Ethiopian Journal of Development Research, Vol.8, No.2, Oct. 1986

A GDP SHARE ANALYSIS FOR ETHIOPIA

Gerhard Kocklaeuner*

ABSTRACT. In this paper a sectoral time series analysis is performed with respect to GDP data of Ethiopia. The Analysis covers GDP at constant factor cost from 1953 E.F.Y. up to 1974 E.F.Y. both including and excluding population growth. Partly due to weak reliability of GDP data for Ethiopia rank regression is introduced as a special means of analysis. Using this method structural changes in nonlinear GDP share developments can significantly be detected at the end of 1965 E.F.Y.. An aggregation of individual share estimates finally provides a modelled series of GDP totals.

1. INTRODUCTION

Descriptive analysis of sectoral contributions to overall GDP are usually performed in terms of percentage shares. Varying percentage shares may be taken as indicators of structural change. An inferential time series analysis of GDP shares, however, has to provide means for modelling structural developments. Econometric theory here suggests the use of macro-economic regression models specified for quantitative variables. Following Shourie [13, p. 34], for forecasting and policy prescription purposes of developing countries 'the apparent faith in these models is unwarranted.' This scepticism to a large part results from limitations in the data base. In developing countries like Ethiopia national accounts are periodically revised. The corresponding figures of sectoral contributions to overall GDP often show a varying and partly low reliability. This issue has been treated thoroughly by

---

[*] Assistant Professor, University of Hannover, Hannover, FR Germany.

Kocklaeuner [9] after gaining experience in specification and estimation of a macro-econometric model for Ethiopia (see [8]). A GDP share analysis for Ethiopia therefore should refrain from using models in purely quantitative variables. A quantitative time series analysis of GDP by sectors should reduce model assumptions as far as possible for gaining statistical conclusions.

Consequently, when it comes to methodological considerations, methods of exploratory data analysis are of special interest. Such methods usually intertwine with robust and nonparametric procedures (cf. [11] for the regression context). Unfortunately, nonparametric statistical methods are usually directed to cross-sectional data. For time series data, with respect to regression problems their scope is restricted to simple linear models (see [6, pp. 200-218]). But a short look on GDP by sector (see Table 1 in the appendix) already reveals the nonlinear but almost monotone nature of sectoral GDP developments in time. So use of the rank transform in regression - as introduced by Iman and Conover [7] - remains to be an appropriate procedure for analysis. Rank regression fits monotone nonlinear trends and thereby provides a bridge between parametric and nonparametric Statistics (see [5] with respect to a general discussion of rank transformations).

In the sections to follow at first the method of rank regression will be presented. Special consideration will be given to its use in share analyses. Then rank regression will be applied to sectoral GDP data for Ethiopia. Separate fits will be provided for the development of each of the four main subsectors of the Ethiopian economy covering the period from 1953 E.F.Y. up to 1974 E.F.Y. Tests against structural changes at the end of 1965 E.F.Y. will be performed. GDP totals will be modelled by aggregating individual share estimates, finally.

2. RANK REGRESSION IN SHARE ANALYSIS

Applying the rank transform in regression simply means replacing the data with their ranks. For a simple linear trend model:

$$(1)\quad Y_t = a + bt + u_t, \qquad t - 1, \ldots, T$$

the rank transform leads to the rank regression model

$$(2)\quad R(Y_t) = a + bt + u_t, \qquad t = 1, \ldots, T$$

where u_t denotes a stochastic disturbance term. Here $R(Y_t)$ is the rank assigned to observation Y_t of variable Y. Due to the exogenous variable t being a time variable ties in assigning ranks only can occur with respect to the endogenous variable Y. Here an assignment of average ranks is proposed in literature (see [7] e.g.). In rank regression now an ordinary least squares analysis is performed for model (2). This model 'fits the monotone nonlinear trend while robust regression is forced to treat nonlinearity as outliers' (cf. [7, p. 503]). The fitted equation based on ranks is given by:

$$(3)\quad \hat{R}(Y_t) = \frac{T+1}{2} + r\left(t - \frac{T+1}{2}\right), \qquad t = 1, \ldots, T$$

where r is Spearman's rank correlation coefficient between the observations of Y and t. For a given time period to the predicted rank $\widehat{R}(Y_{t_o})$ usually will have to be transformed into a predicted value $\hat{Y}_{t_o}$ as follows:

$$(4)\quad \text{If } R(Y_{(i)}) < \hat{R}(Y_{t_o}) < R(Y_{(i+1)}), \text{ then}$$

$$\hat{Y}_{t_o} = Y_{(i)} + (Y_{(i+1)} - Y_{(i)}\ (R(Y_{t_o}) - R(Y_{(i)}).$$

Extrapolations for forecasting purposes, for example, are excluded. So rank regression reduces to a method for analysing *ex post* data.

