Inflation in Europe – Different Measures and their users

Berlin, 4 and 5 October 2001

Theme 3

Market and Bank Analysts/Research Institutes

Peter von der Lippe

Essen University, Germany

Conceptual problems with the harmonised index of consumer prices (HICP)

by

Peter von der Lippe

1. Background to my critique of the HICP

When studying the harmonised index of consumer prices (HICP), we must distinguish between two quite different things:

1. endeavours to produce standardised and comparable methods for compiling national indices of consumer prices (ICPs) or the special index needed for the European HICP (or ICP for short), and

2. the aggregation of national ICPs based on the European Union Member States or countries belonging to the European Monetary Union (EMU) into an HICP that is officially used as the basis for, or measure of, ECB[1] monetary policy.

There are similarities between these two topics in that the methodological approach to the HICP played a decisive role in shaping the contents of the regulations adopted as part of the harmonisation process. This is particularly true of the HICP's main feature, namely the fact that it is a chain index based on national baskets of goods that change over time. Had an entirely different approach been chosen, the harmonisation regulations would have looked very different in terms of their contents.

This is also true of a second feature of the HICP methodology - the role played by enthusiasm (particularly in the USA) for the "economic theory of index figures" and the "True Cost of Living Index" (COLI), which very much shaped the Boskin Commission's thinking[2].

In my critique of the conceptual premises of the HICP I will thus be concentrating on these two points, because:

- the HICP is designed as a chain index - which, whilst consistent with the SNA[3], has a number of drawbacks that clearly receive only scant attention in the SNA, and because

- although it does not profess to be a COLI (unlike the USA's ICP), it is very much influenced by COLI-type thinking.

The HICP may well provide a valuable service, and a solution that is more satisfying from a theoretical point of view might prove quite elusive. Nevertheless, the quality of a statistical measure is difficult to assess if it is not clear what exactly is being measured, and if the result (in this case the measured "rate of inflation") depends on such a *wealth* of factors.

---

[1] European Central Bank.
[2] The Boskin Commission also thought COLI a worthy goal, in theory at least.
[3] System of National Accounts.

2. Harmonisation strategy and the concept of comparability

Notwithstanding criticism of the underlying concepts, the object of harmonisation can be wholeheartedly endorsed - "HICPs shall be considered to be comparable if they reflect only differences in price changes or consumption patterns between countries". The fact that different national CPIs also reflect differences in consumption patterns hardly needs to be pointed out. Another question is whether the common (country-aggregated) HICP rate of inflation should be affected by such differences.

Criticism has recently been levelled at the HICP for the monetary union (hereinafter "MUICP", or Monetary Union Index of Consumer Prices), particularly by Georg Erber of the DIW Berlin, basically because the MUICP rate of inflation, which the ECB also uses as a guideline, has *two* components:

1. the price increase measured using a common European basket of goods (as is done when calculating purchasing power parities [PPPs], comparisons being made for two countries - or one country plus the Community - on the basis of the *same* basket of goods[4]), and
2. differences between the national baskets of goods that make up the MUICP.

Alignment of consumer patterns in the EMU countries will gradually eliminate the second component of the MUICP rate of inflation, which in this respect is naturally "too high". It has been argued that the existence of this structural component should be a reason for the ECB to base itself on a 2.5% rather than 2% annual increase in the MUICP, so that countries with large national weightings (such as Germany) are not unintentionally exposed to deflationary tendencies.

It is certainly right that comparability between countries (which in itself presupposes an ability to make country-based aggregations) is contingent upon there being no differences in inflation rates that can be attributed to divergent *methods*. It is also right that this type of comparability can only be achieved by means of agreements and regulations on methods to be uniformly applied.

The principles of harmonisation drawn up at EU level in this field are laudable, at least in terms of intent - the quest for "best practices" (but how should we define these?) or the laying down of rules, not in detail, but just to the extent necessary to ensure that certain "minimum standards" are observed. The phased introduction of the HICP is, in any case, to be welcomed. It should likewise be accepted that, when regulations are being drawn up, a country may be outvoted. This is, after all, the price of integration and increased international co-operation.

