Econometric Theory, **19**, 2003, 744–753. Printed in the United States of America.
DOI: 10.1017/S0266466603195023

FINITE-SAMPLE INSTRUMENTAL VARIABLES INFERENCE USING AN ASYMPTOTICALLY PIVOTAL STATISTIC

PAUL BEKKER
University of Groningen

FRANK KLEIBERGEN
University of Amsterdam

We consider the K-statistic, Kleibergen's (2002, *Econometrica* 70, 1781–1803) adaptation of the Anderson–Rubin (AR) statistic in instrumental variables regression. Whereas Kleibergen (2002) especially analyzes the asymptotic behavior of the statistic, we focus on finite-sample properties in a Gaussian framework. The AR statistic then has an *F*-distribution. The finite-sample distribution of the K-statistic is, however, affected by nuisance parameters. We consider two extreme cases for the nuisance parameters, which provide tight bounds for the exact distribution. The first case amounts to perfect identification—which is similar to the asymptotic case—where the statistic has an *F*-distribution. In the other extreme case there is total underidentification. For the latter case we show how to compute the exact distribution. We thus provide tight bounds for exact confidence sets based on the K-statistic. Asymptotically the two bounds converge, except when there is a large number of redundant instruments.

1. INTRODUCTION

Instrumental variables inference can be affected rather strongly by the quality of the instruments. Especially in the presence of weak instruments usual inferential procedures such as likelihood-based Wald, likelihood ratio, and Lagrange multiplier tests may show considerable size distortions, and the performance of confidence sets may be abysmally poor (see, e.g., Nelson and Startz, 1990; Staiger and Stock, 1997; Zivot, Startz, and Nelson, 1998; Hausman and Hahn, 1999). However, an exact test in a Gaussian context has been provided by Anderson and Rubin (1949). The resulting Anderson–Rubin (AR) statistic is pivotal and has an *F*-distribution, which does not depend on nuisance parameters and is not affected by the degree of underidentification.

The authors' research documented in this paper has been funded by the NWO Vernieuwingsimpuls research grant "Empirical Comparison of Economic Models." Address correspondence to: Paul Bekker, Department of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands; e-mail: P.A.Bekker@eco.rug.nl; and Frank Kleibergen, Department of Quantitative Economics, University of Amsterdam, Roeterstraat 11, 1018 WB, The Netherlands; e-mail: kleiberg@fee.uva.nl.

 © 2003 Cambridge University Press 0266-4666/03 $12.00

https://doi.org/10.1017/S0266466603195023 Published online by Cambridge University Press

However, the AR statistic has a limiting chi-square distribution with a number of degrees of freedom that equals the number of instruments. This number exceeds, or equals, the number of structural parameters, which affects the power of the test statistic. As has been shown by Kleibergen (2002) it is possible to construct a statistic, the K-statistic, with similar asymptotic pivotal properties but with a limiting chi-square distribution that has a number of degrees of freedom equal to the number of structural parameters. Thus, the K-statistic has an asymptotic distribution with a minimal number of degrees of freedom without being hampered by a poor performance close to points of underidentification, as is the case for the more traditional inference procedures. This conclusion can be drawn with respect to the asymptotic behavior of the test statistic. Kleibergen (2002) does not provide finite-sample evidence. Here we do consider finite-sample properties.

Under Gaussianity the AR statistic has an F-distribution and thus provides an exact finite-sample test. The K-statistic has a finite-sample distribution that depends on nuisance parameters. It thus does not provide an exact test. Based on the asymptotic pivotal behavior, we, however, expect the asymptotic distribution to provide an accurate approximation of the exact finite-sample distribution, but we do not know the degree of accuracy. Therefore, we compute bounds for the exact distribution of the K-statistic by assuming Gaussian distributions for the disturbances.