In Share analysis now an aggregated variable Y will always equal the weighted sum of its aggregates Xj $(j=1, \ldots, p)$ *i.e.*

$$(5) \quad Y_t = \sum_{j=1}^{p} w_j X_{jt}, \qquad t=1, \ldots, T.$$

where w_j denotes some weight constant, appropriately chosen. Here fitted values $\hat{Y}_t$ can be obtained by applying rank regression in two different ways. Once, following equations (3) and (4) a separate rank regression regarding Y could be run. Alternatively, separate rank regressions on all aggregates X_j – after transformations to fitted values – could be aggregated via:

$$(6) \quad \hat{Y}_t = \sum_{j=1}^{p} w_j \hat{X}_{jt}, \qquad t=1, \ldots, T.$$

This identity again should hold in case of $p+1$ separate rank regressions explaining the variables $Y, X_1, \ldots, X_p$. However, in linear regression models prediction of sum-constrained dependent variables will always lead to restrictions on the explanatory variables and on the parameters implied (see */12/*, especially). Consequently, both approaches can not be brought in line with each other. Logical consistency of multivariate share models is assured especially if fitting follows equations *(6)*. Thus the second approach, which has been described by Beckwith */3/* in detail, will be favoured here.

This approach again is applicable if multivariate regressions on ranks are considered instead of model *(2)*. Such models are of special interest in the analysis of structural change. A special case is represented by:

$$(7) \quad R(Y_t) = a + bt + cD_t t + u_t, \qquad D_t = \begin{cases} 0 \ if\, t \leqslant t_1 \\ 1 \ if\, t > t_1 \end{cases}$$

where the dummy-variable D is introduced to model structural change at the end of period t_1 (see Maddala [10, pp. 132-141], who gives a summary of the use of dummy variables for taking qualitative effects into account). Note, that via D different slopes for separate regimes are specified. As indicated above, rank transforms in regression analysis should be restricted mainly to monotone data. Therefore structural shifts hardly can be detected in rank regression analysis.

3. RANK REGRESSION ANALYSIS OF ETHIOPIAN GDP

This section is devoted to the application of rank regression in time series analyses of Ethiopian GDP. Sectoral GDP data are considered both including and exluding population growth. GDP by sector at constant factor cost of 1953 E.F.Y. is summarized in [4, p. 174]. The NRDC—CPSC Secretariate reports population figures for Ethiopia every year. So GDP values by sector at constant population and at constant factor cost of 1953 E.F.Y. can be calculated (see Table 1 in the appendix). In the period under consideration average annual population growth has been around 2 per cent. Discarding the influence of population growth leads to GDP data being interpretable as per capita figures. As Table 1 indicates, these figures show a fairly low degree of nonmonotocity. In the Other Services Sector even there is a strict increase from 1953 E.F.Y. to 1974 E.F.Y.. All time series to analyse exhibit some kind of nonlinearity. This is especially obvious when looking at the period beginning in 1966 E.F.Y.. Regarding the development of Services an exponential trend might give an appropriate fit for the data including population growth not being reported here. However, such a model 'will be only as good as the estimate of the nonlinear function used to fit it' (see [7, p. 501]). Remembering limitations concerning the quality of available data enough arguments for applying rank regression seem to be enlisted. So a comparision of rank regression with analysis of simple linear trends will be given for Ethiopian GDP in the first part of this section. The second part then will be devoted to tests against structural change and to modelling the aggregate series of GDP totals.