However, the contents of individual regulations and their theoretical (and sometimes political) *raison d'etre* are a different matter. Here criticism is perfectly in order.

3. The concept of "pure price comparison" (temporal comparability)

Why not simply compare two specific average prices (values) at two different points in time in order to measure the rise in prices? The reason is that, in addition to price changes, a whole series of other structural changes influence the comparison. This is clear from the reasons that the SNA advances for not using the unit value index. The SNA rejects this approach on the grounds that it is "affected by changes in the mix of items as well as by

---

[4] Much has been written on international comparisons of this type, the best known contributions in this field being the methods devised by Geary and Khamis (GK) and by Eltetö, Köves and Szulc [EKS]. These works clearly played no role in the creation of the HICP. Why not? Though true that the HICP is an aggregation of countries rather than a comparison between countries, why does the comparability of goods and baskets of goods play a role in the latter but not in the former?

changes in their prices. Unit value indices cannot therefore be expected to provide good measures of average price change over time" (para 16.13).

It is one of the abiding secrets of the SNA why it - correctly - acknowledges the disadvantages of the unit value index but then goes on to recommend chain indices, failing to realise that the very same objections can be raised against these. The basic objection here is that chain indices are not based on pure price comparisons. The comparison is not "pure" in the sense that the result is influenced merely by the change in prices, not by other variables too. Nor is it "pure" in the sense that the outcome is influenced only by the periods 0 and t used in the comparison, not by the intervening periods (1, 2, ..., t-1) with their prices and volumes.

In index theory, there are basically two ways of constructing situations such that prices related to these reflect the *actual* increase in prices in a certain sense. We can either:

1. compare expenditure on a *basket of goods* which, on paper at least, is the *same* (in terms of type, volume and quality of goods), or

2. compare the theoretical minimum expenditure required to purchase a selection of goods of the same value in terms of its *utility* during periods 0 and t,

concluding that an increase in prices has taken place (i.e. inflation has occurred) between periods 0 and t if expenditure (defined in one way or another) has increased.

The first approach is the famous Laspeyres' index, $P_{0t}^L = \dfrac{\sum p_{it} q_{i0}}{\sum p_{i0} q_{i0}}$ or, more simply,

$\sum p_t q_0 / \sum p_0 q_0$, where p stands for the prices and q the quantities of goods and services i = 1, 2, ..., n during the base period 0 (in the case of HICP, 1996) and reporting time t. The second approach produces a constant utility index (or COLI), where inflation is lower than the figure generated by P_{0t}^L if and when households make rational substitutions, i.e. consume relatively less (or more) of goods that are becoming relatively more (or less) expensive.

Temporal comparability is thus guaranteed by the uniformity of the basket of goods or of "utility", though a uniform situation may be more readily established in the first instance (the same basket) than in the second (the same level of utility).

This is not the place to compare these two approaches, especially since neither the HICP - nor indeed its country-based aggregated equivalent, the MUICP - really corresponds to either. Although, as has already been said, the HICP does not claim to be a COLI, unlike the American CPI, at least one of the regulations imposed on the Member States as part of the harmonisation programme, namely Commission Regulation No 1749/96 of 9 September 1996 on the inclusion of "newly significant goods and services", smacks of usefulness and constant utility indices.

This is very much in line with the thinking behind the Boskin Commission's "new product bias" (akin to "outlet bias"), the assumption being that a CPI should also reflect the added usefulness of new goods or forms of selling (which the US CPI did not do at the time, hence the mention of "bias"). However, it is precisely this, i.e. the inclusion of new goods (except for the regular re-basings), that the Laspeyres' approach urges us to avoid, as it adversely affects the comparability of the basket of goods.

Including new goods as quickly as possible is an action motivated by considerations of usefulness. Thus we find the Boskin Report[5] saying that petrol is not as expensive as the

---

[5] All the examples included in my critique of the economic-theory-index are taken from the discussions held in the US.

index suggests because credit cards now make paying easier. Likewise, the prices given for eating out are too expensive, as there is now a large selection of frozen food that can be prepared at home. Or the increase in the price of books overestimates inflation, since it is now easier than ever to borrow books. And so the arguments continue - a consumer price index should also reflect a greater selection of goods, a cleaner environment, longer life expectancy and so forth.

