Some conservative comments on hedonic methods

Contribution to the Symposion on Hedonic Methods in Wiesbaden June 2001

by Peter von der Lippe

My remarks focus on two points only. Firstly on the sort of results policy makers possibly expect the hedonic method will produce, and secondly on some aspects concerning the practice of price statistics.

I. The inflation rate and hedonic regression

1.   The problem I have with hedonic regression is not the method itself, but rather the purpose it should serve. Since the Boskin Commission presented several upward biases the US-CPI is alleged to possess, we have experienced an enormous public and political interest in issues of index compilation. There is some distrust that Official Statistics might cause great harm by not moving rapidly enough to the latest methods originating from overseas. The "hedonic" method to make adjustments for quality change is generally seen – in the perception of the many non-statisticians at least – to result in *lower rates of inflation* and higher rates of real growth.
Assume for a moment the opposite were true, more inflation and less growth: I venture the bet that this seminar would not take place and there would not be much left from hedonics that would attract the general public.

2.   From the point of view of the traditional Laspeyres index approach quality adjustment has to be made only in order to ensure that like is compared with like. An attempt is made to measure the *"pure" price movement* and in contrast to the *unit value index* in a *price* index[1] changes (substitutions) in models and outlets are *not* to be regarded as price changes. When they occur such that it is no longer possible to effectively keep all side conditions of the price comparison constant the effects of such events that disturb comparability should at least be discounted appropriately. It is in this context and in such situations (e.g. the disappearance of the model the price of which was hitherto quoted) that the adjustment for quality change gains relevance. In other situations the introduction of new commodities should rather be avoided for the sake of comparability unless a new base of the index is defined.

3.   We notice (not without displeasure) that making comparisons only under conditions which are as similar as possible, is no longer accepted to be a cornerstore of index methodology. We see a renascence of chain indices in which by definition there is no need to take care of comparability over more than just any two adjacent periods and which are explicitly praised for their flexibility in automatically "incorporating" all sorts of new developments. Yet this symposion shows that the problem which is allegedly solved by chaining seems to live on.
In a similar vein the "true" Cost of Living (COLI) model stands in sharp contrast to the Laspeyres concept of price movement described under no. 2 by abandoning the idea that price determining factors like the selection of goods and outlets etc. should be kept constant if ever possible (or neutralized if impossible in the practice of price index compilation). Coming from an "economic theory" approach it is maintained for example, that not only new goods should be introduced as soon as possible, but also

---

[1] A unit value index (UVI) is affected by all sorts of changes in the mix of models, outlets etc. as such changes may enter the formula without adjustments whatsoever. Hence they influence the result and they are therefore implicitly counted as price change. A "pure" price index differs from a UVI in that it makes corrections in order to neutralize the effect of non-price changes.

an estimate made of the extra-utility these goods represent compared with those goods that were observed until now.

4.   The fact that emphasis continues to be given to quality adjustment although the fundamental ideas governing index compilation have changed dramatically (even converted into its opposite in some parts) is in my view most astonishing, and thus deserves being explained and discussed. It is certainly not my job to call into question the chain index or the COLI philosophy here, but nonetheless it should be worked out that the *motivation of quality adjustment* seems to be quite different

- to preserve *comparability* in the traditional framework on the one hand, and

- to account for *welfare gains* on the other.

The last mentioned aspect is clearly addressed in a statement of Silver and Heravi: "If a new good is introduced it is not sufficient to simply wait for two successive price quotations and then incorporate the good. This would ignore the welfare gain to consumers as they substitute from old technology to new technology. Such welfare gains are inseparately linked to the definition of COLI defined as indexes, which measure the expenditure required to maintain a constant level of utility (welfare)"[2].

For those who are primarily interested in lower inflation rates the criterion "constant level of utility" is much more promising than just some minor corrections made to offset changes in the conditions for price quotation[3]. It appears to me that this (in-advertedly) facilitates possibly to downrate inflation significantly. With some irony and exageration one might say that progress in fighting crimes, illiteracy, or diseases now may well imply that less effort is needed on the part of the ECB to fight inflation.

