The GSBGM Working Paper Series

WORKING PAPER 10/94

A study of the distributions underlying business survey responses in New Zealand

Vincenzo Cassino

GRADUATE
SCHOOL
OF BUSINESS &
GOVERNMENT
MANAGEMENT
VICTORIA UNIVERSITY
OF WELLINGTON

ISSN 0114-7420
ISBN 0475-11476-0

The Graduate School of Business and Government Management Working Paper Series 10/94. Oct. 1994

Cassino, Vincenzo 'A study of the distributions underlying business survey responses in New Zealand.'

Enquiries concerning this paper direct to the author *
General enquiries concerning GSBGM publications to:
Monica Cartner
Research Co-ordinator
The Graduate School of Business and Government Management
The Victoria University of Wellington
PO Box 600
Wellington
New Zealand

Telephone: 64-4-495 5085
Facsimile: 64-4-496 5435
E-mail: Monica.Cartner@vuw.ac.nz

* Vincenzo Cassino
Economics Group
Faculty of Commerce and Administration
Victoria University of Wellington

Printed by The Victoria University of Wellington Printers.

A Study of the Distributions Underlying Business Survey Responses in New Zealand

Vincenzo Cassino

CONTENTS

I.	Introduction	2
II.	Ivaldi's Analysis	3
III.	Testing for Bivariate Normality	11
IV.	New Zealand Survey Data	14
V.	Empirical Results	16
VI.	Conclusions	20
	Appendix	22
	References	25

Abstract

This study examines the distributions underlying business survey responses in New Zealand. These distributions are used in procedures which derive quantitative statistics from qualitative survey data. They are also used in an econometric procedure recently developed by the French economist Marc Ivaldi which uses survey data to test structural economic models. These procedures often assume that the distributions are normal. The purpose of this study is to examine the empirical validity of this assumption. The polychoric correlation coefficient procedure is used to test the goodness of fit of a standardised bivariate normal distribution on data from the New Zealand Institute of Economic Research's Quarterly Survey of Business Opinion.

Key words: Business surveys, normal distribution, polychoric correlation coefficient, Quarterly Survey of Business Opinion.

Journal of economic literature classification: C19

A Study of the Distributions Underlying Business Survey Responses in New Zealand

Vincenzo Cassino[1]

I. Introduction

Data from tendency surveys of consumers and firms has been studied by economists since the introduction of the first survey of German industries by the IFO-Institute fur Wirtschaftforschung in 1950. A number of techniques have been developed which use this qualitative data to examine the relationship between unobservable quantitative variables. These techniques are often based on the assumption that the responses to tendency surveys are underpinned by quantitative variables. These quantitative variables trigger the different survey responses when they cross certain thresholds. In addition, it is often assumed that the quantitative variables are normally distributed. A univariate technique developed by Theil (1952) and Carlson and Parkin (1975) estimates the mean and variance of the normally distributed quantitative data underlying the qualitative responses. This technique has been widely applied[2]. Unfortunately, this approach lacks economic foundations, and is simply an application of

---

[1] *Economics Department, Reserve Bank of New Zealand.* This paper is based on research carried out for my MCA thesis at Victoria University in 1993. I would like to express my thanks to my supervisor, Professor Fraser Jackson, for his guidance and support while I was working on the thesis. I would also like to thank to Professor Viv Hall, Bob Buckle and Jacques Poot for their helpful advice and comments. Thanks also to Bob Buckle for initially suggesting this area of research. The work in this paper was carried out before joining the Economics Department of the Reserve Bank, and the results do not necessarily reflect the views of the Reserve Bank of New Zealand.

[2] See for example Danes (1975) and Defris and Williams (1979) on Australian data, Carlson and Ryder (1973), Smith (1978), Smith (1982) and Bennett (1984) on British data, Batchelor (1982) on European data, Carlson and Ryder (1973) and Batchelor (1986) on American data, Seitz (1988) on German data, and Hall and King (1976) and Roseveare and Millar (1988) on New Zealand data.

statistical relationships.

A more economically rigorous procedure was derived recently by the French economist Marc Ivaldi. His procedure (Ivaldi (1990), (1991), (1992)) is based on formulating optimisation-based models for the quantitative data which underlies the qualitative responses to business surveys. One of the assumptions used in Ivaldi's technique is that the variables underlying the discrete responses have a multivariate normal distribution. This is used to derive the correlation between the underlying variables and solve the latent variables estimation problem.

The aim of this paper is to examine the validity of the normality assumption for data from the New Zealand Institute of Economic Research's Quarterly Survey of Business Opinion. In addition to testing the validity of Ivaldi's procedure, extending the analysis to a bivariate framework should provide new insights into the properties of these underlying distributions by providing extra degrees of freedom.

The paper is organised as follows. Section II presents Ivaldi's procedure by summarising the structure of a model of a profit maximising firm, and demonstrates how the model relies on an assumption of multivariate normality. Section III describes the empirical procedures which are used by Ivaldi to implement his model, and are used in this study to test for normality. Section IV provides brief details of the QSBO data examined in this study. Finally, section V summarises the results of the bivariate normal tests carried out.

II. Ivaldi's Analysis

Ivaldi (1991) constructs a model based on a stochastic control problem, which describes the behaviour of a representative firm. The firm chooses its production level (q_t) to minimize its expected discounted costs. Uncertainty arises from the levels of demand (α_t) and costs, both of which follow general ARMA(p,q) processes. The uncertain element of costs (γ_t) forms part of the direct costs of current production (C_{1t}). Non-stochastic costs are also incurred by changing the level of production (C_{2t}), and carrying over inventories (C_{3t}). Given the uncertainty about the levels of demand and costs in future periods, firms will choose a linear decision rule in period 0 which makes q_t a function of the information set available to the firm in period t. This rule is derived from the following constrained optimization problem:

Minimise $$\lim_{T\to\infty} E_0\left(\sum_{t=1}^{T} r^t \sum_{i=1}^{3} C_{it}\right)$$

subject to $x_t - x_{t-1} = q_t - \alpha_t \qquad t=1...T...$

where:

$E_0(\bullet)$ = Expectations conditional on the information set at period 0
r = The discount factor
x = The net inventory level

The constraint is simply an accounting identity which relates changes in the level of inventories each period to differences between the levels of demand and output. Deriving the first order condition for this optimisation problem produces the following Euler equation for each period:

$$E_{t-1}\left[rq_{t+1} + \psi q_t + q_{t-1}\right] = 1/c_2 \left[c_1 E_{t-1} \gamma_t - c_3 E_{t-1}\alpha_t \right]$$

where

$\psi = -\ 1/c_2\ (c_1 + c_3 + (1+r)c_2)$
c_1 , c_2, c_3 = Parameters in the cost functions