We find on one extreme—provided by the case of perfect identification—an F-distribution, similar to the distribution of the AR statistic, albeit with fewer degrees of freedom. On the other extreme—provided by the case of total underidentification—we find a distribution that is complicated analytically but can be simulated easily. That is, both extreme distributions can be computed in practice. We show that these extreme cases provide tight bounds. In many practical cases the two bounds are very close indeed, as we might expect based on the asymptotic pivotal property of the statistic. However, in the case of many redundant instruments, the bounds may differ even asymptotically. In fact, we also compute the limiting distribution in this case by means of the many-instruments asymptotics as described in Bekker (1994). We find it to be different from the large-sample chi-square distribution.

The paper is organized as follows. In Section 2 we describe the model and discuss the AR statistic. The K-statistic is briefly discussed in Section 3 alongside the construction of its bounding distributions. Section 4 provides some Monte Carlo computations.

2. THE MODEL AND THE ANDERSON–RUBIN STATISTIC

Consider a classical instrumental variables regression model in a cross-section context. That is, let

$$y = X\beta + \varepsilon,$$
$$X = Z\Pi + V, \qquad \textbf{(1)}$$

https://doi.org/10.1017/S0266466603195023 Published online by Cambridge University Press

where Z is an $n \times k$ matrix of full column rank that consists of the nonstochastic instrumental variables and X is $n \times m$, which may contain endogenous in addition to exogenous explanatory variables. The latter are assumed to be columns of Z also and $m \leq k$. If interest is restricted to those elements of the parameter vector β that relate to the endogenous explanatory variables, the instruments in Z might be separated into two parts. For ease of exposition we will not do this and consider the m-vector β as a whole. We assume the rows of (ε, V) to be independently normally distributed with zero mean; we denote the covariance matrix by Σ.

Let in general for a matrix H of full column rank the projection matrix be denoted by $P_H = H(H'H)^{-1}H'$. Consider the quantity

$$\mathrm{AR}(\beta) = \frac{(y - X\beta)'P_Z(y - X\beta)}{(y - X\beta)'(I_n - P_Z)(y - X\beta)}. \tag{2}$$

Minimizing $\mathrm{AR}(\beta)$ over β provides the limited information maximum likelihood estimator. However, we are not concerned with estimation but with testing. If the argument in (2) equals the true β, then $y - X\beta = \varepsilon$ and the numerator and denominator of (2) have independent chi-square distributions with k and $n - k$ degrees of freedom, respectively. So, when multiplied by $(n - k)/k$, $\mathrm{AR}(\beta)$ has an $F_{k,n-k}$ distribution. An exact test of H_0: $\beta = \beta^\star$ is found by verifying whether or not $\mathrm{AR}(\beta^\star)$ has a small enough p-value in the $F_{k,n-k}$ distribution. The resulting test is exact and known from Anderson and Rubin (1949). Notice that the AR-statistic has an F-distribution whether or not the model is identified. That is, even if Π has a deficient column rank $\mathrm{AR}(\beta)$ will be distributed as an $F_{k,n-k}$ random variable. This shows that it is a pivotal statistic.

Although the AR-test is exact under Gaussianity, it may have poor power if the number of instruments k exceeds the number of explanatory variables m. That is, in the numerator of (2), $y - X\beta$ is projected on the full space spanned by Z. We know, however, that, if β deviates from the true value, the mean of $y - X\beta$ is located in the subspace spanned by $Z\Pi$. The problem is then that we do not know Π. Still, the dimension of the space onto which we project $y - X\beta$ can be reduced in a way that preserves the pivotal property of the statistic asymptotically.

