IPS HPP Occasional Paper No.14

National Health Accounts

Estimation of Confidence Intervals for Estimates of National Health Expenditures Derived from Health Accounting Studies

Ravi P. Rannan-Eliya

Aparnaa Somanathan

May 1999

Institute of Policy Studies of Sri Lanka
Health Policy Programme

About the IPS HPP Occasional Paper Series

Papers in this series are not formal publications of the Institute of Policy Studies. They are used primarily as a channel to promote timely dissemination of work in progress, or final dissemination of research results which may not warrant formal publication as IPS Working Papers. They may be preliminary and unpolished results of analysis, and have not necessarily been edited or reviewed as thoroughly as formal IPS Working Papers. In certain instances, IPS HPP Occasional Papers are duplicates of papers published elsewhere, but based on IPS HPP work. These papers are intended for free distribution to encourage discussion and comment, and to inform a wider audience of ongoing IPS research. Citation and use of such a paper should take into account its potentially provisional nature.

The findings, interpretations, and conclusions expressed in these papers are entirely those of the author(s), and do not necessarily represent those of the Institute of Policy Studies or of the Institute of Policy Studies Health Policy Programme.

IPS HPP Occasional Papers are made available free, but only as downloads as Acrobat PDF files from the IPS website (www.ips.lk/health). For further information, please contact the Health Policy Programme, Institute of Policy Studies, 99, St. Michael's Road, Colombo 3, Tel: +94-1-431-368, Fax: +94-1-431-395, Email: health@ips.lk, or visit the IPS website at www.ips.lk.

Recommended Citation

Ravi P. Rannan-Eliya and Aparnaa Somanathan. 1999. Estimation of confidence intervals for estimates of national health expenditures derived from health accounting studies. Health Policy Programme Occasional Paper 14. Colombo: Institute of Policy Studies.

Document information

This was originally presented at the PHR NHA Symposium in Rotterdam, The Netherlands on June 4, 1999

ABSTRACT

Confidence intervals are not routinely estimated or published with national health accounting studies. However, the availability of these become more important as the results of NHA estimations are more widely used by policy-makers, and are used increasingly in cross-national analyses of health expenditure variations.

Errors in most NHA studies arise from similar origins. The impact of sampling errors on survey data are well known, but non-sampling errors are often not fully accounted for. While the literature on non-sampling error is less extensive than for sampling error, it is still possible to analyse sources of non-sampling error and to develop methods of approximating estimators of variance for resulting statistics.

It is recommended that health accountants should routinely estimate variances and confidence intervals in NHA estimates. It is feasible to do, despite the frequent limitations of the data sources used. Such an effort would not only assist users of the statistics, but would also help analysts' better target efforts at quality improvement.[1]

[1] The authors draw on work that they carried out when developing Domestic Health Accounts for the Health and Welfare Bureau of the Government of the Hong Kong Special Administrative Region, People's Republic of China. The assistance of the many officials and counterparts in the Hong Kong SAR government are acknowledged, although they bear no responsibility for any opinions and errors contained in this paper.

INTRODUCTION

National Health Accounts (NHA) are the main approach used to estimate national health expenditures in the USA and an increasing number of other countries (OECD, 1998). The estimates from NHAs are the basis of most cross-national health expenditure analyses (e.g., Newhouse; Schieber et al., 1994), as well as the basis for analyses of trends in national health spending within countries. It is expected that recent developments in the adoption of national health accounting methods will lead to a substantial increase in the coverage of countries for which national health accounts estimates are available. In the Asia-Pacific region, more than a dozen territories maintain national health accounts or are in the process of establishing them. With the increased availability of such estimates, comparative analyses of inter-country differences are more feasible for increasing numbers of non-OECD economies, and are more likely to be carried out. This raises important issues regarding the accuracy of such estimates, and the related failure to regularly provide details of the accuracy of estimates published.

The first NHA were first developed in the 1960s by HCFA in the USA following the introduction of Medicare. The USA NHA are considered a gold standard for most NHA-type work, because of their relative comprehensiveness, detail and the length of time that the series has been produced. Although NHA estimates are based on synthesis of many data sources including surveys, it is not common practice to provide estimates of the confidence intervals attached to the estimates of total spending. While confidence intervals for the US statistics are not normally published with the estimates, the estimates can be considered reasonably accurate within the scope of the definitions used. However, NHA estimates in other countries are rarely as accurate, especially since the availability of data is less, and since most private expenditures in other countries are in forms for which accurate data and records are not available. This has importance when efforts are made to compare spending across time or across countries, and if the differences observed are within the likely degree of error in the estimates.[2] Policy decisions and analyses are often made on the basis of such observed differences (e.g., Jeong and Gunji, 1994).

The absence of routine estimations of confidence intervals may inhibit efforts to improve the quality of estimation methods and results by underemphasising the extent of uncertainty in the data, as well as failing to identify areas where efforts should be focussed in the data for future improvement.