Assuming that a transversality condition is satisfied, the above second-order difference equation can be solved for the following general solution:

$$E_{t-1}q_t=\lambda_1 q_{t-1}+\frac{\lambda_1}{c_2}[c_3\sum_{j=0}^{\infty}\beta^j E_{t-1}\alpha_{t+j}-c_1\sum_{j=0}^{\infty}\beta^j E_{t-1}\gamma_{t+j}]$$

where

$\beta = 1/\lambda_2$

λ_1 , λ_2 = The roots of the left-hand side of the Euler equation when written in terms of the lag operator (L):

$rq_{t+1} + \psi q_t + q_{t-1} = (rL^{-1} + \psi + L)q_t$

$= r(L^{-1} - \lambda_2)(1- \lambda_1 L)q_t$

$\lambda_2 > \lambda_1$

Before estimating the structural model developed above, Ivaldi assumes that the variables in the model (the 'true variables') are measured with error by some other variables (the 'measuring variables'). These errors are interpreted as noisy components which prevent the firm from behaving according to the model above, and may be caused by factors such as technical failures or changes in government policy[3] . The estimation procedure is further complicated by the fact that the measuring variables are only indirectly observable through discrete survey data. In Ivaldi's case this is from the Institut National de la Statistique et des Etudes Economiques (INSEE) survey of French industry.

As a result of these features, there are two sections to Ivaldi's

---

[3] Ivaldi (1991) p.53

empirical model[4]. The first section contains the model's structural equations. (1) and (2) represent the exogenous stochastic processes for demand and costs, and the identity defining changes in these variables. (3) is the production plan determination equation converted to differenced form to make it consistent with the survey data. (4) is the unbiasedness condition for rational expectations. Finally, (5) specifies an AR(1) process for the measurement error on the expectation variable for production (ε^{qet}):

Structural Equations

Stochastic Processes:

$$\tilde{\alpha}_t=\theta_1\tilde{\alpha}_{t-1}+u^{\alpha}{}_t \tag{1}$$

$$\tilde{\gamma}_t=\theta_2\tilde{\gamma}_{t-1}+u^{\gamma}{}_t$$

and

$$\Delta\tilde{\alpha}_t=\tilde{\alpha}_t-\tilde{\alpha}_{t-1} \tag{2}$$

---

[4] A tilde (~) signifies a true variable while a star (*) signifies a measuring variable.

$$\Delta\tilde{\gamma}_t=\tilde{\gamma}_t-\tilde{\gamma}_{t-1}$$

Production Plans

$$\Delta\tilde{q}^{e}_{t+1}=\Delta\tilde{q}^{e}_{t}+g_1'\Delta\tilde{q}_t+g_2\Delta\tilde{\alpha}_t+g_3\Delta\tilde{\gamma}_t+u^{e}_{t} \qquad (3)$$

$$\Delta\tilde{q}_{t+1}=\Delta\tilde{q}^{e}_{t+1}+\Delta u_{t+1} \qquad (4)$$

$$\varepsilon^{qe}_{t}=\rho\,\varepsilon^{qe}_{t-1}+u^{qe}_{t} \qquad (5)$$

where u^{α}_{t} , u^{γ}_{t} , u^{e}_{t} , u^{qe}_{t} = independent white noise processes

$$\Delta u_{t+1} = u_{t+1} - u_t$$

The second section of the model contains the measurement equations which describe the relationships between the true variables and the measuring variables. (6) shows the measurement errors in expected and realised output. In (7) Ivaldi uses inventories (L_t) and order-backlogs (S_t) as joint proxies for demand and costs[5]. The empirical relationship between orders and inventories, and demand and costs, is examined more fully using an errors-in-latent-variables

---

[5] Ivaldi (1991) p.55

model similar to the one presented here in Ivaldi (1990).

Measurement Equations

$$\begin{aligned} \Delta q^{e*}_t &= \Delta\tilde{q}^e_t + \varepsilon^{qe}_t \\ \Delta q^*_t &= \Delta\tilde{q}_t + \varepsilon^q_t \end{aligned} \tag{6}$$

$$\begin{aligned} L^*_t &= \lambda\tilde{\alpha}_t + \tilde{\gamma}_t + \varepsilon^l_t \\ S^*_t &= \tilde{\alpha}_t + \mu\tilde{\gamma}_t + \varepsilon^s_t \end{aligned} \tag{7}$$

where: ε^{qe}_t, ε^q_t, ε^l_t, ε^s_t = measurement errors

Equations (1) to (7) represent the full model to be estimated. Under its current specification, all of the model's parameters cannot be identified from the cross sectional data of one survey. This problem can be overcome by using panel data. It can be shown that at least 4 successive surveys are needed to identify all the parameters in the model[6]. The structural equations for these

---

[6] *Ibid* p.60

periods can be collected to form the following system:

$$\eta = B\eta + \zeta \qquad (8)$$

where[7]

η = A vector containing all the latent endogenous and exogenous true variables, and the errors from the measurement equations

ζ = A vector of residuals

B = A non-singular matrix of regression coefficients

Similarly, the measurement equations can be collected to form the following system:

$$y^* = \Gamma\eta \qquad (9)$$

where

y^* = The latent measuring variables which will be indirectly observed through the survey data

Γ = A non-singular matrix of the regression coefficients of y^* on η

Equations (8) and (9) together form the general latent variable model[8]. The parameters of this model are usually derived from the sample covariance / correlation matrix of the observable y^* variables[9]. In the current situation,

---

[7] The precise order of the following vectors and matrices will depend on the number of survey periods examined by the model.

[8] See Aigner et al (1984).

[9] See Ivaldi (1992) pp.230-232 for a simple example.

however, data on the measuring variables is available only indirectly in the form of discrete survey data. To overcome this problem, it is assumed that firms' responses to each survey question are triggered when the measuring variable crosses certain thresholds[10] :

$$y_{it} = \begin{cases} 1 \text{ if } & \Delta y^*_{it} > \delta^1_t(y) \\ 2 \text{ if } & \delta^2_t(y) < \Delta y^*_{it} \leq \delta^1_t(y) \\ 3 \text{ if } & \Delta y^*_{it} \leq \delta^2_t(y) \end{cases} \qquad (10)$$

where

y_{it} = An indicator variable relating a firm's survey responses to the values of y^*_{it}

y^*_{it} = An individual measuring variable in the y^* vector

$\delta^j_t(y)$ = The thresholds triggering the different survey responses

To estimate the model's parameters the unobserved y^* measuring variables are assumed to have a multivariate normal distribution, with a mean vector of zero, and a covariance matrix $\Sigma = E(y^* y^{*\prime})$. An estimate of Σ (S) is then derived using the polychoric correlation coefficient technique. This technique, originally developed by Pearson (1901), derives the correlation between two variables when only qualitative data on them is available.