There is no end to discussions of this nature if the ultimate aim[6] is not to achieve price comparability but to calculate for consumers an increase in usefulness that is not reflected in an increased price. The problem is not how to deliberately under-calculate inflation[7], the problem is whether there is any point in price statistics moving ever further away from observable prices for specifically defined goods.

4. The obsession with weightings that are as up-to-date as possible

Neither of the above-mentioned approaches - pure price comparison or COLI - justifies the use of a chain index, the desired means to the HICP end. At the forefront of attention here is the constant updating of the basket of goods and the ability to alter the basis of comparison (selection of goods and businesses and so forth) at any point in time - something that should *not* happen with the P_{0t}^L Laspeyres' formula with weightings (q_0) that are fixed in the interests of a pure price comparison for the duration of an index with base 0.

It is sometimes said that the P_{0t}^L index is no longer appropriate in an age in which everything is changing so quickly - so much so that weightings may have to be changed yearly. This was clearly the thinking behind Commission Regulation No 2454/97 of 10 December 1997 on "minimum standards for the quality of HICP weightings". For "the quality of", read "recent"!

It should also be borne in mind that, since all the countries aggregated under the MUICP are constantly changing their weighting schemes, the index for all the countries together (i.e. the MUICP) represents not just a wealth of baskets, but also baskets of differing ages. And if the "1996-base" index currently stands at 120, this does *not* mean that something that cost 100 in year 0 (i.e. 1996) now (i.e. in year t) costs 120 (the "something" is, in reality several baskets of goods from 1996 [*not just one* constant basket], and these baskets have been *changed repeatedly*, all of them being k × t baskets of k countries in t periods (1996 to the present, t).

In a chain index, a comparison between two periods 0 and t is produced not directly[8] but indirectly, i. e. by *chaining* (multiplying) the links (the individual factors), e.g.:[9]

$$\overline{P}_{0t}^L = \frac{\sum p_1 q_0}{\sum p_0 q_0} \frac{\sum p_2 q_1}{\sum p_1 q_1} \dots \frac{\sum p_t q_{t-1}}{\sum p_{t-1} q_{t-1}} .$$

---

[6] When taking account of quality changes, for example.

[7] Which is perfectly feasible if the base on which indices are calculated can be changed at will, as was done, for example, by the former GDR where there was virtually never any official inflation. If the price of a product increased, the increase was either explained away as a concomitant increase in quality, or the product was replaced by an item whose price had not increased. Keeping a basket of goods stable for a number of years does have an advantage that should not be underestimated: it throws up a compelling conceptual barrier to manipulation of any sort. By doing away with this, we open the door to manipulation.

[8] We prefer the term "direct" index and thus speak of a "direct Laspeyres' index". The term "fixed-base index" is a misnomer here, quite apart from the fact that "fixed" has negative connations in the view of many people.

[9] As can be seen, this type of index - in this case a Laspeyres' chain index comprises not just one basket but several (t). We will not go into the reasons advanced by supporters of chain indices in favour of this

5. The disadvantages of chain indices

In my book on chain indices[10], I endeavoured to critically review all arguments normally advanced in favour of chain indices[11] and investigated as many characteristics of chain indices as possible. It was primarily the realisation than not one of the arguments in favour of such indices is watertight that made me into more and more of a non-chainer. I will not go into details of this here, though I will list the drawbacks of chain indices that gradually became clear to me:

1. The theory on which they are based is inconsistent.

2. Chain indices violate almost all the axioms generally applied to indices, since

3. They are "path-dependent", which explains numbers 1 and 2;

4. They reflect no less than *four* types of influence (rather than just one, as is ideally the case with a pure price comparison).

5. They do not lend themselves readily to aggregation, and

6. They cannot therefore be used for deflation.

7. Finally, continued calculation of these indices by official statistical bodies will probably prove considerably more troublesome[12].

