

Designing a CPI in a developing country

Some experiences

Version 1

by

Jörgen Dalén

Abstract: This short note presents some experiences from the part of the author when working as a consultant on CPI design for DIGESTYC, Ministry of Economy, El Salvador. It is believed that some of the approaches on index techniques, sampling, product change and special products can be of more general applicability in countries of a similar size and stage of development as El Salvador. In the end, some specific recommendations are offered for wider discussion with the aim of developing a shorter manual for this purpose.

Introduction

The CPI Manual is very useful, when designing a CPI with all its detailed methods and approaches, since it includes a discussion of most aspects of CPI construction. However, it is short on concrete advice and recommendations. This is very apparent when there is a need to design a CPI for a small developing country with limited staff resources. There are probably in the order of 100 such countries in Asia, Africa and Latin America. Hence there is a need for developing a shorter and simpler manual that is relevant for such countries. This brief note puts forward some proposals that may be useful with this objective in mind.

The author is involved in the detailed design of a new CPI in El Salvador in Central America. The project, which started in March 2007, builds on a household budget survey (HBS) carried out in 2005/6 and will result in a new CPI replacing the old series based on an HBS of 1992. The project covers in essence all aspects of CPI design and is planned to continue until early 2010.

This note presents selected issues from the project, where it is believed that the methods used or approaches taken could have a wider interest for countries in a similar stage of development as El Salvador. The issues discussed are:

- Product selection and outlet sampling
- Monthly pseudo-chaining
- Replacements, overlaps and quality change
- Temporary price reductions
- PCs and the monthly chaining method.
- Mobile phones and user profiles
- Central price collection.

The issues are not discussed in any depth. But it is hoped that the ideas put forward here could be contrasted to other similar CPI design projects.

Country background

With a population of about 5.7 million people (July 2008 Census) on 21,000 km², El Salvador is the most densely populated nation in Central America. GDP per person was around \$ 3,500 in 2007 or 8 % of the US level (12-13 % if calculated on a PPP basis). The service sector is the largest component of GDP at more than 60 %, followed by the industrial sector at 30 % and agriculture at 10 % of GDP.

A sharp difference in income and living standards within the population has resulted in a two-tier economy. Traditional markets and modern supermarkets exist in parallel. Food products have a large share (about 26 %) of the household budget but modern products like mobile telephones, computers, cars and flat screen TVs also occupy significant shares.

During the last few years two CPI divisions within DIGESTYC have worked in parallel, one that runs the official CPI (with Dec 1992 as base) and one that sets up the new CPI with weights from the Household Budget Survey of 2005/6. The second group is planned to be the only group once the new CPI has become official during 2009.

The new CPI staff has 33 staff members, of which 6 with academic training and 25 price collectors (including 6 supervisors), who all operate from the central office. There are also three quality controllers.

Product selection and outlet sampling

A fundamental process when designing a new CPI is to determine the set of basic product offers (observation points) to follow over time. This process involves (i) the selection of (representative) products, (ii) the sampling of outlets and (iii) the matching of products and outlets (which products to look for in which outlets) and (iv) the selection/sampling of a specific product offer for a product in an outlet.

Product selection: The HBS of 2005/6 covered around 4,500 households, representing the whole country. Consumption was classified into more than 2,000 products, most of which turned out to have a very small expenditure. One of the first issues was to determine the market basket of the CPI. Around 250 products were deemed to be an appropriate number to include.

We suggested as a rule of thumb to try to include all identifiable products with above 0.1 % of total expenditures (unless there were special reasons to exclude them) but not to include products with less than 0.03-0.05 % (grey zone). Including very small products is a waste of resources, since a certain minimum number of observations would anyway be needed and their contribution to the overall accuracy of the index would be negligible. Complementary criteria suggested by DIGESTYC were (i) to include products with widespread consumption (for example among poor people) even though the weight might be small¹ and (ii) to included products with expected increase in consumption in the future (like internet service). Products, for which price collection is extraordinarily difficult or dangerous could also be excluded.²

Outlet sampling and outlet-product matching: The process of determining in how many and which outlets to price which products is one of the more important but seldom discussed issues in CPI construction. This is more than a simple sampling issue. Ideally this should be based on information on where consumers make their purchases and on the variation in price level and change between outlets. In practice, information on these matters is usually scattered and incomplete. In the Salvadorian case, there is rich information in the HBS on where purchases of food are made but less for other products. Purchases of foodstuffs was done mainly in (i) department stores (ii) supermarkets, (iii) informal outlets, (iv) traditional markets, and (v) from street traders. For purchases of non-food there was no HBS information but ad hoc information was collected in an ongoing household survey.

