January 1994

Diagnostic testing and sensitivity analysis in SAM construction

R. P. Byron

BOND UNIVERSITY
School of Business

DISCUSSION PAPERS

"Diagnostic Testing and Sensitivity Analysis in SAM Construction

R.P. Byron
School of Business
Bond University

DISCUSSION PAPER NO 51

January 1994

University Drive, Gold Coast,QLD,4229
AUSTRALIA

January 1994

Diagnostic Testing and Sensitivity Analysis in SAM Construction

R.P.Byron[1]
Bond University
Queensland

1. Introduction

One of the major issues when constructing a Social Accounting Matrix is that of checking the initial cell estimates provided by the statistician. Statisticians typically protest strongly on this issue, maintaining that all relevant information has been incorporated in the cell estimates and that no further improvement is possible. If the cell estimates are also expected to satisfy various identities then, one obvious check (and adjustment) is to use the identities in testing and balancing modes. Such identities have often been used for balancing - where the adjustment typically depends on producing new estimates which are as close to the initial estimates as possible within the confines of a quadratic norm where the weights are the reliabilities of the initial estimates. Statisticians seem to view such mechanical adjustments with as much enthusiasm as the Royal Navy first viewed the advent of steam. The adjusted estimates satisfy adding up conditions which are accounting identities and the burden of the adjustment falls on the least reliable of the initial estimates.

The dual of constrained estimation is the process of hypothesis testing. The method of classical statistics is to set up a null hypothesis, embodying certain prior information, to test the hypothesis and, if it is not rejected, to incorporate that prior information in the estimation process. The end result is enhanced efficiency. The process of testing is important because hypotheses are often rejected and theory refined, as a result. The procedure of estimating and balancing a SAM has been discussed elsewhere {see Arkhipoff [1969], Stone [1975], Byron [1977] and van der Ploeg [1982] for examples}. However, the process of testing has had no attention in the literature, despite the fact that ad hoc testing and refinement procedures are often used in SAM construction. This paper explores the issue of testing and stems from an earlier paper by Byron, Crossman, Hurley and Smith [1993] which found that constrained estimation or balancing when accompanied by "ad hoc" testing was invaluable in focussing on erroneous initial estimates of cells in a SAM. That experience suggested it might be worthwhile to develop and formalise statistical tests for SAM construction.

2. The Statistical Model

A SAM is simply a table with row and column adding-up conditions, the estimates of the cells (parameters) in the matrix have different levels of reliability, and the adding up conditions are applicable to the true cell values. The row and column adding-up conditions are valid restrictions (identities) applicable to the true cell

[1] The intellectual interest and financial support of the Queensland Treasury is gratefully acknowledged.

values; because the adding up restrictions are valid, they can be used to improve the initial estimates of the cells. If cell estimates are viewed as a random variables with known probability distributions, then traditional statistical theory is applicable and it is easily shown that the use of true prior information must result in more efficient estimates of those cells.

To illustrate, suppose X is a 4x4 SAM, then there are 7 independent row and column restrictions. Let x=vec(X), where the vectorisation is by column and the 4 row restrictions are stacked first, followed by the column restrictions, then the restriction matrix for Gx = h = 0 has one redundant restriction.

$$G=\begin{bmatrix} 1&0&0&0&1&0&0&0&1&0&0&0&-1&0&0&0\\ 0&1&0&0&0&1&0&0&0&1&0&0&0&-1&0&0\\ 0&0&1&0&0&0&1&0&0&0&1&0&0&0&-1&0\\ 0&0&0&1&0&0&0&1&0&0&0&1&0&0&0&-1\\ 1&1&1&-1&0&0&0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&1&1&1&-1&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&1&1&1&-1&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&-1 \end{bmatrix}$$

SAM balancing is just a procedure to ensure that a set of estimates is produced which satisfy the adding up restrictions and are as close as possible to the statistician's initial estimates of the cells in the context of a quadratic norm. If the condition that all cell members be non-negative is imposed, SAM balancing is identical to the quadratic transportation problem of operations research.

Let $\hat{x}$ and $\tilde{x}$ be the initial and constrained estimates of the cell members respectively and let μ be the true values. The constrained objective function (with the redundant equation of G deleted) is

(1) $$(\tilde{x}-\hat{x})'\,V^{-1}(\tilde{x}-\hat{x})+\lambda'(G\tilde{x}-h)$$

The first order conditions are

(2) $$\tilde{x}=\hat{x}-VG'\tilde{\lambda}$$

(3) $$\tilde{\lambda}=(GVG')^{-1}(G\hat{x}-h)$$

(4) $$\tilde{x}=\hat{x}-VG'[GVG']^{-1}G\hat{x}$$

since h=0.

If the data generation process is that the initial estimates of X are iid with $\hat{x}=\mu+e$, then Gμ=0, E(e)=0 and $GE(\hat{x})=0$.

(5) $$\tilde{x}=(\mu+e)-VG'(GVG')^{-1}G(\mu+e)$$

(6) $$E(\tilde{x})=\{I-VG'[GVG']^{-1}G\}\mu=\mu$$

Next,

(7) $$\tilde{x}-\mu = e - VG'(GVG')^{-1}Ge$$

so

(8) $$E(\tilde{x}-\mu)(\tilde{x}-\mu)' = V - VG'(GVG')^{-1}GV$$

which supports the usual result that the constrained estimator is at least as efficient as the initial estimator and because the prior information is correct, $\tilde{x}$ is also unbiased.

