

# More on the F-test under nonspherical disturbances\*

Walter Krämer<sup>†</sup>      Christoph Hanck<sup>‡</sup>

May 29, 2008

## Abstract

We show that the F-test can be both liberal and conservative in the context of a particular type of nonspherical behaviour induced by spatial autocorrelation, and that the conservative variant is more likely to occur for extreme values of the spatial autocorrelation parameter. In particular, it will wipe out the progressive one as the sample size increases.

*Key words:* F-test, spatial autocorrelation.

---

\*Research supported by Deutsche Forschungsgemeinschaft (DFG) under SFB 475. We are grateful to Anurag Banerjee for providing us the Gauss routines for our calculations.

<sup>†</sup>Fakultät Statistik, TU Dortmund, D-44221 Dortmund, Germany, walterk@statistik.uni-dortmund.de

<sup>‡</sup>Department Quantitative Economics, Universiteit Maastricht, NL-6211 LM Maastricht, c.hanck@ke.unimaas.nl

# 1 Introduction and summary

The robustness of the F-test to nonspherical disturbances has concerned applied statisticians for many decades. The present paper considers the F-test in the context of the linear regression model

$$y = X\beta + u = X^{(1)}\beta^{(1)} + X^{(2)}\beta^{(2)} + u, \quad (1)$$

where  $y$  and  $u$  are  $T \times 1$ ,  $X$  is  $T \times K$  and nonstochastic of rank  $K < T$ ,  $\beta$  is  $K \times 1$ , and the disturbance vector  $u$  is multivariate normal with mean zero and (possibly) nonscalar covariance matrix  $V$ . The design matrix is partitioned into  $X^{(1)}(T \times q)$  and  $X^{(2)}(T \times (K - q))$  and the null hypothesis to be tested is  $H_0 : \beta^{(1)} = b^{(1)}$ .

The standard F-test assumes that  $V = \sigma^2 I$  and rejects for large values of

$$F = \frac{(\tilde{u}'\tilde{u} - \hat{u}'\hat{u})/q}{\hat{u}'\hat{u}/(T - K)}, \quad (2)$$

where  $\hat{u} = y - X\hat{\beta}$ ,  $\hat{\beta} = (X'X)^{-1}X'y$ ,  $\tilde{u} = y - X^{(1)}b^{(1)} - X^{(2)}\tilde{\beta}^{(2)}$ ,  $\tilde{\beta}^{(2)} = (X^{(2)'}X^{(2)})^{-1}X^{(2)'}(y - X^{(1)}b^{(1)})$ . Its null distribution is central F with  $q$  and  $T - K$  degrees of freedom and the problem to be studied here is the robustness of this null distribution to deviations from  $V = \sigma^2 I$ .

Vinod (1976) and Kiviet (1980) address this problem for a given disturbance covariance matrix  $V$ , and derive bounds for the size of the test when the design matrix  $X$  varies across all  $T \times K$  matrices of rank  $K$ , while Banerjee and Magnus (2000) and Hillier and King (1982) consider the test statistics themselves. Below we follow Krämer (1989), Krämer et al. (1990) and Krämer (2003) by fixing  $X$  and letting  $V$  vary across certain subsets of possible disturbance covariance matrices which are likely to occur in practice. This seems the more natural approach, as  $X$  is always known in applications, whereas  $V$  is an unknown  $T \times T$  parameter matrix.

The subset of disturbance covariance matrices under study here is implicitly defined by the spatial autoregressive scheme

$$u = \rho Wu + \varepsilon, \quad (3)$$

where  $\varepsilon$  is a  $T \times 1$  normal random vector with mean zero and scalar covariance matrix  $\sigma_\varepsilon^2 I$ , and  $W$  is some known  $T \times T$ -matrix of nonnegative spatial weights with

$w_{ii} = 0$  ( $i = 1, \dots, T$ ). Although there are many other patterns of spatial dependence which have been suggested in the literature (see Anselin and Florax (1995) for an overview), the one defined by (3) is by far the most popular, so it seems worthwhile to investigate the behaviour of parameter estimates and tests when the regression disturbances “misbehaves” according to this particular scheme.

Below we build on Krämer (2003), who shows that the size of the test can tend to both one and zero as the parameter  $\rho$  varies across its allowable range. While Krämer (2003) is silent on the respective empirical relevance of the two extreme cases, we show here that the conservative variant is far more likely to occur in practice, and will wipe out the liberal one as sample size increases.

