Jour. Ind. Soc. Ag. Statistics
43 (2), 1991 : 148—157

A Methodology for Measuring Income Stability and its Application *

Murari Singh
and
Rajendra P. Singh
International Crops Research Institute for the Semi-Arid Tropics, Patancheru, Andhra Pradesh
(Received : December, 1987)

Summary

A non-linear model with first order autocorrelated errors is used to measure income stability of rural households over years. The model has been tested with income data from the village level studies in the semi-arid tropical areas of India.

Keywords : Autocorrelation, Maximum Likelihood estimates, Generalized least square estimate, Autocovariance

Introduction

HOUSEHOLD INCOME is probably the most important single indicator of the impact of new technology on agricultural development and rural welfare. Semi-arid tropical areas are characterized by a very unstable and poor resource environment leading to yearly fluctuations in rural income. A method is suggested to estimate fluctuations in income of the rural households which largely conditions the pace of technical change and agricultural development in the semi-arid tropical areas.

Variance and coefficient of variation are commonly used for measuring fluctuations or stability over years. These parameters, however, are inadequate because they suggest only a relative phenomenon and they do not make any allowance for autocorrelation existing in the income of a household over years. This paper studies the stability of a household in generating incomes over years considering autocorrelation between incomes over years.

2. *Model*

Let Y_{it} be the income of the *i*th farmer in *t*th year (i = 1, 2, . . . , N ; t = 1, 2, . . . , T). A simple model to examine the farmer's performance in case of no interaction between farmer and years is :

$$Y_{it} = \alpha_i + \theta_t + \varepsilon_{it} \quad (1)$$

Where α_i is the *i*th farmer's performance, θ_t the *t*th year effect and ε_{it} random errors with mean zero and variance σ^2. However, if the farmer's performance to generate income does not remain constant over years, the following model can be used.

$$Y_{it} = \alpha_i + \beta_i \theta_t + \varepsilon_{it} \quad (2)$$

When $\beta_i = 1$ then (2) represents the stable performance of the *i*th farmer as in (1). Therefore, β_i estimates stability for each farmer and can be tested by $\beta_i = 1$. Model (2) has been widely used in examining the stable performance of varieties or genotypes of a crop when grown over several environments (see Yates and Cochran [6], Finlay and Wilkinson [3], Perkins and Jinks [5], Eberhart and Russell [2], Digby [1] among others). Unlike the methods of Eberhart and Russell [2], which uses as the mean of observations in the *t*th environment, we shall following [1], estimate from the data using model (2) subject to the constraint $\Sigma_t \theta_t = 0$. In the genotypes x environment interaction study, the errors ε_{it} are independent, but in income study over several years errors are dependent and we assume that they follow a first order autocorrelation.

$$\varepsilon_{it} = \rho\, \varepsilon_{it-1} + \eta_{it} \quad (3)$$

$|\rho| < 1$, is the autocorrelation parameter, η_{it} are independent normally distributed random variables with mean zero and variance σ_η^2.

Note that $\sigma^2 = \sigma_\eta^2 / (1 - \rho^2)$ for the variance of ε_{it}.

Section 3 presents the estimation of the parameters of the non- linear model (2) with error structure (3). This model has been applied to the income series of the farmers of three different regions in the SAT India in section 4.

3. *Estimation of Parameters*

Maximum likelihood estimates of parameters α_i, β_i, θ_t ($i = 1, \ldots, N$; $t = 1, \ldots, T$) in (2), which are the same as the generalized least squares, can be obtained by minimizing the likelihood function.

$$L = \prod_{i=1}^{N} \zeta_0 \exp\left[- (\underline{Y}_i - \alpha_i \underline{J}_T - \beta_i \underline{\theta})' \, \Omega^{-1} (\underline{Y}_i - \alpha_i \underline{J}_T - \beta_i \underline{\theta}) / (2\sigma^2) \right] \quad (4)$$

Subject to $\underline{\theta}' \underline{J}_T = 0$

where

$$\zeta_0 = 1 / \left[(2\pi)^{T/2} \mid \Omega \mid^{1/2}\right]; \ \mid \Omega \mid = (1-\rho^2)^{T-1}$$

$\underline{Y}_i = (Y_{i1}\, Y_{i2} \ldots Y_{iT})'$

$\underline{\theta} = (\theta_1\, \theta_2 \ldots \theta_{iT})'$ and $\underline{J}_T$ is a T-component column vector of unities.