In the traditional statistical model, the role of hypothesis testing is to test the validity of the restrictions. **Here the restrictions are identities and must hold for the population parameters. However, what is at issue is whether or not the initial estimates of the cells are unbiased**. Suppose $\hat{x} = \mu + \delta + e$ where δ is the bias in the initial (unconstrained) estimates of μ. Henceforth, we refer to this is IEB, initial estimate bias. Clearly, if the initial estimate is biased, the constrained estimates of that cell and all linked cells will be biased.

The objective is to isolate IEB by examining the deviations between the restricted and initial estimates of μ and setting up a test statistic based on their deviations using the normality of e. Since

(9) $$\tilde{x}-\hat{x} = -VG'(GVG')^{-1}G(\mu+e)$$

and

(10) $$E(\tilde{x}-\hat{x})(\tilde{x}-\hat{x})' = VG'(GVG')^{-1}GE\{(\mu+e)(\mu+e)'\}G(GVG')^{-1}GV$$

(11) $$\Sigma = E(\tilde{x}-\hat{x})(\tilde{x}-\hat{x})' = VG'(GVG')^{-1}GV$$

under H_O. Furthermore,

(12) $$E(\tilde{x}-\hat{x}) = -VG'(GVG')^{-1}G(\mu + E(e)) = 0.$$

since $G\mu = 0$. However, if any element of $\hat{x}$ is biased, such that $\hat{x} = \mu + \delta + e$ then

(13) $$E(\tilde{x}-\hat{x}) = -VG'(GVG')^{-1}G\delta.$$

Under the null hypothesis (that $\delta = 0$) the covariance matrix is defined above and the quadratic form provides a Chi-squared statistic

(14) $$(\tilde{x}-\hat{x})'\Sigma^{-1}(\tilde{x}-\hat{x}) = \hat{x}'G'(GVG')^{-1}G\hat{x} \sim \chi^2_{2n-1}$$

The components of the quadratic form can also be set up as standard normal deviates by dividing the difference in the constrained and initial estimates by their standard deviations. In other words, the validity of the restrictions $G\hat{x}$ can be tested separately, under the assumption that the other restrictions are correct. This is one way to isolate the source of the bias.

The initial estimates of the elements of X can be used to provide a test of IEB. If the initial estimates of the elements of X are unbiased, their expected values will satisfy the row and column adding up conditions. If a single estimate is biased, then it will not satisfy its row and column adding-up restrictions and **any test statistics on those two restrictions will isolate the offending initial estimate**. For example, the restrictions in the 4x4 case indicate that any bias in $\hat{x}_{11}$ will be observed in restriction 1 and restriction 5, any bias in $\hat{x}_{12}$ will be observed in restriction 2 and restriction 5, and so on.

The statistical model is $\hat{x} = \mu + e$, each restriction contains estimates of different X terms; these initial estimates are statistically independent and since the true values satisfy the adding-up restrictions, if $y = G\hat{x}$, then E(y) = 0 and Var(y) = $GCov(\hat{x})G'$. Since y is normally distributed, a test statistic based on $\frac{y-0}{\sqrt{Var(y)}}$ will have a standard normal or a central t distribution, under H_0 depending on whether Var(y) is known or estimated. The key is that any error in an initial estimate of a cell **will show up in two of the test statistics, enabling that cell to be identified**. This is a decomposed Wald test based on the initial estimates.

An "ad hoc" procedure often used is to test if the restricted estimate is within two standard deviation units of the initial estimate. This procedure uses the initial estimate of the variance to provide the standard deviation. Using (4) and setting the 'ad hoc' test up as a quadratic form it is immediately seen to be a likelihood raio test and by substitution becomes a Wald test, which establishes the legitimacy of the procedure.

(15) $$(\tilde{x}-\hat{x})' V^{-1}(\tilde{x}-\hat{x}) = \hat{x}' G' (GVG')^{-1} G\hat{x}$$

An LM test is also easily constructed based on the distribution of the quadratic form $\hat{\lambda}' \Sigma^{-1} \hat{\lambda}$ under H_0. Substitution yields

(16) $$\hat{\lambda}' \Sigma^{-1} \hat{\lambda} = \hat{\lambda}' (GVG') \hat{\lambda} = \hat{x}' G' (GVG')^{-1} G\hat{x}$$

with the usual equivalence between the LM and Wald tests. However, in the classical context, where the restrictions are at fault, if only one restriction is incorrect, a test statistic can be set up based on the individual Lagrangians. This has only been done occasionally in the literature before {see Byron [1972], for an example}. In the present context, single Lagrangians provide tests which can pinpoint the source of the bias. The Lagrangians will have a standard normal or central t distribution depending on whether V is known or estimated.

