



Pacific
Community
Communauté
du Pacifique

SPC Headquarters

95 Promenade Roger Laroque
BP D5, 98848 Noumea Cedex
New Caledonia

Email: spc@spc.int
Phone: +687 26 20 00
Fax: +687 26 38 18

Siège de la CPS

95 Promenade Roger Laroque
BP D5, 98848 Nouméa Cedex
Nouvelle-Calédonie

Email : spc@spc.int
Tel : +687 26 20 00
Fax : +687 26 38 18

2019 MAY PSMB MEETING

DOCUMENT N° 10:

RMI HIES EXPERIMENT USE FOR CPI

Prepared by
Mr Rens Hendriks,
Statistics for Development Division (SDD)
of the
Pacific Community (SPC)

May 2019

Pacific Community (SPC) Headquarters: Noumea, New Caledonia. Regional Offices: Suva, Fiji;
Pohnpei, Federated States of Micronesia; Port Vila, Vanuatu. Country Office: Honiara, Solomon Islands.

www.spc.int spc@spc.int

Siège de la Communauté du Pacifique (CPS) : Nouméa (Nouvelle-Calédonie). Antennes régionales : Suva (Fidji) ;
Pohnpei (États fédérés de Micronésie) ; Port-Vila (Vanuatu). Bureau de pays : Honiara (Îles Salomon).

www.spc.int spc@spc.int

BACKGROUND

Historically expenditure-related surveys in the Pacific have relied on a combination of household expenditure diaries and some recall questions in order to compile estimators of average and aggregate household expenditures. The data from these diaries has been used, inter alia, to support the rebasing of consumer price indices (CPI) for all PICTs. However, in recent years there has been a move towards a much greater reliance on recall questionnaires and much less attention on the traditional expenditure diaries. This has been supported by the introduction of computer assisted personal interview (CAPI) systems that have led to a significant cost reduction in the conduct of household expenditure surveys.

This paper assesses the comparability of diary and recall data for use in rebasing the CPI.

Recall questionnaires have been generally used to capture infrequent expenditures, such as the acquisition of household durables. The diary has been used to capture day-to-day expenditures; respondents are asked to keep a detailed accounting of all purchases and other disbursements of resources made for a period of one to two weeks. The resulting estimates are scaled to adjust for non-participation biased and to correspond to a common time period, and then summed and/or averaged to provide the estimator for the aggregate in question.

Research by the World Bank (Brzozowski, Crossley & Winter, 2017) has suggested that well-designed recall questionnaires on expenditures yield estimators of similar accuracy as those from diary-type surveys. As a rule, recall questionnaires are far cheaper to implement than surveys based on a full diary or hybrid methodology. Given the lack of resources for statistical collections in the Pacific, this provides a strong incentive to adopt the recall methodology for surveys on household expenditures and living standards.

Besides yielding estimates on expenditure aggregates and incidence rates of (absolute or relative) poverty, household expenditure questionnaires have traditionally also served as the primary source for the derivation of expenditure weights for the consumer price index (CPI). In countries with advanced statistical programmes, the availability and integration of detailed tax records, turnover statistics, supermarket scanner-data, and other data sources into detailed national accounts expenditure estimates have led to a steady decline in the reliance on household expenditure surveys for the derivation of CPI weights.

To some degree, this process is underway in the Pacific as well. Commodity flow techniques are regularly employed to generate alternative expenditure estimates for stigmatised products such as alcohol and tobacco. Furthermore, in traditionally centralised markets such as banking and telecoms, more accurate turnover figures directly obtained from enterprises are often used as a substitute for the survey-based estimates.

Nevertheless, across the Pacific region as a whole, household expenditure surveys remain by far the most important data source for CPI weights. Given the importance of accurate inflation numbers for economic policymaking, it is important to ascertain that a potential change in methodology does not have a significantly negative effect on the accuracy and level of detail of the expenditure estimates.

RESULTS

CHANGE IN LEVEL OF DETAIL

The impact of the methodological break between diary and recall will depend on a number of effects. Most importantly, some loss in the available level of detail is inevitable in the switch to the recall questionnaire. The household diary does not place limits on the descriptions respondents use to track their expenditures. As a result, it allows for the calculation of expenditure estimates up to the finest possible levels of granularity. For recall questionnaires, in order to be able to conduct the survey within acceptable timeframes and budget constraints, expenditures are by necessity grouped into categories.

Intuitively, this would mean that the use of a diary would have a significant advantage for countries that use the expenditure survey mainly as a source for CPI weights. However, there are a number of caveats. For one, the expenditure survey is typically designed to yield accurate estimates of expenditures at the COICOP Subclass (or similar) level. Even though the diary enables the estimation of expenditures at more detailed levels, it is often questionable to what degree these estimates are still accurate from a statistical point of view.

Second, the descriptions used by respondents do not necessarily correspond neatly to the product classification used. Diary entries often include records on purchases such as “fish” or “fuel” without an indication whether the product concerned falls under “tuna” or “flying fish”, or “petrol” or “diesel”. As a result, low-level expenditure estimates can have a tendency to concentrate in generic placeholder products, often described as “other” or “not elsewhere classified”. Perhaps ironically, for important goods and services, recall surveys can actually improve the low-level expenditure estimates by explicitly including the separate products in the questionnaire.

