had little time to pursue it further. Nevertheless, I'll say a few words on this topic, primarily summarizing the work of others, some of whom are here. I'll focus the bulk of my remarks, however, on the post-Commission-Report debate, and conclude with one of the recommendations of the Commission which is relevant to the main topic of the Conference: the measurement of productivity in the Services industries. I will be talking about U.S. statistics. The numbers mayor may not be quite different for Canadian statistics.

The Commission estimated that the bias in the CPI as a Cost of Living Index (COLI) is running currently at about 1.1 percent per year. What does this estimate imply for the measurement of the growth in aggregate productivity? Not all of the biases identified by us are transmitted to the measurement of real GDP. Because real GDP is now a chain-weighted index, it is not subject to what we called upper-level substitution bias. Also, in several areas, especially computers, airfares, and medical costs, the components of GDP are deflated by different and potentially better price indexes. These two effects by themselves would reduce the applicable bias estimates for real PCE expenditures by about 0.3 or so. This leaves, however, unexamined the measurement of price deflators for business and residential investment and the government sector. DRI, in its February 1977 issue, takes our estimates at face value, and concludes that the growth in real GDP and aggregate productivity has been underestimated by about 0.9 percent per year during the last decade and a half.

One argument made against the Commission's conclusions is the possibility that it has overestimated the bias due to quality change, in part by double counting quality changes which

had been already allowed for in the standard BLS procedures (see Madrick, Moulton and Smedley, and Triplett). I will come back and argue this point in a few minutes, but if one were to take their argument seriously and or use our published though subjective range of uncertainty to allow for some such duplication, one might reduce the bias estimate as it applies to the PCE by another 0.2 or so, leaving us with a lower bound of approximately 0.6 percent per year bias in the growth of GDP and productivity. Slifman and Corrado, of the Federal Reserve Board, coming at this question from the other end, question the large negative productivity growth numbers in various service industries implied by the official NIPA numbers, and get 0.5 percent per year as their estimate of this bias by setting the negative productivity growth rates to zero. This seems to me a lower bound on what might be a reasonable range on this bias, i.e., between 0.5 to 1.0 percent per year (our estimates, after all, could also be too small!).

Returning to the post-report CPI-bias debate, I note that it has focused on two rather different topics: 1) Did the BLS already do much "quality change" adjustment, and hence we may have double counted some of it; and 2) would our findings, if implemented, be unfair to specific groups, such as elderly pensioners. The latter complaint extends also to the use of the CPI as a COLI, because it does not allow adequately for changes in the environment in which we all live.

The first argument about how much "quality adjustment" is already in the CPI arose in part because of a semantic misunderstanding about what the Commission and the BLS each meant by quality change. It is based on the forthcoming Brookings paper by Moulton and Smedley, who in their first draft reported that the BLS adjusted about 2.6 percent out of the CPI in 1995. This finding was the basis for some of the cool response by the BLS Commissioner, Katherine Abraham, and a major component of the attack on us by Madrick in the New York Review of Books. In the meantime, however, this number (in their revised version, March 1997) has shrunk to 1.76, and to only 0.3 percent if "outliers", commodity pairs where the implicit price-quality differential exceeds 100 percent, are excluded from the computation and geometric measures are taken (to exclude formula bias) since such large differences are unlikely to reflect what either the BLS or we had in mind in this context. At this point, the argument loses its quantitative import and we could leave it at that, but it is still instructive to discuss it, since it illustrates both the substantive and the communication difficulties in this field.