Re 1: One theoretical aspect (though mention could be made of several other inconsistencies) is as follows: chainability (or transitivity) means that an interval (0, t) can be divided up into subintervals *at will* and, since $P_{0t} = P_{0s}P_{sz}P_{zt}$, the results will be the same. Importantly, this equation should hold for *every* s and z. Thus, two six-month results or four quarters always produce the same *annual* results. Integration can thus be done consistently over time. However, this is definitely not the case with chain indices, basically because chainability and the continuing modification of the weighting scheme (for quantity q) are incompatible. Not only is the product of annual chaining with an

interval of 0 to 4 $\dfrac{\sum p_1 q_0}{\sum p_0 q_0} \dfrac{\sum p_2 q_1}{\sum p_1 q_1} \dfrac{\sum p_3 q_2}{\sum p_2 q_2} \dfrac{\sum p_4 q_3}{\sum p_3 q_3}$ different from the direct index P_{0t}^L, i.e.

$\dfrac{\sum p_4 q_0}{\sum p_0 q_0}$, it is also different from a biennial chaining $\dfrac{\sum p_2 q_0}{\sum p_0 q_0} \dfrac{\sum p_4 q_2}{\sum p_2 q_2}$.

Thus chain indices rely on chaining (multiplication)[13], but they are not chainable.

Re 2: In view of the above, the index will not necessarily read 100% when all prices in 0 and t are exactly the same. Depending on the prices and quantities recorded during the

---

approach. On closer inspection, none of the comparisons are either accurate or sound (a favourite move is to erroneously compare just one of the product's links with the index. It is not the chain component [the link] that is the problem, it is the multiplication of components [i.e. the chain]).

[10] P. von der Lippe, Chain Indices, A Study in Price Index Theory, Vol. 16 of the Publication Series "Spectrum of Federal Statistics", (published by the Statistisches Bundesamt, Wiesbaden), Stuttgart 2001.

[11] op cit., chapter 6 (pp 218).

[12] Particularly in view of the increasing frequency with which baskets of goods must be determined (as a result of, for example, household surveys). On the other hand, chain indices may be more convenient to calculate than direct indices, as no attention need be paid to comparability between more than two adjacent periods at any one time. It is generally accepted that chain indices involve a lot more survey work. However, there are clearly no real cost/benefit analyses. This matter will not be discussed any further under point 7.

[13] Indeed, they can *only* be calculated by multiplication, so errors in one link affect the chain calculated up to that point in all subsequent periods. Direct indices, by contrast, can be compiled either by multiplication *or* "directly" (usual in practice only).

intervening periods, the chain index may be more or less than 100% (in other words, it breaches the identity axiom), whereas a direct index (in fact *any* direct index) where $p_{it} = p_{i0}$ for all values of i will obviously always read 100%. Similarly, chain indices breach almost all the axioms that make indices meaningful[14]. The point is that, whilst a link *is* an index within the meaning of index theory, the chain is *not*. It does not possess the requisite properties (i.e. it breaches axioms), though each individual link does (i.e. it complies with the axiom in question).

Re 3: Chain indices do not just breach the identity axiom, they may also swing up or down at random as a result of even short-term (cyclical) changes in prices. A chain index does not really compare two situations as at 0 and t, it is a composite measure of a time series' shape. As such, it is dependent on what happened before (it is "path-dependent"), whereas a direct index will always provide an independent[15] and pure comparison between two periods. This path-dependency is the direct opposite of chainability (i.e. consistent aggregation over time)[16].

Re 4: the result of a chain index is influenced by no fewer than four variables:
1. variations, during periods 0 and t, in the prices of weighted goods[17] that are as similar as possible,
2. shifts in weightings caused by substitution (households maximising utility) between periods 0 and t, i.e. $q_{i0} \neq q_{it}$, as well as two biases specific to chain indices, viz.:
3. path dependency, and the fact that
4. for the summation of goods i = 1, 2, ..., n over the periods 0, 1, ..., t, the i-th item need not always be identical (i.e. changes in the "domain of definition" of the index function).

For the MUICP, there is a fifth source of variation, namely the country weightings that change over time[18].