There are formal and informal outlets. Formal outlets were covered by an economic census in 2006 and in this case the method of systematic pps sampling (with number of employees as size measure) was used. For informal outlets inside and outside markets, a more subjective sampling procedure had to be used in the absence of more detailed information. Street traders (where food expenditures are large) are difficult to include due to their instability. Since they are often active around traditional markets and can be expected to have similar kinds of goods and

¹ Some fruits like lemons, coconut, pineapple and melons were included for this reason.

² In addition to prostitution and narcotics, which are always excluded from CPIs, this also applied to "agua ardiente", a local alcoholic drink often sold in bars in areas, where price collectors felt insecure.

price levels as traditional markets, these expenditures are to some extent represented by fixed market stands³.

When determining sample sizes (number of product offers) for each product, ideally the sample size should be proportional to the product weight (W_j) multiplied with the standard deviation for price ratios (σ_h). Whereas the product weight is a clear and constant concept, the standard deviation is much more complex. It is different for different time horizons (1-month, 12-month etc. and these differences could be substantial, especially for seasonal products), it changes over time and is not available for new products or products with new specifications. In this situation it is advisable to concentrate on product weight and make sample sizes proportional to product weight, with a minimum number of 10 or so for small products.

Product specification and product offer selection in outlet: There is a traditional approach in CPIs to use a very strict and narrow product specification for each representative product. This approach probably originates from the time when elementary aggregates were calculated as ratios of arithmetic average prices. With this form of calculation there is a need to have a homogeneous set of product offers (see CPI manual para 20.71-80). However, with the adoption of the geometric mean as the elementary aggregate formula (also used in El Salvador), there is no longer a need for homogeneity and the focus can instead be placed on representativity. The logical criterion is then to look for the most sold product offer in each outlet within a wide product specification. This is also the approach used in El Salvador.

Monthly pseudo-chaining

By monthly pseudo-chaining we refer to the calculation procedure that is very briefly described in par. 9.84 of the CPI Manual. A separate calculation is done between each two consecutive months, which is then linked to the index for the last month according to:

$$I^{0:t} = \sum w_i^b I_i^{0:t-1} I_i^{t-1:t}, \text{ where } \sum w_i^b = 1$$

$I_i^{t-1:t}$ denotes the monthly change for product i and is an aggregate (normally geometric mean) of the price changes for all product offers for product i.

Monthly pseudo-chaining provides a convenient environment for handling replacements and overlaps, since there is no need for maintaining an identical set of product offers over time. Under other arrangements, the introduction of new or changed products or outlets is more cumbersome. This is the main advantage of the approach.

The technical calculation algorithm is also fairly simple. Last month's observed price becomes the next month's base price. The only exceptions are when a price was temporarily missing or temporarily reduced last month, when instead next month's base price is calculated as last month's base price multiplied with last month's price index for the product concerned.

³ An idea proposed by the author was to relax the rule of "same outlet" in these cases and ask price collectors to look for the first available street seller who sells a product with a certain specification within a certain, limited geographic area on each instance of price collection.

Without pseudo-chaining, new base prices for the *initial base period* will have to be imputed each time non-comparable replacement occurs, which is a much more complex procedure to implement.

Replacements, overlaps and quality change.

The treatment of product change is a real challenge in countries with small resources. Overly complex methods are not viable with limited staff resources - simple and yet robust and reliable methods are needed. The approach described below is an attempt in this direction and we open up for a debate on whether it is good enough for a recommendation for wider use.

The first distinction used in the Salvadorian system is between replacements and overlaps. A replacement occurs when a product offer is unavailable without previous notice, whereas an overlap is a planned action by CPI staff.

When replacements occur, the price collector has to classify them into *comparable* and *non-comparable*.⁴ For comparable replacements the price change is taken into account without adjustment but non-comparable replacements are excluded from the calculation of price change in the month the replacement occurs.

In special product fiches (an example is included as Annex 1), the product is described together with a list of characteristics to observe. If at least one characteristic that changes between and old and a new product offer is considered price-determining, the replacement will be non-comparable, else it will be comparable. Through these product fiches, subjectivity from the part of price collectors is kept to a minimum.

The other possibility to introduce a new product offer is through an *overlap*. Due to the monthly pseudo-chaining procedure (described above) it is procedurally very easy to introduce a completely new product offer (a new outlet, a new general product specification or just a new detailed specification in an existing outlet). It will automatically be included for price comparison in the second month it is observed.