The dispersion matrix of $\underline{Y}_i$ is $\sigma^2\Omega$, where

$$\Omega = (\rho \mid t-t' \mid), \ t, \ t' = 1, 2, \ldots, T.$$

Since correlation $(Y_{it}, Y_{it'})$ = correlation $(\varepsilon_{it}, \varepsilon_{it'}) = \rho \mid t-t' \mid$ for $i = i'$

$= 0$ otherwise

The inverse of the matrix Ω is

$$\Omega^{-1} = [1/(1-\rho^2)] \begin{bmatrix} 1 & -\rho & 0 & \ldots & 0 \\ -\rho & 1+\rho^2 & -\rho & \ldots & 0 \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \cdot & \ldots & -\rho \\ 0 & \cdot & \cdot & -\rho & 1 \end{bmatrix}$$

$$\Omega^{-1} = [1/(1-\rho^2)]\ C \quad (\text{say}) \tag{5}$$

The maximization of (4) with respect to α_i, β_i, θ_t's is equivalent to minimizing;

$$Q = \sum_i (\underline{Y}_i - \alpha_i \underline{J} - \beta_i \underline{\theta})' \ \Omega^{-1} (Y_i - \alpha_i \underline{J} - \beta_i \underline{\theta}) + \lambda \underline{\theta}' \underline{J}_T \tag{6}$$

or

$$Q = \sum_{i=1} \left[\varepsilon_{i1}^2 + \varepsilon_{ir}^2 + (1+\rho^2) \sum_{t=2}^{T-1} \varepsilon_{it}^2 - 2\rho \sum_{t=1}^{T-1} \varepsilon_{it}\, \varepsilon_{it+1}\right] / (1-\rho^2) + \lambda \sum_{t=1}^{T} \theta_t \tag{7}$$

Where λ is the Lagrangian multiplier and

$$\varepsilon_{it} = Y_{it} - \alpha_i - \beta_i\ \theta_t$$

An alternative way of estimation of parameter could be to note

$$Y_{it} - \rho\, Y_{it-1} = \alpha_i(1-\rho) + \beta_i(\theta_t - \rho\theta_{t-1}) + \eta_{it} \tag{8}$$

Which gives errors η_{it} with constant variance and minimizing

$$\sum_{i=1}^{N}\sum_{t=2}^{T} [\, Y_{it} - \rho\, Y_{it-1} - \alpha_i\,(1-\rho) - \beta_i\,(\theta_t - \rho\,\theta_{t-1})\,]^2$$

with respect to α_i, β_i and ρ for given θ_t's using Cochrane and Orcutt iterative process as described in Johnston [4]. We obtained the equation for α_i, β_i for given θ_t but the equations of estimation of θ_t do not permit the unique estimate of θ_t subject to the above constraint as the matrix in the least square solution for θ_t becomes singular. Therefore, we shall discuss here the estimation procedure for the generalized least squares.

For estimation of parameters α_i, β_i we take (initialize) parameters ρ and θ_t and observe that (2) can be written as

$$\underline{Y} = \mathbf{X}\,\underline{\delta} + \underline{\varepsilon} \tag{9}$$

where

$$\underline{Y} = (\,\underline{Y}'_1\ \underline{Y}'_2\ \ldots\ \underline{Y}'_N)$$

$$\delta' = (\delta'_1\ \delta'_2\ \ldots\ \delta'_N)\,, \quad \underline{\delta_i} = (\,\alpha_i\,,\ \beta_i)$$

$\mathbf{X} = \mathbf{I}_N\ \Theta\ \mathbf{Z}$; $\mathbf{Z} = (\,\underline{J}\,, \underline{\theta}\,)$; Θ stands for Kronecker product.