"Pure price comparison" could be taken to mean that the first source of variation should be isolated. If we attempt to do just this with the direct Laspeyres' index, we are not denying the enhanced dynamics of the world of merchandise. We are not acting out of nostalgia, we are merely acknowledging that the fact that the pace of change has quickened does not mean that we should start comparing chalk with cheese, we are acknowledging that the purpose of a price index is still to measure changes in *prices* and not changes in, say, volume, which is the job of a volume index.

It may be objected that fictitious price developments of this type are not of interest. This can be countered by arguing that, for analytical purposes, it is better to have separate and

---

[14] It is largely up to the individual to decide which axiom violation is most serious. In a review of my book, W. Krämer singles out violation of the mean value property: a chain index may be larger than its largest consitutent link (partial index) or smaller than the smallest. This affects aggregation and deflation, problems discussed in points 5 and 6.

[15] ie. independent of what happened in the intervening periods.

[16] basically, all my objections come down to problems of aggregation, be it in time (i.e. over intervals of time) or for goods (i.e. partial indices).

[17] again, "goods" also includes services.

[18] or even changes to the group of countries as a whole. With the MUICP, it is assumed that the problem of new countries joining the EMU can be circumvented by "linking in" (as if this did not adversely affect comparability with previous periods!). The same applies to methodological changes, e.g. broadening coverage. The index continues to be calculated, with the inclusion of, say, certain insurance services that had previously not featured in it. If, as has happened, people are less than scrupulous when it comes to comparability, what appears to be an inflation rate may be attributable, in part at least, to a simple change in methodology.

"pure" measures for different phenomena, and to correlate these, than to have a single measure that somehow includes everything.

In order to take account of substitution effects when measuring inflation, i.e. the second of the four sources of variation listed above, there may well be micro-economic reasons, which we will not go into here, which call for a COLI approach (as in the USA). The only important thing to realise is that the standard objection to the *fixed* (and possibly outdated) basket of goods P_{0t}^L can also be countered with indices which are under discussion as COLI approximations and which take account of the most recent basket of goods in each case (in much the same way as the direct Fisher or Törnquist indices).

If it were merely a question of taking current volumes (q_{it}) into account in addition to those of the base period (q_{i0}), use could be made of such indices, obviating the need for a chain-index approach and eliminating the third source of variation (path-dependency), which is not only unwanted but also very difficult to justify in theoretical terms.

However, the main difference between the direct Fisher index P_{0t}^F and the chain index, e.g. a Laspeyres - ($\overline{P}_{0t}^{LC}$) or Fisher chain index ($\overline{P}_{0t}^{FC}$) is not just the fact that P_{0t}^F can be interpreted as a COLI, whereas the chain index manifestly cannot, but also the fact that the i-th values for P_{0t}^F in 0 and in t are *for the same commodity*[19] (or at least should be), which in practice is not easy to achieve, whereas such considerations do not apply to chain indices.

However, the price of the much-vaunted "flexibility" with a chain index is another source of variation (the fourth) of unknown magnitude. This means that the definition domain of the index function can be changed at virtually every link. At $\overline{P}_{0t}^{LC}$, say, inflation may thus be greater or smaller (by an unspecified amount) than with the pure price comparison P_{0t}^L, because the $\overline{P}_{0t}^{LC}$ conceals a selection of goods that has been changed repeatedly over time[20].

6. The analytical value of chain indices

If, as reports suggest, the ECB is under pressure to depart from its two-pillar approach with its focus on the volume of money and concentrate on targeting inflation directly, it is obviously important to know which influences in particular the index represents, to what extent they can be called "inflationary" and to what extent they react to the central bank's instruments. Seen in this light, a complex and flexible method of measuring a target parameter is not necessarily advantageous.

Aggregates of partial indices are useful not just as targets, but also for analysing the causes of inflation. The fact that chain indices cannot be consistently aggregated (using partial indices) and do not therefore show price linearity[21] detracts from their analytical value. Poor aggregation properties (cf. point 5 above) mean that laymen and external analysts who do not perhaps have access to the detailed data held by Eurostat or the ECB cannot calculate and compare ad hoc variants of the HCPI with or without certain components (e.g. administered prices, highly volatile prices, etc.) - something that they

---

[19] This is true of all direct indices, but not of chain indices.
[20] Not that this was done with reluctance owing to practical difficulties with meeting the requirements of the Laspeyres formula. Indeed, it was done quite consciously in the light of the obsession with ultra-recent weightings and with the virtually instantaneous inclusion of "newly significant goods".
[21] This is a particular requirement as aggregation consistency.

were previously able to do with conventional mechanisms for measuring inflation (i.e. direct indices).

