IOSR Journal of Economics and Finance (IOSR-JEF)
e-ISSN: 2321-5933, p-ISSN: 2321-5925.Volume 11, Issue 2 Ser. VII (Mar – Apr 2020), PP 16-20
www.iosrjournals.org

Time varying Estimates of Production Function Model

Dr. A.V.S. Raghavaiah[1] Prof M. Subbarayudu[2]
[1]*Associate Professor, Department of Statistics, S.K.S.C. Degree College, Proddatur, Kadapa Dist, A.P*
[2]*Professor, Department of Statistics, S.V. University, Tirupathi, Chittor Dist. A.P*

Abstract: *In production function technology, f(X) where Y is output called endogenous variable, X is data matrix on a set of exogenous variables, f is the function form. In this article an attempt is made to develop an alternative method to find the estimates of time varying parameters of a given model. Validity of the proposed method can be compared with the other existing estimation methods using concept of minimum error mean sum of squares.*

Key words: *Endogenous variables, exogenous variables, Regression analysis, Time varying parameters, Error mean sum of squares.*

Date of Submission: 15-04-2020 Date of Acceptance: 30-04-2020

I. Introduction

Regression analysis is a statistical technique for investigating and modeling the relationship between variables. Applications of regression are numerous and occur in almost every field, including engineering, the physical and chemical sciences, production technology, management, life and biological sciences and social sciences.

Consider a general linear regression model Y=Xβ+€. We know the ordinary last squares estimate of the parameter β which is given by $\hat{\beta}=(X^1X)^{-1}X^1Y$. In this paper our aim is to find the time varying estimates of the parameters of the models. To achieve this we are given a methodology in the next section.

The problem of statistical inference in the multiple regression models with random coefficients has been examined by Forchlich.B.R [1] and Swamy P.A.V.B. and T.S. Mehtha [2]. Several models have been suggested in the literature to deal with the cross section studies there can be heterogeneity in the parameters across different units where as in time series studies there can be variation over time in the parameters. In view of these Manohar Rao M.J and Dilip M, Nachane [3] has given a computational procedure by taking numerical examples for dealing with the problem of varying parameters (in a time series context) via the theory of optimal control. In this paper we consider the estimation of a model with time varying structure.

The parameters of the model are assumed to be subject to permanent and transitory changes over time. Estimation methods are developed, and the asymptotic properties of the estimates are derived by Thomas F. Cooley and Edward C. Prescott [4].

This paper considers the possibility of tracking ‘cohorts’ through such data. A ‘cohort’ is defined as a group with fixed membership, individuals of which can be identified as they show up in the surveys. The most obvious example is an age cohort. e.g. all males born between 1945 and 1950, but there are other possibilities (Korean war veterans or founding members of the Econometric Society). Consider any economic relationship of interest that is linear in parameters (but not necessarily in variables). Corresponding to individual behavior of them will exist a cohort version of the relationship of the same form. But with cohort means replacing individual observations. If there are additive individual fixed effects, there will be corresponding additive cohort fixed effects. Further, the sample cohort means from the surveys are consistent but error-ridden estimates of the true cohort means. Hence provided errors-in-variables techniques are used (and error variances and covariance can be estimated from the surveys), the sample cohort means can be used as panel data for estimating the relationship. Such data are immune to attrition bias and can be extended for long time periods. There is also evidence to suggest that the errors in variables problems may be just as severe for genuine panel data; in the created panels considered here, the problem is controllable. The paper discusses appropriate errors in variables estimators with and without fixed effects by Angus Deaton [5].

In this article a new approach based on the reasoning of averaging individual model gradients is proposed. A real data set of coffee sales in relation to shelf space is used to examine the performance of simple averaging method (SAM) and OLS. Model performance of OLS and SAM is assessed using the root mean square error, mean absolute error and graphical methods. It is shown that both methods yield comparable results based on a one dimensional regression problem. In this case, SAM is recommended since it has less stringent conditions for application as opposed to the OLS method by Kikawa R. Cliff, Kalema M. Billy [6].

The model $Y_{it} = \Sigma_k (\beta_k + \delta_{ik} + \gamma_{tk})x_{ikt} + \epsilon_{it}$ with β_k and γ_{tk} random is considered as a means of pooling the time series of a cross-section sample. The model is placed in a mixed analysis of variance framework. Relationships between various estimation criteria are derived and their asymptotic properties compared. Some implementation problems are also discussed by Cheng Hsiao [7].

In this paper an attempt is made to estimate a regression equation using a time series of cross sections. It is assumed that the coefficient vector is distributed across units with the same mean and the same variance-covariance matrix. The distribution of the coefficient vector is assumed to be invariant to translations along the time axis. A consistent and an asymptotically efficient estimator for the mean vector and an unbiased estimator for the variance-covariance matrix of the coefficient vector have been suggested. Some asymptotic procedures for testing linear hypotheses on the means and the variances of coefficients have been described. Finally, the estimation procedure is applied in the analysis of annual investment data, 1935-54, for eleven firms by P. A. V. B. Swamy [8].