An overlap is used for example when the price collector observes that a product offer is no longer well sold but before it completely disappears, which means that there will be a month in which both product offers are observed. This case is determined in the field by the price collector, whereas other overlaps can be initiated from the central office.

No direct quality adjustment is done. Direct quality adjustment using specialized procedures like hedonics, option prices or other methods are generally beyond reach for small statistical offices without sufficient staff resources, such as the one in El Salvador.⁵

As a first step towards evaluating the methods, implicit quality indexes (IQIs) have been calculated. IQIs are defined as the ratio between the change in average prices and the actual price index for a product. The idea is that, if average prices increase more than the actual index, this implies that the product quality has also increased (and vice versa). Of course, many other things such as sampling procedures can influence the IQI but on average, over many products

⁴ When a product offer is no longer available, the price collector is to inquire as to whether it is permanently or just temporarily unavailable. If it is permanently unavailable a replacement should be sought immediately but if it is temporarily unavailable they should wait until the third month before replacing it (unless it returns).

⁵ Direct quality adjustment for cars using for example option prices or fuel consumption adjustments is still under consideration.

and longer time periods, quality change should be a major factor. Since there is a broad priori understanding about the relative quality changes in different product groups (quality increases for technological products but less so for traditional products like food, furniture and clothing) IQIs are useful as an analytical tool for judging the risk for quality bias in the index.

Table 1 shows average IQIs for the first year of the new Salvadorian CPI.

Table 1: Average Implicit Quality Indexes, December 2007 – December 2008, per product group

Division	Products included	Average IQI
Coicop 1	Food and non-alcoholic drinks	0,9997
Coicop 2	Alcohol and tobacco	0,9799
Coicop 3	Clothing and footwear	1,0122
Coicop 5	Household products	1,0071
(Coicop 6)	Pharmaceuticals	0,9975
(Coicop 7)	Petrol and car accessories	0,9876
Coicop 9	Recreational products	1,0282
Coicop 11	Food away from home	0,9891
Coicop 12	Other products	0,9958

() means that not all products in the division are included in the comparison

We see that in all divisions average IQIs are close to one and for coicop 1, where quality change is least likely to occur, it is almost exactly one. Within recreational products, there is a large upward movement for TVs (1,40) explained by replacements of old TVs for flat-screen TVs.

The IQI analysis in Table 1 is a first indication that no major problems are inherent in the replacement methodology. Since this methodology is also easy to use, we feel that the approach based on the simple comparability distinction could be a useful one for developing countries.

Temporarily reduced prices

The issue of temporarily reduced prices was discussed at some length in the beginning of the project. The approach recommended in the CPI Manual (par. 6.83) is to allow temporary price reductions into the index.

However, in the context of using the simple comparability criterion for replacements as described above, this practice risks to run into problems. This is because, if the price reduction is due to stock clearance or is in the end of the product life cycle, there is likely to be a replacement in the next month to a product offer with a normal price. If this replacement is classified as non-comparable, then the price increase that normally follows after the end of sales will not be recorded in the index. This leads to an under-estimating bias – a problem that has been experienced by many countries for clothing products.

For this reason it was decided not to include temporarily reduced prices into the index. An exception was made for technological products, where the price reduction at a stock clearance can be seen as a measure to make the price/quality relation of the old product offer match that of a new, higher quality variant.

Mobile telephony

Mobile telephony is by now spread all over the world, from the poorest to the richest countries. Its share of private consumption is likely to be around or more than one percent. In the Salvadorian case the CPI weight is around 2 %. It is therefore a very important product in any CPI but at the same time one of the most complex ones with regard to the pricing mechanisms.

The pricing mechanisms in the mobile telephone markets are characterised by a system with multiple tariffs, involving different combinations of price per minute, inside/outside own network, peak/ offpeak, fixed monthly fee (sometimes with a number of free minutes) and other conditions. For a regular weighted index calculation, expenditure information is needed for each tariff component. And this is still not enough in the frequent situations, where new types of tariff components are introduced.

The only viable solution, as has been experienced in many other countries, is a methodology relying on user profiles. It could be asked, however, whether such a methodology is too complex for a developing country and whether sufficient information exists to back up the profiles.