NHA estimates are derived from multiple data sources. In all virtually cases these consist of a mixture of official budgetary records, household surveys, surveys of health care providers, enterprises, and insurance agencies, and other miscellaneous survey sources. The process of compiling NHA estimates usually involves reconciling differing

[2] The other major issue preventing close comparison of NHA estimates in non-OECD territories is the lack of a standardised health expenditure definition and health accounting framework. Greater effort between countries to build consensus on common definitions and standards is needed to currently needed.

estimates. While considerable efforts are made during such reconciliation, little effort is allocated to showing what impact these transformations have on accuracy.

The following discussion covers some commonly encountered problems in NHA estimation, illustrating with examples from Hong Kong's NHA estimates explores how estimations of confidence intervals might be made. At the end, such estimation methods are combined to produce an estimate of the overall accuracy of Hong Kong's NHA estimates, and used to identify which components contribute the most to variance in the estimates.

COMMON PROBLEMS IN NHA COMPILATION

1. Omissions in coverage

It is common when NHA systems are first being established for there to be gaps in the availability of data or estimates. With time these can normally be addressed. However, in many cases, some analysts have chosen to omit these items from the estimation of totals, later counting them in subsequent years (e.g., The Philippines NHA). Other than the fact that this produces misleading time trends if definitions change, it also leads to estimates, which are biased downwards. If a quantity is unknown, then omitting it is equivalent to assigning it a value of zero.

It would probably be better to provide a guesstimate in all cases within an acceptable range, and then treat the number as an estimate with a variance equivalent to the known range of possible values. In most cases, even if there are no data available, the analyst can at least determine an upper and lower bound for the actual value. Often the lower bound can be taken to be zero, while the upper bound is typically not substantial, since the quantities for which no or little data are available generally are not the major components of health spending (e.g., expenditures by employers for employees other than third party insurance or expenditures by non-profit organisations). As a practical measure, if these upper and lower bounds are treated as equivalent to 95% confidence intervals, then this would at least generate an estimate of the variance in the point estimate, albeit an over-estimate, since the upper and lower bounds in this case would represent absolute confidence intervals.

If data are completely absent for a specific item, and even a guess cannot be made as to possible maximum or minimum values, an alternate solution is to explicitly change the scope of coverage of the NHA and exclude the items for which data are minimal. However, in practice this is unlikely since the quantities for which no data are available tend not to be the major components of health spending in a given country.

2. Estimating household expenditures from household survey data

The largest component of health spending after government expenditures, if not the largest, in middle-income and low-income economies consists of out-of-pocket household spending. In most lower and middle-income countries with no social insurance, out-of-pocket spending accounts for 25-80% of total health expenditures. With the exception of small amounts recorded by official providers as user fee revenue, most of these expenditures are not reported in routinely available from provider statistics or other formal data sources. Most NHA estimations therefore must rely on use of household sample surveys recording health care utilisation and expenditures and general household consumption surveys.

Sampling errors

Available household sample surveys consist of two main types:

i) General household consumption surveys, carried out by national statistical agencies mainly for the purpose of estimating national income accounts, price indices and general living standards.

ii) Special health surveys focusing specifically on utilisation and / or expenditures on health services.

Estimates from all such sample survey data are liable to the usual sampling errors common to all surveys, including incomplete or non-comprehensiveness in the sample frame and sampling errors of various types. Examples of incomplete sample frames include surveying only urban households excluding rural households, registered citizens excluding other persons, or the non-institutionalised population excluding those currently institutionalised. This type of problem is generally self-evident, and is well understood by most statisticians. It is most easily reduced by using only nationally representative household surveys, and any remaining sampling errors can always be estimated. E.g., non-coverage of the institutionalised population can be a particular problem, since health expenditures associated with those in hospital for long periods can be substantial even though they account for only a small fraction of the overall population, and in many surveys are not even included in the sample population. It can be dealt with by conducting a separate survey for this sub-population, e.g., the USA).

Sampling errors are generally well understood (Särndal et al., 1992), and most widely used household sample surveys in developing countries routinely report estimates of the sampling errors for key variables, or sufficient ancillary information to permit the informed reader to make such estimates. This class of problem due to sampling errors and deficiencies in the sampling frame will be ignored in the rest of this paper.

Non-sampling errors

Data on health expenditures are collected in surveys in two general approaches. In some surveys, the respondents are asked to report total expenditures on specific goods and services in a given time period; in a few cases retrospective reporting is supplemented by use of diary methods. This approach is typical of most household general consumption surveys carried out central statistical agencies for the purpose of estimating national income accounts or price indices. The second approach, which requires more detailed instruments, records the incidence of illness episodes or visits to providers in a given time period, and the expenditures associated with such visits or illness episodes (e.g., World Bank Living Standards Measurement Survey, US National Medical Expenditure Survey). So the respondent may be asked how often they had been ill during the past month, and then for each occasion whether they visited a provider, and if so how much they spent. Illness episodes or visits to providers are used in these cases as prompts to recall expenditures.