Any discussion of an estimator is incomplete if its large sample properties are not included. In the present context this poses a difficulty. The initial estimates are an informed guess which are assumed to be unbiased. Any errors are independently, identically (and perhaps normally) distributed. The test statistics are based on linear combinations of the underlying random variables. As the size of the system increases the number of random terms increases by the square, whereas the number of restrictions to be tested increases linearly (n^2 versus 2n-1). The Wald and LM tests can thus be expected to be well behaved as the size of the system increases because they are linear functions of an increasing number of errors. The Difference test, based on the distribution of $(\tilde{x}-\hat{x})$ is also well behaved as the size of the SAM increases because, following (4), the test is a linear function of an increasing number of errors. The upshot is that the three tests should be well behaved as the size of the system increases; under the null hypothesis they should have the assumed distributions with their hypothetical rejection rates and under the alternative hypothesis they should display reasonable power.

Consider the Lagrangian (3), $\tilde{\lambda} = (GVG')^{-1}(G\hat{x} - h)$. If the null hypothesis is true $G\mu = 0$ and $\tilde{\lambda} = (GVG')^{-1}Ge$. However, given G is a matrix whose row vectors consist of unit elements and zeros plim $\frac{Ge}{n} = 0$. It is easily shown that lim $\frac{GVG'}{n} = K$, where K is a constant matrix, so the estimator of λ is consistent.

To see this, consider the form of GVG' when n=4 and V is diagonal with, for convenience of exposition, all diagonal elements equal. Referring to G on page 2

$$GVG' = \left[\begin{array}{cccc|ccc} 4v & 0 & 0 & 0 & v & v & v \\ 0 & 4v & 0 & 0 & v & v & v \\ 0 & 0 & 4v & 0 & v & v & v \\ 0 & 0 & 0 & 4v & -v & -v & -v \\ \hline v & v & v & -v & 4v & 0 & 0 \\ v & v & v & -v & 0 & 4v & 0 \\ v & v & v & -v & 0 & 0 & 4v \end{array}\right]$$

with the notation that the common diagonal elements are v. Since n=4, GVG' when divided by n cannot increase in magnitude as the size of the SAM increases. Hence λ is consistent.

Given this consistency, it appears that test statistics based on a linear combination of the random variables (the Wald, LM and Diff procedures) will be valid whatever the size of the SAM. The large sample properties of test statistics can be described by the behaviour of their confidence intervals and a consistent estimator provides a well behaved test [see Bickel and Docksum, ch.6]. The simulations below offer further evidence of the validity of the test procedures.

3. Empirical Evidence

In what follows an exploratory Monte Carlo experiment is set up with 1000 replications. In the first pass, the estimates of μ are simulated to vary normally around the true μ with a variance consistent with the initial assumptions made. There are 16 parameters to be estimated in the 4x4 case resulting in 8 test statistics corresponding to the row and column restrictions.

The true values of the cell members and their (true) variances are

$$X = \begin{pmatrix} 10 & 2 & 5 & 17 \\ 5 & 2 & 8 & 15 \\ 9 & 3 & 9 & 21 \\ 24 & 7 & 22 & 53 \end{pmatrix} \text{ and } V = \begin{pmatrix} 5 & 1 & 2 & 4 \\ 2 & .01 & 1 & 3 \\ 2 & 1 & 2 & 4 \\ 3 & 1 & 4 & 4 \end{pmatrix}.$$

The first row in Table 1 provides the N ratio for $G\hat{x}$ when the null hypothesis is true, and all the cell estimates are unbiased. The average of these ratios is always close to zero.

Table 1

Mean Values of Wald Test Statistics
Resulting from Biased Cell Estimates - 1000 Replications

bias	n1	n2	n3	n4	n5	n6	n7	n8
H_0 true	-0.0308	-0.0316	0.0239	0.0096	-0.0159	-0.0679	-0.0101	-0.0148
H_0 false								
x1	**5.7427**	-0.0316	0.0239	0.0096	**5.7576**	-0.0679	-0.0101	-0.0148
x2	-0.0308	**3.2317**	0.0239	0.0096	**2.2935**	-0.0679	-0.0101	-0.0148
x3	-0.0308	-0.0316	**2.6906**	0.0096	**2.2935**	-0.0679	-0.0101	-0.0148
x4	-0.0308	-0.0316	0.0239	**3.4737**	**-3.4800**	-0.0679	-0.0101	-0.0148
x5	1.1239	-0.0316	0.0239	0.0096	-0.0159	**2.2377**	-0.0101	-0.0148
x6	-0.0308	-0.0152	0.0239	0.0096	-0.0159	-0.0448	-0.0101	-0.0148
x7	-0.0308	-0.0316	1.3573	0.0096	-0.0159	**2.2377**	-0.0101	-0.0148
x8	-0.0308	-0.0316	0.0239	1.1643	-0.0159	**-2.3734**	-0.0101	-0.0148
x9	**2.2786**	-0.0316	0.0239	0.0096	-0.0159	-0.0679	**2.6566**	-0.0148
x10	-0.0308	1.6001	0.0239	0.0096	-0.0159	-0.0679	1.3233	-0.0148
x11	-0.0308	-0.0316	**2.6906**	0.0096	-0.0159	-0.0679	**2.6566**	-0.0148
x12	-0.0308	-0.0316	0.0239	**4.6284**	-0.0159	-0.0679	**-5.3434**	-0.0148
x13	**-4.6496**	-0.0316	0.0239	0.0096	-0.0159	-0.0679	-0.0101	**4.1164**
x14	-0.0308	**-4.9265**	0.0239	0.0096	-0.0159	-0.0679	-0.0101	**3.0836**
x15	-0.0308	-0.0316	**-5.3094**	0.0096	-0.0159	-0.0679	-0.0101	**4.1164**
x16	-0.0308	-0.0316	0.0239	**-4.6092**	-0.0159	-0.0679	-0.0101	**-4.1460**