In the Salvadorian case the Household Budget Survey of 2005/6 had information on the expenditure on mobile telephony for each sampled household, which could be used for establishing consumption levels for a low, a medium and a high budget for the two products included in mobile telephony – prepaid cards and postpaid invoices. The 2x3 profiles were estimated so as to approximately conform to the expenditures at the 20th, 50th and 80th percentile of the household distribution for each product. This resulted in the following expenditure profiles:

- For prepaid cards: Low=\$5, medium=\$10, high=\$19
- For post-paid invoices: Low=\$14, medium=\$30, high=\$50

Certain assumptions and simplifications were needed

- The consumer chooses the best tariff for his profile (this is a basic assumption in the user profile method)
- The consumer requires an always activated phone
- Temporary discounts were not taken into account
- Only the most applied tariffs were accounted for
- Only main services were included so far (not SMS, some international calls, Internet browsing from mobile and other low-expenditure services)

Based on these simplifications it was possible to specify more detailed profiles including number of calls/minutes to fixed/mobile, same/other operator, national/international calls etc. These specifications had to be based on some further assumptions, since the major operators were not willing to share their detailed call structure with DIGESTYC.

Nevertheless, it has proved possible to include new tariffs according to the given model. The HBS information, which is likely to exist in most countries, was a key information component.

PCs

PCs occupy a weight of 0.34 % in the Salvadorian CPI. This weight is of the same order as in developed countries (weights can vary widely also in developed countries).

Hedonics is out of reach in a small developing country like El Salvador. It is also clear that it is not acceptable to treat PCs just as any other product due to the continual quality improvements and rapid model turnover in this market. This would lead to an over-whelming number of non-comparable replacements with quality improvements, which would go unrecognised.

The (monthly) chaining approach, described in the CPI manual, para 7.153-158, provides a natural way out of this dilemma. We will describe how this procedure is applied in El Salvador.

Five major outlets (PC sellers) in the capital of San Salvador are included in the sample. In each outlet the five most sold PCs or PC systems were selected in the first month of observation (March 2008). Each PC model is followed over time according to the exactly same specification until it is no longer sold. Then a new PC (the most sold of the new ones) is brought into the sample but without comparison backward (in the first month it does not contribute to the index). In this way there are always some models in each outlet that contribute to the index. Out of a maximum of 25 (5x5) models, the average number included in the index in the period March 2008 – March 2009 was 18.4. The PC index for the same period was estimated to 85.2.

This method requires expert price collection, i.e., a staff person dedicated to this task and with a special interest in this field and the special requirements of the method. In El Salvador this is so far one of the academic staff persons.

Central price collection and the Internet

Certain products are not sold in normal outlets or else the price collection approach needs to be different from standard price collection in the field. Every country likely applies central price collection for a number of products. Information on micro-weights for different categories within such products are often available, sometimes on websites, sometimes after requesting the information from companies or relevant government agencies. Below, it is briefly described how central price collection was applied in El Salvador. (For mobile phones and PCs the description is above.)

For each product there are special calculation schemes, usually included in special purpose spreadsheets.

Rents

A special rent survey is carried out by price collectors for dwellings that were rented according to an ongoing household survey of 2007. An important principle in a CPI rent survey is that the rent for a fixed dwelling is followed over time regardless of change of tenant. It is ensured that the price concept is unchanged (with/without electricity water, telephone etc.)

(Owner-occupied housing is excluded from the Salvadorian CPI.)

Education

Information on prices is available at the Ministry of Education, which has to approve price changes. For pre-university education, all schools report their price changes to the Ministry, which makes them available to DIGESTYC. A calculation is done (on a sample basis since data on

1000 schools are received on paper format) knowing that the remaining schools did not change their prices).

Three of the major universities are included and here information on student enrolment in different courses is known along with the fees applied. Calculation can take on a weighted form with number of students as quantity weight.

Electricity

All price information for the five suppliers operating in El Salvador is available at the regulatory authority for telecom and electricity. For small consumers, prices are subsidized and after some effort it was also possible to obtain detailed enough information on the subsidy system. Since tariffs depend on amount of consumption, it was necessary to include information on households in different consumption brackets into the calculation. This information was found without difficulty in the website of the regulatory authority.

Hospitals

A mail survey to hospitals asks for prices for a bed one night in the hospital, normally including a nurse.

Bus fares

Bus fares are regulated by the Government and do not change often. They changed in the beginning of 2009. It was then possible to obtain prices for all 1100 bus routes in the country. A sample of these were drawn and it was manually ascertained that the routes matched those of an earlier price list.

Conclusions

Wrapping up this short paper we would like to offer some tentative recommendations for discussion. The scope for the recommendations is CPI design in small and medium-sized developing countries. A shorter manual for this purpose is certainly needed and the below proposals could then be compared and contrasted to experiences by other consultants and statistical agencies.

