ONCE AGAIN: TESTING FOR REGIONAL HOMOGENEITY*

Peter M. Schulze†

ABSTRACT. Summarizing the foregoing discussions in this journal on testing for regional homogeneity the present note shows that in the model of Zellner's seemingly unrelated regressions *one* test statistic may be used not only to test for overall homogeneity but also to examine for individual coefficient homogeneity. This aim is achieved by varying the linear restrictions in the test statistic according to different problems. To illustrate these tests regional consumption functions for the 11 Bundesländer (States) of the Federal Republic of Germany are used.

1. INTRODUCTION

Previous discussions of intraregional economic homogeneity in this journal were concerned with the testing of overall homogeneity for the entire vector of population coefficients [Johnson (1975), Schulze (1977)] and of a particular population parameter [Lin (1985)]. Both questions, due to different problems, are of relevance to the regional scientist.

Lin (1985, p. 131) has derived the test statistic to test for one parameter. This note will complete the discussion and show that in the model of Zellner's seemingly unrelated regressions (SUR) the test statistic derived by Zellner (1962) may be used not only to test for overall homogeneity but also to examine for individual coefficient homogeneity (Section 2). In Section 3 the conclusions are illustrated by an empirical example.

2. MODELS AND TESTS

The starting point of our discussion is Zellner's well-known SUR model:

$$(1) \qquad \underset{(RT\times 1)}{\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_R \end{bmatrix}} = \underset{(RT\times RK)}{\begin{bmatrix} \mathbf{X}_1 & 0 & \ldots & 0 \\ 0 & \mathbf{X}_2 & \ldots & 0 \\ \vdots & & & \\ 0 & 0 & \ldots & \mathbf{X}_R \end{bmatrix}} \underset{(RK\times 1)}{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_R \end{bmatrix}} + \underset{(RT\times 1)}{\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_R \end{bmatrix}}$$

with region $r = 1, \ldots, R$, period $t = 1, \ldots, T$, and explanatory variable $k = 1, \ldots, K$. This approach is extended by Parks' (1967) assumption of a first-order autoregressive process for the latent variables in each regional regression equa-

*The author wishes to thank J. Schmidt for computations and helpful comments.

†Professor, Institute of Statistics and Econometrics, Johannes Gutenberg-University, Mainz, West Germany.

Received January, 1986; in revised form April, 1986; accepted May, 1986.

tion:

(2) $$\epsilon_{rt} = \rho_r \epsilon_{rt-1} + \omega_{rt}$$

Details of the estimation procedure shall not be discussed here. [See, for example, Kmenta (1971, Chap. 12).]

Model (1) assumes that in each equation the coefficients are fixed and constant over time but may differ among regions. It serves as the unrestricted model for the alternative hypothesis H_1. The null hypothesis H_0—as a restricted model—may be formulated in different ways according to the specific question of interest. The hypotheses are included in the appropriate formulations and are considered by the linear restrictions

(3) $$\mathbf{C}\,\beta = \mathbf{c}$$

with **c** a known vector with J elements and **C** a known matrix with full row rank and of order $J\,[=K(R-1)] \times RK$.

Under the assumption that ϵ is normally distributed and that the null hypothesis (3) is true, the test statistic is given by [Zellner (1962, p. 355)]

(4) $$F^* = \frac{(\mathbf{c} - \mathbf{C}\,\hat{\beta})'\,\{\mathbf{C}\,[\mathbf{X}'(\hat{\Omega}^{-1} \otimes \mathbf{I})\mathbf{X}]^{-1}\mathbf{C}'\}^{-1}(\mathbf{c} - \mathbf{C}\,\hat{\beta})/J}{\mathbf{e}'(\hat{\Omega}^{-1} \otimes \mathbf{I})\mathbf{e}/(RT - RK)}$$

where J is the number of restrictions (number of rows in **C**) and **e** the estimators of ϵ ($\otimes$ is the Kronecker product). Equation (4) possesses an F-distribution with J and RT-RK degrees of freedom.[1] If Ω and β are replaced by their estimators this will only be asymptotically valid.

In the following we shall formulate four different null hypotheses which may be of value to the regional researcher.