The procedure assumes that the pairs of unobserved continuous variables (y^*_i , y^*_j) which underlie the discrete data each have a bivariate normal distribution, with a zero mean vector, unit variances and correlation ρ_{ij}. When all the correlation coefficients have been estimated to form a sample correlation

---

[10] Ivaldi (1991) p.71

matrix, it can be expressed in terms of the parameter matrices:

$$\Sigma = \Gamma(I - B)^{-1} \Psi (I - B')^{-1} \Gamma' \qquad (11)$$

where:

$$\Psi = E(\zeta\zeta')$$

If the model is identified, this equation can be solved for a unique set of parameters in B, Γ, and Ψ.

Clearly, this estimation procedure rests on the assumption of bivariate normality. Ivaldi rejected the normality hypothesis in 56 out of 136 tests at a 5% level of significance on French survey data. This rejection rate was attributed primarily to skewness in the distributions[11].

III. Testing for Bivariate Normality

A general description of the polychoric correlation procedure for a r x s contingency table is provided in Olsson (1979). A simple application of the general procedure to a 3 x 3 table is provided in Tallis (1962), and forms the basis of the following exposition. Let X^* and Y^* be two standardised random variables with a bivariate normal distribution. Assume that these variables are only observable indirectly through the trichotomous discrete variables X and Y, whose values depend on whether the underlying continuous variables cross certain thresholds, as described in section II.

If the combinations of X and Y are summarised on a contingency table

---

[11] Ivaldi (1991) p.113

as shown in Figure 1, it is easy to see that the following properties hold for the marginal distributions of X* and Y*:

$$\Pr(X=x_i) = \Phi(a_{i+1}) - \Phi(a_i)$$

$$\Pr(Y=y_j) = \Phi(b_{j+1}) - \Phi(b_j)$$

$$i=j=0,1,2$$

where:

$\Phi(\bullet)$ = The standardised univariate normal distribution function

a_1, a_2, b_1, b_2 = The threshold values for X* and Y* which trigger the different discrete responses in X and Y

$a_0=b_0=-\infty$ and $a_3=b_3=+\infty$

In addition, the expected probabilities for each cell in the contingency table can be expressed as functions of the joint distribution of X* and Y* according to the following general formula:

$$P_{ij} = \Phi(a_{i+1},b_{j+1} \mid \rho) - \Phi(a_i,b_{j+1} \mid \rho) - \Phi(a_{i+1},b_j \mid \rho) + \Phi(a_i,b_j \mid \rho)$$

where:

$\Phi(\bullet, \bullet \mid \rho)$ = The standardised bivariate normal distribution function of X* and Y* conditional on the correlation coefficient ρ

The likelihood function for the contingency table is given by:

$$L=C\prod_{i=0}^{2}\prod_{j=0}^{2}P_{ij}^{n_{ij}} \qquad (12)$$

Taking logs, this produces the log likelihood function:

$$l=\ln L=\ln C+\sum_{i=0}^{2}\sum_{j=0}^{2}n_{ij}\ln P_{ij} \tag{13}$$

where:

C = a constant which does not depend on the parameters to be estimated

A maximum likelihood approach can be used to derive estimates of the correlation coefficient and the four thresholds. This is done by differentiating the log likelihood with respect to each of the five parameters and then deriving the values of ρ, a_1 , a_2, b_1 and b_2 which set the first order derivatives equal to zero. Full details of the derivatives are presented in Tallis (1962) and Olsson (1979). A brief summary is in the appendix. Since the functions are non-linear in parameters, solving this five equation system requires some iterative estimation technique, such as the Newton-Raphson procedure:

$$\hat{\theta}_{t+1}=\hat{\theta}_t-\lambda(G_t)^{-1}g_t$$

where:

$\hat{\theta}$= Vectors of the new (t+1) and previous (t) parameter estimates of (ρ, a_1 , a_2, b_1, b_2)

$(G_t)^{-1}$ = Estimate of the inverse of the Information matrix
$= -E(\partial^2 l(\theta) / \partial \theta \partial \theta')^{-1}$

g_t = Vector of first order partial derivatives = $\partial l(\theta) / \partial \theta$

λ = Step size in each iteration ($\lambda=1$ was used here)

Once the parameters have been estimated, the expected frequencies associated with each cell in the contingency table can be compared to the actual frequencies, to determine the goodness-of-fit of the bivariate normal model. One commonly used goodness-of-fit measure is the likelihood ratio test:

$$LRT=2\sum_{i=1}^{9} o_i \ln(\frac{o_i}{e_i})$$

where:

o_i = The observed number of observations in each cell of the contingency table

e_i = The expected number of observations in each cell of the contingency table

This statistic will asymptotically have a chi-square distribution with 3 degrees of freedom.

IV. New Zealand Survey Data

The data used for this study is from the Quarterly Survey of Business Opinion (QSBO) conducted by the New Zealand Institute of Economic Research (NZIER). The format of the QSBO is similar to that of overseas business tendency surveys, such as those conducted by the Ifo-Institute and the INSEE. There are general questions on economy-wide conditions, and questions concerning the past and expected movement of variables related to the individual firms which require 'up' / 'same' / 'down' / 'N/A' responses. The responses to the surveys of Manufacturers and Builders provide the data for the empirical analysis in this study. The sample involves data from survey n.58 (September Quarter 1975) to survey n.129 (June Quarter 1993), making a total of 72 surveys. Responses to the questions which most closely match the variables used by Ivaldi were used[12] :

> Question 14. Changes in the level of new orders realised during the past three months, and expected during the next three months
>
> Question 15. Changes in the level of output realised during the past three months, and expected during the next three months
>
> Question 23. Changes in the level of stocks (raw materials) realised during the past three months, and expected during the next three months
>
> Question 24. Changes in the level of stocks (finished goods) realised during the past three months, and expected during the next three months

For each survey, the responses to the questions above were summarised into 28 contingency tables which showed the joint distribution of the responses to each combination of questions.

---

[12] The data was extracted from the data file containing the firms' individual responses to each survey held by Professor Fraser Jackson.

V. Empirical Results

All of the empirical analysis described above was carried out using the *APL* (version 10) programming language[13].

A summary of the polychoric correlation coefficient results is presented in Table 1. These figures show the strength of the linear relationship between changes in each variable. The highest mean values for the correlation coefficients across the entire sample period are found to be between realised changes in new orders and realised changes in output, and between the equivalent expected variables, which are regularly 70-80% in the individual surveys. This suggests that output is extremely responsive to demand. The second highest mean correlations are between realised changes in raw stocks and realised changes in finished stocks, and their expected equivalents, which range between 50% and 70%. The remaining 24 bivariate models are generally found to have low correlations (less than 30% in absolute value).

The results of the bivariate normal hypothesis testing are summarised in Table 2. It is clear that the usefulness of the normal distribution as an approximation to the actual bivariate relationships varies considerably between the different models. The relationships best modelled by a joint normal distribution are those involving combinations of realised finished stocks and realised raw stocks, with expected orders and expected output. The worst models are generally found to be those involving combinations of realised and expected stocks. There is a strong positive relationship ($r=62.7\%$) between the models' polychoric correlation coefficient and their rate of rejection of normality.