Non-sampling errors are universal for these surveys, and their impact can be substantial (Fowler, 1995; Cannell and Fowler, 1965; Cannell et al., 1965; Cannell et al., 1977; Cannell et al, 1987). In most cases, the magnitude of the non-sampling errors may be greater than the known sampling errors. Typically encountered non-sampling errors include:

i) Under-reporting of visits or illness episodes owing to recall loss. As found in the mostly US literature, underreporting of health care utilisation or illness is positively associated with longer recall periods, use of proxy respondents, with chronic versus acute illness, and insufficient training or supervision of field interviewers. For the same recall period, underreporting is greater for ambulatory care than for inpatient care. The empirical research suggests a one to two week recall period is best for ambulatory visits, and a six to 12 month recall period for hospitalisation. Shorter recall periods are even better, but typically will result in small numbers of non-zero observations in each cell, unless the sample is very large.

ii) Over-reporting of visits or illness episodes owing to telescoping, where incidents occurring before or after the recall period are reported.

iii) Under-reporting of visits or illness episodes when such incidents are considered embarrassing, or are associated with social stigma. Examples include mental illness, sexually transmitted diseases, abortions where this is illegal, and visits to traditional medical providers when the respondent believes this to be disapproved by the interviewer.

In general in our experience, the problem of under-reporting is greater than over-reporting, so the net bias tends to be downwards. Strategies that might be employed to handle this include improvements in survey design to reduce recall bias. Nevertheless, it is virtually impossible to conduct a survey with no recall bias in the final results, and recognising this the ultimate strategy must be to use other independent data sources to verify the survey data, estimate the recall bias, and adjust for it. In fact, the utility of using NHA to estimate health expenditures lies especially in the informational gains from reconciling differing data sources. In practice, while all surveys can be improved, analysts must eventually deal with survey data with some degree of recall bias, usually downwards.

Adjusting household sample survey data

The development of methods to deal with non-sampling error in survey data is not as well developed as for sampling errors. Nevertheless, the available literature on sources and types of non-sampling error does suggest possible solutions. Use of other independent data sources to determine the degree of bias, and to suggest levels of adjustment would seem to be in most cases the most appropriate approach. In doing so, it should be remembered that other data sources might also have their deficiencies, which is why they cannot completely replace the survey data in question. Such deficiencies include lack of comprehensiveness (e.g., few data sources cover all types of household health expenditures), or lack of accuracy and reliability themselves.)

In adjusting available survey data to control for non-sampling errors, it is worth keeping in mind available empirical research on recall bias in surveys. This suggests not only the general existence of such bias, but that the level of such bias varies depending on the length of recall period, identity of respondent (whether proxy or not), nature of subject being queried, demographic characteristics of respondent, etc. Given this, it would be expected that the extent of such bias may vary within the same survey data set, and this suggests that if other data are available that the level of adjustment itself should vary depending on the item under consideration, as well as the availability of secondary data. So for example, for a given recall period, ambulatory expenditures should be expected to be subject to greater recall loss than inpatient expenditures. Similarly, purchases of goods have found to be better reported than those of services, and less frequently purchased items better than commonly purchased items (Mahalanobis, 19__). Any process of adjustment should take this into account; otherwise a simple across the board adjustment will introduce or retain the bias caused by differential recall bias across individual items.

In some surveys data on household expenditures comprises two variables: one of volume or the rate of utilisation of providers, and the second of price. Total expenditure is the product of both variables. Each of these may be independently verifiable using secondary data. However, it does not follow that the level of under-reporting will be the same for each. Where expenditures are recorded having been prompted by recall of visits, then the degree of net under-reporting or recall bias for the price of each visit might be expected to be less than that for the rate of utilisation. The literature suggests general under-reporting of utilisation, but not of the price of episodes, and actual experience is consistent with this (Rannan-Eliya, 1995, 1998; Data International, 1999). If this were so, then a first-cut solution would be to adjust the data for under-reporting of volume while retaining the unadjusted price data.

Adjusting the data for under-reporting controls for one type of non-sampling error. In addition to this the confidence intervals expected from the known sampling errors should be calculated. The following illustrates this approach using the Hong Kong health accounts (Rannan-Eliya and Somanathan, 1999) as an example.

Example: Estimation of household spending at doctors' clinics in Hong Kong[3]

Most outpatient care in Hong Kong takes place at private doctors' clinics. These are staffed by full-time private doctors, who are paid on a fee-for-service basis and who are permitted to dispense medicines. Payments are nominally made in most cases for the drugs dispensed, with the price of drugs dispensed implicitly covering the cost of the doctors' time. Most patients pay out-of-pocket using their own funds. Some consultations are covered by employer schemes, which negotiate a set rate (usually lower than the market rate) for each visit with the doctor by covered employees. In the case of visits covered by private insurance, the patient customarily pays first, and then applies for reimbursement from the insurance firm.