5.   To bring this more general part of my remarks to an end I would like to mention an observation which fundamentally changed my thinking concerning price statistics. According to official statistical figures we never had an even slight inflation in *the former GDR* though the general public was well aware of rising prices everywhere. The trick to achieve this feat was to consistently find *the* suitable amount of quality improvement to exactly offset the rise of a price and if this could not be done with some plausibility a intentional rearrangement of the selection of goods and services was made, more or less overtly in such a way that the commodity the price of which went up was replaced by another one where the price remained constant or was even declining[4].

---

[2] Mick Silver and Saeed Heravi, The Measurement of Quality-Adjustment Price Changes.

[3] It is for example by no means suprising that the appearance of generica in addition to branded pills may yield widely differing results. To treat them as different goods and continue to quote prices of branded pills of course leads to the highest possible inflation while the lowest reflects the opposite extreme assumption of "a pill is a pill" and results from taking an average in the unit-value-index manner (see F. M. Fisher and Zvi Grilliches, Aggregate Price Indices, New Goods and Generics, Quarterly Journal of Economics. Vol. 110 (1995), pp. 229-244). This once more shows that it is far more important to decide whether to follow the (traditional) price index logic or the unit value index logic than to decide how to measure quality change (with or whithout making use of regression) when an assessment of a quality change has to be made. It appears to me that there is a tendency to mix up both issues. It is not regression which lowers inflation, but rather the idea that a move to lower priced goods and outlets should be regarded as deflationary as it is (implicitly) done in the unit value framework but conspicuously *not* in the traditional Laspeyres concept of inflation.

[4] It is of course true that there was not really a theory-driven search for some gains in utility to com-mensurate the observed rise of prices in the former GDR, but a rather blunt desire to do away with inflation. The situation is of course different in democratic countries where no such political influence may be wielded. Yet the example at least shows which aspects of index methodology deserve attention when we suspect some kind of manipulation in price statistics.

It is possibly no coincidence that the criticism of the traditional Laspeyres ideology in price statistics focuses precisely on these two elements of index compilation, viz. the allegedly insufficient and belated accounting for utility gains of new goods and the constancy of the basket.

The idea that there might be some welfare gain in new products compared with a hypothetic price they might have had in the base period when they were not yet existent or a surplus utiliy owing to the mere increase in choice did not occur to the GDR statisticians as it was apparently beyond the scope of socialist economic theory to account for differences in utility which are not necessarily incorporated in tangible goods.

What makes the socialist tactics in dealing with quality change and selecting goods dishonest is of course the fact that the amount of quality change and the principle governing the re-selection were exclusively result-oriented, i.e. determined by the desire to lower inflation.

6.   Notwithstanding – and that is in my view the lesson to be learned from the GDR experience - it remains to be recognized that the departure from the Laspeyres ideas might open floodgates and hence Official Statistics should be mindful to all in-novations which might be result-motivated and arouse suspicion of manipulation. It is not only the application of regression analysis which characterizes "hedonics" and leads to less inflation,[5] but it is also the "when" and "why" of making adjustments.

II. Cost-benefit-considerations and some practicalities concerning the hedonic approach.

7.   It is clear that the hedonic method to evaluate quality change is *costlier* than some less sophisticated traditional alternatives. We have to collect not only much more data[6] but also data for the mere purpose of making adjustments that is for *indirect use* only. Thus additional data is needed and we call for such an extension of statistics in a time when statistics is urged to minimize response burdens and to avoid unnecessary data collection.

8.   In my knowledge there is not much *cost-benefit-analysis* done in this field. The problem of course is the "benefit" let alone its "amount". Even if we could be sure that the result of the traditional matching method were incorrect (erroneous by which criterion?) or biased (to what extent?) it is not under all circumstances clear that there should be a substantial damage inflicted on the overall result of the index compilation. We witness for example discussions about the correct measurement of the quality of computer software, (or parts thereof such as PC data base products), irrespective of the weight they have in the CPI. Given that the rate of inflation is the all-items-CPI and not the specific price level of computer and IT products the question seems justified what we are willing to pay for that exercise, however interesting and path breaking it may be from a theoretical point of view.