## 2 The null distribution under spatial autocorrelation

The coefficient  $\rho$  in (3) measures the degree of correlation, which can be both positive and negative. There is no disturbance autocorrelation where  $\rho = 0$ . Below we focus on the empirically more relevant case of positive disturbance correlation, where

$$0 \leq \rho < \frac{1}{\lambda_{max}}$$

and where  $\lambda_{max}$  is the Frobenius-root of  $W$  (i.e. the unique positive real eigenvalue such that  $\lambda_{max} \geq |\lambda_i|$  for arbitrary eigenvalues  $\lambda_i$ ). The disturbances are then given by

$$u = (I - \rho W)^{-1} \varepsilon, \tag{4}$$

so  $V := Cov(u) = \sigma_\varepsilon^2 [(I - \rho W)(I - \rho W)']^{-1}$  and  $V = \sigma_\varepsilon^2 I$  whenever  $\rho = 0$ .

The behaviour of the test statistic (2) when disturbances are given by (3) is best seen by first rewriting it as

$$F = \frac{u'(M^{(2)} - M)u/q}{u'Mu/(T - K)}, \tag{5}$$

where  $M = I - X(X'X)^{-1}X'$  and  $M^{(2)} = I - X^{(2)}(X^{(2)'}X^{(2)})^{-1}X^{(2)'}$ . Let  $F_{q,T-K}^\alpha$  be the  $(1 - \alpha)$  quantile of the central F-distribution with  $q$  and  $T - K$  degrees of freedom, respectively, where  $\alpha$  is the nominal size of the test. Then

$$\begin{aligned}
P(F \geq F_{q,T-K}^\alpha) &= P(u'(M^{(2)} - M)u - \frac{q}{T-K}F_{q,T-K}^\alpha u'Mu \geq 0) \\
&= P(u'(M^{(2)} - dM)u \geq 0) \\
&\quad (\text{where } d = 1 + \frac{q}{T-K}F_{q,T-K}^\alpha) \\
&= P(\eta'(I - \rho W)'(M^{(2)} - dM)(I - \rho W)\eta \geq 0) \\
&\quad (\text{where } \eta = \frac{1}{\sigma_\varepsilon}\varepsilon \sim N(0, I)) \\
&= P\left(\sum_{i=1}^T \lambda_i \xi_i^2 \geq 0\right) \\
&= P\left((1 - \rho\lambda_{max})^2 \sum_{i=1}^T \lambda_i \xi_i^2 \geq 0\right), \tag{6}
\end{aligned}$$

where the  $\xi_i^2$  are iid  $\chi_{(1)}^2$  and the  $\lambda_i$  are the eigenvalues of  $(I - \rho W)'(M^{(2)} - dM)(I - \rho W)$ , and therefore also of  $V(M^{(2)} - dM)$ .

The limiting rejection probability as  $\rho \rightarrow 1/\lambda_{max}$  depends upon the limiting behaviour of  $(1 - \rho\lambda_{max})^2 V$ . We confine ourselves to the case where  $W$  is symmetric, which appears to be the more important one in practice. This will for instance occur if spatial dependence follows the  $j$ -ahead-and- $j$ -behind or the equal-weight criteria (see section 3 below). Then  $W$  admits a spectral decomposition

$$W = \sum_{i=1}^T \lambda_i \omega_i \omega_i', \tag{7}$$

where we have without loss of generality arranged the eigenvalues  $\lambda_i$  in increasing order, and

$$V = \sum_{i=1}^T \frac{\sigma_\varepsilon^2}{(1 - \rho\lambda_i)^2} \omega_i \omega_i' \tag{8}$$

is the resulting spectral decomposition of  $V$ , which always exists as  $V$  is symmetric. The point of our argument now is that

$$\lim_{\rho \rightarrow 1/\lambda_{max}} (1 - \rho\lambda_{max})^2 V = \sigma_\varepsilon^2 \omega_T \omega_T', \tag{9}$$

a matrix of rank 1. Therefore, all limiting eigenvalues of

$$(1 - \rho\lambda_{max})^2 V(M^{(2)} - dM) \quad (10)$$

are zero except one, which is given by

$$c_T = tr(\omega_T \omega_T' (M^{(2)} - dM)) = \omega_T' (M^{(2)} - dM) \omega_T. \quad (11)$$

This constant  $c_T$  is crucial for our analysis. It determines whether the F-test will eventually be conservative or liberal. If  $c_T$  is positive, the rejection probability of the F-test will tend to 1 as  $\rho$  approaches  $1/\lambda_{max}$ . The test is then liberal in the extreme, at least for values of  $\rho$  close to the edge of the parameter space.