The dispersion of $\underline{Y}$ is

$$\mathbf{D}\,(\,\underline{Y}) = (\,\mathbf{I}_N\,\Theta\,\Omega\,)\,\sigma^2 = \psi$$

The generalized least square estimator of $\underline{\delta}$ is given by

$$\hat{\underline{\delta}} = (\,\mathbf{X}'\,\psi^{-1}\mathbf{X}\,)^{-1}\ \mathbf{X}'\,\psi^{-1}\,\underline{Y} \tag{10}$$

and the dispersion matrix of $\hat{\delta}$ as

$$D(\,\hat{\underline{\delta}}\,) = (\,\mathbf{X}'\ \psi^{-1}\ \mathbf{X}\,)^{-1} \tag{11}$$

After simplification we obtain

$$(\,\mathbf{X}'\ \psi^{-1}\ \mathbf{X}\,)^{-1} = [\,\mathbf{I}_N\ \Theta\ \ (\mathbf{Z}'\ \Omega^{-1}\ \mathbf{Z}\,)^{-1}\,]\,\sigma^2$$

$$(X' \ \psi^{-1} X)^{-1} \ X' \ \psi^{-1} \underline{Y} = \begin{bmatrix} (Z' \ \Omega^{-1} Z)^{-1} & Z' \ \Omega^{-1} \underline{Y}_1 \\ (Z' \ \Omega^{-1} Z)^{-1} & Z' \ \Omega^{-1} \underline{Y}_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ (Z \ \Omega^{-1} Z)^{-1} & Z' \ \Omega^{-1} \underline{Y}_N \end{bmatrix}$$

$$(Z' \ \Omega^{-1} Z) = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} / (1-\rho^2) = A/(1-\rho^2) \text{ , say}$$

where $a_{11} = (1-\rho) \ [T(1-\rho) + 2\rho]$

$$a_{12} = \rho(1-\rho) \ (\theta_1 + \theta_T)$$

$$a_{22} = \sum_{t=1}^{T} \theta_t^2 + \rho^2 \sum_{t=2}^{T-1} \theta_t^2 - 2\rho \sum_{t=1}^{T-1} \theta_t \theta_{t+1}$$

$$Z' \ \Omega^{-1} \underline{Y}_i = [1/(1-\rho^2)] \begin{bmatrix} (1-\rho)(Y_{i1} + Y_{iT}) + (1-\rho)^2 \sum_{t=2}^{T-1} Y_{it} \\ \sum_{t=1}^{T} \theta_t Y_{it} + \rho^2 \sum_{t=2}^{T-1} \theta_t Y_{it} - \rho \sum_{t=1}^{T-1} (Y_{it} \theta_{t+1} + Y_{it+1} \theta_t) \end{bmatrix}$$

$$= [1/(1-\rho^2)] \ \underline{b}_i \text{ (say).}$$

Thus, the estimate of $\underline{\delta}$ is :

$$\hat{\underline{\delta}} = (\hat{\underline{\delta}}'_1 \ \hat{\underline{\delta}}'_2 \ \ldots \ \hat{\underline{\delta}}'_N)'$$

where

$$\hat{\underline{\delta}}_i = A^{-1} \underline{b}_i \tag{12}$$

and dispersion of $\hat{\underline{\delta}}_i$

$$D(\hat{\underline{\delta}}_i) = (Z \ \ \Omega^{-1} Z)^{-1} \sigma^2 \text{ for all i} \tag{13}$$

