Measurement of States' Taxable Capacities
Ninth Finance Commission's Methodology

Dipankor Coondoo
Robin Mukherjee

The efforts made by the Ninth Finance Commission, as required by its terms of reference, to develop a methodology for estimating the taxable capacity of states should be welcomed. However, there are a number of problems, many of them technical in nature, with the methodology adopted by the commission and it should therefore be subjected to critical examination.

THE First Report of the Ninth Finance Commission (NFC) has already evoked considerable discussions in government circles, particularly at the state level, as well as among economists. While the report has been complimented for a number of unique features, criticisms have also been made that the scheme recommended for 1989-90 is regressive and the "approach to its main award for 1990-95 and its target for phasing out revenue deficits at all levels by 1994-95 are bound to adversely affect equity, welfare and growth" [Guhan, 1989].

Perhaps the most significant departure of the present report from the earlier ones was prompted by the terms of reference of the commission requiring it to adopt a normative approach for assessing the receipts and expenditures on the revenue account of the states and the centre. There should not be any dispute that *ad hocism* should be· dispensed with as far as possible and therefore the attempts made by the commission, as required by the terms of reference, to develop a methodology for estimating the taxable capacity of the states should be welcome. It seems, however, that there are a number of problems, many of which are technical in nature, with the methodology and it should therefore be subjected to critical examination. The purpose of this note is to draw the attention of economists and econometricians to this important feature of the report and to put forward some observations on it.

THE METHODOLOGY

The methodology used by the Finance Commission for measuring the taxable capacity for 14 major states for the year 1989-90 is the regression technique. It recognises two types of factors which are likely to affect the level of per capita tax revenue of a state, viz, the factors relating to taxable capacity and those relating to tax effort, and implicitly assumes that these two types of factors are mutually independent. The observed data on per capita tax revenue are therefore regressed on a set of explanatory variables (like the per capita SDP, the proportion of non-primary sectoral SDP in SDP and the Lorenz ratio of the consumption distribution of the state) to obtain an estimated taxable capacity function for the individual states, and the discrepancy ·between the observed value of per capita tax revenue and the corresponding estimate is

taken to be primarily due to the combined effect of the tax effort variables which are omitted in the regression specification. The pooled cross-section and time series data for the states have been utilised to estimate a single function for taxable capacities by introducing statewise and yearwise dummy variables. On observing the non-homogeneity of the estimated taxable capacity functions across all the 14 states considered in the analysis, the states have been classified subsequently into three groups (in terms of the level of per capita SDP) in an effort to obtain three distinct groups of states each having a homogeneous taxable capacity function. The relative taxable capacities and the corresponding efficiency/inefficiency of individual states in terms of tax effort have been measured with reference to these estimated group-specific taxable capacity function.

The NFC methodology summarised above and the empirical results obtained by following this methodology (which are presented in the NFC report) give rise, as stated earlier, to some questions essentially of a methodological nature. These questions are important and deserve to be resolved, for the nature and the quality of the empirical results are likley to be crucially dependent on whether or not these questions can be resolved. In what follows, we briefly mention these methodological questions.

(1) *Specification of the Regression Model*

The regression model considered in the empirical analysis includes three explanatory variables, viz, the per capita SDP, share of non-primary sector in SDP and the Lorenz ratio of consumption, taken to capture taxable capacity of a state, and a Cobb-Douglas type functional form relating per capita state tax revenue to these variables has been employed. As already mentioned, the methodology presupposes that the variables relating to tax effort are independent of the explanatory variables included in the regression, so that the coefficients of the explanatory variables are not biased due to the omission of the variables relating to tax effort and other determinants of taxable capacity of a state, if there be any.

Is this assumption of independence empirically justifiable? If it is not, then the consequence of omission of the variables would essentially be to bias the estimates of the included explanatory variables (taken to reflect

the taxable capacity of the states) and, as a result, the estimated taxable capacity function would no longer measure taxable capacity of the states. Doubts about the assumption have already been expressed by Bagchi as he writes "there is also the possibility that tax effort itself is a function of the level of development of a state with increasing pressures for public expenditures and may not be an entirely exogenous factor" [Bagchi, 1989].

Another major difficulty with the specification of the taxable capacity function relates to its aggregative nature. The states collect qualitatively different types of taxes (taxable capacities in respect of which may be dependent on different types of factors, not all of which can be proxied by the three explanatory variables considered). Moreover, the composition of the state tax revenue in terms of the different tax types may be widely different across states. In such a case, the existence of an aggregate tax capacity relation as used in the NFC exercise may be questionable. Further, one would normally seek a causal justification of a regression equation used in an empirical economic analysis. Is it possible to give such a causal justification in the present case?

(2) *Analysis of Homogeneity of the Taxable Capacity Functions:*

Presumably for want of enough data, a fixed effect model of regression is taken as the basis for empirical exercise in the NFC report and an attempt has been made to estimate a homogeneous taxable capacity function across states. Given the sample size, the task of estimating a meaningful fixed effect model appears rather formidable. For example, in the case of the unrestricted model estimated by pooling the data for six years for all the 14 states, there are 84 sample observations available for estimating 62 parameters, including the coefficients of the statewise dummies in the slopes and the intercept. Appendix to the NFC report mentions that the period covered for the regression analysis was 1980-81 to 1985-86. In the main body of the report (Sec 3.5 of Ch III) however it is stated that the pooling of cross-section observations with time series was done for the period from 1980-81 to 1984-85. This would mean that the number of observations was actually 72. Is it possible to estimate any stable regression relationship

on the basis of such a scanty sample of observations?