Then there is the problem of the usefulness of chain indices as deflators (cf. point 6 above), i.e. the properties the resultant "volumes" possess. The SNA identifies this as a shortcoming (though it plays it down), as the *volumes*[22] are not "additive", i.e. do not satisfy the same definitions as the *values*[23] from which they are derived by deflation. There are other disadvantages too[24].

The volumes generated by chain indices are not, for example, proportional in terms of volume, which is why the identity axiom does not hold: even if all quantities in periods 0 and t are identical, the "volume" (i.e. the proxy for the sum of individual quantities which is not defined) may well be larger or smaller in t than in 0, which is again a corollary of path-dependency but obviously impairs interpretation of the "volumes".

7. Country-based aggregation

In addition to chaining (for which there is no real justification[25]), the HICP also features country-based aggregation (e.g. in the case of the MUICP[26], this means using the EMU countries to produce a European inflation rate)[27]. The simple idea was had of constructing the MUICP links as a weighted[28] mean of national HICPs, with country weightings (c_m) for k member states (m = 1, ..., k) thus:

$$M_{0t} = \left(\sum_{m=1}^{k} c_{m1} H_{m1} \right) \cdot \left(\sum_{m=1}^{k} c_{m2} H_{m2} \right) \cdot ... \cdot \left(\sum_{m=1}^{k} c_{mt} H_{mt} \right) = M_1 \cdot M_2 \cdot ... \cdot M_t,$$

where H_{mt} is the link (from t-1 to t) for country m (i.e. its HICP increase).

The problem with this index is as follows: if the annual national inflation rates are similar, the MUICP is basically redundant, but if they are different - which is likely to be the case, since the constituent countries (k) will have very different standards of living, and some goods cannot be traded across borders (housing rents, for instance) - what is the MUICP actually measuring? Nor is it clear exactly how the aggregation (or more prcisely: the averaging) of price level based on countries with changing weightings should be interpreted in terms of economic theory.

The MUICP is *influenced by a number of factors* that may be more or less interpreted as manifestations of "inflation":

- aggregation of the **country** price levels,

- aggregation over **goods**, the type and weighting (in the baskets) varying from country to country and from period to period, and

---

[22] i.e. as aggregates "at constant prices", in so far as this has any meaning in connection with chain index deflators.

[23] aggregates at constant prices.

[24] For details of these, cf. chapter 5 of my book "Chain Indices".

[25] as can be seen from the concepts discussed above - pure price comparison and COLI (i.e. the idea of the need for income compensation so that the household (what household?) can maintain the same level of utility). The formula contains not just one, but kt baskets of goods, nor is there a constant utility function for a "European" household living everywhere simultaneously.

[26] Monetary Union Index of Consumer Prices (MUICP).

[27] Politicians are never happy to look at more than one figure at a time. The tendency is always therefore to take several figures and combine them into one, even if they measure completely different things.

[28] True to the obsession with weightings generally shown in the field of indexing, experts concentrated on topics such as how to produce country weightings that were as up-to-date as possible, and how to prevent a high-inflation country disproportionately affecting the overall inflation rate.

- aggregation over **time**, which because of path-dependency, always has a deleterious effect in a chain index, and finally

- the constantly changing **calculation bases**, e.g the inclusion of new goods or sales outlets.

We do not share the widespread obsession with ultra-topical index weightings. Instead, we think the most important thing is to produce a pure price comparison and let the indices (e.g. expenditure ratio or mean increase in prices) speak for themselves. After all, experience shows that the more opaque a measuring approach is, the greater the tendency will be to manipulate findings to suit political ends and to attack the method itself[29]. Nor should we forget that the ECB and the euro-curreny have yet to undergo the acid test.

---

[29] cf. the example of the Boskin commission.