Private doctors are in theory supposed to pay taxes on their practice income, but tax evasion is assumed to be universal, and incomes reported to the tax departments are assumed to be grossly underreported. The only available credible sources of data on fees paid by households to private clinics therefore consist of household survey data. There are three main such sources:

i) The Household Expenditure Survey (HES) – a nationally representative household general consumption survey used to construct the retail price index, can conducted every five years. The sample is typically 3,000 households. The survey uses a one week diary for regular expenditures, and records total expenditures coded by category of product and by source of retail outlet. The survey does not record frequency of purchase, merely the total amount spent in the recall period.

ii) The General Household Survey (GHS) – a permanent labour force survey, conducted on a continuous basis using a rotating panel design. One round is conducted each month using a nationally representative sample frame, with an approximate sample size of 9,500 households (or ~36,000 individuals). In addition to the regular employment and demographic questions, supplemental questionnaires are added in each round to ask about selected additional topics. Supplementary questions have been asked on a regular basis (every 2-3 years) on outpatient doctors' consultations, inpatient use, and dental care utilisation. The special topic questionnaire schedules are brief, taking typically two minutes of respondent time. GHS topic surveys typically ask for the number of visits to a doctor / hospital and the price paid per visit.

iii) Household health care utilisation and expenditure survey (HHCUES) – a specially commissioned telephone survey focusing on health care utilisation and expenditures carried out once in the first quarter of 1998. The sample consisted of

3 The discussion here takes Hong Kong's Domestic Health Accounts estimates as an example. However, it should be noted that the numbers and calculations given are for illustration only, and that some adjustment of the actual data has been made to for the purposes of exposition.

7,000 households. Telephone penetration in Hong Kong is the highest in the world, so non-coverage of individuals who reside in households without telephones is considered to be low. Nevertheless, the sampling frame might be considered less comprehensive than for the GHS/HES, and the response rate is actually lower. The detailed questionnaire focussed on health care use and expenditures.

Both the first two surveys suffer from substantial under-reporting. According to the Census and Statistics Department (CSD), who conducts both surveys, the HES generally under-reports all expenditures by 40-50%.[4] This is the difference between household spending estimated using the national income accounts and production-based approaches and that reported in HES. GHS appears to also suffer from under-reporting. This non-sampling error due to underreporting is substantially greater than the estimated sampling errors in the HES.

GHS reports two variables: the utilisation rates at different health care providers, and the price paid per visit. The utilisation rates for public sector providers can be compared with official records, while one survey estimate exists for private doctors clinics. The price per visit can be compared with an estimate from the HHCUES and data from insurance schemes and employers, but the time periods for each survey are not identical.

Estimation of volume

The HES data are known to be under-estimates, but cannot be directly compared with alternate data sources. When the GHS data are compared with other data sources, in particular administrative records of the two main public health care providers (Department of Health and the Hospital Authority), the utilisation rate appears to be significantly under-reported,[5] while the price data appear to be unbiased or at the very least not significantly different from collected in other population surveys. This is consistent with theoretical expectations. Adjustment for underreporting therefore consists mainly in adjusting the utilisation rate. The only comparison is with utilisation at public sector providers, and it is assumed that the level of under-reporting is the same for both private and public sector providers.[6] If this assumption is made, the adjustment factor (AF) can be estimated as the ratio of actual visits to public sector providers to the number reported in the sample survey. The number of private provider visits is then the number reported in the survey multiplied by the adjustment factor, AF. The error in the final estimate will then consist of the sampling error in the estimate of public sector providers visits in the survey.

---

[4] Personal communication from Elizabeth So, Senior Statistician, Hong Kong Census and Statistics Department.

[5] Unpublished data from a 1998 survey of private clinic doctors carried out by Harvard University and the Hong Kong Medical Association suggests a level of underreporting consistent with these other data sources (authors' analysis).

[6] There are no available data or *a priori* expectations on theoretical grounds, which would indicate either the direction of such a bias if it existed, or its magnitude.

To complicate matters in this case, there was considerable uncertainty amongst the analysts as to which official data to use for the number of public sector visits, since the GHS survey technically only requested respondents to report visits to medical doctors, while there is reasonable doubt as to whether visits involving other medical staff such as nurses were correctly categorised and excluded by respondents. Two numbers could be used: one for visits involving doctors, and a second higher one including all visits to public sector providers, including visits for dressings and injections and treatment by nurses. Although the problem of miscategorization of visits by respondents is probably real, it is not possible to say how substantial this is. Depending on how the visits are counted an official figure of 13.0 and 14.0 million visits can be taken, the first number assuming no miscategorization by respondents, and the higher number assuming that all visits to public providers not involving doctors were misreported as doctors' visits. A compromise might have been to take the average of both numbers, but the analysts in this case determined that the higher number should be used (Table 1).

Table 1: Estimate of public sector doctor consultations in 1996 (millions)

	GHS survey estimate	Actual figure (low estimate)	Actual number (medium estimate)	Actual number (high estimate)
Annualised number per year	8.41 (7.9 – 8.9)	13.0	13.5	14.0
Implied adjustment factor	1.0	1.55 (1.46 – 1.65)	1.61 (1.52- 1.71)	1.66 (1.57 - 1.77)

Note: Parentheses contain 95% confidence intervals as estimated using published data.