9.   It is well known that we not only ought to collect a vast amount of additional information on prices and quality characteristics to guarantee reliable results. More-over the sell-by-date of this information is unfortunately mostly in the near future. The

---

[5] In fact it is difficult to see why the use of the regession technique alone should consistently produce a lower level of quality adjusted prices and thus a lower rate of inflation.

[6] When in the traditional framework product A is replaced by B the quality of A and B is explored while the hedonic approach necessarily requires an examination of many more products, say C, D, etc. in order to arrive at reasonable estimates of regression coefficients of various characteristics.

hedonic function has to be *re-estimated at rather short intervals*. It is precisely because of the speedy progress as regards quality that we cannot use "shadow prices" once established by hedonic regression for many subsequent years.

It is not only the lag between the time the regression was run and the time in which use is made of its result which matters, it is also the market situation prevailing in the time interval that constitutes the data base.

10. In the hedonic approach it is generally assumed that consumers makie well informed choices maximizing their utility such that they are willing to pay more to the extent that the quality of the product in question is "better". Yet there are of course *market situations* (on the demand *and supply* side) in which we cannot validly infer underlying consumers' valuations from observed prices. There is, however, nothing that will prevent us from regressing prices (or quantities) on "characteristics" that is from running a hedonic regression in such situations.

Interestingly we may encounter such situations more frequently in times in which the problem in question arises, that is when consumers have to adapt to new possibilities and to abandon old choices as they are no longer available or relevant.

We cannot possibly disapprove of traditional methods on the grounds that prices do not always convey the sort of information they should have in order to validate the method, and at the same time forget that this may well apply to the new methods too, i.e. to methods of infering "shadow prices" for parts, components or "characteristics" of goods.

11. It is maintained for example that for various traditional methods the conditions of their applicability might be impaired in those time of price observations in which quality adjustment is in particular required. There may be reasons given, to doubt that the ratio of prices in fact represents the ratio of values:

- the old product may be discounted while the new one may be introduced at an unusually high price such that the difference in prices clearly overstates the difference in quality;

- on the other hand the new item may be introduced at an unusually low price in order to induce demand (understating the quality change), and the subsequent increase of the price is primarily indicating a return to normal market conditions rather than a true rise of the price; and

- finally purchasers may have difficulties in assessing the relative qualities as the new item is incomparable to all old ones because it has completely new features or serves an entirely new purpose (see below no. 15).

Hence differences in prices may not necessarily reflect only differences in value and utility but also abnormal situations in the marketplace due to exactly the sort of events that are under investigation, e.g. the appearance of new and disappearance of old products. Why should this only apply to traditional methods and not also to at least some items and prices of the data base of a hedonic regression ?

12. It is in general assumed that differences in prices reflect different degrees of utility the *consumers* enjoy, and of their willingness to pay for it. But why should such differences in practice (not in theory) really mirror the underlying consumer preferences alone? It is conspicuous that here as well as in the case of the COLI-theory we hear much about buyers, and their utility maximizing behavior but what about the producer his costs, the supply conditions he faces or his access to the

market etc.[7] "A large and statistically significant coefficient for a particular quality attribute in an estimated hedonic price function may reflect, not consumers' high valuation of that attribute, but rather the difficulties or high costs which producers have in achieving that attribute per se or in relation to other attributes." Needless to say that precisely in this case a quality correction of the observed price should *not* be made, as this is an example of a pure cost-push inflationary event which the CPI ought to reflect.

13. Hedonic regression is applicable primarily whenever "quality" depends on some measurable "characteristics" as given above all in the case of high-tech (or simply "tech") durables such as automobiles, computers, television sets and the like. In the case of services or nondurables described mainly in terms of "taste" or emotional effects, or in general in the absence of *observable criteria* the method appears less suitable. The fact that in such applications sometimes the experts' judgement proved insignificant[8] a regressor or that the most part of the price variation not infrequently was explained by dummy regressors, denoting brands or so, may cast some doubts on the reliability and validity of the statistical results.[9]

14. A well known problem with hedonics is that usually all observations entering the regression equation have *equal weight* in determining the regression results as if they were of equal importance. This is due to the fact that mostly catalogue prices are used. Analysts often neither account for the true *transaction prices* nor for the trans-action volumes[10].