In attempting to estimate the parameters of a linear regression system obeying two separate regimes; it is necessary first to estimate the position of the point in time at which the switch from one regime to the other occurred. The suggested maximum likelihood estimating procedure is based upon a direct examination of the likelihood function. An asymptotic and a small-sample test are suggested for testing the hypothesis that no switch occurred against the single alternative that one switch took place. The procedure is illustrated with a sampling experiment in which the true switching point is correctly estimated by Richard E. Quandt [9].

II. Methodology:

1. Let us consider the model $Y=X\beta+\in$, where Y is dependent vector of order (n x 1); X is (n x k) data matrix of independent variables; β is (k x 1) vector of parameters and $\in$ is (n x 1) vector of disturbances; n is the size of the given data set; K is the number of parameters involved in the model; n > k.
2. Divide the given set of data as s=(n-k) subsets with equal number of data points (k+1) each, such that the first subset consists of the first (k+1) data points of the given data, the second subset consists of the second (k+1) data points by omitting the first data point, the third subset consists of the third (k+1) data points by omitting the first two data points and at the end the s^{th} subset consists of the s^{th} (k+1) data points by omitting the first (s-1=n-k-1) data points of the given set of data.
3. Fit the regression model for each subset of the given data and obtain OLS estimate of vector of regression coefficients $\hat{\beta}_i$, i=1, 2, ….
4. Place the elements of $\hat{\beta}_i$ corresponding to each of the (k+1) time points in i^{th} subset, i = 1, 2, ……..s. in a tabular form as shown in table-1.
5. Then we have the following system of vectors of regression coefficients corresponding to each data point of the given system of data.

The data points (1, n), (2, n-1),… (k, n-k+1) has regression coefficients respectively as 1, 2, ….. k. And the remaining middle data points i.e., from (k+1) to (n-k) has (k+1) regression coefficients each.

6. Now the estimates of time varying parameters of each data point can be considered as the mean of the corresponding regression coefficients of that time point.
7. Validity of the proposed method is tested using the concept of error mean sum of squares (ems) as follows, where

$$ems = \sum_{i=1}^{n} ei^2 / n, \ \ ei = Y_i - \hat{Y};$$

i) Find $\hat{Y}_i$ i=1,2, …. n corresponding to proposed method (using time varying regression coefficients) as well as usual OLS method,

ii) Find ems for both the methods.

iii) If ems of the proposed method is lesser than usual OLS method we conclude that the proposed method is more consistent than usual OLS method.

For example, let us consider the model $Y=a+bX+\in$, a time series data with n=20 time points. Hence the number of parameters in the model is k=2; the number of subsets of given data is n-k=18, each having k+1=3 data points. Following table explains the method of calculating the estimates ($\hat{T}_i$) of time varying parameters for each time point in the given set of data.

Table-1:

Time Point s n	**OLS Estimates of Regression Coefficients of Subsets**																		$\hat{T}_i$
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	
1	a_1																		$\hat{T}_i$
2	a_1	a_2																	$\hat{T}_2$
3	a_1	a_2	a_3																$\hat{T}_3$
4		a_2	a_3	a_4															$\hat{T}_4$
5			a_3	a_4	a_5														$\hat{T}_5$
6				a_4	a_5	a_6													$\hat{T}_6$
7					a_5	a_6	a_7												$\hat{T}_7$
8						a_6	a_7	a_8											$\hat{T}_8$
9							a_7	a_8	a_9										$\hat{T}_9$
10								a_8	a_9	a_{10}									$\hat{T}_{10}$
11									a_9	a_{10}	a_{11}								$\hat{T}_{11}$
12										a_{10}	a_{11}	a_{12}							$\hat{T}_{12}$
13											a_{11}	a_{12}	a_{13}						$\hat{T}_{13}$
14												a_{12}	a_{13}	a_{14}					$\hat{T}_{14}$
15													a_{13}	a_{14}	a_{15}				$\hat{T}_{15}$
16														a_{14}	a_{15}	a_{16}			$\hat{T}_{16}$
17															a_{15}	a_{16}	a_{17}		$\hat{T}_{17}$
18																a_{16}	a_{17}	a_{18}	$\hat{T}_{18}$
19																	a_{17}	a_{18}	$\hat{T}_{19}$
20																		a_{18}	$\hat{T}_{20}$

Where $\hat{T}_1 = a_1$

$\hat{T}_2 = (a_1 + a_2) / 2$

$\hat{T}_3 = (a_1 + a_2 + a_3) / 3$

$\hat{T}_4 = (a_2 + a_3 + a_4) / 3$

$\hat{T}_5 = (a_3 + a_4 + a_5) / 3$

⋮

$\hat{T}_{18}$ = $(a_{16} + a_{17} + a_{18}) / 3$

$\hat{T}_{19}$ = $(a_{17} + a_{18}) / 2$

$\hat{T}_{20}$ = (a_{18})

$a_1, a_2, \ldots\ldots\ldots, a_{18}$ are the regression coefficients of $1^{st}, 2^{nd}, \ldots\ldots, 18^{th}$ subsets of the given set of data. **Note that, we have to construct this type of table for each coefficient of the model to calculate the estimate of time varying parameter.**

Later we can find $\hat{Y}_i$ and hence error mean sum of square for the proposed method as well as usual OLS method for comparison.

