Catalogue no. 11-522-XIE

Statistics Canada International Symposium Series - Proceedings

Symposium 2005 : Methodological Challenges for Future Information needs

2005

Statistics Canada Statistique Canada

Canada

Proceedings of Statistics Canada Symposium 2005
Methodological Challenges for Future Information Needs

VARIANCE INFORMATION FOR DATA USERS

John Wood and Markus G. Šova[1]

ABSTRACT

The traditional approach to presenting variance information to data users is to publish estimates of variance or related statistics. However, such simple statistics do not convey a full understanding of how users should interpret variance and what they might do to mitigate its effects or allow for its impact through the interpretation of statistics. This paper addresses some of these wider issues. It examines the effects of variance on the interpretation of published statistics and what users and producers of official statistics might do to assist analysis in the presence of variance. It also examines potential sources of variance, and considers what might best be done to reduce variance. Finally, it looks at the financial costs to producers and users of reducing or not reducing variance and how the costs of producing more accurate statistics could potentially be offset by the financial benefits of greater accuracy.

KEY WORDS: Analysis of Variance; Cost-Benefit Analysis; Costs of Variance; Measures of Variance.

1. INTRODUCTION

Recent years have seen much development in the production and publication of quality measures for official statistics. Although there are several dimensions to quality, such as relevance, timeliness and coherence, this paper will concentrate on the dimension of accuracy. More specifically, the discussion will cover issues regarding sampling variance and will only briefly allude to other aspects of accuracy, such as bias or non-sampling errors. This is not to negate the importance of these other aspects of accuracy or other dimensions of quality. On the contrary, the arguments presented here could and should be extended to cover these aspects but doing so would complicate the presentation a great deal. We concentrate on variance partly because of its central role in determining the accuracy of published statistics but doing so also helps to simplify and clarify the exposition and some illustrative, numerical examples are readily available.

The issue of what variance information to present to users, and how to present it, is not as straightforward as it at first seems. The traditional approach to presenting variance information to data users is to publish estimates of variance or related statistics, such as standard errors, coefficients of variation, confidence limits or simple grading systems. However, these simple statistics do not convey a full understanding of how users should interpret variance and what they might do to mitigate its effects or, at least, allow for its impact on the interpretation of statistics. This paper addresses some of these wider issues. Section 2 considers the traditional variance information provided by producers, which users this information might be useful for and what use might be made of the information. In section 3, we use a specific example to examine the effects of variance on the interpretation of published statistics and what users and producers of official statistics might do to assist analysis in the presence of variance. Section 4 examines potential sources of variance and considers what might best be done to reduce variance by controlling sample size, sample allocation and response rates. Section 5 considers the possible use of variance information for the allocation of resources. Section 6 examines the financial costs and benefits of reducing variance and how we might trade off the costs of producing more accurate statistics against the financial benefits of greater accuracy. Finally, section 7 provides a brief summary of our arguments for alternative ways of presenting variance information.

Throughout the paper, we approach this topic from the point of view of economic statistics produced by the Office for National Statistics (ONS) in the United Kingdom (UK) but we hope that the general principles underlying the

[1] John Wood, Office for National Statistics, Government Buildings, Cardiff Road, Newport, NP10 8XG, United Kingdom (e-mail john.wood@ons.gov.uk); Markus Šova, Office for National Statistics, Government Buildings, Cardiff Road, Newport, NP10 8XG, United Kingdom (e-mail markus.sova@ons.gov.uk)

examples presented will be more widely applicable, internationally and across all fields of official statistics. Our intention is not to provide definitive answers but to raise questions on the most appropriate ways to present variance information and to stimulate debate and discussion on this topic.

2. TRADITIONAL MEASURES OF VARIANCE

Official statistics published in the UK have included a wide variety of the traditional measures of variance, such as standard errors, coefficients of variation, confidence intervals, grading systems and the identification of "statistically significant" differences. These measures are all very familiar to the professional statistician but not all users find them useful or even comprehensible. Users of statistical outputs possess a wide range of knowledge of statistical concepts, from those who may be familiar with them (such as producers of statistics, policymakers, academics and researchers) through those who may have some limited understanding (such as supranational bodies like the European Union, government departments and local government bodies) to those who may have little or no knowledge of statistical concepts (such as businesses, the media, politicians and the general public). Note that we include producers of statistics among the users of statistics - for example, National Accounts are compiled from a wide range of official statistics.