Most of the reported "quality adjustment" by the BLS, 1.65 out of 1.76 in the unadjusted-for-outliers numbers, comes from "linking" procedures, where no explicit quality adjustment is made, but rather a missing item is replaced by another and the price change during the link period is imputed using either the general estimated inflation rate in the index or the rate of price change of other commodities in the particular class. Most of such quality adjustments have very little to do with the change in the qualities of the goods available in the market but are rather the consequence of the BLS sampling procedures, which focuses on pricing a very specific item in a particular store and city: Granny Smith apples in 2-lbs. bags, a grey poplin men's raincoat, or a 12 cubic feet refrigerator by a particular manufacturer, with the freezer on top. There are thousands upon thousands such commodities in the market but only a small fraction of them is in a particular store at any time. The pricing agent has to deal with the rapid turnover and high probability of stockout. Roughly one out of two items disappear sometime during the year and have to be replaced by a different item in the same general class: a larger versus a smaller package of yogurt, a blue raincoat versus black, a refrigerator with its freezer at the bottom rather than at the top. But this churning is not what we had in mind by "quality change" and the appearance of new goods: the increased variety and freshness of vegetables and fish due to the improving transport facilities and the globalization of trade, the substitution of less invasive surgical procedures for gallstones operations, and more.

The whole BLS procedure which tries to price a "fixed" market basket of goods and services is biased against including significant quality changes in such comparisons. If the pricing agent cannot find a particular 12-inch TV set in the store, she is instructed to look for the closest "comparable" item, say a 13-inch one, rather than to consider the newly available version with a VCR built in or look for the antibiotics procedure which replaces ulcer surgery. Yes, the BLS does "quality adjustments" forced upon it largely by its sampling framework and the product turmoil in the markets, but it is not looking for the "quality change" that we were worried about. (We were well aware of the "linking" procedures followed by the BLS and discussed some of their potential shortcomings.) Its focus on pricing a "fixed basket" inclines it to avoid it. In principle we are both right: Yes, the BLS does much "quality adjustment". It has to. And no, it does not adjust much for quality change, as we understand it (except in the case of automobiles, apparel, and possibly rental apartment units), and these two facts do not contradict each other.

Moreover, such explicit quality adjustments account for only about 6 percent of the total "treatment of substitutions" effect, and amount to only 0.08 percent per year in the "outlier-cleaned" recomputations. While the Commission's estimates can be questioned on many grounds, there is very little overlap between them and the recent numbers produced by the BLS.

Even though it may turn out that some of our estimates of quality change in particular items may be too high, others are likely to be too low. That is why we included a range of uncertainty in our estimates, from 0.3 to + 0.5 for the overall estimate of bias. There are at least two sources of underestimation in our numbers: The first is that, except for a few cases, with low overall weight, we did not allow for the additional gain (consumer surplus) of the numerous new commodities in the economy. Nor did we try to quantify all of the intangible aspects of quality change, such as the improved safety of home power tools or the improved quality of stereo sound and TV pictures. The second is implied by recent evidence that we may have underestimated the biases in some of the areas we did examine: an alternative measure of consumer prices, the PCE (personal consumption expenditures) deflator, has been rising by about half a percent less per year (since 1992) than the CPI. An unpublished examination of this difference by the BLS indicates that most of it arises from the use by the BEA of alternative price indexes for hospital expenditures and airfares. These indexes do not adjust for any of the quality changes mentioned by us. Rather, they appear to be doing a better job of measuring the transaction prices in these markets.

The second major complaint about the Commission's recommendations is that they are unfair to the poor and the elderly. I don't think that we were unfair to them, but life is indeed unfair to them! That has, however, little to do with the CPI. The CPI is a measure designed for the average, unaging household, in an unchanging socio-economic and natural environment. Clearly, some of the poor, elderly or not, should be provided with a better safety net by a society as well off as ours. But that is not what the CPI is for. The major complaint of many of our critics is really about the low levels of entitlement and is not really about indexation. That complaint can be right without the index story being wrong. Currently, the best evidence (see the article by Garner et al. in the September Monthly Labor Review) indicates that CPI type price indexes computed for different groupings of the poor differ very little from the overall CPI over the

1984-94 period examined by them. Another BLS study (Pamela B. Hitschler, Monthly Labor Review, May 1993) reports that older consumers where much better off in 1990 than in 1980, without any of our bias adjustments.