Next, a bias was introduced, by adding $4\text{x}Var(x_{ij})$ to each cell estimate, one at a time. For the first cell, this should result in a rejection or high N-value for the first and fifth restrictions; a bias in the estimate of the second cell (vectorised columnwise) should result in a rejection of restrictions 2 and 5, and so on. The mean values of the test statistics on the restrictions with the offending parameter estimates tends to be much larger than 2, and the pattern supports the previous conjecture. In some cases the tests do not appear to isolate offending initial estimates, particularly in relation to x_{22}; however, an examination of the rejection rates associated with the test statistics will be more meaningful.

The mean results for the LM tests are given in Table 2 and are similar to the Wald test results in Table 1. The well known equality W>LR>LM is observed once again [see Berndt and Savin(1977)]. Usually the performance of a test can be improved by increasing the sample size or improved variance estimation; but it is not obvious how either can be exploited here.

Table 2

Mean Values of LM Test Statistics
Resulting from Biased Cell Estimates - 1000 Replications

	11	12	13	14	15	16	17
x1	**2.9162**	-1.0963	-0.8588	1.0062	**3.7092**	-0.1389	0.1651
x2	-0.9296	**2.4322**	-0.5830	0.6515	1.8593	0.4632	0.3008
x3	-0.8697	-0.5772	1.8593	0.5323	**1.9778**	-0.0503	0.1048
x4	1.0937	0.7266	0.4524	**2.2083**	**-2.6688**	0.1722	0.4991
x5	0.6673	0.0407	-0.3886	0.2810	-0.1327	**2.1389**	0.0925
x6	0.0076	0.0306	-0.0223	-0.0050	-0.0004	0.0742	0.0375
x7	-0.3543	-0.0335	0.7181	0.3487	0.1147	**2.1262**	0.1011
x8	0.2975	0.0057	0.3421	0.6429	-0.0165	**-2.0346**	0.1594
x9	1.5363	-0.2561	-0.4942	0.7700	-0.2836	0.0045	**2.2507**
x10	-0.2776	1.3105	-0.3703	0.5077	0.0458	0.2718	1.2499
x11	-0.5068	-0.4045	1.7191	0.9054	0.2111	-0.0208	**2.2679**
x12	0.7287	0.6179	0.8894	**2.2014**	-0.0247	0.1536	**-3.6728**
x13	**-5.8402**	-1.3438	-1.7067	**2.0755**	**3.2949**	1.9027	**2.4367**
x14	-1.2227	**-5.2956**	-0.9408	1.1725	1.9067	0.7074	1.5165
x15	-1.7540	-1.0472	**-6.1333**	1.8047	**2.3055**	1.9533	**2.4022**
x16	**2.0968**	1.3182	1.7186	**-5.8627**	**-2.7080**	-1.8107	**-3.0860**

Table 3 presents the N-ratios for the differences between the constrained and the initial estimates divided by their standard errors when H_0 is false; that is, as a result of the biases introduced on the 16 coefficients. The approach works, the diagonal N-values should the large and significant, the off-diagonal elements insignificant. This is the observed pattern.

Further results on the same 4x4 SAM are given below. Firstly, based on 1000 replications, it is verified in Table 4, that the test statistics are (individually) standard normal under the null hypothesis. The Kolmogorov-Smirnov test is satisfied at the 5% level in all cases except one, that being one of the difference tests, and even there the rejection is only marginal. Empirical critical values are calculated for these distributions based on 2.5% of the distribution under H_0 being under either tail. These critical values are then used to establish the rejection rates under H_1 in Table 5. The perturbations are those mentioned previously and as can be seen, when a bias is introduced into the (1,1) element, the rejection rates on the corresponding row and column of the Wald and LM tests increase dramatically. The results for the difference tests (expressed in terms of the coefficients rather than the restrictions) also increase substantially for the offending coefficient and the accompanying row and column. These elements are highlighted in Table 5.