The large variety of different users, with different needs and different levels of statistical competence, means that variance information must be tailored to meet their different needs and levels of understanding. Even for users who understand statistical concepts, such as policymakers and academics, the traditional measures of variance may not provide them with the information best suited to their particular needs and purposes. Later sections of this paper present examples of such cases.

We suggest, therefore, that official statisticians need to consider using a wider range of possible measures of variance so as to provide users with measures which actually help them to interpret and use our statistics. This may require measures which have greater analytical content than simple statements of numerical accuracy.

Some of the traditional measures already go some way to providing this analytical content. For example, grading systems help to remove the statistical mystique surrounding standard errors and coefficients of variation by converting them to broad bands indicating the general quality of the statistics. These are more easily understood by non-statistical users. However, little is known about whether users find them helpful or use them meaningfully or whether the bandings presented suit users' needs.

Some publications on social statistics explicitly mention "statistically significant" differences. These typically involve comparisons of population characteristics between different sections of the population and are covered by a statement along the lines of "differences cited in the text are statistically significant unless otherwise stated". This approach could be adopted more often in economic statistics. Every month statistics on inflation and economic output are published to a precision of 0.1 percentage points. However, they are usually not accurate to this level of precision. Unfortunately, for most of these statistics there are no estimates of variance to assess their accuracy, although ONS is working on this for UK statistics.

We contend, however, that the analytical content of traditional measures of varaince does not go far enough. The examples in the following sections present possible, alternative measures of variance which, we think, might provide information of more direct use to users of official statistics.

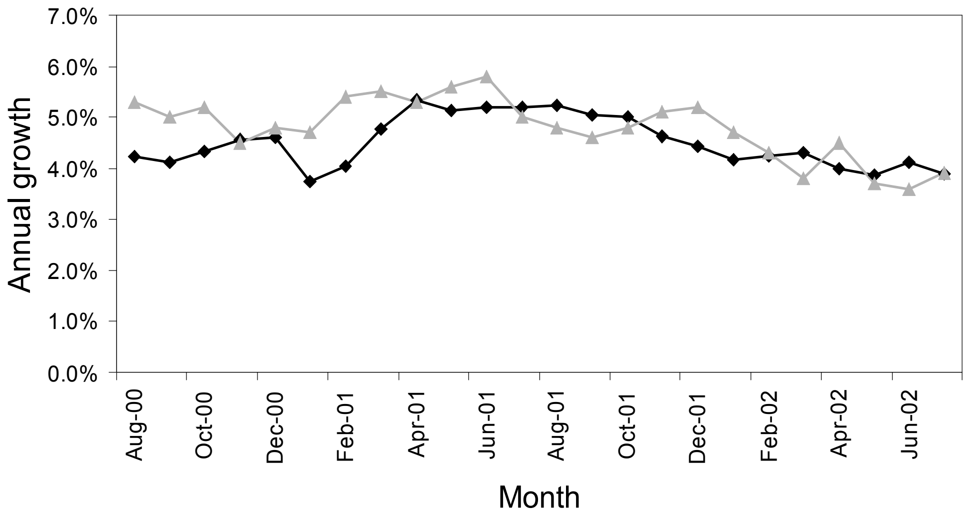
3. EFFECTS OF VARIANCE

We start our analysis by examining the effect of variance on published statistics for a specific example. A few years ago, ONS finished a project to develop estimators for the sampling variance of movements in the UK Average Earnings Index (AEI – see Youll, 2002). We found that these variances could be represented reasonably well by a first-order auto-correlation model of monthly growth rates. The variance (σ^2) of the monthly growth rate of the AEI is roughly constant over time. The correlation coefficient (ρ) for the covariance between monthly growth rates for adjacent months is also roughly constant and negative. The covariance between monthly growth rates for non-adjacent months is negligible. In short, we may represent the monthly movement g_i in the AEI for month i by the following AR(1) process: $E[g_i]=\mu$; $Var[g_i]=\sigma^2$; $Cov[g_i, g_j]=\rho\sigma^2$ ($|i-j|=1$); $Cov[g_i, g_j]=0$ ($|i-j|>1$) .