3. BOUNDING THE EXACT DISTRIBUTION OF THE K-STATISTIC

Consider a regression of the columns of V on ε. That is, let

$$V = \varepsilon\lambda' + W,$$

$$\lambda = \frac{\sigma_{V,\varepsilon}}{\sigma_\varepsilon^2}, \tag{3}$$

https://doi.org/10.1017/S0266466603195023 Published online by Cambridge University Press

where W is independent of ε and where $\sigma_{V,\varepsilon}$ and σ_ε^2 denote the covariance of V and ε and the variance of the latter, respectively. Consequently, $X - \varepsilon\lambda' = Z\Pi + W$ is like Z independent of ε but has a smaller column dimension. If λ were known, we could replace the projection matrix P_Z in the numerator of (2) by $P_{\widetilde{X}}$, with

$$\widetilde{X}(\lambda) = P_Z(X - (y - X\beta)\lambda'). \tag{4}$$

For the true value of β, $\widetilde{X}(\lambda)$ is stochastically independent of ε. Thus, this adaptation of the AR statistic would still be pivotal but now distributed, when multiplied by $(n-k)/m$, as $F_{m,n-k}$. Of course, λ is unknown. However, Kleibergen (2002) simply uses a consistent estimator of λ that is stochastically independent of $Z'y$ and $Z'X$ under the null where β is specified. This estimator, $\hat{\lambda}$, is specified as in (3) with $\sigma_{V,\varepsilon}$ and σ_ε^2 replaced by

$$\hat{\sigma}_{V,\varepsilon} = \frac{X'(I_n - P_Z)(y - X\beta)}{n-k},$$

$$\hat{\sigma}_\varepsilon^2 = \frac{(y - X\beta)'(I_n - P_Z)(y - X\beta)}{n-k}. \tag{5}$$

The resulting K-statistic is given by

$$K(\beta) = \frac{(y - X\beta)'P_{\widetilde{X}(\hat{\lambda})}(y - X\beta)}{(y - X\beta)'(I_n - P_Z)(y - X\beta)/(n-k)}. \tag{6}$$

Asymptotically, it is distributed as a chi square with m degrees of freedom. Further discussions are given by Kleibergen (2002).

The denominator of the K-statistic (6) is given by $\hat{\sigma}_\varepsilon^2$, so it results from projecting the first column of the following matrix onto the space that is spanned by its last m columns:

$$P_Z(y,X)\begin{pmatrix} 1 & 0 \\ -\beta & I_m \end{pmatrix}\begin{pmatrix} \hat{\sigma}_\varepsilon^{-1} & -\dfrac{\hat{\sigma}'_{V,\varepsilon}}{\hat{\sigma}_\varepsilon^2} \\ 0 & I_m \end{pmatrix}$$

$$= (0, Z\Pi) + P_Z(\varepsilon, V)\begin{pmatrix} \hat{\sigma}_\varepsilon^{-1} & -\dfrac{\hat{\sigma}'_{V,\varepsilon}}{\hat{\sigma}_\varepsilon^2} \\ 0 & I_m \end{pmatrix}. \tag{7}$$

Theorem 1 states an alternative expression of the K-statistic that results from (7).

https://doi.org/10.1017/S0266466603195023 Published online by Cambridge University Press

THEOREM 1. *An alternative expression of the K-statistic (6) is*

$$K(\beta) = \frac{\tilde{\varepsilon}' P_{\widetilde{X}} \tilde{\varepsilon}}{\hat{\sigma}_{\tilde{\varepsilon}}^2},$$

$$\widetilde{X} = P_Z \left(\widetilde{W} - \tilde{\varepsilon} \frac{\hat{\sigma}'_{\widetilde{W},\tilde{\varepsilon}}}{\hat{\sigma}_{\tilde{\varepsilon}}^2} \right) + Z\Pi\Sigma_W^{-1/2}, \tag{8}$$

where the matrix $(\tilde{\varepsilon}, \widetilde{W})$ *is defined by*

$$(\tilde{\varepsilon}, \widetilde{W}) = \left(\frac{\varepsilon}{\sigma_\varepsilon}, W\Sigma_W^{-1/2} \right),$$

with W as in (3) and

$$\hat{\sigma}_{\widetilde{W},\tilde{\varepsilon}} = \frac{\widetilde{W}'(I_n - P_Z)\tilde{\varepsilon}}{n-k}, \qquad \hat{\sigma}_{\tilde{\varepsilon}}^2 = \frac{\tilde{\varepsilon}'(I_n - P_Z)\tilde{\varepsilon}}{n-k}.$$

Proof. See the Appendix.