2.1 At first it would be interesting to know whether there are any significant differences between regions, that is

$$H_0 : \beta_{1.} = \beta_{2.} = \ldots \beta_{R.} = \beta$$

H_1 means that at least one $\beta_{r.} \neq \beta$. Equation (3) is then specified as follows:

$$\begin{bmatrix} 1\,0\ldots0 & -1\,0\ldots0 & 0\,0\ldots0 & \ldots & 0\,0\ldots0 & 0\,0\ldots0 \\ 0\,0\ldots0 & 1\,0\ldots0 & -1\,0\ldots0 & \ldots & 0\,0\ldots0 & 0\,0\ldots0 \\ \vdots & & \vdots & & & \vdots \\ 0\,0\ldots0 & 0\,0\ldots0 & 0\,0\ldots0 & \ldots & 0\,0\ldots-1 & 0\,0\ldots0 \\ 0\,0\ldots0 & 0\,0\ldots0 & 0\,0\ldots0 & \ldots & 0\,0\ldots1 & 0\,0\ldots-1 \end{bmatrix} \begin{bmatrix} \beta_{11} \\ \vdots \\ \beta_{1k} \\ \vdots \\ \beta_{R1} \\ \vdots \\ \beta_{RK} \end{bmatrix} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

C is, in general, a $J\,\{=K(R-1)] \times RK$ matrix.

[1]Formula (4) corresponds to Lin's formula (5).

2.2 If under 2.1 the test statistic (4) shows significant differences, researchers will be interested in finding out which regions are responsible for this result. For this purpose a pairwise test may follow which is a special case of that in 2.1. Matrix **C** is reduced to a $K \times RK$ matrix.

2.3 It may also be interesting to know whether certain coefficients $\beta_1, \beta_2, \ldots,$ (for example growth rates, propensities to consume) in the observed regions are significantly different:

$$H_0 : \beta_{1k} = \beta_{2k} = \ldots = \beta_{RK}$$

H_1 means that one regression coefficient differs from at least one other one. If for example β_2 is to be tested on significance, (3) is formulated as follows:

$$\begin{bmatrix} 0\,1\ldots0 & 0\,{-1}\ldots0 & 0\,0\ldots0 & \ldots & 0\,0\ldots0 & 0\,0\ldots0 \\ 0\,0\ldots0 & 0\,1\ldots0 & 0\,{-1}\ldots0 & \ldots & 0\,0\ldots0 & 0\,0\ldots0 \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ 0\,0\ldots0 & 0\,0\ldots0 & 0\,0\ldots0 & \ldots & 0\,1\ldots0 & 0\,{-1}\ldots0 \end{bmatrix} \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \vdots \\ \beta_{1k} \\ \vdots \\ \beta_{R1} \\ \beta_{R2} \\ \vdots \\ \beta_{RK} \end{bmatrix} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

Here in general **C** has dimensions $(R - 1) \times RK$.

2.4 This is a special case of 2.3. If under 2.3 significant differences of a special regression coefficient have been shown, in general one will perhaps examine for which regions these coefficient-specific differences are valid. Matrix **C** in 2.3 reduces to a $1 \times RK$ matrix.

3. APPLICATION

To illustrate the tests noted above, the consumption functions

$$\left(\frac{C}{Y}\right)_{rt} = \beta_{r1} + \beta_{r2}\left(\frac{\Delta Y}{Y}\right)_{rt} + \beta_{r3}\left(\frac{C}{Y}\right)_{rt-1} + \epsilon_{rt} \qquad (5)$$

for the 11 Bundesländer (States) of the Federal Republic of Germany are estimated by annual data from 1960–1976, where

C = private consumption,
Y = disposable income, and
ΔY = absolute difference of Y between period t and $t - 1$.

This formulation of the consumption function is used by Evans (1969, Chap. 3) for macroeconomic analysis and discussed by Schulze (1982) for regional analysis.

The Parks (1967) estimation procedure leads to the following results reported in Table 1.