The standard errors for the estimates in GHS are not reported in the published tables. However, an estimate can be estimated from the raw data. The standard error can then be estimated to give a 95% confidence interval for the survey estimate of the number of public sector visits of ~ ±0.52 millions.

The number of actual visits to public sector providers used was adjudicated by a process of consensus amongst the analysts. However, for purpose of estimating confidence intervals, it would be best to take both the upper and lower estimates under different assumptions, as these represent those most plausible maximum and minimum estimates. Applying the estimated 95% confidence intervals to these yield estimates of the confidence intervals for the adjustment factor under different assumptions. The lower bound of the lowest and the upper bound of highest of these can be taken as appropriate confidence intervals, i.e., the sample survey estimate of private doctor visits is 14.0 millions, and the estimated adjustment factor, AF, is 1.66 [95% confidence interval 1.46 – 1.77].

This procedure controls for the overall level of underreporting and incorporates a confidence interval taking into account expected sampling errors. However, it does not account for any differential recall bias between private sector and public sector visits.

To estimate actual doctors' revenues this estimate of the number of visits must be multiplied into the estimate of the average price. The price estimate is itself is a sample survey estimate with an associated sampling error: HK$ 198 (SE = $5). The 95% confidence interval for the price is therefore $ 188 – 208.

The resulting variance in the estimator for total doctors' revenues is then related to the variances of the two subcomponents, price and volume, since revenues are the product of price and volume. Although an exact calculation of the variance of the revenue estimate in this case is not possible, an approximation is possible using the Taylor Linearisation technique for variance estimation (Särndal et al, 1992). Applying this technique to this case, yields the formula for a variance in the product, Z, of two independent totals, X and Y:

$$\boldsymbol{Var}\ (Z)\ /\ Z^2 \quad = \quad \boldsymbol{Var}\ (X)\ /\ X^2 \quad + \quad \boldsymbol{Var}\ (Y)\ /\ Y^2$$

where ***Var*** (θ) is the estimator for the variance of a quantity θ.

Applying this method gives an estimation of the variance in the estimate of total revenues, as illustrated in Table 2.

Table 2: Estimation of confidence intervals for doctors' revenues

Item	**Estimate**	**Variance**	**95% CI**
Price of doctors' visit (A)	$197.8	25.0	178.2 – 217.4
Number of visits (B)	39.9 million	6.0	35.1 - 42.5
Total revenues (A*B)	$7,889 million	293,264	6,828 - 8,950

Note: Since confidence interval for B is not symmetrical, the lower bound taken as limit of symmetrical confidence interval.

Example: Estimation of other household health expenditures

Other household expenditures consist largely of spending at pharmacies, shops, traditional Chinese medicine practitioners and stores, and other retail outlets. Virtually no secondary data exist for any of these items (other than spending at government providers and at private hospitals), and the only available data source is the HES. However, the HES suffers from under-reporting as noted. To adjust the HES data on these items of spending, an estimate of the general level of underreporting must be made. This was done by examining available secondary data for specific items of spending and comparing them with the expenditures reported by the HES. Based on this, and the information from the research literature about the differences in recall bias for different types of expenditures, an estimate was made of the overall level of underreporting (Tables 3, 4).

	A	B	C	D	E	F	G	H
5	***Table 3: Analysis of underestimation of household expenditures derived from HES data***							
6		Reported in HES 94/95			**Alternative estimates**			
7	Item	Expenditure per capita	Annual total expenditure in HK	Estimated level of underestimate	Actual	Low estimate	High estimate	Source
8		(HK$ per month)	(HK$ millions)		(HK$ millions)	(HK$ millions)	(HK$ millions)	
9	Foreign medicines/medical supplies	9.3	681	78.3%	1,000-1,400	1,040	1,387	IMS data assuming 20% to 60% retail margin
10	HA charges	4.0	293	76.9%	518			HA DHA data
11	Private hospitals	18.8	1,376	25.9%	1,733	1,792	2,389	Estimated private hospital revenues adjusted for insurance payments
12	Private physicians/other private medical char	48.4	3,543	90% - 110%	7,157	7,061	7,157	(1) Low estimate based on IMS/Tax data, (2) High estimate based on DHA estimate

	B	C	D	E
19	***Table 4: Correction for underreporting in HES 1994/95 survey***			
20	Item	Expenditure per capita	Adjustment	Final total estimate
21		(HK$ per month)	factors	
22	Dentists	9.1	100%	1,332
23	Chinese medicines	16.2	100%	2,372
24	Herbalist's charges	8	100%	1,171
25	Other medical charges	2	100%	293
26	Health food/spectacles/optical goods/therape	9.7	100%	1,420
27	NGO clinics	0.4	100%	59
28				
29	Private doctors	46.6	120%	$ 7,504
30	Public institutions and private hospitals	24.1	100%	$ 3,528
31	Other goods and services	47.2	100.0%	$ 6,910
32	Total	117.9	107.9%	$ 17,943