If the matrix $\Pi\Sigma^{-1/2}$ from Theorem 1 were known, the distribution of $K(\beta)$ could be simulated, because $(\tilde{\varepsilon}, \widetilde{W})$ can be simulated, as their elements are independent standard normal distributed. However, the matrix $\Pi\Sigma_W^{-1/2}$ is not known. Still we can distinguish between two extreme cases.

THEOREM 2. *When* $\Pi = \theta\Pi^\star$, *where* $\Pi^\star$ *is a fixed finite* $k \times m$ *matrix and* θ *is a scalar,* $K(\beta)$ *is a stochastically decreasing function of* $|\theta|$, *so its finite-sample distribution moves to the left when* $|\theta|$ *increases. Hence, the finite-sample distribution of* $K(\beta)$ *is bounded from above by the distribution that results in the case of a zero value of* θ *and from below by the distribution that results in the case of an infinite value of* $|\theta|$.

Proof. See the Appendix.

The bounds on the finite-sample distribution of $K(\beta)$ in Theorem 2 result from extremal values of the nuisance parameters in the expression of the exact finite-sample distribution of $K(\beta)$. Hence, they do not result from bounding the statistic of interest directly as, for example, in Sargan and Bhargava (1983). Sargan and Bhargava (1983) obtain upper and lower bounds on the distribution of their statistic of interest by constructing lower and upper bounds of the statistic itself using properties of ratios of quadratic forms. We obtain the bounds in an entirely different manner and construct them from the finite-sample distribution.

The upper bound on the finite-sample distribution of $K(\beta)$ results from the extreme case of total underidentification: $\Pi = 0$. In this case the distribution can be simulated easily. The distribution depends only on the dimension param-

https://doi.org/10.1017/S0266466603195023 Published online by Cambridge University Press

eters k, m, and n. In the case of an infinite value of θ, which forms the lower bound on the finite-sample distribution of $K(\beta)$, $P_{\bar{X}}$ converges in probability to $P_{Z\Pi}$, and so the distribution of $K(\beta)/m$ converges to $F_{m,n-k}$.

The preceding discussion implies that we can compute bounds for the finite-sample distribution of the K-statistic using simulation. Conservative tests or confidence sets can be based on the critical value found for the distribution simulated at $\Pi = 0$. However, if n is large then $\hat{\sigma}_{\hat{W},\hat{\varepsilon}}$ is small, in probability, and the correlation between $\tilde{\varepsilon}$ and $\tilde{X}$ becomes negligible. In that case the bounds are very close. Indeed, if $n \to \infty$, then $K(\beta) \overset{A}{\sim} \chi^2_m$ (see Kleibergen, 2002).

Would this limit distribution also provide an accurate approximation if there are many redundant instruments? To answer this question we consider the limit distribution of $K(\beta)$ for an asymptotic parameter sequence where the number of instruments increases with the number of observations: $k/n \to \alpha > 0$.

For the lower bound we found a finite-sample distribution given by $F_{m,n-k}$. Clearly, the lower bound converges to a χ^2_m-distribution whether or not $\alpha > 0$. For the upper bound, where $\Pi = 0$, the results are summarized in Theorem 3.

THEOREM 3. *The limiting distribution of the K-statistic when* $\Pi = 0$ *and* $k/n \to \alpha > 0$ *is characterized by*

$$(1 - k/n)K(\beta) \overset{A}{\sim} \chi^2_m. \tag{9}$$

Proof. See the Appendix.

If $\alpha > 0$ we thus find a simple difference between the asymptotic upper and lower bounds. The difference is of order k/n. These practical results are confirmed numerically in the next section.