For the Whole Economy index, excluding bonus payments, σ is about 0.15 percentage points and ρ is about -0.3. Earnings growth has been fairly stable over the last few years at about 4% per year, or μ=0.33% per month. Using these parameters, we can produce a random realisation of this simple model for comparison with the actual index growth.

Figure 1 below shows, for the AEI, 12-month inflation rates (that is, the increase in the current index on the index 12 months previously), which is the statistic of most interest to users, for the period August, 2000 to July, 2002. The data are a little out of date but this does not affect the principle being demonstrated. One line in figure 1 shows earnings inflation as measured by the AEI, the other is a random realisation of the auto-correlation model described above.

Figure 1. Estimated monthly growth of the UK AEI (for the whole economy) compared with a realisation of an AR(1) model.



We suggest that even the most statistically competent observer would have difficulty in identifying, with confidence, which line is which. The important point to note from this example is that, under some circumstances, it is almost impossible to distinguish between published statistics and random noise around a stable reference level.

In October 2003, the Governor of the Bank of England made a speech referring to the “statistical fog hanging over the British economy” (King, 2003). This created much interest in the financial press and some thinly-veiled criticism of official statisticians. Less noticed, but extremely pertinent, was another statement in that same speech, indeed in the same paragraph: “After a decade of unparalleled stability of both growth and inflation, it is time to take stock.” This “decade of unparalleled stability” provided just the conditions under which our statistics look like random noise around some stable reference level. Thus the question arises: in such circumstances, how do we impart to our data users that the message of our statistics is in the stable signal and not in the noise?

To consider this issue, we can borrow some ideas from Steel & Smith (2005), although the analysis here is much less sophisticated than their's. Using our simple auto-correlation model, we can assess how long it would take to identify a sudden jump in earnings inflation from 0.33% per month to 0.5% per month (that is, from 4% to 6% per annum). As the basis for this assessment, we apply a simple decision rule: if the observed 12-month inflation rate exceeds the 95% upper confidence limit for our stable state of 4% per annum, we infer that earnings inflation has increased. Table 1 shows, for two different values of σ, the probabilities of breaching this limit after different numbers of months at the higher inflation rate.

For σ =0.15, corresponding to earnings excluding bonuses, we need to wait 6 months before we have an acceptably high probability (89%) of identifying this sudden, large change in underlying inflation.

For σ =0.3, corresponding to earnings including bonuses, the situation is even worse and we need to wait for up to 12 months before being reasonably certain of identifying the change.

Table 1: Probabilities (in %) of breaching the 95% upper confidence limit after n months of 2% higher annual inflation.

n: number of months at 2% higher annual inflation	σ=0.15, ρ=-0.3 (earnings excluding bonuses)	σ=0.30, ρ=-0.3 (earnings including bonuses)
1	12	8
3	42	18
6	89	42
9	100	69
12	100	89

Part of the problem here is the preferred target measure: the 12-month inflation rate is not well suited to quick identification of sudden, recent changes. Two questions arise from this example. Would measures of the kind proposed, quantifying the impact of variance on decision-making, be more useful to users than simple standard errors or confidence intervals? Can we use our knowledge of the variance-covariance structure of our statistics to suggest better target measures, which would more quickly or easily identify economic turning points or other characteristics important to users?

4. SOURCES OF VARIANCE

One of the reasons for ONS's work on producing variance estimators for the Average Earnings Index was to be able to assess the impact of sample changes on the accuracy of the index. These changes were: a sample re-allocation; a possible increase in sample size; and the effect of increasing response rates. Table 2 shows the estimated standard errors for each of these potential changes relative to the sample current at the time.

Table 2: Average (over August 2001 to July 2002) estimated standard errors (in percentage points) of AEI 12-month growth rates.

	Excluding bonuses	Including Bonuses
Previous sample	0.35	0.66
Re-allocated sample	0.32	0.65
50% larger sample	0.29	0.61
100% larger sample	0.27	0.57
Fully responding sample	0.24	0.39

Looking at the three potential changes to the sample in turn, table 2 shows the following results.

a. <u>Sample Allocation</u>
The proposed new allocation produced only a slight reduction in estimated standard errors. This is quite reassuring because it indicates that the original allocation, which was based on limited information, was already close to optimal.

b. <u>Sample Size</u>
Increasing the sample size by 50% does very little to reduce the standard errors. Even a doubling of the sample size does not reduce them by much (no more than by about 15%-20%).

c. <u>Non-response</u>
Increasing response rates is the best potential source of a substantial reduction in standard error. Most of the benefit from a fully responding sample comes from the largest businesses.