3.1 Rank Regression Versus Fitting a Linear Trend

Following Iman and Conover [7, p. 509] rank regression should not 'be regarded as a blanket solution but is best used in conjunction with other methods of analysis. That is, a standard regression should be used and the residuals compared to those using the rank transform'. Accordingly least squares fits were obtained both regarding models presented in equations (1) and (2). Concerning sectoral GDP data including population growth a summary of results is given as follows:

$$
\begin{aligned}
(8) \quad \hat{R}(GDP_{AS,\,t}) &= 11.5 + 0.968\,(t-11.5), && SAR=275(769), \\
& && RVN=2.03 \\
\hat{R}(GDP_{OCS,\,t}) &= 11.5 + 0.928\,(t-11.5), && SAR=451(81[illegible]), \\
& && RVN=0.92 \\
\hat{R}(GDP_{DS,\,t}) &= 11.5 + 0.981(t-11.5), && SAR=158(674), \\
& && RVN=2.08 \\
\hat{R}(GDP_{OS,\,t}) &= t\,.
\end{aligned}
$$

Here the sector under consideration is indicated by using appropriate subscripts. SAR denotes the sum of absolute residuals obtained after transforming predicted ranks into predicted actual shares of GDP according to equation (4). In brackets the corresponding SARs after fitting simple linear trends are given. Thus a clear superiority of rank regression over fitting linear trends can be stated. Of course, in linear trend analysis a consideration of the Durbin-Watson test statistic always shows significant serial correlation due to misspecification (see Maddala [10, p. 274] e.g.). Due to strict monotonicity rank regression provides an exact fit in the Other Services sector. Misspecification, however, again can be present in rank regression models. The von Neumann-Ratio RVN calculated from fitted ranks is indicating misspecification for the Other Commodity Sector (cf. [10, p. 287] e.g.).

At this point one may wonder why tests against serial correlation or misspecification are not performed on ranks of residuals in the given context. Bartels [1] provides a rank version of the von Neumann's Ratio test for randomness, for instance. This test outperforms the normal theory version of the test in non-normal populations. However, as shown in [2], when applied in regression models this test is strongly affected by non-normality. Normality of disturbance terms again can not be assumed to hold when specifying model (2) to the data on hand. This view is supported by looking at the rank version of RVN for data from Agriculture and Distributive Services Sectors. In both cases contrary to RVN its rank version leads to the conclusion of misspecification.

Turning now to sectoral GDP data excluding population growth i.e. to data of Table 1, rank regression results in:

$$
\begin{aligned}
(9)\quad & GDP_{AS}: r = -0.692, && SAR = 509(635), && RVN = 0.70 \\
& GDP_{OCS}: r = 0.763, && SAR = 603(757), && RVN = 0.56 \\
& GDP_{DS}: r = 0.921, && SAR = 275(642), && RVN = 0.70 \\
& GDP_{OS}: r = 1.0 && &&
\end{aligned}
$$

Again rank regression clearly outperforms fitting a linear trend, where obviously misspecification reduces values of the Durbin-Watson statistic always below one. Contrary to the original data in (9) the RVN values indicate misspecification with respect to all sectoral fits. Consequently instead of model (2) alternative rank regression models should be considered. One possible specification already is given in equation (7). This model allows for testing against structural change and shall be analysed in the sequel.

3.2 Structural Change and Aggregation

The dummy-variable model (7) has been confronted both with sectoral GDP data including and excluding population growth. Rank regressions were always performed using the ordinary least squares approach. Structural change was assumed to take place at the end of 1965 E.F.Y. i.e. with the beginning of decisive changes in the political and economic system of Ethiopia. The original data — not adjusted for changes in population — provided fits as follows:

$$(10) \qquad \hat{R}(GDP_{AS,t}) = -0.70 + 1.15t - \underset{(0.08)}{0.15}D_t t, \quad R^2 = 0.95, \; DW = 2.31$$

$$\hat{R}(GDP_{OCS,t}) = -1.19 + 1.28t - \underset{(0.11)}{0.28}D_t t, \quad R^2 = 0.90, \; DW = 1.20$$

$$\hat{R}(GDP_{DS,t}) = 0.04 + 1.01t - \underset{(0.06)}{0.02}D_t t, \quad R^2 = 0.96, \; DW = 2.18$$

The estimated standard deviations of coefficients' estimates, given in brackets, suggest results being in line with those enlisted in equations (8). At a level of significance $a = 0.05$ a t-test leads to insignificant contributions of $D_t t$ to the explanation of ranks of $GDP_{AS,t}$ and $GDP_{DS,t}$. Structural change can not be observed here. Remember, that in (8) correct specification was indicated with respect to rank regressions concerning these variables. However, with regard to ranks of $GDP_{OCS,t}$ a significant structural change can be observed. For this equation the Durbin-Watson coefficient DW is lying in the indifference region ($a = 0.05$, test against positive autocorrelation, cf. [10, p. 508] e.g.). This result again corresponds to (8), where the value of RVN has indicated some kind of misspecification.