4. MONTE CARLO RESULTS

To determine the applicability of the bounds on the exact distribution of $K(\beta)$, we construct the bounds for $m = 1$ and some values of k and n. Figures 1 and 2 show the lower and upper bounds on the exact distribution function of $K(\beta)/m$. Figure 2 shows that there is a considerable difference between the lower and upper bounds when k/n is large. Figure 1 shows that when k/n gets smaller, the difference decreases and becomes negligible, as expected. The figures also contain a limiting approximation of the upper bound. Instead of the asymptotic approximation χ^2_m/m, based on (15), we used the $F_{m,n-k}$ distribution, which is more accurate in small samples and more convenient too, because it is also used for the lower bound. The figures show that this limiting approximation is accurate even for small values of n. Hence, it suffices to use critical values from the $F_{m,n-k}$ distribution both as lower bound and, by multiplying them by $1/(1 - k/n)$, as upper bound of the exact critical values of $K(\beta)/m$.

https://doi.org/10.1017/S0266466603195023 Published online by Cambridge University Press

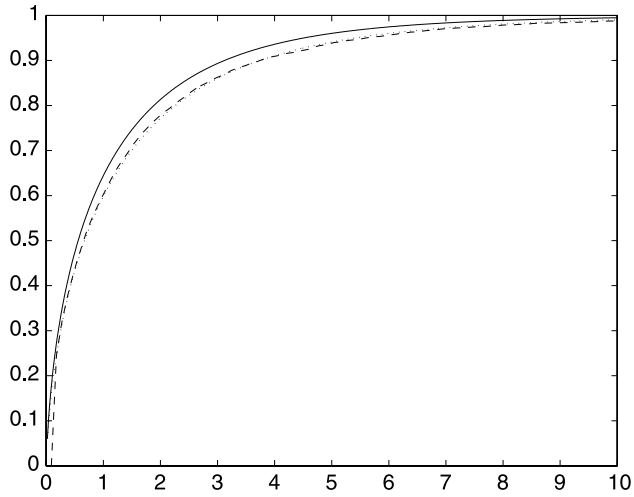


FIGURE 1. Lower bound (—) and upper bound (- -) and limiting approximation of upper bound (...) of the distribution function of $K(\beta)/m$ for $n = 25$, $k = 5$, and $m = 1$.

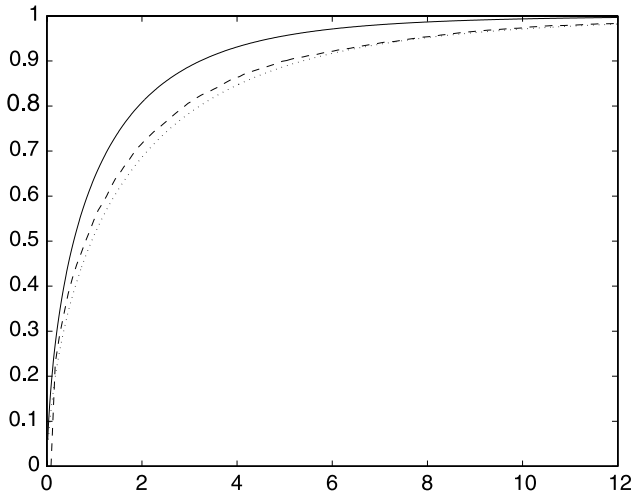


FIGURE 2. Lower bound (—) and upper bound (- -) and limiting approximation of upper bound (...) of the distribution function of $K(\beta)/m$ for $n = 25$, $k = 10$, and $m = 1$.

https://doi.org/10.1017/S0266466603195023 Published online by Cambridge University Press

REFERENCES

Anderson, T.W. & H. Rubin (1949) Estimators on the parameters of a single equation in a complete set of stochastic equations. *Annals of Mathematical Statistics* 21, 570–582.

Bekker, P.A. (1994) Alternative approximations to the distributions of instrumental variable estimators. *Econometrica* 62, 657–681.

Hausman, J. & J. Hahn (1999) A New Specification Test for the Validity of Instrumental Variables. Working paper, Massachusetts Institute of Technology.

Kleibergen, F. (2002) Pivotal statistics for testing structural parameters in instrumental variables regression. *Econometrica* 70, 1781–1803.