At the time, there was the possibility of receiving more funds from the UK Treasury to increase the sample size, if that would improve the accuracy of earnings inflation statistics. Because of the limited potential impact on standard errors, ONS received no extra money from the Treasury. However, ONS did implement a programme of response-chasing, targeted on the largest businesses. Whether this has achieved the expected reduction in standard errors is currently under investigation.

The point we are making here is that information on the sources of variance may help users with their own decision-making. Knowing the effect on variance of changes in the resources allocated to a survey may help those who finance the survey to decide funding levels. Looking more widely to the suppliers of survey data, can information on the effect of non-response be used to encourage responses from survey respondents, especially those with a large impact on the resultant estimates?

We can also look at the sources of variance more widely. ONS has recently completed a project to estimate standard errors for movements in the UK Index of Production (IoP), specifically for the 12-month growth rate from September 2003 to September 2004. The estimation process allows us to extract the contributions to the total variance from the four contributing surveys, as shown in table 3.

Table 3: Estimated contributions to the variance of the 12-month growth (September 2003 to September 2004) of the UK Index of Production.

Component survey	Contribution to variance ($\times 10^4$)	Contribution to variance (%)
Turnover	0.591	93.8
Inventory Movements	0.021	3.3
Producer Price Indices	0.009	1.4
Export Price Indices	0.009	1.4

The very small contributions from Inventory Movements, Export Price Indices and even Producer Price Indices are rather surprising but they accord with several coherence checks we applied to confirm the validity of the results. However, this illustrates one practical use of such a high level variance analysis: the extreme dependence of the accuracy of the IoP on the associated turnover survey raises the question whether the turnover estimates from this survey are sufficiently accurate for the intended purpose. This, in turn, depends on the relative importance of IoP in the wider set of ONS's economic statistics.

Unfortunately, variance information is not readily available for this wider set of statistics. We can, instead, use the IoP as an example, in microcosm, of how we might use variance information to determine the allocation of resources between different surveys. We discuss this in the next section on the uses of variance information.

5. USES OF VARIANCE INFORMATION

Statisticians are well accustomed to using knowledge of the likely variance-covariance structure of survey data to try to optimise sample design and sample allocation. We normally think of this as a purely professional responsibility but this need not always be the case. Users also have an interest in the design and allocation of a sample because decisions on these affect the scope and accuracy of the outputs produced. For example, a recent project to re-allocate the sample for UK Producer Price Indices identified the possibility of obtaining a considerable gain in efficiency by collapsing a large number of low-level indices into a much smaller number of higher level indices. ONS is not, yet, instigating this reduction in the number of indices because of the potential impact on users. Such a major re-design and re-allocation might lead to the withdrawal of popular indices, so extensive user consultation would be required. In such a context, would the costs of implementing an inefficient sample allocation be a helpful measure of variance for users? Would a cost saving of, say, £0.5m be sufficient justification for withdrawing some popular indices? Would users be willing to pay such a sum to retain their favoured indices? The willingness to pay for inefficiencies in sample design and allocation might be a useful indication of how important to users these favoured indices actually are.

Users also have a wider interest in the resources allocated to statistical outputs. Although they may expect government statistical offices to make efficient use of taxpayers' money, they are also likely to want as much resource as possible allocated to the statistical outputs they have most interest in. They may even require justification from statistical offices for cases when users' favoured statistical outputs do not receive as much funding as users would like.

Clearly, determining the optimal allocation of resources across all statistical outputs would be very complicated but we can illustrate a possible approach by using the variance results obtained for the Index of Production, as discussed in the preceding section. Table 4 shows the current annual budget assigned to the four surveys involved and a notional “optimal” allocation, derived by assuming that the contribution to variance is inversely proportional to the resources allocated.

Table 4: The annual costs (current and “optimised”) of the surveys contributing to the UK Index of Production, with the corresponding impact on the standard error of 12-month index growth.