3. Estimating facility/unit expenditures from facility survey data

In some instances, accurate data on revenues or expenditures are available from some types of institutional providers. The most common example of this are private hospitals or employer schemes. However, it is unusual outside the USA for financial records from most private hospitals to be made available, so the analyst must work from sample survey data. If the total size of the hospital or provider population is known, then a sample survey can yield an estimate of total revenues in that population. However, two approaches present themselves in handling the data. One is to estimate the total using the simple π estimator. However, the ratio estimator can also be used in many cases, since additional information other than financial data may be available for all such units, as well as for those surveyed. For example, in many countries private hospitals may have to report to official agencies their number of beds or other data such as number of discharges. Clearly if there is a strong linear relationship between some of these other ancillary characteristics and total revenues, then it would be more efficient to use the ratio estimator, especially if the sample size is large (Särndal et al., 1992; Wu and Deng, 1983).

Example: Estimation of private hospital revenues in Hong Kong

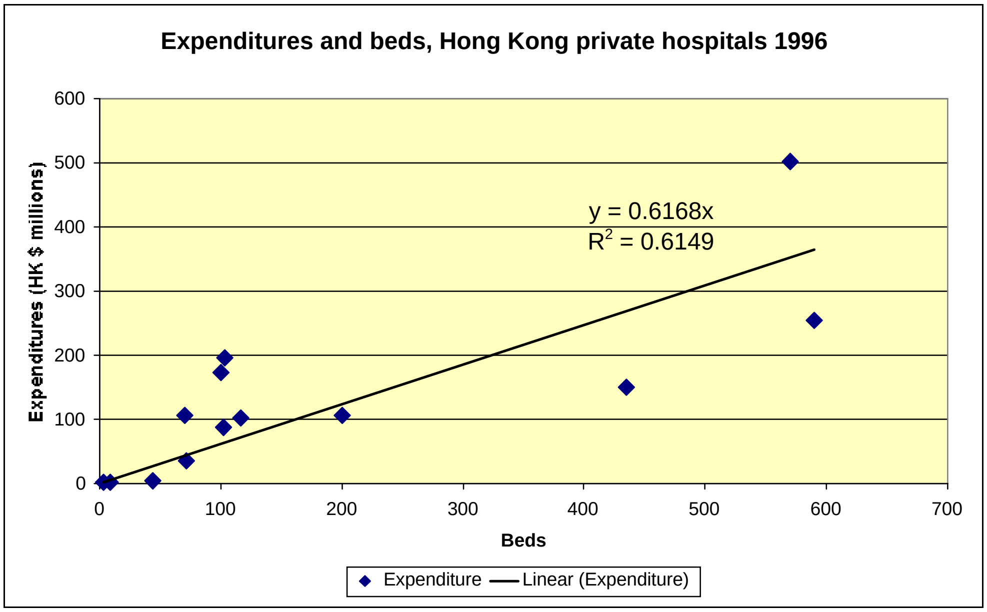
Private hospitals do not routinely report financial data to the health agencies, and are reluctant to accept any mandatory requirements to do so. However, all private hospitals must report bed numbers and discharge data to the Department of Health (DOH). For the purpose of this exercise, DOH conducted a survey of all private hospitals asking for data on their financial expenditures and revenues. Reporting was voluntary, and the response rate was approximately 60%.[7]

Chart 1 shows the relationship between expenditures and bed numbers in the sample. Given the relatively high correlation coefficient ($r>0.5$), a ratio estimator would be expected to be more efficient than the π estimator for estimating total revenues. However, some caution may be needed since the total number of hospitals in the sample is small ($n=12$), which suggests that the bias associated with the ratio estimator is more significant. Using Stata to calculate a ratio estimator for revenues of private hospitals outside the sample, total private hospital revenues is estimated as HK$ 2,356 millions (95% confidence interval $ 2,164 - 2,549 millions).

---

[7] The actual response rate was higher than this, but for purposes of exposition a lower rate is used here. In practice, for most other countries, a response rate of 60% or less in surveys of private hospitals involving voluntary completion would not be uncommon.

This estimate would be improved in terms of reduced bias if the number of hospitals in the sample were increased towards 20. Although data on discharges by hospital were not available at the time, it would be worth examining such data to determine whether the number of discharges are better correlated with overall expenditures.



A difference estimator could be conceivably used to estimate time trends if the sample of hospitals reporting expenditures was substantially less in other years. However, this is not so in this case.

In many cases, the analyst does not have access to survey data collected through systematic sampling, and the available data may be non-random in origin and small in number. The following example illustrates the estimation of a variance and confidence interval is still possible in these circumstances.

Example: Estimation of employer expenditures

Many large employers in Hong Kong provide medical benefits to their employees. These consist typically of insurance schemes to cover inpatient care, and negotiated arrangements with panels of doctors to provide ambulatory care to employees at a set price, usually lower than their customary rates. In some cases, the employer arranges to pay the clinic doctors an additional amount per visit for each consultation.