---

[13] The user-defined functions utilised were written mainly by Professor Fraser Jackson, although a number of alterations and improvements were made.

It is clear from the standard errors of the correlation coefficients and the likelihood ratio tests that many of the sample distributions are extremely volatile. Bootstrap simulation analysis was conducted to determine whether this movement reflected sampling variability or changes in the actual distribution. The bivariate distributions from a survey drawn at random (survey n.73, June 1979) were used as comparison distributions for each model. Two hundred samples from the individual responses to this survey were then drawn. By fitting bivariate normal distributions to each simulated drawing, the sampling distributions of the models' parameters under the survey n.73 distributions could be determined. These results were then compared to the actual distributions of the statistics calculated. Table 3 presents the bootstrap results for the correlation coefficients of a number of models.

If the bivariate distributions over the sample period were from a population with the same structure as the sample in survey n.73, 10% of the 72 correlation coefficients derived for each model (approximately 7) would be expected to be below the 5% percentile or above the 95% percentile. As Table 3 demonstrates, the actual number of coefficients in this range is usually far more than expected. Consequently, the deviation of the correlation coefficients over the sample period from the pattern in survey n.73 is statistically significant. This suggests that the volatility observed in Tables 1 and 2 is at least partly due to changes in the actual distributions. It also appears that the volatility of the bivariate distributions is positively related to the strength of the variables' correlation.

The volatility of the bivariate distributions is confirmed by the fluctuations in the proportion of models rejecting normality in each survey over the sample period. This is demonstrated in Graph 1. The rejection rate for the 28 models tested on each survey's responses was found to fluctuate from a high of 78.6% to a low of 25%, with a mean of 49.76%.

The results considered so far suggest that the usefulness of the bivariate normal distribution as an approximation varies considerably between models and across surveys. This was examined formally using analysis of variance procedures on the likelihood ratio statistics. A simple additive linear model was fitted to investigate whether the deviations of individual likelihood ratios (L_{ij}) from their overall mean (a_0) could be explained by the model (M_i) and survey (S_j) under consideration:

$$L_{ij} = a_0 + a_1M_i + a_2S_j + e_{ij}$$

where e_{ij} = a random error term

The results are presented in Table 4. This demonstrates that the variation in the mean likelihood ratio statistics of different models and surveys is statistically significant. This provides evidence confirming that the validity of the bivariate normal distribution as an approximation does vary between models and across surveys.

To determine the reason for the widespread rejection of normality, the relationship between the actual and expected frequencies in each cell was examined. For all the models in each survey, the deviation between the actual frequency and the expected frequency was calculated, to determine the number of instances of under-prediction in each cell-position. If the deviations from the expected frequency were random, with over- and under-prediction equally likely in each cell, 14 instances of under-prediction should be expected for each cell-position out of 28 models. The actual average number of instances of under-prediction for each cell across the sample period is presented in Table 5.

The figures in Table 5 demonstrate that persistent under-prediction occurs in the middle 'no change' / 'no change' cell, and in the corner cells which involve a combination of change responses. Conversely, cells involving a change and a no change response show a systematic pattern of over-prediction. These results are consistent with leptokurtosis (excessive peakedness and fat tails) arising from too many respondents answering 'no change' to the questions. A possible explanation for this phenomenon is provided by Theil (1966) and Ronning (1990), who suggest that some of the no change responses may actually be don't know's. This finding is at odds with Ivaldi's conclusion, which attributed the rejection of normality to skewness.

Another issue worth investigating is whether the rejection of normality by the models is independent. This means examining whether the rejection of bivariate normality by one model provides any predictive information about the conclusion on other models in the survey. Two sets of tests for independence were carried out. Firstly, the correlation matrix for the likelihood ratio tests of each model was derived. Secondly, chi-square tests for independence were performed on tables displaying the joint distribution of acceptance and rejection of normality for each pair of models.

For the 28 bivariate models per survey, 378 inter-model relationships were examined. The results of the correlation matrix of the likelihood ratio tests for each model are summarised in Table 6. Clearly, there is almost no relationship between the likelihood ratio test statistics derived for most pairs of models.

These findings were confirmed by the chi-square tests, in which only 9% of the test statistics exceeded the 5% critical value and rejected independence. This result is not altogether surprising given the volatility of many of the likelihood ratio tests observable in Table 2. It is also not surprising that in the few cases where the bivariate normality of models is not independent, the two models usually have a common variable. For example, a correlation of 71.37% was obtained between the likelihood ratio tests of the expected orders vs expected finished stocks model and the expected raw stocks vs expected finished stocks model. A similar result (r=71.32%) was found between the realised finished stocks vs expected orders model and the realised finished stocks vs expected output model.

VI. Conclusions

The study appears to provide a blow to Ivaldi's justification for using the bivariate normal distribution in his procedure for estimating structural parameters from qualitative survey data. The polychoric correlation coefficient procedure was used to fit standardised bivariate normal models to data from the NZIER's Quarterly Survey of Business Opinion for the period September 1975-June 1993. On average, 50% of the models derived in each survey rejected the normality hypothesis. However, the rejection rate of individual models varied considerably, and was positively related to the strength of the variables' correlation. The distributions of these highly correlated variables were also found to change significantly over time.

Despite this high rejection rate, the anaysis did provide new insights into the survey data. The rejection of normality was due primarily to leptokurtosis (peakedness and fat tails), resulting from an excessive number of observations in the central 'no change' cell. This is consistent with the observation made by a number of researchers in the univariate literature that some 'no change' responses may actually be 'don't know's. However, it is at odds with Ivaldi's finding that the rejection of normality is due to skewness. It was also found that the distributions of the different bivariate models are generally independent, unless they contain a common variable.

These findings should prove useful for applications of the Carlson and Parkin procedure. It appears that for real variables, a symmetric distribution more peaked than the normal should be used to estimate the mean and variance of the variables underlying the responses to survey questions.

This raises the question of further research possibilities. An obvious first step would be to test the bivariate normality of other variables. So far only the within-survey relationships between real variables have been examined. As a result, no investigation of the dynamic relationships which must surely exist in cross-survey models has been made. It would also be interesting to examine whether the leptokurtosis found in the distributions of these real variables is also present in the distributions of nominal variables such as prices and costs. If prices are set on an administered basis, and firms are unwilling to reduce prices as numerous models suggest, it is possible that there may also be skewness present in the bivariate distributions.

An alternative extension would be to carry on with Ivaldi's methodology to derive full correlation matrices for New Zealand data using the polychoric procedure in order to estimate the structural parameters for one of the models applying the technique. Comparing the results with the findings of other studies which test the same hypotheses with different techniques may give some indication of the operational significance of using the bivariate normal distribution in Ivaldi's methodology. For example, the findings of a test for Rational Expectations could be compared to the results of Buckle et al (1990), which uses the approach of Kawasaki and Zimmermann (1986).

