

Challenges of collecting prices from the internet - case Statistics Finland

Key words: Internet sales, price collection, concept of HICP/CPI, quality adjustment

Abstract:

The concept for the European Harmonised Index of Consumer Prices and therefore the Finnish HICP/CPI is a domestic concept. It should cover sales in all outlet types within the territory of the Member State. The traditional way of buying can be covered easily - if you are in a country and go to an outlet to buy a product/service, the consumption and the price is covered by the CPI/HICP of that country. If you cross a border, the CPI of the other country covers the consumption. The share of shopping from the internet has doubled in Finland in five years time.

The other issue in collecting prices from the internet is quality adjustment, price surveillance and quick change of the products. The price collection from the internet requires a different kind of approach than that from the "normal" outlets. In traditional price collection the price collector has knowledge from the previous product and she/he can use the staffs knowledge in the case of forced change of the product. When collecting prices from the internet the price collector has to make the decisions by her/himself and there is no expert assistance in the quality adjustment available.

In addition, the old way of buying from the (mail order) catalogues has moved into the web order. For example, the package holiday catalogues have disappeared in the last years and the holidays are bought from the internet. If the package under the price collection has been sold out, you are not able to collect the price. From the quality adjustments point of view this is a new kind of challenge.

This paper brings up the questions/challenges that the CPI staff of Statistics Finland has faced in the recent past on this subject.

1. INTRODUCTION OF CASE FINLAND	2
2. SHOULD THE CROSS-BORDER PURCHASES BE INCLUDED ..	3
3. PRICING QUESTIONS	4
3.1. Is there difference where you buy?	4
3.2. What is the final price for the product?	5
4. THE PRICE COLLECTION FROM THE INTERNET IN PRACTICE	6
5. CHALLENGES OF COLLECTING PRICES FROM THE INTERNET	7
6. CONCLUSIONS	8

1. Introduction of Case Finland

Each mid-month, Statistics Finland interviewers collect altogether around 50,000 prices on 483 commodities from approximately 2,700 outlets for the Consumer Price Index. In addition, some 1,000 items of price data are gathered by centralised collection. The centralized collection covers also the internet collection.

The market situation from the point of internet sales has developed vastly in recent years as the number of internet lines have increased among consumers. In past five years time the purchases on the consumer goods and services has increased in Finland from approximately 2.7 billion in 2007 to 5.1 billion in 2011¹. This is approximately 6.1 per cent from the total consumption. If we cover also the purchases made from foreign outlets and some services, the amount raises to 10 billion².

The biggest groups in the internet sales are travelling (accommodation, travelling tickets, package tours), transportation (cars, boats, motorbikes, snowmobiles) and, electronics and computers. The growing sector is purchase of banking services and groceries. On the other hand, the purchase of contents, i.e. loading music, software, games etc. has decreased².

When studying the internet sales in domestic outlets and foreign outlets the situation changes. There are differences in the consumption habits in different categories, e.g. travelling is not in the top 10 of the purchasing from abroad - and actually, this purchase should not even be included in the CPI concept as the consumption happens also abroad. The biggest groups in the internet purchasing from abroad are medicines, sports equipment and spare parts for cars².

The problematic issue is how to get the information on the cross-border purchases. There are some studies on the subject at least in Finland, but they base on questionnaires and for the consumer it is not necessarily clear if the outlet is domestic or not - even the language does not tell the information. In addition, it is not easy task to receive the data of the purchases from foreign

¹ Statistics Finland: Use of information and communications technology statistics release 2011.

² TNS Gallup: The statistics on net purchases 2011

outlets. For example if a Finnish person buys contact lenses from a Portuguese company, the purchase is only in the registers of this Portuguese company and the purchasers knowledge. Therefore, there should be a separate survey on the subject and this causes costs and response burden. The question is whether this can be executed in a cost-efficient way.