To investigate the impact of the change in detail, we calculated detailed expenditure estimates at the product level. Using this list, we investigated if there are items with significant levels of expenditure (>1%) in the diary-based surveys, but which are missing from the recall questionnaires. The resulting list is shown in the table 1 below.

Table 1: Expenditure shares of items unique to the diary with significant expenditures

Item code	Item description	arm_2 CAPI diary	arm_3 PAPI diary	arm_4 PAPI diary
0113023099	Other sea food	0.14%		
0116042006	Banana, raw, not further specified	0.03%	0.11%	0.10%
0116042099	Other fruits	0.03%	0.03%	0.28%
0117045099	Vegetables, not further specified	0.17%		0.02%
0119063099	Sauce, not further specified	0.21%	0.38%	0.34%
0119066010	Food, unspecified	0.28%	0.14%	0.03%
0119066098	Infant formula, water added, n.f.s.		0.54%	0.05%

Compared to our a priori expectations, the list of affected items is surprisingly short. Furthermore, for most of the food items, it appears that their listing is simply the result of not enough detail being available to be included in the pre-coded commodity classification used for the recall questionnaire. As such, we expect that these expenditures were not ‘missed’ in the recall questionnaires, but instead were included in their more specific product categories.

Of the entire food item list, only two items merit some specific commentary. The expenditures for *infant formula* are somewhat confusing, with a comparatively large expenditure share for arm_3, and negligible expenditures in the other two arms. We expect that this is the result of the smaller sample sizes utilised for diary surveys. Traditionally, in most countries around the Pacific, this product shows up as highly significant. While in theory it is possible that customs in RMI are different, it might be prudent to add this specific item to the items asked after in the recall module – potentially in the dairy section.

Second, *Sauce, not further specified* registers significant expenditures on all three diary based surveys. This however does not appear to be the result of an important missing product in the recall item list. Instead, it appears that a large number of transactions of *baking powder* were erroneously processed with the code for sauces.

In sum then, to our surprise, we cannot identify a single product with significant expenditures in the diary, which was not included in the recall questionnaire. Therefore, we conclude that the required level of detail available for rebasing the CPI in the survey does not obstruct the adoption of recall questionnaires for the purpose of measuring aggregate expenditures in the Pacific.

This conclusion comes with one important qualifier however. The diary is designed to be exhaustive, it (should) capture *all* expenditures over a certain period. There is an inherent risk in full recall-based questionnaires that an important product could be missed or be forgotten. It is therefore imperative that every country that employs a full recall questionnaire makes a significant effort to create a specific item list that includes all important consumption goods and services regularly acquired in the country.

LEVEL EFFECTS

The methodological break from the move to full recall questionnaires could conceivably affect the level of expenditures measured. Normally, this type of change is not particularly problematic for the CPI. As long as the effect is somewhat evenly distributed over all expenditure categories, expenditure shares remain largely unchanged. In this case, however, the methodological break concerns a specific group of product categories. As a result, there could be a shift of the expenditure shares of categories at high levels of the classification.

To determine whether a noticeable shift is occurring, we calculate weights at various levels of the COICOP classification. We assume for this purpose that RMI wishes to calculate an index targeting monetary consumption expenditures. We therefore remove from the dataset all cash transfers/gifts, investments, imputed transactions (i.e. imputed rents) and measures of home production. However, we included cash purchases of products for the purpose of giving to another household. Finally, we exclude purchases of vehicles. The infrequent nature of vehicle purchases, combined with small sample sizes of the diary surveys has led to large differences in expenditure estimates for this category which disproportionately affect the expenditure shares of other categories.

Table 2: HH average expenditures by COICOP division - annual US\$

	COICOP Division	arm_1 CAPI recall	arm_2 CAPI diary	arm_3 PAPI diary	arm_4 PAPI diary	arm_5 CAPI recall bounded
01	Food and non-alcoholic beverages	3,365.52	3,903.12	2,769.82	2,578.97	3,784.19
02	Alcohol, tobacco and narcotics	1,482.72	1,124.13	781.10	1,107.75	792.85
03	Clothing and footwear	1,127.48	927.23	724.60	700.90	735.42
04	Housing, water, electricity, gas...	1,054.65	1,981.80	1,074.68	1,241.72	1,530.50
05	Furnishings, household equipment...	529.71	866.39	342.22	437.16	287.87
06	Health	15.11	41.47	25.07	6.27	13.22
07	Transport	946.96	1,306.01	836.92	999.02	737.14
08	Communication	223.23	495.45	220.28	440.36	366.27
09	Recreation and culture	208.61	779.84	203.85	302.21	236.12
10	Education	149.57	147.58	136.83	74.23	107.07
11	Food away from home	1,624.49	2,065.53	1,476.90	1,545.99	1,514.89
12	Misc. goods and services	670.26	491.41	242.93	137.77	588.65
00	Total	11,398.31	14,129.96	8,835.19	9,572.34	10,694.20