The complaint by the elderly and their advocates does bring out two problems with the use of the CPI for indexation and productivity measurement. An individual's COLI can diverge from the average and the average COLI may change for reasons other than just changes in commodity and service prices. In particular, natural and economic environments may change. We all age and the cost of keeping us on the same "level of satisfaction", at least the medical cost, rises over time. At some point it actually becomes too high and we will reach our individual transversality condition. But that is not built in to our indexation procedures, where it is assumed that resources have been allocated optimally over our lifetime. Changes in environments are even more problematic. They do affect a "full" COLI. Consider the appearance of AIDS. It surely increased health insurance rates and the "true" COLI, but I think that it would be reasonable for the price agencies to rule that this represents an extension of coverage and hence is a "quality improvement" in the insurance policy and not a rise in its price. It is reasonable to try and separate the "Price Index of Living" from other changes in the cost of living, but such changes do affect the "level of living" that is achievable at any point of time with a given amount of nominal resources. People can and do feel richer or poorer even when prices do not change.

But not all price changes can or should be indexed. When OPEC raises energy prices, when AIDS appears, we all become poorer. Indexing was developed to protect workers and pensioners from monetary inflation on the assumption that there were gainers from inflation that could be taxed to compensate the losers. But many changes in prices and the cost of living occur in contexts where there are no gainers who could be justly taxed. If energy prices rise why should not all sectors of the society share in this disaster? Does it really make sense to tax one group even more so that another could escape the consequences of this disaster entirely? Because of such considerations, the CPI is ultimately not the right instrument for indexation. I have previously suggested the use of the median wage instead (Griliches, testimony). This would have provided a somewhat lower level of indexation in the past decade, but has the virtue that it would allow the pensioners and other recipients to benefit also from the average growth in the economy

and not just keep their "baskets" fixed.

The same conundrums carry over also to productivity measurement issues. A rise in the resistance to antibiotics by certain strains of bacteria resident in our hospitals will reduce the productivity of the health care system by increasing the time needed by patients to recover from various existing diseases and also increase some of the inputs required for achieving the same final goals. This should show up both as an increase in the price of such services to the consumer and a decline in their quality. On the other hand, a new strain of influenza will increase the output and not reduce the productivity of the health care industry as we should measure it, even though the achievable "level of living" of the consumer population as a whole will have declined. We are still very far from being able to sort out such changes in technology and environment in the medical care area, even though much progress has been made along these line recently (see the work of Cutler, McClellan, Newhouse and others).

As Peter Hill has already emphasized, to do the measurement of productivity of services right we will have to recognize that a major input has been omitted, the consumer herself. Consumers participate in the production of most services, either actively, as in shopping or extracting cash from the banking system, or passively, when being "worked on" by the medical industry (or its historical precursors, the barbers). To get such measurements right will require much more extensive information on how we use our time, both in production and consumption, and also on the "quality" of this time. One of the Commission's major recommendations was to urge the statistical agencies to develop an appropriate survey design and start collecting data on time use by consumers, a topic that has been largely ignored by them. Progress in the measurement of productivity growth in the services requires it! (Applause)

Questions

Question: Where can we obtain a copy of the CPI Commission report?

Griliches: The Washington Post website has it.

Question: There is the question of whether the wealthy are included in the CPI. There was, I think, a National Bureau of Economic Research paper on consumption of the wealthy compared to others. It would be useful to know your views on how the poor are affected.

Griliches: There are many things said here. I will try to answer a few. The current procedures use a consumer expenditures survey to get the weights for the index. These weights include the wealthy to the extent that they answer the questionnaires and by and large, if anything, they probably answer <u>more</u> than the poor. Second, in some sense, one of the complaints against the current CPI is that every dollar gets equal weight. In that sense, the wealthy get more weight than just their share if you wanted to have a democratic index rather than what's called a plutocratic index. Our index currently is a plutocratic index and the whole GDP, in that sense, is plutocratic because it weights everything by dollars and not by individuals.

Now what do we know about the poor? Well, at some level we know quite a bit, probably more than about the rich, because I think the poor have been studied more, for various reasons. The computations that I referred to are faulty, in a sense. What they deal with is a composition of consumption in relatively large groups which are significantly different. Do these weight differences make a difference when you use them to re-compute the price index? It turns out that - at least over the period we considered - the relative price changes were not large enough and the correlation between weight differences and price changes is not high enough to produce a significant difference in re-computation when you use a set of different weights from the average.