If  $c_T$  is negative, the rejection probability will tend to zero, and the test will eventually be extremely conservative. And if  $c_T = 0$ , the limiting behaviour of the test cannot be determined from the limiting behaviour of the eigenvalues of  $(1 - \rho\lambda_{max})^2 V(M^{(2)} - dM)$  (which are all zero). Section 3 now sheds some light on which of these cases is more likely to occur in empirical applications.

### 3 Exact rejection probabilities in finite samples

The first important point to make is that the crucial constant  $c_T$  depends only on  $X$  and  $W$  and the nominal size of the test, all of which are known. Therefore,  $c_T$  is known as well and can guide the user in interpreting a test: If  $c_T < 0$ , one has to beware of a loss in power, and if  $c_T > 0$ , one has to beware of spurious rejections.

The following argument shows that the first problem is far more likely to occur in practice: Rewrite the critical constant as

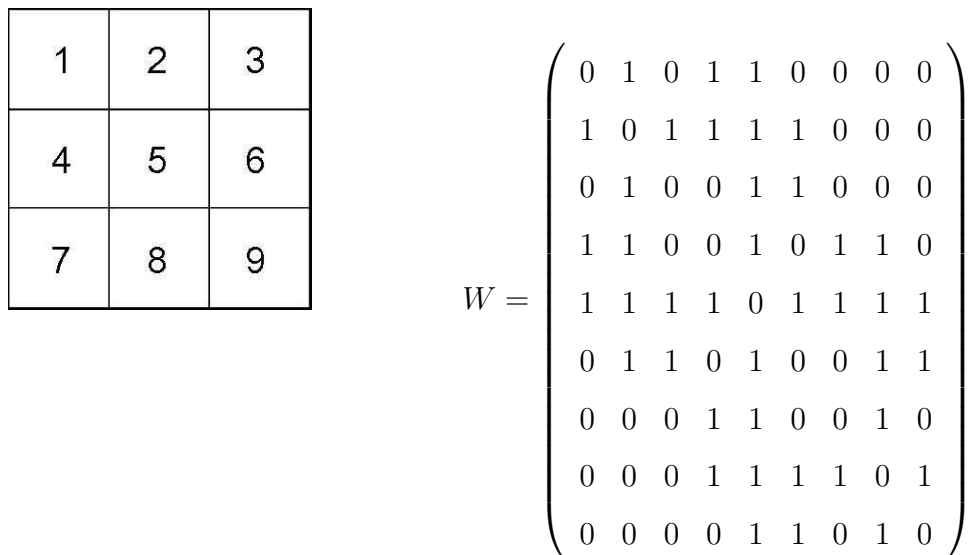
$$c_T = \omega_T' M^{(2)} \omega_T - \omega_T' \frac{T - K + q}{T - K} F_{q, T-K}^\alpha M \omega_T. \quad (12)$$

Then it is easily seen that in general  $c_T < 0$  (i.e. except in very contrived cases). This follows from the fact that

$$\frac{T - K + q}{T - K} F_{q, T-K}^\alpha \rightarrow \chi_q^{2, \alpha} / q \quad (13)$$

as  $T \rightarrow \infty$ , which is larger than 2 for moderate values of  $\alpha$  and  $q$ . (It takes the values 3.84, 2.99 and 2.60 for  $\alpha = 0.05$  and  $q = 1, 2$  und 3, respectively). This will in general be more than enough to counterbalance the fact that  $\omega_T' M^{(2)} \omega_T > \omega_T' M \omega_T$ .

Figure 1: An example of the queen matrix



Of course one can always construct a weighting matrix and regressor matrices  $W$ ,  $X$  and  $X^{(2)}$  such that  $\omega_T' M^{(2)} \omega_T = 1$  and  $\omega_T' M \omega_T = 0$  and therefore  $c_T > 0$ . For instance, let  $\iota = (1, \dots, 1)'$  be a  $(T \times 1)$ -vector and choose  $X^{(2)}$  orthogonal to  $\iota$ . E.g., for  $T$  even, pick  $X^{(2)} = (1, -1, 1, -1, \dots, 1, -1)'$ . Let