Quite understandably, the empirical results indicate heterogeneity of the taxable capacity function across states. However, the subsequent steps that are performed to identify sub-groups of states for which the taxable capacity functions are homogeneous give rise to further questions. The principal component analysis which was performed to devise a composite development index (to be used for grouping the states) was based on eight variables. Of these, three of the first -four variables (*including* per capita SDP) were earlier considered as explanatory variables in the regression analysis based on the data for all the 14 states and were seen not to explain significantly the tax revenue. As reported, the first principal component would explain 60 per cent of the observed variability of the eight development indicators, and the composite development index was constructed by using the weights implied by the first principal component. Given the fact that a reasonably large part of the variance could not be explained by the first principal component, the uniqueness of the grouping of the states based on the composite development index becomes a debatable issue. Thus, the question arises whether one could have any alternative grouping of states with homogeneous taxable capacity function. If such alternative grouping could be possible, then the sanctity of the groupwise homogeneous taxable capacity function would become questionable. Another question, which the results of the principal component analysis gives rise to concerns the specification of the functional form of the basic regression model of the Cobb-Douglas form. It is reported that the ranking of the states by the level of per capita SDP is highly correlated with that based on the composite development index mentioned above. Further, it is found that the regressions for the three sub-groups of states (classified by the level of per capita SDP) are homogeneous in respect of the slope coefficients. This possibly implies that the slope coefficients in the basic model are independent of the level of per capita SDP (i e, the elasticities of per capita tax revenue with respect to the explanatory variables are not constant). Therefore, the question arises whether the Cobb-Douglas specification of the basic regression model is adequate.

(3) *Linear Restriction on the Coefficients*

In the final stage of the regression analysis, groupwise homogeneous taxable capacity functions have been estimated by incorporating a linear restriction that the coefficients of the statewise intercept dummies within each group add up to unity. It appears that the reported regression results are related to the linear-logarithmic specification of the basic regression model. If this be the case, it is not clear how a linear restriction of the type mentioned above is warranted. In the situation where the factors affecting taxable capacity are independent of

the factors relating to tax effort, the effect of omission of the set of variables reflecting tax effort from the regression specification may be as follows: the mean effect of omission of these variables would result in a statewise intercept term. However, in the event some of the variables reflecting tax capacity are also omitted in the specification (which is likely to have happened in the present case), no efficiency connotation should be attached to the multiplicative state-wise intercepts. Hence, the question of normalising these intercepts does not possibly arise. Even in the case of omission of tax effort variables alone, one may, at best, postulate that the product of the state-wise intercepts (in the original Cobb-Douglas specification) is unity. In such a case, the relevant restriction in the linear-logarithmic set-up would be that the statewise intercepts added up to zero. From the reported results, it may be seen that the intercept for some of the states is negative. Thus, these are the restricted estimates of the linear-logarithmic specification (since it is impossible to have a negative intercept term in the multiplicative Cobb-Douglas specification). Now, if the linear restriction is indeed irrelevant, it is not clear whether the claim of efficiency of the estimates would be legitimate. More importantly, the effect of introducing an irrelevant coefficient restriction may very well be quite damaging so far as the estimates of the other parameters of the regression equation are concerned, and thus may significantly affect any policy conclusion based on the estimated regression equation with such an improper linear restriction.

(4) *Regression Results*

We may now pass on to the regression results presented in the appendix. It has been mentioned that a large number of functions with several combinations of variables were tried out before choosing the ones presented on consideration of their having the most satisfactory statistical properties. It is stated in particular that the reported "equations have been chosen for the purpose of making projections' on three considerations": (i) the explanatory power of the equations should be high; (ii) the residuals should not have a high degree of autocorrelation; and (iii) the signs of the regression coefficients should be on expected lines.

The results show very clearly that the explanations achieved are indeed remarkably high. It turns out, however, that among the economic variables only per capita SDP is significant. In other words, SDP is found to be the only significant determinant of per capita tax revenue. But still no explanation has been provided for retaining the other non-significant variables. In fact, it is not even clear whether the regressions were re-estimated by dropping the non-significant variables and, if so, what kind of results were obtained. Secondly, since the values of the D-W statistic have been presented, one may presume that autocorrelation of the residuals was tested by using the Durbin-Watson test.

The problem is how would one apply and interpret the Durbin-Watson test in case of pooled cross-section time series data. Thirdly, nothing has been stated about the expected signs of the regression coefficients. It is found that the estimated regression coefficient for Lorenz ratio of consumption is positive for high and middle income states · but negative for the low income group. How can it be said then that the three considerations listed above were actually followed?

CONCLUSION

On the whole it seems that the methodology followed by the NFC is quite innovative. But there are a number of serious limitations and one may doubt if the purpose has been served. The moot question then would be to decide whether the results of the NFC methodology, as it stands now, should be accepted particularly since it has serious practical implications. We would only like to state in this context that we do not quite agree with Bagchi when he says "those who still have reservations about the normative approach based on econometric models (and some of them may be valid especially when there are severe limitations of data) have to come up with alternative approaches which would be more objective or less questionable" [Bagchi, 1988]. The fact that one does not have a better alternative at hand should not debar one from voicing objections against an existing or proposed methodology.

References

Bagchi, Amaresh (1988): 'First Award of the Ninth Finance Commission', *Economic and Political Weekly*, Vol XXIII, No 49, December 3, 1988.

First Report of the Ninth Finance Commission (for 1989-90), July 1988.

Guhan, S (1989): 'The Norm and the Tilt: First Report of the Finance Commission', *Economic and Political Weekly*, Vol XXIV, No 2, January 14, 1989.

Gurumurthi, Sitaram (1988): 'First Report of the Ninth Finance Commission: Analysis', *Economic and Political Weekly*, Vol XXIII, No 50, December 10, 1988.