2. *Should the cross-border purchases be included*

In Finnish CPI/HICP internet purchases are included into the index, but only from the outlets operating in the Finnish territory, i.e. they have to be considered as Finnish businesses and have at least a branch in Finland. The question is whether the purchases from the other countries via internet should be included into the Finnish CPI and thus in any other correspondingly. In the European case this is an important and relevant decision as the purchases should be treated in the same way in all EU-countries. Otherwise the same products would be included in several CPIs or if neglected, the price development would be excluded from the European level HICP figures.

In Jan Walschots paper³ presented in the Ottawa group meeting 2011 there is a good point brought to discussion on this matter. The coverage is often (and at least in European Harmonised Index of Consumer Prices, HICP) defined to domestic concept of consumption expenditure.

One way of considering the domestic concept can be found from the point of consumption. As in services the HICP regulations and standards say that the service should be entered into the month index when the consumption can commence, e.g. package tours are normally bought and paid at least one month before the journey, but they are included into the index for the month the journey begins. The similar approach could be implemented in purchases: the purchase should be in the index of the country where the consumption occurs.

This causes another question as mentioned in the Jan Walschots paper, the traditional cross-border consumption. Walschot would exclude this consumption, but still there is the question of consistency in treating the consumption. For example, a person lives in France and wants to buy a computer from the Luxembourgish outlet. In the case he drives to Luxembourg for the purchase, the consumption would be included in the Luxembourgish CPI. But if the same purchase is done from the internet, the consumption should be included into the French CPI. In both cases the money goes to Luxembourg and the consumption occurs in France, but the treatment differs. The question is should it differ?

In this paper we do not even try to go to the more complex cases as Walschot already well covers these in his paper. The similar and consistent approach should be found to all cases regardless.

³ Jan Walschots: The treatment of internet purchases in Consumer price statistics Harmonising their treatment in the HICP

3. Pricing questions

3.1. Is there a difference in where you buy?

The target of the CPI is to describe the price development as well as possible. The outlets should reflect the universe of purchase pattern, i.e. all relevant outlet types should be covered. Increasingly the outlets operate in two or three ways: they have an outlet where consumer comes and shops, and additionally there still exists mail order catalogues and thirdly the internet pages for shopping. The mail order catalogues are maybe the recessive channel for purchasing in the traditional way but they still can be used as helpful tools for the internet shopping.

If the outlet works only in the internet or only as a traditional outlet, the pricing question is easier. But if the outlet operates in both ways there is question whether the pricing differs - or merely whether the price development differs in the outlets and the internet. Other costs are also a question, but they are covered in the chapter 3.2.

If the prices are same, the question is whether the collection from both outlet types is necessary. In Finland for some outlets the pricing is the same for products sold in outlets and in the internet and therefore the development is same. In these cases it doesn't matter whether the prices are collected by price collectors from the outlets or by CPI staff from the internet. The only point is the weight the product gets.

Let us take an example: women's trousers in Finnish CPI. Currently there are 168 prices collected from the outlets and four from the internet outlets. Only one of these internet outlets operate solely in the internet. The other three have a similar pricing strategy in the outlets as they have in the internet.

In the Finnish CPI the index for a single product is a weighted average of regional indices. The regional indices are geometrical means of price relatives. There are six regions and in one region the number of prices for women's trousers varies from 9 to 39. And then, there is the internet sales, where four prices are collected centralized and considered as one price observation in each region.

If the price collector would collect the prices from these outlets and only one price would be treated as internet price the weight of this internet price, thus its effect to the index would increase, it would exceed the weight of the internet purchases, and internet purchases would be over represented in the index. If the pricing stays consistent, this does not have that big impact to the overall index. But if there is a change in pricing manners, this could cause bias into the index.

Table 1. The change of weights if the treatment of internet pricing would be changed for women's trousers in Finnish CPI.