Table 3

Standardised Differences resulting from Perturbations
Coefficients

Pert.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	**-7.09**	-2.02	-2.64	2.77	-1.52	0.83	0.67	-0.75	-2.52	0.88	0.63	-0.85	2.91	-1.11	-0.82	0.99
2	-1.08	**-3.79**	-1.2	1.26	0.11	-1.84	-0.1	0.07	0.59	-2.22	0.26	-0.37	-0.94	2.41	-0.54	0.63
3	-1.28	-1.08	**-3.47**	1.49	0.6	0.44	-1.06	-0.39	0.73	0.49	-1.63	-0.47	-0.88	-0.59	1.9	0.51
4	1.83	1.54	1.99	**-4.69**	-0.77	-0.55	-0.43	-1.14	-1.32	-0.93	-0.81	-1.62	1.08	0.71	0.49	2.19
5	-0.53	0.09	0.43	-0.38	**-2.52**	-1.85	-1.91	1.99	-0.56	-0.02	0.3	-0.25	0.66	0.02	-0.35	0.26
6	0.0	-0.01	-0.01	0.02	-0.04	-0.04	-0.05	0.05	0.06	0.03	0.04	-0.04	0	0.01	0.02	-0.03
7	0.25	-0.05	-0.77	-0.19	-1.91	-1.79	**-2.56**	1.94	0.3	0.04	-0.65	-0.3	-0.36	-0.05	0.76	0.33
8	-0.28	0.03	-0.32	-0.6	1.94	1.82	1.86	**-2.47**	-0.31	-0.05	-0.39	-0.51	0.29	-0.01	0.38	0.62
9	-1.26	0.49	0.66	-0.98	-0.86	0.19	0.31	-0.47	**-3.41**	-1.53	-1.66	1.67	1.53	-0.27	-0.45	0.75
10	.25	-1.24	0.25	-0.41	-0.07	-0.99	-0.04	-0.05	-0.9	**-2.04**	-0.82	0.82	-0.29	1.29	-0.33	0.49
11	.30	0.22	-1.75	-0.61	0.36	0.31	-1	-0.58	-1.71	-1.42	**-3.56**	1.57	-0.52	-0.42	1.76	0.88
12	-0.71	-0.54	-0.8	-2.05	-0.54	-0.47	-0.68	-1.16	3.1	2.57	2.79	**-6.22**	0.72	0.6	0.93	2.18
13	2.4	-1.45	-1.51	1.37	1.52	-0.8	-0.87	0.7	2.63	-0.76	-0.81	0.64	**-5.85**	-1.36	-1.67	2.05
14	-0.83	3.39	-0.93	0.83	0.04	2.64	-0.13	0.01	-0.36	3.34	-0.59	0.48	-1.23	**-5.31**	-0.9	1.15
15	-0.73	-0.91	3.3	0.63	-0.92	-1.03	1.75	0.91	-0.79	-0.99	2.99	0.86	-1.76	-1.06	**-6.09**	1.78
16	0.84	1.02	0.9	2.78	0.66	0.83	0.8	1.56	1.37	1.5	1.52	2.06	2.09	1.3	1.76	**-5.88**

Table 4

Simulation Results under Ho

Wald Test

means and standard deviations

1	2	3	4	5	6	7	8
-0.0308	-0.0316	0.0239	0.0096	-0.0159	-0.0679	-0.0101	-0.0148
1.0089	1.0268	1.0029	0.9882	1.0001	0.9801	1.0406	1.0293
type I error							
44	54	41	52	47	49	58	51
Kolmogorov-Smirnov Test							
0.0281	0.0413	0.0187	0.0171	0.0341	0.0361	0.0238	0.0263

critical value at 95% = 0.0430

LM Test

means and standard deviations

1	2	3	4	5	6	7
-0.0022	-0.0206	0.0431	-0.0151	-0.0175	-0.0723	-0.0194
1.0486	1.0092	1.0138	1.0011	1.0158	1.0056	1.0380
type I error						
68	54	43	49	55	46	62
Kolmogorov-Smirnov Test						
0.0212	0.0404	0.0266	0.0165	0.0203	0.0377	0.0288

Difference Test

means and standard deviations

1	2	3	4	5	6	7	8
0.0216	0.0337	-0.0223	-0.0033	0.0749	0.0757	0.0468	-0.0648
1.0061	1.0344	0.9960	0.9866	0.9681	0.9850	0.9906	0.9914
9	10	11	12	13	14	15	16
0.0210	0.0334	-0.0182	-0.0071	-0.0022	-0.0206	0.0431	-0.0151
1.0116	1.0501	1.0289	1.0255	1.0486	1.0092	1.0138	1.0011
type I error							
41	62	44	42	46	46	52	47
53	67	57	54	68	54	43	49
Kolmogorov-Smirnov Test							
0.0429	0.0412	0.0273	0.0210	0.0374	0.0483	0.0289	0.0348
0.0259	0.0343	0.0202	0.0261	0.0212	0.0404	0.0266	0.0165

Table 5

Rejection Rates on Test Statistics resulting from Biased Initial Estimates (H1)

Wald and LM Tests

	1	2	3	4	5	6	7	8
11 Wald	**1000**	49	49	49	**999**	49	49	49
LM	**790**	188	161	211	**951**	62	46	
21 Wald	49	**870**	49	49	**626**	49	49	49
LM	134	**649**	112	130	**429**	92	49	
31Wald	49	49	**796**	49	**626**	49	49	49
LM	118	88	**502**	104	**479**	54	47	
41Wald	49	49	49	**940**	**946**	49	49	49
LM	164	104	97	**641**	**744**	59	55	
12 Wald	**208**	49	49	49	49	**660**	49	49
LM	**95**	50	80	65	51	**619**	47	
22 Wald	49	49	49	49	49	50	49	49
LM	50	50	51	49	50	52	50	
23 Wald	49	49	**298**	49	49	**660**	49	49
LM	56	52	**146**	73	46	**616**	47	
24 Wald	49	49	49	**215**	49	**642**	49	49
LM	56	51	83	**129**	50	**560**	46	
13 Wald	**629**	49	49	49	49	49	**745**	49
LM	**319**	55	93	153	52	51	**494**	
23 Wald	49	**286**	49	49	49	49	**262**	49
LM	57	**229**	79	98	49	67	**165**	
33 Wald	49	49	**796**	49	49	49	**745**	49
LM	67	66	**440**	188	48	54	**506**	
34 Wald	49	49	49	**997**	49	49	**1000**	49
LM	108	86	176	**637**	52	61	**948**	
14 Wald	**997**	49	49	49	49	49	49	**987**
LM	**999**	268	439	597	885	528	578	
24 Wald	49	**997**	49	49	49	49	49	**864**
LM	205	**998**	179	250	452	137	225	
34 Wald	49	49	**999**	49	49	49	49	**987**
LM	368	175	**1000**	496	619	545	564	
44 Wald	49	49	49	**996**	49	49	49	**982**
LM	513	230	440	**1000**	756	466	876	