$$= (1-\rho^2) A^{-1} \sigma^2$$

Now consider the estimation of the parameter $\underline{\theta}$ for given values (estimates) of $\underline{\alpha}$ and $\underline{\beta}$. The model (2) can be expressed as :

$$\underline{Y}^* = Z^* \,\underline{\theta} + \underline{\varepsilon} \tag{14}$$

with $\underline{\alpha} = (\alpha_1 \ \alpha_2 \ \ldots \ \alpha_N)'$; $\underline{\beta} = (\beta_1 \ \beta_2 \ \ldots \ \beta_N)'$;

$$\underline{Y}^* = \underline{Y} - \underline{\alpha} \ \Theta \ \underline{J}_T \ ; \ Z^* = \underline{\beta} \, \Theta \ I_T$$

where I_T is identity matrix of order T.

The generalized least square estimate of $\underline{\theta}$ subject to constraint $\underline{\theta}' \underline{J}_T = 0$ is obtained by minimizing

$$\Phi = (\underline{Y}^* - Z^* \underline{\theta})' \ W^{-1} (\underline{Y}^* - Z^* \ \underline{\theta}) + 2\lambda \, \underline{\theta}' \underline{J}_T$$

where $W = I_N \Theta \ \Omega \ \sigma^2$

This yields (after differentiating Φ with respect to θ and equating to zero and simplification)

$$\hat{\underline{\theta}} = (Z^{*\prime} \ W^{-1} Z^*)^{-1} \ (Z^* \ W^{-1} \underline{Y}^* - \lambda \, \underline{J}_T)$$

Constraint $\underline{\theta}' \underline{J}_T = 0$ results in

$$\lambda = \underline{J}'_T (Z^{*\prime} \ W^{-1} \ Z^*)^{-1} Z^{*\prime} \ W^{-1} \underline{Y}^* / [\underline{J}'_T (Z^{*\prime} \ W^{-1} Z^*)^{-1} \underline{J}_T]$$

We get after algebraic simplification, the following

$$(Z^{*\prime} \ W^{-1} Z)^{-1} = (\underline{\beta}' \underline{\beta})^{-1} \sigma^2 \ \Omega$$

$$\underline{J}'_T (Z^* \ W^{-1} Z^*)^{-1} \ \underline{J}_T = (\underline{\beta}' \underline{\beta})^{-1} \underline{J}'_T \Omega \ \underline{J}_T \sigma^2 = (\underline{\beta}' \underline{\beta})^{-1} S \sigma^2 \tag{i}$$

where $S = (1+\rho)\, T/(1-\rho) - 2\rho\, (1-\rho^T)/(1-\rho)^2$

$$(Z^* W^{-1} Z^*)^{-1} Z^{*\prime} W^{-1} \underline{Y}^* = (\underline{\beta}' \underline{\beta})^{-1} \sum_{i=1}^{N} \beta_i (\underline{Y}_i - \alpha_i \underline{J}_T) \tag{ii}$$

$$\lambda = T \sum_{i=1}^{N} \beta_i (\overline{Y}_i - \alpha_i)/(S\sigma^2) \quad \text{where} \quad \overline{Y}_i = \sum_t Y_{it}/T \tag{iii}$$

$$\Omega\ \underline{J}_T = [1/(1-\rho)]\,(1-\rho^T,\ 1+\rho-\rho^2-\rho^{T-1},\ldots,\ 1+\rho-\rho^t-\rho^{T-t+1},\ldots,\ 1-\rho^T)' \qquad \text{(iv)}$$

Thus the estimate of $\underline{\theta}$ simplifies to

$$\hat{\underline{\theta}} = (\underline{\beta}'\,\underline{\beta})^{-1}\,[\sum_{i=1} \beta_i\,(\underline{Y}_i - \alpha_i\,\underline{J}_T) - (T \sum_{i} \beta_i\,(\bar{Y}_i - \alpha_i)/S)\ \Omega\ \underline{J}_T]$$

and its dispersion matrix

$$D(\hat{\underline{\theta}}) = (\,\underline{\beta}'\,\underline{\beta})^{-1}\ \sigma^2\ \Omega$$