$$X = [\iota : X^{(2)}] \tag{14}$$

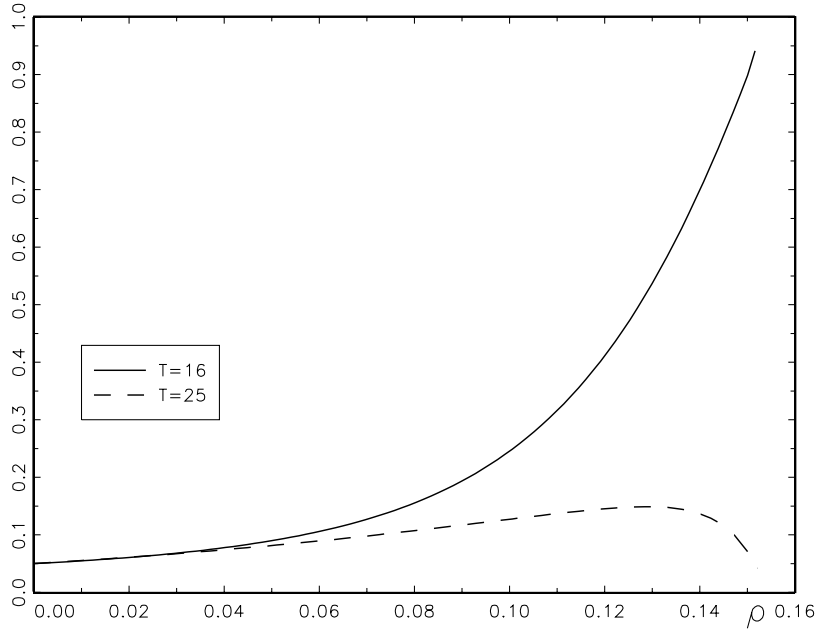
(that is, test  $H_0 : \beta^{(1)} = 0$ ) and let

$$W = W^{EW} = (w_{ij}^{EW}) = \begin{cases} 1 & \text{for } i \neq j \\ 0 & \text{for } i = j \end{cases}, \tag{15}$$

the equal weight matrix. This and similar cases will however rarely happen in the natural course of events, and will become ever more unlikely as sample size increases.

Figure 2 gives an example where  $W$  is derived from the queen-criterion (see Figure 1 for an illustration of the criterion with  $N = 9$ ): There is a square of cells, and all cells around a given cell obtain a weight of 1. The sample size is then a square number. We choose  $X$  to have  $K = 2$  and  $T = 16$  or  $25$  such that, for  $T = 16$ , the second column is the (normalized) eigenvector corresponding to the largest eigenvalue of  $W$ , (which happens to be  $\lambda_{max} = 5.85$ ), and the first column is any  $(16 \times 1)$ -vector

Figure 2: Rejection probabilities for the queen matrix



orthogonal to  $\omega_{16}$ . For  $t > 16$ ,  $x_{t_1} = 1$  and  $x_{t_2} = t - 16$ . Then we have  $c_{16} = 0.316$  and  $c_{25} = -0.850$ , and so our theoretical result predicts that the rejection probabilities will tend to one as  $\rho \rightarrow 1/\lambda_{max}$  for  $T = 16$  and will tend to zero as  $\rho \rightarrow 1/\lambda_{max}$  for  $T = 25$ . Figure 1 shows that this is indeed the case.

The case  $c_T = 0$ , where our analysis does not apply, will occur for instance whenever  $\omega_T$  is in the column space of  $X^{(2)}$ . The most important special case is when  $W$  is row-normalized and therefore  $\omega_T = \frac{1}{\sqrt{T}}(1, \dots, 1)'$  and where in addition  $X^{(2)}$  contains an intercept. However, row-normalization will often destroy the symmetry of  $W$ , so this case is not covered by our discussion above.

## References

- Anselin, L. and Florax, R. (eds.) (1995):** *New directions in spatial econometrics*. Berlin (Springer).
- Banerjee, A. N. and Magnus, J. (2000):** “On the sensitivity of the usual t- and F-tests to covariance misspecifications.” *Journal of Econometrics* 95, 157 – 176.
- Hillier, O.H. and King, M.L (1982):** “Linear regression with correlated errors: bounds on coefficient estimates and t-values” In: King, M. L. and Giles, D. E. A. (editors): *Specification Analysis in the Linear Model*. London (Routledge and Kegan-Paul), 74 – 80.
- Kiviet, J.F. (1980):** “Effects of ARMA errors on tests for regression coefficients: comments on Vinod’s article, improved and additional results.” *Journal of the American Statistical Association* 75, 333 – 358.
- Krämer, W. (1989):** “On the robustness of the F-test to autocorrelation among disturbances.” *Economics Letters* 30, 37 – 40.
- Krämer, W. (2003):** “The robustness of the F-test to spatial autocorrelation among regression disturbances” *Statistica* 63, 435-440.
- Krämer, W.; Kiviet, J. and Breitung, J. (1990):** “The null distribution of the *F*-test in the linear regression model with autocorrelated disturbances.” *Statistica* 50, 503 – 509.
- Vinod, H.D. (1976):** “Effects of ARMA errors on the significance tests for regression coefficients.” *Journal of the American Statistical Association* 71, 929 – 933.