The Difference Test

	1	2	3	4		1	2	3	4
11	**1000**	**442**	**754**	**824**	12	**93**	51	77	82
	360	127	110	104		**744**	**463**	**448**	**526**
	698	148	92	134		**88**	53	62	56
	790	188	161	211		**95**	50	80	65
21	**219**	**949**	**242**	**281**	22	50	49	49	49
	47	**454**	50	52		51	50	51	51
	88	**536**	61	62		51	50	50	50
	134	**649**	112	130		50	50	51	49
31	**283**	**150**	**933**	**371**	32	65	51	**132**	58
	89	73	**150**	61		**523**	**439**	**705**	**509**
	111	84	**361**	69		59	50	**100**	53
	118	88	**502**	104		56	52	**146**	73
41	**502**	**369**	**503**	**999**	42	67	50	64	**105**
	126	90	51	**18**		**473**	**429**	**459**	**678**
	244	131	126	**361**		64	52	72	**74**
	164	104	97	**641**		56	51	83	**129**
13	**282**	73	109	178	41	**694**	262	342	322
	146	46	63	62		**322**	133	112	105
	920	**274**	**371**	**408**		**749**	104	126	112
	319	55	93	153		**999**	**268**	**439**	**597**
23	66	**187**	58	84	42	160	**914**	172	144
	51	**172**	50	48		46	**757**	50	49
	149	**453**	**127**	**137**		70	**896**	91	90
	57	**229**	79	98		**205**	**998**	**179**	**250**
33	66	64	**432**	111	43	126	122	**913**	106
	62	53	**138**	79		169	177	**416**	146
	383	**241**	**933**	**365**		128	140	**821**	148
	67	66	**440**	188		**368**	**175**	**1000**	**496**
34	124	71	135	**565**	44	157	187	150	**824**
	79	81	81	**189**		100	127	139	**355**
	857	**723**	**763**	**1000**		271	327	295	**536**
	108	86	176	**637**		**513**	**230**	**440**	**1000**

The results for Tables 6 and 7 are based on a 4x4 SAM with all initial cells and their variances of the same magnitude (equal to 1 and 0.2 respectively). The final row and column variances were then the sum of their rows and columns and the weighted quadratic form respectively. The story is much the same as observed previously

Table 6

Simulation Results under Ho: Type 1 Error

Wald Test

1	2	3	4	5	6	7	8
43	51	45	43	43	47	56	67

Kolmogorov-Smirnov Test

1	2	3	4	5	6	7	8
0.0286	0.0286	0.0160	0.0185	0.0111	0.0240	0.0314	0.0397

K-S Critical Value 0.0430

LM Test

1	2	3	4	5	6	7
46	56	53	48	55	56	52

Kolmogorov-Smirnov Test

0.0299	0.0290	0.0313	0.0303	0.0217	0.0201	0.0471

Diff Test

48	51	46	49	44	53	45	48
40	53	45	55	46	56	53	48

Kolmogorov-Smirnov Test

0.0263	0.0185	0.0196	0.0164	0.0283	0.0288	0.0140	0.0208
0.0238	0.0269	0.0317	0.0348	0.0299	0.0290	0.0313	0.0303

Table 7

Wald and LM tests: Rejection Rates under H1

	1	2	3	4	1	2	3	4
11 Wald	**320**	49	49	49	**357**	49	49	49
LM	**198**	54	55	62	**242**	49	49	
12 Wald	49	**310**	49	49	**357**	49	49	49
LM	70	**187**	55	62	**242**	49	49	
31 Wald	49	49	**339**	49	**357**	49	49	49
LM	70	54	**187**	62	**242**	49	49	
41 Wald	49	49	49	**722**	**992**	49	49	49
LM	91	97	88	**180**	**959**	49	49	
21 Wald	**320**	49	49	49	49	**297**	49	49
LM	**198**	54	55	62	49	**213**	49	
22 Wald	49	**310**	49	49	49	**297**	49	49
LM	70	**187**	55	62	49	**213**	49	
23 Wald	49	49	**339**	49	49	**297**	49	49
LM	70	54	**187**	62	49	**213**	49	
24 Wald	49	49	49	**722**	49	**993**	49	49
LM	91	97	88	**180**	50	**963**	49	
13 Wald	**320**	49	49	49	49	49	**295**	49
LM	**198**	54	55	62	49	49	**226**	
23 Wald	49	**310**	49	49	49	49	**295**	49
LM	70	**187**	55	62	49	49	**226**	
33 Wald	49	49	**339**	49	49	49	**295**	49
LM	70	54	**187**	62	49	49	**226**	
43 Wald	49	49	49	**722**	49	49	**992**	49
LM	91	97	88	**180**	50	49	**964**	
14 Wald	**992**	49	49	49	49	49	49	**648**
LM	**997**	104	110	180	242	213	226	
24 Wald	49	**989**	49	49	49	49	49	**648**
LM	120	**996**	110	180	242	213	226	
34 Wald	49	49	**992**	49	49	49	49	**648**
LM	120	104	**994**	180	242	213	226	
44 Wald	49	49	49	**1000**	49	49	49	**1000**
LM	532	498	481	**1000**	959	963	964	