In a study that I alluded to earlier which examined this at the micro-level, a price index was computed separately for each household. It indicated that there is a tremendous amount of variability, but that variability is as large or larger within identifiable classes as it is between classes. The average differs very little, but within any group there are households for whom price changes move quite differently.

Now I think there is a separate story that's running around which somehow says a lot of the new

goods that you're talking about are technological frontier toys which are consumed only by the rich; the poor have not benefitted from them.

I think there's a small view and a large view; the small view is about particular things and many of the things that have been pointed to used widely in the economy, including products like VCRs and television sets. Also, the construction of any quality change measure done correctly is going to be weighted by the consumption magnitudes of the individuals. The high estimates of the original quality changes valued at the original time of introduction are going to be affected by relatively small weights. As you walk down the demand curve, the same quality change for the same new goods is going to be valued at a lower price as its consumption expands. So in some sense, if done correctly, part of this adjustment is built in. I do want to say though, that if you take the longer view it is misleading to say that quality change or technical change has been an upper income phenomenon rather than a lower income phenomenon. Consider:

- health standards
- life expectancy
- aging
- public health
- the impact of the telephone on the options available to people
- the impact of the decline in transport costs which have been available to people to move and escape from wherever they were stuck with in some form

I think the total contribution of the above has been very important for welfare measurement, even if we don't really know how to measure it right.

That doesn't really answer the question whether our particular number is too high by a few decimal points or too low by a few decimal points. But I want to put it in perspective that real gains have been had. I think [these gains] have been very important at the lower edges of the society also - it's not just a "trickle down".

Question: Zvi, you mentioned that the overestimate of personal consumption deflators in the U.S. might explain, or lead to, an underestimate of real GDP growth of about 0.9 percent with a

lower bound of 0.6 percent. Part of what is before this conference is whether we can explain the productivity slowdown after the mid-1970s with those kinds of numbers? Would you like to speculate how that underestimate of real GDP growth might compare to what went on in the decades <u>before</u> the mid-1970s? Has there been an acceleration here, or not?

Griliches: This is a tough one and, frankly, I don't think there has been much research on this topic that I can put my fingers on. There is one major component which clearly increased the bias between 1978 and 1996 or so, which was a formula bias that crept into the CPI due to otherwise a very important intellectual and substantive improvement, which was the introduction of random sampling into the CPI. We had this in the past and presumably we won't see it again. Beyond that, we have to argue the magnitude of changes of today versus yesterday. I think if you take the absolute story of new goods, it is hard to argue that, for example, the VCR was more important than the telephone, or Walmart was more important than the supermarket, or discount fares are more important than the jet engine itself or the introduction of the automobile earlier. So it's hard for me to think the integral is much larger. On the other hand, the rate at which it has been growing in the last decade has accelerated. Alternatively, the difficulties faced by the statistical agencies have accelerated. And that has to do partly with the decline in communication costs and the decline in the cost of changing things. So, for example, prices used to be fixed in a store because it was expensive to change prices. Now, if price information is in a computerized scanner, one stroke can change prices much more easier. The menu costs have declined. This has also led to the introduction of new goods. I think the number of goods that turnover and the pace of the turnover has increased. When we started looking at personal computers about ten years ago; we were doing a hedonic analysis of computer prices. The average computer model was staying in the sample about three to four years. Now, the average computer model doesn't live for a year. Very few models last more than a year.

Now if you're saddled with a system in which you're pricing a fixed well-specified good, you're faced with a very difficult problem of measurement. And I think that in this sense, things have gotten harder. It's also related to the fact that we haven't really talked extensively about the growth in the service sector. We vaguely know how to do hedonics on computers - never mind

that it's not that easy either - but you think you might be able to do it. How do you do hedonics on doctors? On HMOs - I mean, the generations change, the knowledge that's contained changes. The heterogeneity that is implicit in those things is very, very large and I think that our ability to cope with that may not be growing as fast as the problem is.

Question: Zvi, I remember listening to the same kind of stories about quality change from Richard Ruggles at Yale, who was working on that issue in the late 1950s. What I remember is that there was also a lot of talk about the offsetting deterioration in quality. You didn't mention that at all.