Region	# of price obs + 1 internet obs.	weight of a single internet obs. %	# of price obs + 3 extra obs. + 1 internet obs.	weight of internet obs. %
1	36	0.69	39	2.56
2	37	0.68	40	2.50
3	22	1.14	25	4.00
4	40	0.63	43	2.33
5	25	1.0	28	3.57
6	10	2.5	13	7.69

The second case is maybe easier from the price collections point of view. If the prices in the outlet differ from the prices in the internet, the prices should be collected from both in the proportion of the consumption. At least in Finland some electronics outlets give the information that the prices in their internet pages are valid only for the internet purchases and they differ from the outlet prices.

3.2. What is the final price for the product?

One question in the internet purchase is the final price of the product. The extra costs vary from a product group to another. For clothing there normally are the handling and postal costs that are added into the price, for tickets and package holidays there are the ticket costs and for holiday rentals some producers have service charges for example. Each of these costs are paid by the consumer and should be included into the index.

The easiest part is the extra charges that can be allocated to a single product or service. If the package holiday provider charges for example EUR0 for the print ticket and EUR10 for posted ticket these charges should be included in the consumption proportion into the ticket price. The same applies for any other product from delivery costs of fridge-freezers to service charges for holiday cottages in the cases where you normally buy just one product or service and/or the charge is per product or service.

The more complicated case is the internet purchases of several products with one handling charge. How should this be treated? The ideal way would be to divide it with the average number of products bought from the outlet. Say in the internet sales outlet the normal customer buys 10 items in one purchase action and the handling fee for his purchases is EUR17.50 so it would be EUR1.75 per product.

The handling of postal charges is another question - should they be included into the product price or as a separate item under postal charges? On one hand, the charge is relevant for the consumer and related to the price of the product and he has to pay the postal charges to receive the product. On the

other hand, it is not the charge the outlet receives but the postal service charge for its service. If the second choice is followed, how should the “free of postal charges” -offers be treated?

4. The Price collection from the internet in practice

The outlets from which the data for the Consumer Price Index are collected represent the structure of the retail trade as closely as possible with regards to the size of central retail corporations and outlets. The aim is to describe regional differences as well. For the time being, statistical sampling can only be used to select stores for daily consumer goods because a sufficiently reliable and accurate outlet-specific frame is not available for other types of outlets.

The data for the Consumer Price Index are collected direct from about 2,700 outlets by Statistics Finland interviewers. There are altogether 44 interviewers and they work throughout Finland collecting prices at 6 regions. The interviewers collect prices for the Consumer Price Index monthly between the 10th and 20th of the month. These outlets are daily consumer goods stores, petrol stations and specialised stores. Specialised stores comprise retail outlets (e.g. clothes, household appliances and furniture) and service establishments (e.g. hairdressers, photographer’s studios and optician’s).

In addition to the price, collection made by Statistics Finland interviewers’ CPI staff gathers information by centralised collection of prices. These prices are the same throughout the country (e.g. tobacco, train journeys, mobile phone calls, television licences, electricity and interest on housing loans), are based on other statistics (e.g. alcohol, electricity and housing) or for which price collecting is otherwise most appropriate by a centralised method (e.g. new cars, holiday cottages and package holidays). The data providers comprise large enterprises, organisations and public authorities. The CPI staff collect centralised prices in a different frequency depending on the commodity. Collection frequencies vary from monthly to twice a year. For example, housing, mobile telephone services, cruises on ferry and holiday cottages are collected monthly, medicines and different kind of fees are collected quarterly and taxi and bus journeys and postal charges are collected twice a year.

Consumer prices are also drawn from the internet and from the mail order catalogues (internet sales). These prices contain shoes, household textiles, consumer electronics (e.g. DVD’s, mobile phones, digital cameras), music downloads, rental of the movie, contact lenses, photographic services and clothes (children, men and women). CPI staff collect internet sales -data monthly.

Both the centralised and internet collection comprise together around 250 outlets or other collection targets (Table 2).