- (i) *Use amount of expenditures as the main guiding principle for product selection. Products with more than, say, 0.1 % of total expenditure should normally be included whereas those below, say, 0,05 % should normally be grouped together to be represented by one of its members, with the relatively largest expenditure.*
- (ii) *Sample sizes proportional to product weight, is a robust rule of thumb, if standard deviations are not fully analysed. Over-allocation of products with known volatile prices (fresh foods such as fruits and vegetables) and under-allocation of products with more stable prices (household durables) could complement this rule .⁶*

⁶ Note that petrol prices are normally stable from the point of view of CPI allocation. This is because petrol stations tend to change prices in the same proportion at about the same time. Standard deviations of price ratios will thus be small for petrol although the prices are volatile over time. On the other hand, fresh food prices may depend on the daily supply from different wholesalers which can result in large standard deviations.

- (iii) *When the geometric mean is used as the elementary aggregate formula, the most sold criterion should be used in each outlet within a wide product specification in order to achieve maximum representativity of the sample.*
- (iv) *Monthly pseudo-chaining is a useful calculation technique that facilitates the handling of non-comparable replacements and overlaps.*
- (v) *The simple distinction between comparable and non-comparable replacements, where the former go into the index without adjustment and the latter are excluded, is easy to criticize on theoretical grounds. Yet, for most products (not cars and PCs) it appears to work comparatively well in practice and it is not easy to devise a better method for a country with limited staff resources.*
- (vi) *Introduction of new outlets and product variants by overlap should be encouraged, since it serves to maintain the representativity of the sample and is easy to accomplish with monthly pseudo-chaining.*
- (vii) *The exclusion of temporarily reduced prices for non-technological products may be a way to avoid the well-known clothing bias that is caused by the end-of-sale reductions followed by non-comparable replacements.*
- (viii) *Simple user profiles are a useful method for covering mobile phone tariffs, also in developing countries. It is hard to see any other possibility, since the tariff complexity is at least as great as in developed countries.*
- (ix) *The monthly chaining method is a practical and robust method for PCs. It needs to be carefully implemented, however. In each outlet a number of models needs to be followed. New models have to be introduced each time an old model disappears. The basis for the introduction of new models must be large sales value and not similarity with outgoing models.*
- (x) *Central price collection for certain products, where prices are not easily observed in outlets should be used. It is important to realize the presence of weight and other useful information existing in websites of companies, regulatory agencies etc.*

Annex 1: Field codes

Type code (never changes)	Internal code (price collector does not fill in)	Situation code (concerning product offer status)	Price code (concerning type of price)
Blank not allowed	Blank not allowed	Blank not allowed	Blank=No price available
1=Normal	1=Normal	1=unchanged (or first month)	1= Normal
2=Seasonal product		2=comparable replacement	2=Temporary price decrease
3=Technical product	3=Overlap (first month)	3=non-comparable replacement	3= Recovery from temporary price decrease
		4=Temporary non-availability	4= Temporary price decrease for technical products (list to be produced)
		5=Permanent non-availability	5= Recovery from temporary price decrease for technical products

Annex 2 Product Fiche

Ficha de Productos de la Canasta de Mercado. IPC

División: Prendas de vestir y calzado

Producto: Falda

CCIF: 03122



FALDA: Prenda de vestir que cae desde la cintura para abajo. La primera falda, de piel, apareció hace ya 600.000 años y desde entonces esta prenda nunca ha abandonado a la mujer. En el año 1915 la moda enseñó los tobillos femeninos, pero la auténtica revolución llegó en 1965 gracias a Mary Quant, con el lanzamiento de la minifalda. La **falda, pollera, enagua o saya** es una prenda de vestir que cuelga de la [cintura](#) y cubre las [piernas](#), al menos en parte. Suele ser de tela y de forma cilíndrica o tronco-cónica, y a diferencia de los [pantalones](#) no está dividida. Según el largo las faldas pueden ser: largas, cortas o a la rodilla.

Posee 20 características en la omega, son las siguientes:

Características	Determina precio:	
	Sí	No
Talla	X	
Marca	X	
Tela		X
Composición de la tela		X
Diseño de la tela		X
Color de la tela		X
Estilo/Edición	X	
Modelo/Referencia	X	
Características		X
Unidad de medida		X
Cantidad		X
Accesorios		X
Presentación		X
SKU		X
Botones al frente		X
Tipo de lavado		X
Bolsillos		X
Tipo de falda	X	
País de procedencia	X	
Detalles		X