There is no available systematic data on the prevalence and expenditures of these employer schemes. In the absence of these, a ratio estimator of sorts was used to estimate total employer expenditures. First a non-random sample (n=4) of a few large employers was obtained on an opportunistic basis. A random sample was not possible as most employers were reluctant to release financial data. This small sample yielded some estimates of total employer expenditures per employee, and the estimate was taken as the midpoint of the lowest and highest costs. Then a survey estimate of the total number of employees in Hong Kong covered by such schemes was obtained from two different household surveys, conducted in 1991 and 1998. These suggested that the level of such coverage was 18% in 1991 and 21% of the population in 1998. The difference between the two estimates was assumed to be due to growth in coverage, which was known to have taken place during the time period. Estimates for coverage for intervening years were derived by simple linear interpolation between the two estimates.

An estimator for total employer expenditures was calculated by multiplying the estimate of population coverage by the estimate of the cost per employee based on the employer sample. While there is a known sampling error associated with the survey estimates of population coverage, the sampling error associated with the employer sample cannot be estimated, as this was a non-random sample. However, in the absence of any other data, the lowest and highest costs observed might be used as a lower and upper bound for the estimate. The alternative to estimate the sample variance, but the sample is very small, and non-random. A true 95% confidence interval is not possible. Table 5 gives the confidence intervals calculated accordingly for each estimator.

Table 5: Estimation of employer expenditures

Variable	Source	Estimate	95% CI	Variance
Population coverage by employer schemes (1991)	GHS 1991	18.0%	17.6 – 18.4%	
Population coverage by employer schemes (1998)	HHCUES 1998	20.6%	19.5 – 21.7%	
Population coverage by employer schemes (1997)	Interpolation from GHS 1991 and HHCUES 1998	1.33 millions	1.26 -1.41 millions	0.001666
Employer cost per employee (1997)	Sample of 4 firms	HK$ 3,204	HK$ 2,413 – 3,995	162,870
Total employer expenditures (1997)	(Population coverage * Employer cost)	HK$ 4,274 millions	HK$ 3,114 – 5,434 millions	350,010

COMBINING SUBTOTALS

By applying these and other appropriate methods, confidence intervals and variances can be estimated for each sub-component of national health expenditures. Assuming that the total is obtained by simply adding each sub-component of expenditure, then the confidence interval and standard error for the total estimate is easily derived as the square root of the sum of the implied variances.

Making such an estimate of the confidence interval in the estimate of total spending has value beyond merely indicating the reliability of the final estimate, as illustrated by the following example.

Example: Estimation of confidence interval for Hong Kong health expenditure estimate

Table 6 gives the individual components comprising the estimate for total health expenditures for Hong Kong, and their associated variances, as calculated for the calendar year 1996 estimates of total domestic health expenditures (TDHE) in Hong Kong.

Table 6: Estimation of confidence interval for TDHE (CY 1996)

Item	Estimate	Contribution to total	Variance	95% Confidence interval	Contribution to variance
Public expenditures	29,146	53%	65,077	500	3%
Household spending at private clinics	7,889	14%	293,034	1,061	15%
Household spending at public providers	1,404	3%	0	0	0%
Private hospital revenues	2,356	4%	9,696	193	0%
Employers expenditures not counted above	3,321	6%	211,495	901	11%
Household expenditures not counted above	9,253	17%	1,392,957	2,313	70%
Insurance firm administrative expenses	867	2%	2,707	102	0%
Non-profit expenditures	305	1%	10,412	200	1%
TOTAL DOMESTIC HEALTH EXPENDITURE	54,541	100%	1,985,380	2,762	100%

The overall estimate of error in the estimate of total spending is equivalent to a confidence interval of approximately HK$ 2,800 millions, or 5% in the final result. This can be regarded as relatively low and encouraging, as discussion in Hong Kong about the results have largely focussed on much larger differences across years or between Hong Kong and other countries. This suggests that such comparisons are valid, and need not be questioned on the grounds of data quality.

Analysis of this type of display of the variances contributing to the total variance can be quite useful. As can be observed, 70% of the variance in the final estimate in this case is contributed by the item for household expenditures on services other than doctors or hospitals. The other two items contributing significantly to the variance are private clinic doctors' revenue and employer expenditures. As the employer expenditure estimate was derived from only four companies data, additional effort to obtain data from a larger number of companies would be productive in terms of reducing the sampling error in this subcomponent, and cost-effective in reducing the overall variance of the total estimate.

Other items for which there was even greater uncertainty, such as non-profit expenditures, contribute little to the overall error, as they amount to only small proportion of total spending. This indicates that if the results are to be substantially improved in terms of quality, that efforts should be focused on improving estimates of household spending at retail outlets and for traditional Chinese medicine.

IMPLICATIONS

These calculations demonstrate that is feasible to easily derive reasonable estimations of the error associated with NHA estimates incorporating both sampling and non-sampling error components, and that it may be fruitful in terms of indicating priorities for concentrating subsequent work to improve the quality and reliability of the estimates.