Difference Test

	1	2	3	4		1	2	3	4
11	**482**	**129**	**132**	**220**	21	**117**	55	67	45
	153	57	51	49		**537**	**131**	**141**	**205**
	152	49	45	53		**152**	49	45	53
	198	54	55	62		**198**	54	55	62
21	**117**	**499**	**132**	**220**	22	53	**129**	67	45
	57	**131**	51	49		**153**	**450**	**141**	**205**
	64	**158**	45	53		64	**158**	45	53
	70	**187**	55	62		70	**187**	55	62
31	**117**	**129**	**487**	**220**	23	53	55	**132**	45
	57	57	**141**	49		**153**	**131**	**504**	**205**
	64	49	**164**	53		64	49	**164**	53
	70	54	**187**	62		70	54	**187**	62
41	**815**	**765**	**811**	**997**	24	67	69	79	**82**
	85	**77**	82	119		**829**	**805**	**788**	**997**
	90	**95**	100	97		90	95	100	**97**
	91	**97**	88	180		91	97	88	**180**
13	**117**	55	67	45	14	**815**	69	79	106
	153	57	51	49		**829**	77	82	104
	507	**158**	**164**	**184**		**818**	95	100	101
	198	54	55	62		**997**	**104**	**110**	**180**
23	53	**129**	67	45	24	67	**765**	79	106
	57	**131**	51	49		85	**805**	82	104
	152	**528**	**164**	**184**		90	**746**	100	101
	70	**187**	55	62		**120**	**996**	**110**	**180**
33	53	55	**132**	45	34	67	69	**811**	106
	57	57	**141**	49		85	77	**788**	104
	152	**158**	**511**	**184**		90	95	**778**	101
	70	54	**187**	62		**120**	**104**	**994**	**180**
43	67	69	79	**82**	44	390	382	398	**564**
	85	77	82	**119**		427	391	382	**547**
	818	**746**	**778**	**996**		442	341	391	**505**
	91	97	88	**180**		**532**	**498**	**481**	**1000**

The single test results in Tables 6 and 7 suggest that examination of the Lagrange or Wald test associated with the individual restrictions can enable the identification of biased initial estimates. However, a more thorough approach would be to examine rejection rates for the row-column interaction when both tests **simultaneously** reject. In Table 8 this is done and the joint rejections turn out to be quite low. This highlights the point that the actual Wald or LM test should be a χ^2 test (with two degrees of freedom) on the joint restrictions. The test, of course, is just (16), ie. $\hat{\lambda}' \hat{\Sigma}^{-1} \hat{\lambda} = \hat{\lambda}' (G\hat{V}G')\hat{\lambda} = x' G' (G\hat{V}G')^{-1} G x$ where the matrix G refers to only the two restrictions involved.

Table 8

Rejection Rates under H_1 (4 sector SAM)

Wald, LM and Diff tests

	W1	W2	W12	L1	L2	L12	Diff
1	320	291	117	193	213	21	464
2	325	291	121	201	213	20	447
3	361	291	129	240	213	24	526
4	712	991	707	166	958	150	994
5	320	338	132	193	265	26	521
6	325	338	143	201	265	31	490
7	361	338	137	240	265	41	559
8	712	995	708	166	978	153	998
9	320	296	115	193	233	26	494
10	325	296	130	201	233	31	472
11	361	296	135	240	233	31	504
12	712	992	708	166	961	152	997
13	994	753	751	0	0	0	995
14	990	753	751	0	0	0	997
15	995	753	749	0	0	0	1000
16	1000	1000	1000	0	0	0	1000

Rejection Rates under H1 (10 sector SAM)

Wald, LM and Diff tests

1	135	114	18	116	128	9	287
2	140	114	17	127	128	11	208
3	140	114	7	126	128	6	255
4	130	114	12	137	128	9	209
5	121	114	19	105	128	11	204
6	165	114	17	146	128	9	256
7	146	114	19	130	128	12	223
8	141	114	17	126	128	11	259
9	151	114	22	132	128	13	237
10	704	1000	704	200	1000	200	1000
11	135	141	25	116	135	14	203
12	140	141	21	127	135	16	238

W1 and W2 refer to the individual Wald (row and column) tests. W12 refers to the joint outcome of the two individual tests. The surprising result is that the joint tests (done individually) reject much less freqently than the Difference test. The W12 and L12 columns refer to the rejection count on both restrictions simultaneously. (Note the LM test was not calculated for the last 4 restrictions.) In the second half of the table, the first 12 rows (from X11 to X22) illustrate the same point in the context of a 10 sector SAM. However, the comparisons really are irrelevant unless the appropriate χ^2 tests are done.