Table 2. Key indicators of the Consumer Price Index 2010=100

Key indicator	Quantity
Commodities	483
Outlet collection	2,700
Centralised collection and Internet sales collection	250
Price data per month	51,000
Interviewers collecting prizes	44
CPI team members participating centralised collection from internet and internet sales collection	7

5. Challenges of collecting prices from the internet

Price collection from the internet brings its own challenges to the work of a price collector. If examining the centralised collection several remarks can be made.

5.1. Centralised collection

There are no catalogues to use and prices of the products and services are more often available only at the internet. Examples of these are rental of holiday cottages or package holidays both domestic and international. The consequence from this is that prices of cottages or holidays must be collected at a very early stage, even several months before compiling the Consumer Price Index. After the cottage is rented from the internet, the price of a certain cottage or even all information about the cottage cannot be found from the internet pages. That means that price collector must systematize price collection in a different way than earlier. Price collector is compiling the index on the current month but at the same time he is collecting prices for next summer/winter etc. Moreover, it cannot supposed either that the collector could ask the prices of products or services from the producer because of the response burden caused to the producer.

Its own challenge is given of course by the fact that every product or service producer have their own kinds of web sites. Some of them are easy to use from the price collectors point of view and vice versa. Also the changes are more difficult from the quality adjustment point of view. When utilising catalogues if the package holiday changed in the season t , the quality adjustment could be done by choosing the new holiday from the catalogues that existed in both seasons $t-1$ and t . So, the price change taken into account in the CPI was the real price change. Now it is not possible to reflect the real price change but make an estimate somehow as there is no longer information available from $t-1$.

In the past Statistics Finland interviewers collected the prices of computers and laptops but because price collection was quite challenging at the outlets, at present the collection of computers and laptops belong to the centralised collection. Changing of models and types and fast progress makes the collection of computers and laptops very challenging. It would simplify the

work if the price collector has up to date expertise about the features of the computers and laptops. As other challenges it is difficult to follow the price of same computer because of the fast development of computers and collection is quite time-consuming.

5.2. Internet sales collection

Shoes, household textiles, consumer electronics, clothes, contact lenses and photographic services are collected from the internet sale. The CPI staff collects this data monthly.

Products and services, which are sold at the internet, usually change and vary at high intervals. The collections changes several times a year. For example, if a certain amount of different sizes and colours of a certain women's shirt to be sold in the internet pages of a particular seller. After the colour or size or even the whole shirt is sold out, the shirt cannot be found from the internet anymore. It depends a lot of how fluently price collector can collect prices from internet pages. Some pages are very easy to use at the collectors point of view but sometimes it is very hard to find same products or services as in the last collection round. Usually the internet link of a certain product or service is saved for the next time to ease work. Sometimes collectors have to get a username in order to collect prices. Seasonal products are a special case. For example in February it might be quite challenging to find prices of winter coats because collections may have changed and there is only a new spring collection available. Excellent service for the customers, but challenging to the price collector.

6. Conclusions

Purchasing from the internet is an increasing way of consumption and therefore it should be taken into account when making the outlet samples for the CPI price collection. At least the domestic internet sales should be covered in the CPI for each country when it is a relevant part of the consumption.

The CPI staff has to follow the market as the market share proportions change over time. Some product groups might lose their importance in the internet market while some, even surprising product groups might become important from the internet sales point of view. The CPI staff should have up to date information and as to the changes in the consumption habits and to follow media and public conversations.

The role of the CPI staff as a specialist in price collection methods increases in the future. The most efficient way of collecting prices, at least in near future, is to visit websites for the price collection. This is the way to receive the representative and authentic prices, the consumers have to pay. In some cases, you even have to go almost through the process of purchasing to receive all the relevant price information.

In this paper there are maybe more questions raised than answers given from the coverage point of view for discussion. Some solutions for practical issues can be given from the Finnish experience on the internet sales. We have no doubt that the internet sales shouldn't be included from the scope of out-

let representativeness of the CPIs. But the biggest question still remains - if purchased from abroad, should the consumption be included into the domestic index or not. The most important thing is that these purchases are handled in the same way in all countries, especially in Europe to ensure the quality of the overall European index.