TABLE 1: Estimation Results

State	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	R^2
Schleswig-Holstein	0.24	−0.37	0.63	0.65
Hamburg	0.31	−0.42	0.58	0.74
Niedersachsen	0.49	−0.38	0.25	0.52
Bremen	0.17	−0.53	0.82	0.86
Nordrhein-Westfalen	0.28	−0.43	0.56	0.60
Hessen	0.17	−0.42	0.73	0.72
Rheinland-Pfalz	0.15	−0.43	0.77	0.67
Baden-Württemberg	0.22	−0.42	0.62	0.75
Bayern	0.32	−0.35	0.53	0.64
Saarland	0.23	−0.61	0.81	0.98
Berlin-West	0.06	−0.55	0.97	0.95

Except for the West-Berlin $\hat{\beta}_1$ all coefficients are significant at $\alpha = 0.05$. All coefficients also have the expected sign. In the following β_1 is excluded from the analysis because it is much less interesting than the other coefficients.

At first, according to case 2.1, it is interesting to test for overall regional homogeneity. Here we have $J = 20$ and $RT - RK = (11)(15) - 22 = 132$ degrees of freedom. Expression (4) results in

$$F^* = \frac{(81.46)/20}{(135.38)/132} = 3.97$$

which exceeds $F_{20,132;0.01} = 2.02$ and leads to the conclusion that the regional consumption functions differ significantly. The next question is which coefficients and regions are responsible for these differences.

Case 2.3 leads to the result that all three regression coefficients differ significantly, due to the test statistic (4). For 10 and 132 degrees of freedom and $\alpha = 0.01$ the critical F-value is 2.46 and the computed F-values are: for β_2, $F^* = 2.54$ and for β_3, $F^* = 5.66$. It can be seen that especially β_3 contributes to the regionally different consumption behavior. Referring to the size of effects it seems that reactions of private consumption depend more on the consumption-income-ratio C/Y of the previous period than on changes in income $\Delta Y/Y$ in the current period.

Furthermore, it is possible to analyze which regions (Bundesländer) have led to the significant differences in 2.1 (case 2.2). With $J = 2$ and $RT - RK = 132$, a critical value of $F_{2,132;0.01} = 4.77$ follows. A pairwise test between Nordrhein-Westfalen and Hessen, for example, results in a significant value with $F^* = 11.74$. Checking β_2 and β_3 for these two regions (case 2.4) leads to $F^* = 0.01$ for β_2 and to $F^* = 20.32$ for β_3. The corresponding critical value $F_{1,132;0.01} = 6.82$ shows—as supposed—that β_2 is not significant but β_3 is highly significant.

On the other hand, a comparison—for example between Nordrhein-Westfalen and Baden-Württemberg—shows an F^*-value of 0.60 (case 2.2) and accordingly the F^*-value for β_2 is 0.16 and the F^*-value for β_3 is 1.09 (case 2.3). This confirms the not significant result for these two regions.

4. CONCLUSION

This note shows that in the SUR model—allowing for temporal and spatial autocorrelation—only one test statistic is required. Through corresponding changes of the restriction matrix **C** it is possible to deal with tests not only on homogeneity of individual coefficients but also on overall homogeneity according to the individual questions of the regional researcher.

REFERENCES

Evans, Michael K. *Macroeconomic Activity*. New York: Harper & Row, 1969.

Johnson, Lester W., Jr. "A Note on Testing for Intraregional Economic Homogeneity," *Journal of Regional Science*, 15 (1975), 365–369.

Kmenta, Jan. *Elements of Econometrics*. New York: Macmillan, 1971.

Lin, An-Loh. "A Note on Testing for Regional Homogeneity of a Parameter," *Journal of Regional Science*, 25 (1985), 129–135.

Parks, Richard W. "Efficient Estimation of a System of Regression Equations when Disturbances are Both Serially and Contemporaneously Correlated," *Journal of the American Statistical Association*, 62 (1967), 500–509.

Schulze, Peter M. "Testing for Intraregional Economic Homogeneity: A Comment," *Journal of Regional Science*, 17 (1977), 473–477.

———, "Zur Schätzung regionaler Konsumfunktionen," *Jahrbuch für Regionalwissenschaft*, 3 (1982), 116–128.

Zellner, Arnold. "An Efficient Method of Estimating Seemingly Unrelated Regressions and Test for Aggregation Bias," *Journal of the American Statistical Association*, 57 (1962), 348–368.