15. Another objection against hedonic regression is that it is unworkable just in those situations were it is urgently needed, that is in the case of *completely new commodities* with never before seen characteristics upon their initial introduction. Hedonics is practicable mainly when innovations consist in a recombination of known features but it fails when the commodity (or service) serves an entirely new purpose (e.g. the ability to communicate almost everywhere in the case of mobile phones) or

---

[7] There are not many articles that explicitly address this problem. An example is Nerlove's study of the Swedish wine prices in 1995 from which the following quotation is taken.

[8] The interpretation given to such a finding is generally this: consumers do not consider this characteristic important. However in fact they may well value it and appreciate differences in this respect, but unfortunately the varieties selected for the regression simply are too similar. The result of a hedonic regression not only depends on the sample of descriptive characteristics chosen to determine the price but also on the sample of products representing this "universe" of similar products we have in mind but which is conspicuously not defined with some precision. Moreover the sample of goods entering the hedonic regression may well represent goods belonging to different strata of consumers such that is not legitimate to infer the preference of "the" representative consumer from such data. The insignificance of some jury grades is again not necessarily a proof of its irrelevance, but most likely rather due to the fact that their judgement might be implicitly represented in the other regressors (characteristics).

[9] The brand name "Mercedes" covers cars widely differing in prices and there is definitely much less resemblence between a high and low price Mercedes than for example between a Mercedes and BMW of the same "class".

[10] It is clear that the use of scanner data may be a great help in detecting the respective importance of the varieties and their effective transaction prices (unit values). To quote once more the example of automobiles, it is beyond dispute that e.g. Ferrari cars are less frequently purchased than Renaults cars and that the policy of car producers as regards rebates will differ widely. It is therefore unsatisfactory to collect data from price lists, mail order catalogues, specialist magazines, reports of market research institutes etc. where the volume of transactions is not accounted for. Furthermore without even having to look at the commodites such methods should not necessarily be superior to the traditional decentralized assessment of quality made by price collectors in cooperation with the sales managers of the reporting outlets.

when goods are able to satisfy given needs in a new manner (e.g. airbags serving safety which cannot properly be compared with seat belts).

It is of course clear that alternative methods would run into difficulties as well in such situations. Hence this quite popular argument is possible not fair as *all* hitherto known methods of quality adjustment rely on comparisons of prices between two or more *related* products. The treatment of a novelty without any predecessor whatsoever will necessarily be difficult. It is true, however, that the applicability of the hedonic method will be even more limited for the simple reason that it does not make sense to examine in detail a scatter diagram with two or three data points only.

16. There are many more aspects of the hedonic method which are well known and discussed in detail in the relevant literature, as for example

- the type of the regression function (choice of the functional form);

- to decide which and how many "characteristics" to take in order to capture enough aspects of quality (or aspects in which the selected varieties differ) on the one hand and to avoid multicollinearity on the other;

- the detection of the appropriate source of data[11];

- the use of dummy variables for time periods;

- the constancy of regression coefficients over time, not allowing for changes in the marginal utility assigned to the selected "characteristics"

and many more "practical" problems.

17. To sum up: in my view all we need is a more pragmatic approach in this field. I would be happy if religious zeal would give way to pragmatizm. Hedonic regression is most definitely powerful and useful in that it aims at a more objective and unequivocal quality assessment. "Hedonics" is, however, not above criticism, it may even yield poor results and there are situations in which it is not superior to other more traditional methods like "matching" for example. Contrary to a popular belief it is not consistently "better" (by which criteria by the way?) than the less ambitious methods we used to apply in practice. It is *not* without problems and it is *costly* at that. Therefore it is much better to think of a bunch of methods than to believe there is *the* ultimate method that should replace all other methods as soon as possible.

Above all hedonics should primarily be seen as an application of the regression technique, and it should be strictly kept distinct and separated from the index philosophy (COLI, Laspeyres, chain or so) we wish to follow. I know of course that this is highly controversial and that unfortunately I probably represent in this respect (defending Laspeyres' idea of "pure" price comparison) a conservative (see the title of this paper) if not outdated minority position in this symposium.

---

[11] In their study of Bordeaux wines Combris, Lecocq and Visser for example gave several reasons for not having taken their data from the readily available wine guides.