Component survey	Current annual costs	“Optimised” annual costs
Turnover	£0.8m	£2.1m
Inventory Movements	£0.9m	£0.4m
Producer Price Indices	£1.1m	£0.3m
Export Price Indices	£0.2m	£0.2m
Standard error (percentage points)	0.79	0.56

On this basis, we could obtain an almost 30% reduction in the standard error of the 12-month growth of the Index of Production by a more efficient allocation of resources to the component surveys. Alternatively, under this notional “optimal” allocation, we could obtain the same standard error at a cost of only £1.5m – a 50% saving.

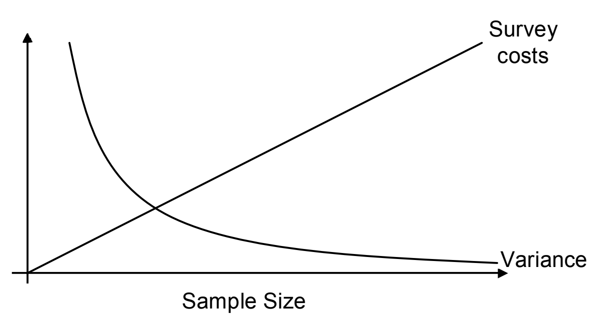
Of course, this is an enormous over-simplification. The component surveys are used in other contexts than IoP. Price indices are used in the monitoring of inflation for macro-economic management and for price escalation clauses in legal contracts. Inventory movements are used in other sectors of the National Accounts, so their impact is not confined to IoP. A thorough analysis would take account of the aggregate contribution of each survey over all outputs, a daunting task and beyond the scope of this expository paper.

Nonetheless, given the analysis above, we can raise the question: would users prefer a 30% reduction in standard error or a saving of £1.5m or some combination of reduction in standard error and reduction in cost? To answer this, we need to be able to assign a monetary value to the reduction in standard error. We address this issue in the next section.

6. COSTS OF VARIANCE

As a survey’s sample size increases, the variances of its estimators typically fall in roughly inverse proportion but the survey costs increase almost linearly. Figure 2 illustrates this schematically.

Figure 2. Schematic illustration of the relationship to sample size for survey costs and variance



For small sample sizes, we can obtain substantial reductions in variance for relatively small additions to absolute cost but as the sample size grows we experience the law of diminishing returns and need to allocate more and more resources to obtain the same amount of variance reduction. The question is: where do we stop? What is the trade-off between variance and cost? How much is a reduction of 20% in the standard error worth?

To address this question we indulge in a (very simplified) thought experiment.

Annual Gross Domestic Product (GDP) for the UK is approximately £1,000bn. Every month, quarter and year, the Office for National Statistics produces a wide variety of economic and financial statistics which the relevant authorities use to manage the economy. We concentrate on one mechanism for doing so: the management of interest rates.

We examined the progress of interest rates and GDP over the last six years (see BoE, 2005 and ONS, 2005a for data sources), from which we derived this simple rule-of-thumb: a change in the Bank of England base rate of 1% leads to a change in GDP of 1% in the opposite direction about one year later. This is very rough-and-ready and could undoubtedly be improved on but our aim here is merely to examine a possible approach, so perfect accuracy is not essential.

Suppose, now, that we were able to reduce the variance of our statistics so that the Bank of England could improve its timing of interest rate changes by one month. Normally, the Bank changes the base rate in steps of a quarter of a percentage point. Accelerating an appropriate decrease by one month would lead to a one-off benefit of 0.25% of one month's GDP a year later, giving a one-off benefit of about £200m. For simplicity, we assume that a similar benefit would also arise from improved timing of interest rate increases, although we acknowledge that it would be more difficult to demonstrate this.

This benefit occurs only once, when we make the change from the old, higher variances to the new, lower variances. But we need to compare it against the annual cost of maintaining our reduced variances. Investing this one-off benefit at an interest rate of 5% would give an equivalent annual benefit of £10m. What would be the annual cost of achieving this benefit?

Using results from the AEI model discussed in section 3, we obtain a (very rough) guide that the necessary reduction in variance could be obtained by doubling the sample size. This might cost about £80m, which is the current annual amount assigned by ONS to the production of economic and financial statistics (ONS, 2005b). This is clearly not worth doing. On the other hand, remembering the findings from section 3, it may be worth putting more resources into increasing response rates, because that is likely to be a more cost-effective way of reducing variance.