National health accounting estimates, as with any empirical estimates, are subject to some error. It is certainly feasible to obtain approximate estimates of the accuracy of the final estimates, even when data are non-systematic. It is recommended that this effort be routinely incorporated into NHA studies, and that all NHA studies report at least a confidence interval for the final total numbers, such as total national health expenditure, and total public and total private expenditures. It would not only help users to judge the appropriateness of conclusions based on the data, but it would also assist in identifying areas where future efforts should be concentrated in improving the quality of estimations.

REFERENCES

Cannell, C and F Fowler. 1965. *Comparison of hospitalization reporting in three survey procedures.* Vital and Health Statistics. Series 2, No. 8. Washington, DC: Government Printing Office.

Cannell, C, G Fisher, and T Bakker. 1965. *Reporting of hospitalization in the Health Interview Survey.* Vital and Health Statistics. Series 2, No. 6. Washington, DC: Government Printing Office.

Cannell, C, K Marquis and A Laurent. 1977. *A Summary of Studies.* Vital and Health Statistics. Series 2, No. 69. Washington, DC: Government Printing Office.

Data International. Forthcoming in 1999. *Bangladesh National Health Accounts.* Report of study conducted for Health Economics Unit, MOHFW, Dhaka and Asian Development Bank, Manila.

Fowler, FJ. 1995. *Improving Survey Questions: Design and Evaluation.* Applied Social Research Methods Series, Vol. 38. Thousand Oaks, CA: Sage Publications.

Jeong, Hyoung-Sun, Atsuaki Gunji. 1994. *The influence of system factors upon the macro-economic efficiency of health care: implications for the health policies of developing and developed countries.* Health Policy. 27: 113-140.

Rannan-Eliya, Ravi P., and Aparnaa Somanathan. 1999. *Estimates of Domestic Health Expenditures 1989/90 to 1996/97, Hong Kong Special Administrative Region.* Special Topic Report No. 1 prepared by Institute of Policy Studies, Sri Lanka, for Health and Welfare Bureau, Hong Kong SAR, under contract to Harvard University.

Särndal, Carl-Erik, Bengt Swensson and Jan Wretman. 1992. *Model assisted survey sampling.* New York: Springer-Verlag New York, Inc.

Schieber, GJ, JP Poullier and LM Greenwald. 1994. *Health System Performance in OECD countries.* Health Affairs. 13: 101-112.

World Bank. 1993. *World Development Report 1993: Investing in Health.* Oxford: Oxford University Press.

Wu, C.F.J., and L.Y. Deng. 1983. *Estimation of variance of the ratio estimator: an empirical study.* In G.E.P. Box et al. (Editors), *Scientific Inference, Data Analysis and Robustness.* New York: Academic Press.

IPS HPP Occasional Paper Series

Number	***Title***
1	Results of Private Health Insurance Study. *Ravi P. Rannan-Eliya and Nishan de Mel.* 1996
2	Results from the IPS/Harvard Public Opinion Poll on User Fees. *Ravi P. Rannan-Eliya* 1996
3	Analysis of Private Health Insurance in Sri Lanka: Findings and Policy Implications. *Ravi. P. Rannan-Eliya* 1997
4	Responses to Population Ageing: A Review of International Experience. *Ravi. P. Rannan-Eliya, Nishan de Mel, Easha Ramachandran and Danajanee Senagama* 1997
5	Ageing and Pensions. *Ravi. P. Rannan-Eliya, Nishan de Mel, Easha Ramachandran and Danajanee Senagama.* 1998
6	Unit Cost Analysis of Public and Private Health Facilities in Sri Lanka in 1992. *Aparnaa Somanathan* 1998
7	Framework for Developing Sri Lanka Health Expenditure Estimates using Health Accounting Approach. 1998 (Revised May 2000)
8	Review of international NHA approaches and proposed conceptual framework for Sri Lanka National Health Accounts. 1998 (Revised May 2000)
9	Provisional Functional Classification System for Domestic Health Accounts of Hong Kong SAR, People's Republic of China. 1998
10	Estimates of Domestic Health Expenditures 1989/90 to 1996/97, Hong Kong SAR, People's Republic of China. 1998
11	Economic impacts of demographic ageing: with special emphasis on Sri Lanka and old-age income security. *Ravi P. Rannan-Eliya* 1999
12	Bangladesh Health Facility Efficiency Study Report. *Ravi P. Rannan-Eliya and Aparnaa Somanathan.* 1999
13	Expenditures for Reproductive Health Services in Egypt and Sri Lanka. *Ravi. P. Rannan-Eliya, Peter Berman, Eltigani E. Eltigani, Indralal de Silva, Aparnaa Somanathan, Varuni Sumathiratne* 2000

Number	***Title***
14	Estimation of Confidence Intervals for Estimates of National Health Expenditures derived from Health Accounting Studies. *Ravi. P. Rannan-Eliya and Aparnaa Somanathan, 1999*
15	WHO Fairness in Financing Study. Estimates for Sri Lanka 1995/96 using WHO Methodology. *Tamara Dorabawila, Suharshini De Silva, Jehan Mendis and Ravi P. Rannan-Eliya, 2001*