Concerning sectoral time series of GDP shares in the per capita case of Table 1, in equations (9) simple rank regressions did not give appropriate fits. So 'better' results have to be expected when the possibility of a change of slope is included now. Fitting model (7) with these data leads to:

$$(11) \quad \hat{R}(GDP_{AS,t}) = 13.47 + 0.37t - \underset{(0.16)}{0.84}D_t t, \quad R^2 = 0.79, \; DW = 1.52$$

$$\hat{R}(GDP_{OCS,t}) = 2.29 + 1.65t - \underset{(0.16)}{0.71}D_t t, \quad R^2 = 0.80 \; DW = 1.10$$

$$\hat{R}(GDP_{DS,t}) = 0.34 + 1.02t - \underset{(0.13)}{0.81}D_t t, \quad R^2 = 0.85, \; DW = 0.77.$$

Here the DW—values of the first two equations no longer suggest existing serial correlation of disturbance terms ($a = 0.01$, test against positive autocorrelation, cf. [10, p. 509] e.g.). In both cases the variable $D_t t$ turns to be significant at $a = 0.05$. So misspecification has been removed by allowing for structural change. With respect to explanation of ranks of $GDP_{DS,t}$ still serial correlation is indicated at the level of $a = 0.01$. Nevertheless the estimated standard deviation in this equation may give some hints for a changing slope at the beginning of 1966 E.F.Y..

When it comes to aggregation now, a summary of fitted rank regression from equations (8) to (11) has to be taken into consideration. For modelling a series of GDP totals equation (6) in its modification:

$$(12)\quad \hat{GDP}_t = \hat{GDP}_{AS,t} + \hat{GDP}_{OCS,t} + \hat{GDP}_{DS,t} + \hat{GDP}_{OS,t}$$

shall be used. According to this identity predicted values of GDP shares – obtained from separate rank regression – have to be added. A separate rank regression analysis is not necessary with regard to GDP totals. Due to strict monotonicity of $GDP_{OS,t}$ values these always will equal $\hat{GDP}_{OS,t}$. Therefore equation (12) reduces to prediction via:

$$(13)\quad \hat{GDP}_t - GDP_{OS,t} = \hat{GDP}_{AS,t} + \hat{GDP}_{OCS,t} + \hat{GDP}_{DS,t}.$$

As an example period $t=22$ corresponding to 1974 E.F.Y. shall be considered for the per capita case (see Table 1). Here predicted ranks should depend on the dummy-variable modelling structural change. Following equations (11) and (4), then:

$$\begin{aligned}(14)\quad \hat{R}(GDP_{AS,22}) &= 3.13, & \hat{GDP}_{AS,22} &= 1419.14\\ \hat{R}(GDP_{OCS,22}) &= 18.39, & \hat{GDP}_{OCS,22} &= 497.13\\ \hat{R}(GDP_{DS,22}) &= 4.96, & \hat{GDP}_{DS,22} &= 317.88.\end{aligned}$$

A comparison with actual values again shows that only the rank regression fit concerning $GDP_{DS,t}$ will have to be improved. In the present version the predicted value $\hat{GDP}_{22} = 2933.85$ heavily underestimates GDP total for 1974 E.F.Y.. This kind of misfit disappears if structural change in Distributive Services is introduced with a delay of three years compared to both of the production sectors. Such a time span is sustained by the data on hand and can be the result of some economic reasoning. Contrary to equation (11) rank regression now leads to:

$$(15) \qquad \hat{R}(GDP_{DS,t}) = -1.51 + 1.29t - \underset{(0.07)}{0.34D_t t}, \qquad R^2 = 0.93, \quad DW = 1.17.$$

In this equation the DW statistic no longer indicates misspecification. Consequently a significant contribution of $D_t t$ has to be stated. For period $t = 22$ predicted rank as well as fitted value of $GDP_{DS,22}$ are revised in:

$$(16) \qquad \hat{R}(GDP_{DS,22}) = 19.39, \qquad \widehat{GDP}_{DS,\,22} = 470.56.$$

Corresponding to this fit a predicted value $\widehat{GDP}_{22} = 3086.53$ near to the original GDP total is obtained.

4. SUMMARY

In this paper rank regression has been introduced as an appropriate means of GDP share analysis for Ethiopia. Using this procedure a structural change in both the per capita time series of GDP from Agriculture and Other Commodity Sectors has been detected at the end of 1965 E.F.Y. A fairly exact fit of GDP totals has been obtained by assuming a similar change with respect to the per capita series of Distributive Services at the end of 1968 E.F.Y. In the context of analysis especially advantages of rank regression compared to fitting a linear trend have become obvious. The use of rank regression in aggregation has been discussed for the first time.