Nelson, C.R. & R. Startz (1990) Some further results on the exact small sample properties of the instrumental variables estimator. *Econometrica* 58, 967–976.

Sargan, J.D. & A. Bhargava (1983) Testing residuals from least squares regression for being generated by the Gaussian random walk. *Econometrica* 51, 153–174.

Staiger, D. & J.H. Stock (1997) Instrumental variables regression with weak instruments. *Econometrica* 65, 557–586.

Zivot, E., R. Startz, & C.R. Nelson (1998) Valid confidence intervals and inference in the presence of weak instruments. *International Economic Review* 39, 1119–1144.

APPENDIX

Proof of Theorem 1. The second term in (7) can be expressed as

$$P_Z(\varepsilon, V)\begin{pmatrix} \hat{\sigma}_\varepsilon^{-1} & -\dfrac{\hat{\sigma}'_{V,\varepsilon}}{\hat{\sigma}_\varepsilon^2} \\ 0 & I_m \end{pmatrix} = P_Z(\tilde{\varepsilon}, \widetilde{W})\begin{pmatrix} \sigma_\varepsilon & \dfrac{\sigma'_{V,\varepsilon}}{\sigma_\varepsilon} \\ 0 & \Sigma_W^{1/2} \end{pmatrix}\begin{pmatrix} \hat{\sigma}_\varepsilon^{-1} & -\dfrac{\hat{\sigma}'_{V,\varepsilon}}{\hat{\sigma}_\varepsilon^2} \\ 0 & I_m \end{pmatrix}$$

$$= P_Z(\tilde{\varepsilon}, \widetilde{W})\begin{pmatrix} \left(\dfrac{\hat{\sigma}_\varepsilon}{\sigma_\varepsilon}\right)^{-1} & \dfrac{\sigma'_{V,\varepsilon}}{\sigma_\varepsilon} - \dfrac{\sigma_\varepsilon \hat{\sigma}'_{V,\varepsilon}}{\hat{\sigma}_\varepsilon^2} \\ 0 & \Sigma_W^{1/2} \end{pmatrix}$$

$$= P_Z(\tilde{\varepsilon}, \widetilde{W})\begin{pmatrix} \hat{\sigma}_{\tilde{\varepsilon}}^{-1} & -\dfrac{\hat{\sigma}'_{\widetilde{W},\tilde{\varepsilon}}}{\hat{\sigma}_{\tilde{\varepsilon}}^2} \\ 0 & I_m \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \Sigma_W^{1/2} \end{pmatrix},$$

where the last equality follows from

$$\hat{\sigma}_{\widetilde{W},\tilde{\varepsilon}} = \frac{\Sigma_W^{-1/2}}{\sigma_\varepsilon}\hat{\sigma}_{W,\varepsilon}$$

$$= \frac{\Sigma_W^{-1/2}}{\sigma_\varepsilon}\left(\hat{\sigma}_{V,\varepsilon} - \frac{\hat{\sigma}_\varepsilon^2 \sigma_{V,\varepsilon}}{\sigma_\varepsilon^2}\right),$$

$$\hat{\sigma}_{\tilde{\varepsilon}}^2 = \frac{\hat{\sigma}_\varepsilon^2}{\sigma_\varepsilon^2}.$$

https://doi.org/10.1017/S0266466603195023 Published online by Cambridge University Press

Consequently, the K-statistic $K(\beta)$ is given by

$$K(\beta) = \frac{\tilde{\varepsilon}' P_{\bar{X}} \tilde{\varepsilon}}{\hat{\sigma}_{\tilde{\varepsilon}}^2},$$

$$\bar{X} = P_Z\left(\bar{W} - \tilde{\varepsilon}\,\frac{\hat{\sigma}'_{\bar{W},\tilde{\varepsilon}}}{\hat{\sigma}_{\tilde{\varepsilon}}^2}\right) + Z\Pi\Sigma_W^{-1/2}. \quad \blacksquare$$