The estimate of σ^2 can be obtained from weighted sum of squares of residual (Johnston, [4], page 259).

$$\hat{\sigma}^2 = (\,\underline{Y} - \underline{\alpha} \ominus \underline{J}_T - \underline{\beta} \ominus \underline{\theta})'\ \Omega^{-1}(\,\underline{Y} - \underline{\alpha} \ominus \underline{J}_T - \underline{\beta} \ominus \underline{\theta})/(NT - 2N - T + 1)$$

The parameters ρ can be estimated using first order autocovariance between residuals and is same as obtained by minimizing the numerator in (7).

$$\hat{\rho} = \sum_{i=1}^{N} \sum_{t=2}^{T} \hat{\varepsilon}_{it-1}\ \hat{\varepsilon}_{it} \Big/ \sum_{i=1}^{N} \sum_{t=2}^{T} \hat{\varepsilon}^2_{it-1}$$

where

$$\hat{\varepsilon}_{it} = Y_{it} - \hat{\alpha}_i - \hat{\beta}_i\ \hat{\theta}_t$$

It can be noted that $\hat{\rho}$ is a consistent estimate of ρ i.e. plim $\hat{\rho} = \rho$.
Test of significance of ρ can be done using Durbin-Watson statistics.

$$d = \sum_{i}^{N} \sum_{t=2}^{T} (\hat{\varepsilon}_{it} - \hat{\varepsilon}_{it-1})^2 \Big/ \sum_{1}^{N} \sum_{t=1}^{T} \varepsilon^2_{it}$$

4. *Results and Discussion*

The model (2) discussed in the section 2 was used to measure the income stability of rural households in India's semi-arid tropics (SAT). Information used in the study is based on the data collected for nine cropping years from 104 households in three villages namely Aurapalle, Shirapur and Kanzara respectively contrasting agroclimatic environments. The results of the model are presented in Table 1 and 2.

Table 1 presents the estimated values of the parameters α_i (mean per person

Table 1. Estimated values of α's and β's of per person income and their standard errors in three villages of India's SAT.

Household number	Village					
	Aurepalle		Shirapur		Kanzara	
i	$\hat{\alpha}_i$	$\hat{\beta}_i$	$\hat{\alpha}_i$	$\hat{\beta}_i$	$\hat{\alpha}_i$	$\hat{\beta}_i$
1	193	0.09	440	–0.32*	532	1.40
2	617	–1.03*	273	0.64	474	0.30
3	269	0.38	515	1.59	514	0.82
4	240	0.01	428	0.20	374	0.44
5	247	0.87	270	0.13	649	1.06
6	180	–0.05	317	0.15	789	0.77
7	244	0.28	985	–1.58*	1204	2.33*
8	328	0.20	379	0.53	356	0.37
9	832	–0.14	484	–1.37*	496	1.19
10	782	0.67	743	0.55	839	1.25
11	237	0.01	522	1.19	585	–0.75*
12	418	0.72	326	0.07	666	2.10
13	245	–0.03	319	–0.52*	703	1.17
14	316	0.08	1856	3.40*	701	1.25
15	264	0.18	580	0.38	612	0.87
16	210	0.05	621	1.95	759	–0.56*
17	313	–0.15	339	0.42	656	–0.14*
18	707	0.87	642	1.58	767	0.15
19	556	0.29	349	0.72	804	1.92
20	776	1.20	324	0.10	624	0.66
21	467	–0.80*	387	0.05	684	–0.23*
22	942	0.42	306	–0.42*	781	2.37*
23	451	0.49	678	2.54*	501	0.71
24	354	0.10	502	0.91	418	0.32
25	232	–0.40	540	0.62	420	1.00
26	888	–1.02*	263	0.59	396	–0.04
27	690	0.28	901	–0.06	532	–0.10*
28	2223	9.32*	700	1.64	1523	2.31
29	820	0.77	475	1.58	338	–0.53
30	758	–0.07	634	1.18	311	–0.02
31	1366	–1.17*	906	–1.09*	3574	0.39
32	2723	2.15	615	–0.08	2044	1.62
33	1139	–0.03	1415	7.97*	901	–0.34*
34	1775	–0.87*	—	—	789	1.60
35	676	1.35	—	—	1348	3.75*
36	—	—	—	—	1877	4.07*
SE	118	0.83	61	0.65	76	0.55