The nature of this study has meant that it has few counterparts other than the work conducted by Ivaldi himself. It is clear, then, the findings reported here are little more than a brief glimpse of the new econometric research possibilities opened up by Ivaldi's attempts to model the distributions of the variables underlying the responses to business surveys. Consequently, there are as many questions raised by the findings as there are answered.

Appendix

The Derivatives of the Log Likelihood for the Polychoric Correlation Coefficient

The objective of this optimisation problem is to simultaneously choose the correlation coefficient (ρ) and the 4 thresholds (a_1,a_2,b_1,b_2) which maximise the log likelihood function:

$$l=\ln L=\ln C+\sum_{i=0}^{2}\sum_{j=0}^{2}n_{ij}\ln P_{ij}$$

Using the joint distribution of X* and Y*, the expected probability that an observation falls into each cell of the contingency table can be expressed in the following way:

$$P_{ij} = \Phi(a_{i+1},b_{j+1} \mid \rho) - \Phi(a_i,b_{j+1} \mid \rho) - \Phi(a_{i+1},b_j \mid \rho) + \Phi(a_i,b_j \mid \rho)$$

where $a_0=b_0= -\infty$ and $a_3=b_3= +\infty$

Using the Chain Rule on the log likelihood function, the general formula used to derive the first order partial derivatives is:

$$\frac{\partial l}{\partial \theta_s}=\sum_{i=0}^{2}\sum_{j=0}^{2}\frac{n_{ij}}{P_{ij}}\frac{\partial P_{ij}}{\partial \theta_s}$$

where:

$\theta_s = \rho,\ a_1,\ a_2,\ b_1,\ b_2$

The key result used to derive the first order partial derivative with respect to the correlation coefficient is :

$$\frac{\partial \Phi(a,b|\rho)}{\partial \rho}=\phi(a,b|\rho)$$

where:

$\phi(\bullet,\ \bullet\ |\ \rho)$ = The standardised bivariate normal density function conditional on the correlation coefficient ρ

Applying this result to the formulae for the expected probabilities in each cell:

$$\frac{\partial l}{\partial \rho}=\sum_{i=0}^{2}\sum_{j=0}^{2}\frac{n_{ij}}{P_{ij}}[\phi(a_{i+1},b_{j+1}|\rho)-\phi(a_i,b_{j+1}|\rho)-\phi(a_{i+1},b_j|\rho)+\phi(a_i,b_j|\rho)]$$

To derive the partial derivatives with respect to the four thresholds, the key result is :

$$\frac{\partial P_{ij}}{\partial a_k}=\begin{cases} 0 & \text{if k} \neq \text{i and k} \neq \text{i+1} \\ \dfrac{\partial\Phi(a_k,b_{j+1}|\rho)}{\partial a_k}-\dfrac{\partial\Phi(a_k,b_j|\rho)}{\partial a_k} & \text{if k = i+1} \\ -\dfrac{\partial\Phi(a_k,b_{j+1}|\rho)}{\partial a_k}+\dfrac{\partial\Phi(a_k,b_j|\rho)}{\partial a_k} & \text{if k=i} \end{cases}$$

where k=1,2

Applying this result to the formulae for the expected probabilities in each cell:

$$\frac{\partial l}{\partial a_k}=\sum_{j=0}^{2}[\frac{n_{k-1j}}{P_{k-1j}}-\frac{n_{kj}}{P_{kj}}]\phi(a_k)[\Phi(B_{j+1,k})-\Phi(B_{jk})]$$

where:

$$B_{jk}=\frac{b_j-\rho a_k}{\sqrt{1-\rho^2}}$$

Similarly, by symmetry:

$$\frac{\partial l}{\partial b_k}=\sum_{i=0}^{2}[\frac{n_{ik-1}}{P_{ik-1}}-\frac{n_{ik}}{P_{ik}}]\phi(b_k)[\Phi(A_{i+1,k})-\Phi(A_{ik})]$$

where:

$$A_{ik}=\frac{a_i-\rho b_k}{\sqrt{1-\rho^2}}$$

In addition, the elements of the Information matrix of second order partial derivatives can be estimated using the following relationship:

$$I_{st}=N\sum_{i=0}^{2}\sum_{j=0}^{2}\frac{1}{P_{ij}}\frac{\partial P_{ij}}{\partial\theta_s}\frac{\partial P_{ij}}{\partial\theta_t}$$

where:

N = The sample size

I_{st} = Element s,t in the Information matrix

References

Aigner, D.J. , Hsiao, C. , Kapteyn, A. and Wansbeek, T. (1984) " Latent Variable Models in Econometrics ", in Griliches, Z. and Intrilligator, M.D. (eds) *Handbook of Econometrics* vol.II, Amsterdam, North-Holland, 1321-1393.

Batchelor, R.A. (1982) " Expectations, Output and Inflation - The European Experience ", *European Economic Review*, 17, 1-25.

Batchelor, R.A. (1986) " Quantitative v. Qualitative Measures of Inflation Expectations ", *Oxford Bulletin of Economics and Statistics*, 48, 99-120.

Bennett, A. (1984) " Output Expectations of Manufacturing Industry ", *Applied Economics*, 16, 869-879.

Buckle, R.A. , Assendelft, E.W. and Jackson, L.F. (1990) "Manufacturers' Expectations of Prices and Quantities: New Zealand Experience, 1964-87", *Applied Economics*, 22 , 579-598.

Carlson, J. and Parkin, M. (1975) " Inflation Expectations ", *Economica*, 42, 123-138.

Carlson, J. and Ryder, H.E. (1973) " Quantitative Expectations from Qualitative Surveys: A Maximum-Likelihood Approach ", *Mimeo*.

Danes, M. (1975) " The Measurement and Explanantion of Inflationary Expectations in Australia ", *Australian Economic Papers*, 14 , 75-87.

Defris, L.V. and Williams, R.A. (1979) " Quantitative versus Qualitative Measures of Price Expectations- The Evidence from Australian Consumer Surveys", *Economics Letters*, 2 169-173.

Hall, V.B. and King, M.L. (1976) " Inflationary Expectations in New Zealand: A Preliminary Study ", *New Zealand Economic Papers*, 10 , 118-151.

Ivaldi, M. (1990) " An Empirical Study of Inventory and Order Appraisals from Business Surveys ", in Florens, J. , Ivaldi, M. , Laffont, J. and Laisney, F. (eds) *Microeconometrics: Survey and Applications*, Oxford, Basil Blackwell.

Ivaldi, M. (1991) " A Structural Analysis of Expectations Formation based on Business Surveys of French Manufacturing Industries " *Lecture Notes in Economics and Mathematical Systems* 354 ,Berlin, Springer-Verlag.

Ivaldi, M. (1992) " Survey Evidence on the Rationality of Expectations ", *Jurnl of Applied Econometrics*, 7 , 225-241.