Obviously, a cost/benefit analysis of this kind would require much further elaboration before it could provide a reliable means of determining, for the production of official statistics, how much financial resources should be provided and where they should be allocated. The point of this relatively simple example is to demonstrate that cost/benefit analyses of this kind may provide a useful tool for communicating variance information to users because they translate abstract statistical concepts, which may be difficult for users to understand or interpret, into equivalent financial terms, which everyone can understand.

There are, of course, many problems with cost/benefit analyses. They are difficult and complicated to produce and are often based on dubious assumptions. There may be multiple benefits from improvements to a single survey and ensuring that all the benefits are factored in and not double counted is an onerous task. Alternatively, there may be multiple costs for a benefit based on the amalgamation of several different statistics (for example, the IoP, which as based on the amalgamation of data from four different surveys). Over all surveys, ensuring that all costs are factored in and offset against the appropriate benefits, is a major problem and much work will be needed to progress such analyses.

As a further complication, the costs and benefits are dynamic and require continual monitoring and review. National statistical institutes continually face multiple new demands on limited resources. For example, in the UK, macro-economic statistics in the service sector are of relatively poor quality, despite the fact that the service sector accounts for about 70% of GDP. A recent review (Allsopp, 2004) recommended a programme of improvement in these statistics. The same review also recommended a programme for the development of improved regional statistics, although whether the quality of the results will justify the necessary resources applied is not clear and was not addressed by the review. Assessing the relative merits of these multiple demands, in terms of costs and benefits, is not easy.

Finally, we have not yet addressed the issues of bias and non-sampling errors. These also need to be considered when analysing costs and benefits, even though the assessment of these sources of error can be extremely challenging. Nonetheless, we suggest that official statisticians do need to consider the costs and benefits of producing official statistics, in order to justify the financial resources applied to them and to provide users with an understanding of the value of these statistics.

7. SUMMARY

To summarise, we return to the question: what variance information do users need? We have argued that the traditional, statistical measures such as standard errors and confidence intervals are not enough to meet users' needs. We suggest that the official statistics community needs to think more creatively, to provide variance information which is of direct, practical benefit to users. This is likely to require measures of variance which are more analytical in content. At an elementary level, we could be more forthright, especially with economic statistics, in telling users when differences are statistically significant or, just as importantly, when they are not. But we also need to go further and devise measures which address directly the concerns of users. What is the effect of variance on the statistics we publish? Are the target measures preferred by users sensible? Can we use the information we have on variance to devise alternative measures which are more accurate indicators of what users really want to know? Is there a more efficient or more effective way of meeting users' needs by allocating resources differently? How much are users willing to pay for more accurate statistics? How can we assess the financial benefit of reducing variance?

This paper has only skimmed the surface of these deep and difficult questions. The final question is: to what extent can these deep and difficult questions be answered?

REFERENCES

Allsopp, C. (2004), *Review of Statistics for Economic Policymaking – Final Report*, Norwich: HMSO, *http://www.hm-treasury.gov.uk/allsopp*

BoE (2005), Repo Rate History, Bank of England, http://213.225.136.206/mfsd/iadb/Repo.asp?Travel=NIx

King, M. (2003), Speech to the East Midlands Development Agency/Bank of England Dinner on 14th October, 2003, Bank of England, http://www.bankofengland.co.uk/publications/speeches/2003/speech204.pdf

ONS (2005a), *GDP Growth*, Office for National Statistics, *http://www.statistics.gov.uk/CCI/nugget.asp?ID=192&Pos=&ColRank=1&Rank=374*

ONS (2005b), *National Statistics Annual Report 2004/05*, Office for National Statistics, *http://www.statistics.gov.uk/about_ns/downloads/NSAnnualReport0405.pdf*

Steel, D. and Smith, P. (2005), Communicating Variances for Interpretation of Changes and Turning Points in Repeated Surveys, *Proceedings of Statistics Canada Symposium 2005*

Youll, R. (2002), Quality of the estimates of earnings growth from the AEI, *Labour Market Trends*, Vol. 110, No. 4, April 2002, pp 207-213, Office for National Statistics