Tukey's Multiple Comparison Test

To test the equality of OLS estimates (a_i, i=1, 2, ……, 18) between the subsets, we may use Tukey's multkple comparison test. Equality of estimates ($\hat{T}_i$) of time varying parameters is also tested using the same test.

Analysis of variance (ANOVA) is useful to test the equality of the population means. If the null hypothesis is accepted then the problem is essentially finished. If the null hypothesis is rejected the question still remains as to which means differ. We can further analyse the differences with a multiple comparison test like student, Neuman-Kaul's test, Duncan's test, Scheffe's test or Tukey's test.

Procedure of Tuke's Multiple Comparison Test

For a specified level of significance α, calculate $\omega = q_\alpha (k, \nu) \sqrt{MSE / n}$.

Where

n	=	Number of observations in each sample
MSE	=	Mean square error from ANOVA table
K	=	Number of different population means
ν	=	Degrees of freedom associated with MSE

$q_\alpha (k, \nu)$ = Upper tail probability of studentized range found in tables of percentage points of studentized range - q(k, α).

To conduct Tukey's procedure completes the following steps:

1. Rank the sample means from highest to lowest and arrange them in the same order.
2. Compute the difference between the largest and the smallest sample means $\left(\overline{Y}_{l\arg est} - \overline{Y}_{smalest}\right)$. If the difference exceeds ω the corresponding means are declared significantly different. Proceed to compute the difference between the largest and the next smallest sample mean, $\left(\overline{Y}_{l\arg est} - \overline{Y}_{2nd\ \ smalest}\right)$. As previously indicated if the difference exceeds ω then declare the corresponding means are different. Continue to make comparisons with the largest sample mean $\left(\overline{Y}_{l\arg est} - \overline{Y}_{3rd\ \ smalest}\right)$ and so on until a difference fails to exceedω. Once a difference between two sample means is less thanω, the difference between the corresponding means and all other means are declared not significant.
3. Next make comparisons with the next largest sample mean $\left(\overline{Y}_{2nd\ \ l\arg est} - \overline{Y}_{smalest}\right)$ and so on; using the same procedure in step 2, continue till all the possible comparisons are made.
4. Summarize the results by drawing a line under the means that are declared not significant.

III. Conclusion:

One of the basic assumptions of the general linear regression model is that the parameters are consistent over all the observations. It has often been suggested that this may not be valid assumption to take. In cross section studies, there can be heterogeneity in the parameters across different units, where as in time series, there can be variation over time in the parameters. On the basis of this idea, the present method of computation of time varying estimates of production function model is developed.

References:

[1]. Forchlich, B.R. (1973), "Some Estimations for RCR model", JASA, vol-6, pp: 329-335.

[2]. Swamy, P.A.V.B. and T.S. Mehtha (1977): "Estimation of linear model with varying coefficients", Journal of American Statistical Association vol.72, pp: 890-898.
[3]. Manohar Rao, M.J and Dilipm M.Nachane (1988): "Varying parameters models- An optimal control formation", Journal of quantitative Economics vol.4, pp-59-79.
[4]. Thomas F. Cooley and Edward C. Prescott, "Estimation in the Presence of Stochastic Parameter Variation", Econometrica, Vol. 44, No. 1 (Jan., 1976), Pp. 167-184.
[5]. Angus Deaton, "Panel Data From Time Series of Cross-Sections", Journal Of Econometrics 30 (1985) 109-126. North-Holland
[6]. Kikawa R. Cliff, Kalema M. Billy, "Estimation of The Parameters of a Linear Regression System Using The Simple Averaging Method", Global Journal of Pure and Applied Mathematics. ISSN 0973-1768 Volume 13, No: 11 (2017), Pp. 7749–7758.
[7]. Cheng Hsiao, "Some Estimation Methods for a Random Coefficient Model", Econometrica, Vol. 43, No. 2 (March, 1975) Pp: 305-325.
[8]. P. A. V. B. Swamy, "Efficient Inference In A Random Coefficient Regression Model", Econometrica, Vol. 38, No. 2. (Mar., 1970), Pp. 311-323.
[9]. Richard E. Quandt, "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes", Taylor & Francis, 03 October 2014, Pp: 873-880.

Dr. A.V.S. Raghavaiah,etal. "Time varying Estimates of Production Function Model." *IOSR Journal of Economics and Finance (IOSR-JEF)*, 11(2), 2020, pp. 16-20.