Kawasaki, S. and Zimmermann, K.F. (1986) " Testing the Rationality of Price Expectations for Manufacturing Firms ", *Applied Economics*, 18 , 1335-1347.

Olsson, U. (1979) " Maximum Likelihood Estimation of the Polychoric Correlation Coefficient ", *Psychometrika*, 44 , 443-460.

Pearson, K. (1901) "Mathematical Contributions to the Theory of Evolution, VII: On the Correlation of Characters not Quantitatively Measurable", *Philosophical Transactions of the Royal Society of London*, Series A, 195, 1-47.

Ronning, G. (1990) " The Informational Content of Responses from Business Surveys ", in Florens, J. , Ivaldi, M. , Laffont, J. and Laisney, F. (eds) *Microeconometrics: Survey and Applications*, Oxford, Basil Blackwell, 123-144.

Roseveare, D. and Millar, C. (1988) " Testing the Rationality of Manufacturers' Inflationary Expectations ", *New Zealand Economic Papers*, 22 , 3-13.

Seitz, H. (1988) " The Estimation of Inflation Forecasts from Business Survey Data ", *Applied Economics*, 20, 427-438.

Smith, G.W. (1978) " Producers' Price and Cost Expectations ", in Parkin, M. and Sumner, M.T. (eds) *Inflation in the United Kingdom*, Manchester, Manchester University Press.

Smith, G.W. (1982) " Inflation Expectations : Direct Observations and Their Determinants " in Artis, M. , Green, C.J. , Leslie, D. and Smith, G.W. (eds) *Demand Management, Supply Constraints and Inflation*, Manchester, Manchester University Press.

Tallis, G.M. (1962) " The Maximum Likelihood Estimation of Correlation from Contingency Tables ", *Biometrics*, 18, 342-353.

Theil, H. (1952) " On the Time Shape of Economic Microvariables and the Munich Business Test ", *Revue de L'Institut International de Statistique*, 20, 105-120. Reprinted in Raj, B. and Koerts, J. (eds) (1991) *Henri Theil's Contribution to Economics and Econometrics*, Volume I, Boston, Kluwer Academic, 281-302.

Theil, H. (1966) *Applied Economic Forecasting*, Amsterdam, North-Holland.

Table 1.
Summary of the Polychoric Correlation Coefficient Results

Bivariate Model	Mean of Correlation Coefficients (Standard Deviation)
Rlsd Orders v Rlsd Output	0.8189 (0.0486)
Rlsd Orders v Rlsd Raw Stocks	0.1107 (0.1005)
Rlsd Orders v Rlsd Fnshd Stocks	-0.0802 (0.1313)
Rlsd Orders v Exp Orders	0.3646 (0.1216)
Rlsd Orders v Exp Output	0.4013 (0.1191)
Rlsd Orders v Exp Raw Stocks	0.1618 (0.1144)
Rlsd Orders v Exp Fnshd Stocks	0.0652 (0.1124)
Rlsd Output v Rlsd Raw Stocks	0.1432 (0.0950)
Rlsd Output v Rlsd Fnshd Stocks	0.0066 (0.1194)
Rsld Output v Exp Orders	0.2681 (0.1194)
Rlsd Output v Exp Output	0.3561 (0.1316)
Rlsd Output v Exp Raw Stocks	0.1609 (0.1080)
Rlsd Output v Exp Fnshd Stocks	0.0957 (0.1137)
Rlsd Raw Stocks v Rlsd Fnshd Stocks	0.5676 (0.1104)
Rlsd Raw Stocks v Exp Orders	0.0434 (0.0890)
Rlsd Raw Stocks v Exp Output	0.0638 (0.0900)
Rlsd Raw Stocks v Exp Raw Stocks	0.1541 (0.1304)
Rlsd Raw Stocks v Exp Fnshd Stocks	0.1104 (0.1094)
Rlsd Fnshed Stocks v Exp Orders	0.0002 (0.0934)
Rlsd Fnshd Stocks v Exp Output	-0.0283 (0.1041)
Rlsd Fnshd Stocks v Exp Raw Stocks	0.0951 (0.1297)
Rlsd Fnshd Stocks v Exp Fnshd Stocks	0.1727 (0.1428)
Exp Orders v Exp Output	0.8587 (0.0448)
Exp Orders v Exp Raw Stocks	0.1725 (0.0974)
Exp Orders v Exp Fnshd Stocks	0.0322 (0.1163)
Exp Output v Exp Raw Stocks	0.2062 (0.0899)
Exp Output v Exp Fnshd Stocks	0.0903 (0.1039)
Exp Raw Stocks v Exp Fnshd Stocks	0.6197 (0.1442)

Table 2
Summary of the Bivariate Normal Hypothesis Tests

Bivariate Model	Reject Rate (%)	Mean of Likelihood Ratio Tests (Standard Deviation)
Rlsd Orders v Rlsd Output	81.9	15.5248 (8.5087)
Rlsd Orders v Rlsd Raw Stocks	33.3	6.4098 (4.2956)
Rlsd Orders v Rlsd Fnshd Stocks	40.3	7.5499 (5.4333)
Rlsd Orders v Exp Orders	83.3	15.3732 (8.5739)
Rlsd Orders v Exp Output	41.7	8.3618 (5.5803)
Rlsd Orders v Exp Raw Stocks	18.1	4.3221 (3.3495)
Rlsd Orders v Exp Fnshd Stocks	16.7	5.0888 (3.7662)
Rlsd Output v Rlsd Raw Stocks	45.8	7.8721 (4.8773)
Rlsd Output v Rlsd Fnshd Stocks	38.9	7.7053 (4.6134)
Rlsd Output v Exp Orders	36.1	7.5560 (5.5727)
Rlsd Output v Exp Output	83.3	16.4325 (8.6023)
Rlsd Output v Exp Raw Stocks	18.1	4.6667 (4.4992)
Rlsd Output v Exp Fnshd Stocks	16.7	4.8369 (3.4898)
Rlsd Raw Stocks v Rlsd Fnshd Stocks	97.2	24.2506 (11.435)
Rlsd Raw Stocks v Exp Orders	8.3	4.1430 (2.7912)
Rlsd Raw Stocks v Exp Output	15.3	4.6357 (3.8194)
Rlsd Raw Stocks v Exp Raw Stocks	98.6	39.0604 (15.4885)
Rlsd Raw Stocks v Exp Fnshd Stocks	72.2	13.0755 (8.1957)
Rlsd Fnshed Stocks v Exp Orders	12.5	4.2439 (3.4216)
Rlsd Fnshd Stocks v Exp Output	15.3	4.5706 (3.8451)
Rlsd Fnshd Stocks v Exp Raw Stocks	73.6	13.5662 (8.2641)
Rlsd Fnshd Stocks v Exp Fnshd Stocks	98.6	34.0810 (13.5889)
Exp Orders v Exp Output	87.5	18.3584 (11.3853)
Exp Orders v Exp Raw Stocks	29.2	6.4007 (5.2409)
Exp Orders v Exp Fnshd Stocks	40.3	8.2947 (6.1904)
Exp Output v Exp Raw Stocks	40.3	7.4259 (5.1196)
Exp Output v Exp Fnshd Stocks	52.8	7.9176 (5.1463)
Exp Raw Stocks v Exp Fnshd Stocks	95.8	23.7922 (11.6569)