APPENDIX

Table 1: GDP by Sector at Constant Population* and at Constant Factor Cost of 1953 E.F.Y. (Million Birr)

E.F.Y.	Agriculture Sector	Other Comm. Sector	Distribution Services	Other Services	GDP
1953	1,504.5	286.2	217.7	323.3	2,331.7
1954	1,515.3	306.7	236.3	350.9	2,409.2
1955	1,530.2	318.9	249.5	360.9	2,459.5
1956	1,528.5	340.9	278.9	371.4	2,519.7
1957	1,568.4	366.5	319.5	409.8	2,664.2
1958	1,551.3	403.8	338.2	428.7	2,722.0
1959	1,567.0	437.5	338.8	444.2	2,787.5
1960	1,547.2	441.3	370.0	463.4	2,821.9
1961	1,538.8	462.7	380.7	479.2	2,861.4
1962	1,536.3	453.8	412.2	490.0	2,892.3
1963	1,529.9	485.7	431.2	503.0	2,949.8
1964	1,561.1	500.2	451.9	531.7	3,044.9
1965	1,533.3	500.8	460.2	555.0	3,049.4
1966	1,488.1	486.0	469.9	574.9	3,018.9
1967	1,459.3	478.8	469.1	614.4	3,021.6
1968	1,492.6	446.8	471.6	639.1	3,050.1
1969	1,455.2	448.5	461.7	648.6	3,014.0
1970	1,397.4	423.9	413.9	679.5	2,914.7
1971	1,397.6	466.6	453.0	680.3	2,997.5
1972	1,427.0	498.9	466.7	691.2	3,083.8
1973	1,426.1	503.1	479.4	696.7	3,105.3
1974	1,418.1	496.0	485.7	699.7	3,099.5

Source: Central Statistical Office, Statistical Abstract.

* Population figures from NRDC–CPSC Secretariat.

REFERENCES

[1] Bartels, R., "The Rank Version of von Neumann's Ratio Test for Randomness", *J. Amer. Statist. Assoc.*, Vol. 77, No. 377, (1982), pp. 40-46.

[2] Bartels, R., "The Rank von Neumann Test as a Test for Autocorrelation in Regression Models", *Commun. Statist.– Theor. Meth.*, Vol. 13, No. 20, (1984), pp. 2495-2502.

[3] Beckwith, N.E., "Concerning the Logical Consistency of Multivariate Market Share Models", *J. Marketing Research*, Vol. 10, (1973), pp. 341-344.

[4] Central Statistical Office, *Ethiopia, Statistical Abstract 1980*, (Addis Ababa, 1982).

[5] Conover, W.J. and R. L. Iman, "Rank Transformations as a Bridge Between Parametric and Nonparametric Statistics", *American Statistician*, Vol. 35, No. 3, (1980), pp. 124-129.

[6] Hollander, M. and D.A. Wolfe, *Nonparametric Statistical Methods*, (New York: Wiley, 1973).

[7] Iman, R.L. and W.J. Conover, "The Use of the Rank Transform in Regression", *Technometrics*, Vol. 21, No. 4, (1979), pp. 499-509.

[8] Kocklaeuner, G. and Asmerom Kidane, *A Macro Econometric Model for Ethiopia: Specification, Estimation, Forecast and Control*, (Research Paper, Addis Ababa University, 1983).

[9] Kocklaeuner, G., "Oekonometrische Modelle in der Entwicklungsplanung Ethiopiens: Datenprobleme und ihre Auswirkungen", *afrika spectrum*, Vol. 19, No. 2, (1984), pp. 135-148.

[10] Maddala, G.S., *Econometrics,* (Tokyo: McGraw-Hill, 1979).

[11] Mosteller, F. and J.W. Tukey, *Data Analysis and Regression,* (Reading, Massachusetts: Addison-Wesley, 1977).

[12] Naert, P.A. and A. Bultez, "Logically Consistent Market Share Models", *J. Marketing Research,* Vol. 10, (1973), pp. 334-340.

[13] Shourie, A., "The Use of Macro-Economic Regression Models of Developing Countries for Forecasts and Policy Prescription: Some Reflections on Current Practice", *Oxford Econ. Papers,* Vol. 24, (1972), pp. 1-35.