Table 3: Expenditure shares by COICOP division

	COICOP Division	arm_1 CAPI recall	arm_2 CAPI diary	arm_3 PAPI diary	arm_4 PAPI diary	arm_5 CAPI recall bounded
01	Food and non-alcoholic beverages	29.53%	27.62%	31.35%	26.94%	35.39%
02	Alcohol, tobacco and narcotics	13.01%	7.96%	8.84%	11.57%	7.41%
03	Clothing and footwear	9.89%	6.56%	8.20%	7.32%	6.88%
04	Housing, water, electricity, gas...	9.25%	14.03%	12.16%	12.97%	14.31%
05	Furnishings, household equipment...	4.65%	6.13%	3.87%	4.57%	2.69%
06	Health	0.13%	0.29%	0.28%	0.07%	0.12%
07	Transport	8.31%	9.24%	9.47%	10.44%	6.89%
08	Communication	1.96%	3.51%	2.49%	4.60%	3.42%
09	Recreation and culture	1.83%	5.52%	2.31%	3.16%	2.21%
10	Education	1.31%	1.04%	1.55%	0.78%	1.00%
11	Food away from home	14.25%	14.62%	16.72%	16.15%	14.17%
12	Misc. goods and services	5.88%	3.48%	2.75%	1.44%	5.50%
00	Total	100.00%	100.00%	100.00%	100.00%	100.00%

Table 2 above shows the expenditure estimates by COICOP division for each of the survey arms. As can be seen, there are significant differences at the aggregate level, as well as for individual COICOP divisions. In division 12, the highest estimate (arm_1) is almost five times higher than the lowest estimate (arm_4).

Nevertheless, when looking at table 3, it is clear that similar consumption patterns occur. Perhaps more importantly, there does not appear to be a clear pattern of over- or underestimation of expenditure shares for specific divisions in the recall questionnaires compared to the diary-based questionnaires (with the exception of division 12). This would suggest that the large variation in expenditures is more likely the result of sampling variance, rather than the change to recall questionnaires. At lower levels in the classification, similar patterns occur. There is nothing in the data to suggest that the expenditure shares for the recall questionnaire are consistently over or underestimated compared to the diary-based questionnaire.

However, the expenditure shares for the low-level items vary significantly, even between the recall questionnaires themselves. Table 4 shows the expenditure shares of the ten most important items (based on arm_1). In particular *Flour*, *Tobacco* and *Butane* show large differences in expenditure shares between the two recall questionnaires.

It would seem reasonable to assume that this is largely caused by sampling variance rather than the different survey approaches. Nevertheless, the differences seem large for questionnaires with a sample size of almost 200 households. Given the importance of these estimates for the CPI, countries might find it a good idea to utilise some of the savings from adoption of the recall questionnaires to increase the sample size of their surveys instead.

Table 4: Expenditures shares of top 10 items (base arm_1)

	Item	arm_1 CAPI recall	arm_2 CAPI diary	arm_3 PAPI diary	arm_4 PAPI diary	arm_5 CAPI recall bounded
0111001099	Rice, not further specified	4.18%	4.30%	4.22%	4.87%	4.62%
0111002099	Flour, not further specified	2.51%	0.99%	1.45%	2.26%	3.72%
0112015099	Chicken, not further specified	3.74%	3.56%	3.49%	2.80%	2.78%
0213080099	Beer, (?% alcohol) not further specified	2.50%	2.43%	2.03%	2.34%	2.41%
0221082002	smoking tobacco	7.09%	1.64%	3.45%	3.52%	2.45%
0451125001	electricity grid	5.12%	7.10%	6.81%	5.14%	5.45%
0452127001	Butane (small can)	2.93%	2.81%	4.55%	6.34%	7.59%
0732201002	bus/taxi service - international	3.07%	3.95%	4.82%	5.29%	3.28%
1111320001	Food away from home - ceremonies	2.92%	2.79%	3.77%	3.73%	2.94%
1111401002	Lunch away from home - not further specified	3.66%	2.81%	3.97%	2.95%	3.64%

CONCLUSIONS

The move to a full recall questionnaire inevitably involves the loss of some detail in the expenditure aggregates. However, we conclude that the granularity of the data produced by the recall questionnaires is more than sufficient for the purpose of rebasing the CPI. Furthermore, we find that there is no (obvious) evidence that the recall questionnaires significantly and consistently over or underestimate expenditures on specific product groups. We conclude therefore that the use of expenditure data for deriving CPI weights does not represent an obstacle for the adoption of recall-based expenditure surveys in the Pacific.

However, we note that there is significant variation in the expenditure shares generated by the different surveys. We do not believe that this is caused by the different survey methodologies and think that instead it is largely the result of the comparatively small sample sizes utilised for the various experimental survey arms. Nevertheless, countries that use the survey data to rebase the CPI may want to choose to forego some of the savings involved from the adoption of the recall methodology to increase the sample size of their survey.

REFERENCES

Brzozowski, M., Crossley, T.F. and Winter, J.K., 2017. A comparison of recall and diary food expenditure data. Food policy, 72, pp.53-61.