Table 3.
The Distributions of the Correlation Coefficients for the Simulated Samples

Model	5th Percentile	95th Percentile	Number of r's Outside
Rsld Orders v Rsld Output	0.8318	0.9322	41
Rsld Orders v Exp Output	0.3393	0.6212	26
Rsld Orders v Exp Fnshd Stocks	-0.1976	0.1279	18
Rsld Output v Rsld Fnshd Stocks	-0.1922	0.2155	5
Rsld Output v Exp Raw Stocks	0.0519	0.3786	12
Rsld Raw Stocks v Rsld Fnshd Stocks	0.2127	0.5032	52

Table 4.
Analysis of Variance of the Likelihood Ratio Tests

Source	Sum of Squares	Degrees of Freedom	Mean Squares	F
Model	161,125.40	27	5967.60	126.03
Survey	19750.10	71	278.17	5.87
Residual	90769.10	1917	47.35	
Total	271644.60	2015		

$F_{.05}(27,1917)=1.4927$, $F_{.05}(71,1917)=1.2990$

Table 5.
Average Number of Instances of Under-Prediction per cell.

	Up	**No Change**	**Down**
Up	17.63	4.92	18.60
No Change	3.29	25.68	3.08
Down	20.63	2.89	21.00

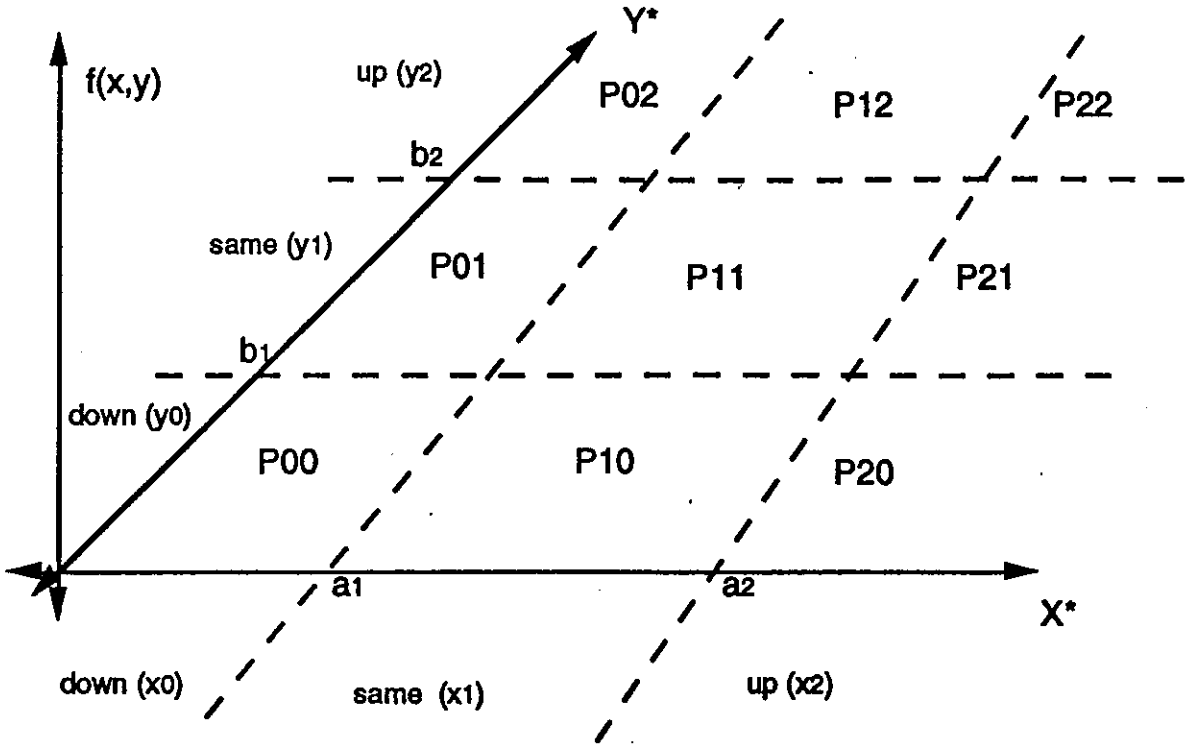
Table 6.
Summary of the Correlation Matrix for the Likelihood Ratio Tests

Interval	Frequency
$-0.2 < r_{mimj} \leq 0$	69
$0 < r_{mimj} \leq 0.2$	183
$0.2 < r_{mimj} \leq 0.4$	87
$0.4 < r_{mimj} \leq 0.6$	34
$0.6 < r_{mimj} \leq 0.8$	5

(r_{mimj} denotes the correlation coefficient between the likelihood ratio tests for bivariate model i and bivariate model j)

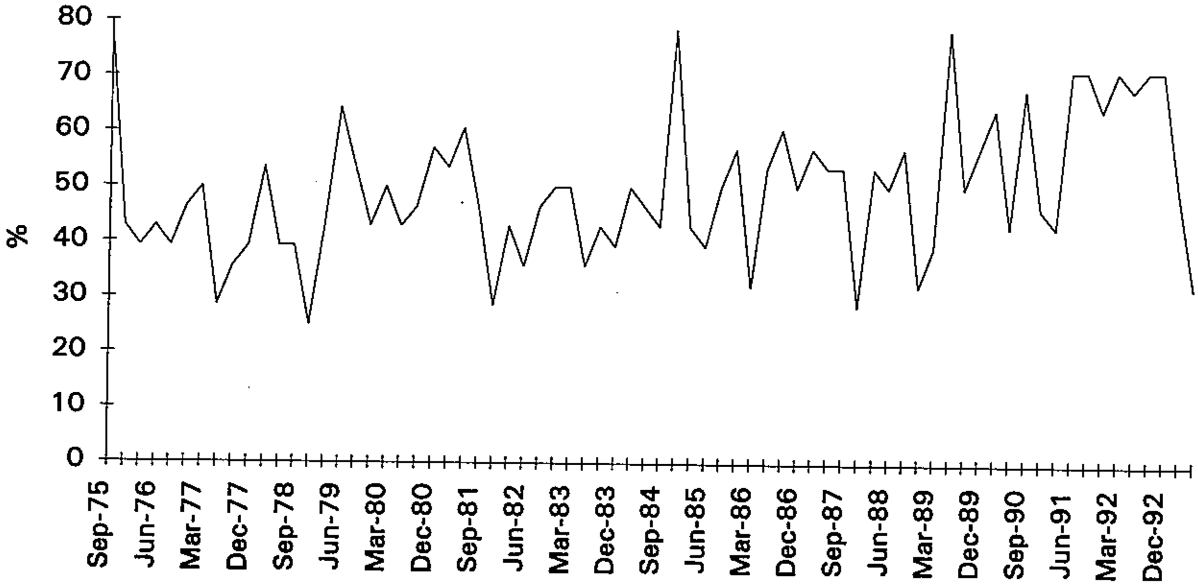
Figure 1.