* Indicates significantly different from unity at 0.05 level of significance.

Table 2. Proportion of stable households, year effects and autocorrelation in per person income in three villages of India's SAT

Particulars	Village		
	Aurepalle	Shirapur	Kanzara
Proportion of stable (stagnant) households	0.86	0.76	0.64
Cropping year effect ($\hat{\theta}$)			
1975/76	–180	–105	–163
1976/77	–96	–128	–107
1977/78	45	–142	9
1978/79	11	–28	–69
1979/80	219	124	–97
1980/81	153	26	–79
1981/82	17	98	70
1982/83	–78	92	161
1983/84	–91	63	275
SE	27	22	25
	277	233	220
Autocorrelation $\hat{\rho}$	0.27**	–0.26	0.04**
D.W. STATISTICS	1.13	1.81	1.55

** indicates that autocorrelation is significant at 0.01 level of significance.

real income in rupees) and β_i (stability measure) with their standard errors for individual households. Table 2 provides the summary of the results indicating proportion of stable households, autocorrelations in the income series over years and cropping years effects etc.

The information in Table 2 indicates that about 64 to 86% of the rural households were stable in these regions. In other words, they did not experience significant change in their income during nine years. They can also be called stagnant households. This shows that in assured rainfall environment more farmers were stable than in undependable, low rainfall areas. In time series analysis, normally one expects positive autocorrelation indicating that current year's income is strongly influenced by previous year's income. But in the semi-arid tropics where environment in relatively unstable the theory of 'success breeds success' does not hold true in all the villages. However, Table 2 does indicate that in Shirapur with relatively undependable rainfall income over years is negatively correlated but not significant. This suggests that in harsh environment like this previous year's income does not necessarily influence current year's income and the effects of bad climate continue to adversely affect the coming years. In contrast the situation in the two other villages are different. In Aurepalle and Kanzara the incomes of the households are positively and significantly correlated over years indicating thereby strong effect of previous year's income on current year's income.

It was also found that mean level of income and stability coefficients are positively correlated in Aurepalle and Kanzara but not in Shirapur. This suggests that households with higher level of income are able to adjust climatic variability and reduce their income variability over years.

REFERENCES

[1] Digby, P.G.N., 1979. Modified Joint Regression Analysis for Incomplete Variety X Environment Data. *Journal of Agricultural Science,* **93**, 81-86.

[2] Eberhart, S.A. and Russell, W.A., 1966. Stability Parameters for Comparing Varieties.. *Crop Science,* **6**, 36-40.

[3] Finlay, K.W. and Wilkinson, G.N., 1963. The Analysis of Adoption in a Plant-Breeding Programme. *Australian Journal of Agricultural Science,* **14**, 742-754.

[4] Johnston, J., 1972. *Econometrics Methods, Second Edition,* McGraw–Hill Kogakusha, Ltd.

[5] Perkins, J.M., and Jinks, J.L., 1973. The Assessment and Specificity of Environmental and Genotype - Environmental Components of Variability. *Heredity,* **30**, 111-126.

[6] Yates, F. and Cochran, W.G., 1938. The Analysis of Groups of Experiments. *Journal of Agricultural Science,* **28**, 556-580.