Griliches: O.K., that's the other question. Have the bads out-weighed the goods? And are there a variety of new bads? I think there is no doubt that there are some. I think, by and large, the quality of many goods that we are using has improved rather than deteriorated. Just to give you an example
of something that you are probably only vaguely aware of, but when is the last time you've had to stop and change a tire? I mean that's the sort of thing that was a relatively common experience when I was a student. It's just not there anymore. That's only one example.

One of the major things that people have worried about in the economic side is the environment. By all measures of quality of environment - we are, by and large, at least in continental North America, much better off in terms of (for example) the quality of the air, than we were when we were students. Again, go back and think about the rate at which you had to wash the shirts and the rate at which they got <u>black</u>.

There are some things that have gone the other way. I think the perception of security has declined. It is not entirely clear whether and how much crime has actually risen. Crime reporting has risen, and crime perception has risen. In a sense perceptions are also real, in the sense that if one feels insecure, one's level of satisfaction is reduced, whether or not the probability of harm has increased. And that affects measurement issues.

One of the things we discuss is the issue of quality change. Take something like an alarm system

on your car. Now today's alarm system is much more sophisticated than it was five years ago, and costs half as much. So there is a quality improvement if you measure it correctly. On the other hand, for all I know, the technology of burglars has been improving roughly at the same rate and the probability of your car being stolen may not have declined. So there is an issue in how to handle environmental changes.

I would say that many goods have disappeared entirely. They disappear eventually because the markets become too small to sustain them. There are fixed costs of entering the market and there are also fixed costs of staying in the market. Whenever a new good comes in, there is a gain in consumer surplus and whenever a new good goes out, there's some loss of consumer's surplus, so there is the problem of doing the accounting right. Just to give you an example - LPs. You might think they have disappeared, but they haven't entirely. In fact, there is a small group of people who think that LPs are better than CDs. If you want a piece of music on LP, it's possible. New LPs are still being made, they are just being made at twice the price of CDs. But they're still there.

Andrew Sharpe: Professor Griliches, on behalf of the Centre I would like to thank you for your masterful exposition of the issues. I found your presentation extremely articulate on everything you've said. In fact, we are really honoured today to be here in many ways. I don't think there is anyone else who could have made such an excellent presentation on these really crucial issues. So again, I would like to thank you for coming to Ottawa to make this presentation, and we look forward to hearing from you again tomorrow. Thank you.

References

Boskin, M. J., et al., Final Report to the Senate Finance Committee from the Advisory Commission to Study the Consumer Price Index, S. Prt. 104-72, Washington, DC: Government Printing Office, 1996.

Cutler, D. and McClellan, M., "The Determinants of Technological Change in Heart Attack Treatment," NBER working paper No. 5751, National Bureau of Economic Research, Cambridge, MA, September, 1996.

Cutler, D., McClellan, M., Newhouse, J. and Remler, D., "Are Medical Prices Declining?" NBER working paper #5750, Cambridge, MA, September, 1996.

Garner, T. I., Johnson, D. S. and Kokoski, M. F., "An Experimental CPI for the Poor," Monthly Labor Review, September 1996, 32-42.

Hitschler, P. B., "Spending by Older Consumers: 1980 and 1990 Compared," Monthly Labor Review, May 1993, 3-13.

Madrick, J., "The Cost of Living: A New Myth," New York Review of Books, March 6, 1997.

Michael, R., "Variation Across Households in the Rate of Inflation," Journal of Money, Credit, and Banking 11(1), 1979, 32-46.

Montgomery, M., Pope, L. and Brinner, R., "Reassessing Real Wages, Productivity, and Growth," DRI/McGraw Hill, February 1997, 29-51.

Moulton, B. and Smedley, K., "Addressing the Quality Change Issue in the CPI," BLS March, 1997, forthcoming in Brookings Papers on Economics Activity.

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

Triplett, J. E., "Measuring Consumption: The Post-1973 Slowdown and Research Issues," Federal Reserve Bank of St. Louis, 1997, in press.