In Table 9 the joint test results are presented for the 4 and 10 sector models. The results for the 10 sector model are not reported under H_0; briefly, there are no problems with the distributions of the test statistics under H_0 for either the 4 or 10 sector models. Only the first three rows are reported for the 4-sector SAM and only the first row is reported in the 10 sector case. The results are fully representative. The Difference test (size adjusted) has better power than the Wald test in every situation, thought the joint test now performs at a similar level to the Difference test. The fact that the Difference test (size adjusted) correctly rejects more frequently than the Wald (or LM) tests, in every case, suggests the existence of an underlying inequality favouring the Difference test. .

Table 9
Wald $\chi^2(2)$ and Difference Tests (4 sector SAM)
1000 replications

Distribution under Ho
Rejection Rate at 5% level: type 1 error

	11	12	13	14	21	22	23	24	31	32	33	34
W	50	54	49	54	47	52	49	55	45	53	49	55
D	41	52	46	52	59	48	54	49	47	55	56	48

Kolmogorov Smirnov test

	11	12	13	14	21	22	23	24	31	32	33	34	
W	.0272	.0272	.0302	.0185	.0217	.0169	.0291	.0228	.0396	.0191	.0285	.0327	cv(1 sided) = .0386
D	.0460	.0329	.0252	.0355	.0372	.0263	.0232	.0218	.0332	.0229	.0181	.0296	cv(2 sided) = .0430

Distribution under H1: Rejection Rate at 5% level

	11	12	13	14	21	22	23	24	31	32	33	34
W	367	349	376	988	373	348	375	985	411	378	390	990
D	515	479	562	996	453	485	471	997	556	490	501	998

Wald $\chi^2(2)$ and Difference Tests (10 sector SAM)
Distribution under H1: Rejection Rate at 5% level

	11	12	13	14	15	16	17	18	19	110
W1	163	173	178	166	186	180	185	171	179	1000
D1	249	242	239	197	259	249	206	203	226	1000
W2	152	158	163	159	169	141	177	195	176	1000
D2	182	196	218	179	201	213	221	207	245	1000
W3	195	178	184	163	174	173	176	176	175	1000
D3	244	188	270	240	206	226	217	214	213	1000
W4	167	179	171	177	172	160	175	186	172	1000
D4	227	206	190	262	242	224	224	248	222	1000
W5	173	176	158	170	162	153	158	162	167	1000
D5	206	238	197	216	222	193	215	230	185	1000
W6	157	154	158	179	177	162	168	168	169	1000
D6	214	190	228	211	229	225	218	221	176	1000
W7	159	152	153	154	167	168	160	176	153	1000
D7	206	185	227	192	251	238	197	223	205	1000
W8	182	181	182	160	182	163	202	182	172	1000
D8	261	209	226	240	257	248	254	190	264	1000
W9	148	149	157	146	159	149	152	151	163	1000
D9	179	226	226	273	170	252	187	202	248	1000

A further Monte Carlo study allowed the initial variances (V) to vary randomly as well as X. The same results were observed as above. Recognising that V has to be estimated and allowing for randomness in that operation (a variation of about 40% based on a uniform distribution) made little difference to the results under either the null or the alternative hypothesis.

4. Conclusions

The results are a vindication of common sense. The Difference test works well and does not appear to be adversely affected by the size of the problem. The statistical theory for the test procedures is straightforward and the tests are properly behaved under the null and alternative hypotheses. Monte Carlo simulations bear out the intuition and the creator of a SAM can choose between Wald, LM and Difference tests in trying to pinpoint initial estimate bias. The simplest procedure, the Difference test, works well, and its use for diagnostic investigation in data construction should be encouraged.

The importance of this is a validation of procedures being used in the field and ammunition to those who have been arguing against those government statistical offices which are unwilling to conceed that the figures they release are only estimates and are subject to random errors, as all estimates are. It has implications for national accounting, trade flow data, input-output analysis as well as SAM construction.

5. References

1. Arkhipoff, O., (1969), "Les Circuits Financiers au Cameroun", Compatilite Nationale 1965/66, etude speciale no 2; Direction de la Statistique, Yaounde.

2. Berndt, E.R. and Savin,N.E., (1977), "Conflict among Criteria for Testing Hypotheses in the Multivariate Regression Model", *Econometrica,* 45, pp.1263-1278.

3. Bickel,P.J. and Doksum, K.A., (1977), *Mathematical Statistics: Basic Ideas and Selected Topics,* Holden-Day, Oakland.

4. Byron, R.P., (1972), "Testing for Misspecification in Econometric Systems using Full Information", *International Economic Review,* 13, pp.745-756.

5. Byron, R.P., P.J.Crossman, J.E.Hurley and S.C.E.Smith, (1993), Balancing Hierarchical Regional Accounting Matrices, *International Conference in memory of Sir Richard Stone*, Certosa di Pontignano, Siena, ITALY, October 17-20, 1993.

6. Stone, R. (1975), Direct and Indirect Constraints in the Adjustment of Observations, in P.J.Bjerve (ed.), *Nasjonalregnskap, Modeller og Analyse,* Statistisk Sentralbyra, Oslo.

7. van der Ploeg, F. (1982), Reliability and the Adjustment of Sequences of Large Economic Accounting Matrices, *Journal of the Royal Statistical Society, Series A* 145, pp. 169-194.