The Contingency Table and the Underlying Quantitative Variables



Graph 1.

The Rejection Rate of Bivariate Normality per Survey at a 5% Level of Significance



THE GSBGM WORKING PAPER SERIES

The main purpose of this series is to reach a wide audience quickly for feedback on recently completed or in progress research. All papers are reviewed before publication.

A full catalogue with abstracts and details of other publications is available, for enquires and to be included in our distribution list, write to:

Monica Cartner
Research Programmes Co-ordinator,
GSBGM, Victoria University of Wellington,
PO Box 600, Wellington, New Zealand
Tel: (04) 495 5085; Fax: (04) 712 200

Code in bold denotes order number, eg: WP 1/91

--- Group denotes the author's academic discipline Group (note this does not necessarily define the subject matter, as staff's interests may not be confined to the subjects they teach).

1990 - 1991 titles available on request.

WP 1/92 **Money and Finance Group**
Burnell, Stephen J. and David K. Sheppard `Upgrading New Zealand's competitive advantage: a critique and some proposals.' 1992 pp 26.

New Zealand Economic Papers Vol 26(1), 1992 pp101-125.

WP 2/92 **Quantitative Studies Group**
Poot, Jacques and Jacques J. Siegers `An economic analysis of fertility and female labour force participation in New Zealand.' 1992 pp 27.

New Zealand Economic Papers Vol. 26, No. 2, December 1992, pp. 219-248

WP 3/92 **Money and Finance Group**
Lally, Martin `Project valuation using state contingent claim prices.' 1992 pp 9.

WP 4/92 **Economics Group**
Kim, Kunhong, R.A. Buckle and V.B. Hall `Key features of New Zealand Business Cycles.'

The Economic Record, Vol. 70, No 208, March 1994, pp56-72

WP 5/92 **Management Group**
McLennan, Roy `The OD Focus Group: A versatile tool for planned change.'

WP 6/92 **Information Systems Group**
Jackson, Ivan F. `Customer-oriented strategic information systems.'

`A customer-oriented IS model: from factory to "Information Factory".' In Proceedings of the Third Australian Conference on Information Systems. Woollongong, Australia. 1992 pp 406-420.

WP 7/92 **Quantitative Studies Group**
Krawczyk, Jacek B. and Boleslaw Tolwinski `A cooperative solution for the three-nation problem of exploitation of the southern blue tuna.'

> `A cooperative solution for a three-agent southern bluefin tuna management problem' In *System Modelling and Optimisation, Lecture Notes in Control and Information Sciences. No. 180.* P. Kahl, ed. Springer Verlag, 1992. pp 747-756.

WP 8/92 **Marketing Group**
Thirkell, Peter and David Stewart `A description of the advertising and direct marketing decision processes of New Zealand marketing managers.'

WP 9/92 **Quantitative Studies Group**
Jorgensen, H.P. and J.B. Krawczyk `Application of optimal control to the determination of an environmental levy.'

WP 10/92 **Economics Group**
Kim, Kunhong `A stochastic overlapping generations real business cycle model of a small open economy'

WP 11/92 **Quantitative Studies Group**
Wu, Ping X. `Testing fractionally integrated time series.'

WP 12/92 **Quantitative Studies Group**
Wu, Ping X. `A test for fractionally integrated time series.'

WP 13/92 **Quantitative Studies Group**
Nijkamp, Peter and Jacques Poot `Endogenous technological change, innovation diffusion and transitional dynamics in a nonlinear growth model.'

WP 14/92 **Management Group**
Cavana, R.Y. `Railway system in New Zealand: Case study in strategic change.'

WP 1/93 **Economics Group**
Bertram, I.G. `Economy-wide effects of a major increase in the wholesale price of electricity: New results from the JOANNA Model.'

WP 2/93 **Economics Group**
Michael Williams and Geert Reuten `The political-economic transformations in central and eastern Europe: A tragic neglect of civil society.'

WP 3/93 **Information Systems Group**
Pak Yoong `Computer-Supported Environmental Scanning: A case study.'

WP 4/93 **Management Group**
Everett Jr., Adam E., Lawrence M. Corbett and Boo Ho Rho `A comparison of Quality Improvement Practices in Korea, New Zealand and the United States of America.'

WP 5/93 **Management Group**
Campbell-Hunt, Colin, David Harper and Bob Hamilton `National renewal and strategic change - First lessons from an early-mover in deregulation.'

WP 6/93 **Management Group**
Cavana, R.Y. `Coastal shipping policy in New Zealand: economy wide implications.'
Also in: Maritime Policy and Management 1994, Vol 21, No 2, 161-172.

WP/793 **Economic History Group**
Mulcare, Tim `Gross Capital Formation and Improved Estimates of Real Gross and Net Capital Stocks to 1990 for the New Zealand Non-Market Production Sector.'

WP 8/93 **Management Group**
Knight, Russell M. and Gilbertson, David W. `Criteria used by venture capitalists: a cross country analysis.'

WP 1/94 **Economics Group**
Nana, G. Hall, V.B. and Philpott, B.P. `Trans-Tasman CGE modelling: Some illustrative results from the Joani model.'

WP 2/94 **Econometrics Group**
Krawczyk, Jacek B. 'Management of effluent discharges: A dynamic game model.'

WP 3/94 **Public Policy Group**
Boston, Jonathon 'The future of cabinet government in New Zealand: The implications of MMP for the formation, organization and operations of the cabinet.'

WP 4/94 **Economics Group**
Kim, Kunhong and Yongyil Choi 'Business cycles in Korea: Is there a stylised feature?'

WP 5/94 **Accountancy Group**
Dunmore, Paul . 'The cross-sectional distributions of financial ratios: theory, evidence, and implications.'

WP 6/94 **Economics Group**
Kunhong Kim, R.A. Buckle and V.B. Hall 'Dating New Zealand Business Cycles.'

WP 7/94 **Management Group**
Brocklesby, John 'Strategic cultural interventions in systems science - Examining the prospects for the further development of methodological complementarism.'

WP 8/94 **Economics Group**
Goodson, Matthew C. 'Irreversible investment, uncertainty and hysteresis: A New Zealand Investigation.'

WP 9/94 **Economics Group**
Boles de Boer, David and Lewis Evans 'Government department to public corporation in a deregulated economy: The economic efficiency of New Zealand telecommunications.

WP 10/94 **Economics Group**
Cassino, Vincenzo 'A study of the distributions underlying business survey reponses in New Zealand.