Proof of Theorem 2. To prove that $K(\beta)$ is stochastically decreasing in $|\theta|$, where $\Pi = \theta\Pi^\star$, we first condition on $(I_n - P_Z)(\tilde{\varepsilon}, \bar{W})$. This does not affect the distribution of $P_Z(\tilde{\varepsilon}, \bar{W})$. Let

$$H = (Z'Z)^{1/2}\Pi^\star\Sigma_W^{-1/2},$$

$$u = (Z'Z)^{-1/2}Z'\tilde{\varepsilon},$$

$$U = (Z'Z)^{-1/2}Z'(\bar{W} - \tilde{\varepsilon}\tilde{\lambda}'),$$

$$\tilde{\lambda} = \frac{\hat{\sigma}_{\bar{W},\tilde{\varepsilon}}}{\hat{\sigma}_{\tilde{\varepsilon}}^2}.$$

We may regress u on U so that $u = U\gamma + \nu$, where ν is independent of U. Next we also condition on U and consider the numerator of $K(\beta)$ given by

$$\tilde{\varepsilon}' P_{\bar{X}} \tilde{\varepsilon} = (U\gamma + \nu)' P_L (U\gamma + \nu),$$

$$L = H\theta + U.$$

Let Q be an orthogonal $m \times m$ dimensional matrix such that $U\gamma$ is orthogonal to the last $m - 1$ columns of $L(L'L)^{-1/2}Q$, such that it has a positive inner product with the first column: $(\gamma' U' P_L U\gamma)^{1/2}$. Furthermore let $\xi = Q'(L'L)^{-1/2}L'\nu$, whose elements are independently normally distributed; then

$$\begin{aligned}\tilde{\varepsilon}' P_{\bar{X}} \tilde{\varepsilon} &= (U\gamma + \nu)' L (L'L)^{-1/2} Q Q' (L'L)^{-1/2} L' (U\gamma + \nu) \\ &= [Q'(L'L)^{-1/2}L'(U\gamma + \nu)]'[Q'(L'L)^{-1/2}L'(U\gamma + \nu)] \\ &= (\xi_1 + (\gamma' U' P_L U\gamma)^{1/2})^2 + \sum_{i=2}^{m} \xi_i^2.\end{aligned}$$

Consequently, we only have to show that, conditional on U, $(\xi_1 + (\gamma' U' P_L U\gamma)^{1/2})^2$ stochastically decreases as $|\theta|$ increases. Here we find an analogy with the noncentral χ^2 distribution that decreases as the noncentrality parameter decreases. So we only have to prove that the nonrandom function

$$\gamma' U' (H\theta + U)\{(H\theta + U)'(H\theta + U)\}^{-1}(H\theta + U)' U\gamma$$

decreases as $|\theta|$ increases. This can be shown to hold true by considering the derivative with respect to θ, which completes the proof. $\blacksquare$

https://doi.org/10.1017/S0266466603195023 Published online by Cambridge University Press

Proof of Theorem 3. We use the following limits, which can be easily verified:

$$\hat{\sigma}_{\tilde{\varepsilon}}^2 \xrightarrow{p} 1, \qquad \hat{\sigma}_{\widetilde{W},\tilde{\varepsilon}} \xrightarrow{p} 0, \qquad \widetilde{W}'P_Z\widetilde{W}/k \xrightarrow{p} I_m,$$

$$\tilde{\varepsilon}'P_Z\tilde{\varepsilon}/k \xrightarrow{p} 1, \qquad \text{and} \quad \widetilde{W}'P_Z\tilde{\varepsilon}/k \xrightarrow{p} 0,$$

so

$$\widetilde{X}'\widetilde{X}/k \xrightarrow{p} I_m.$$

Furthermore,

$$\widetilde{W}'P_Z\tilde{\varepsilon}/(k^{1/2}) \xrightarrow{A} \mathcal{N}(0, I_m),$$

$$(n-k)^{1/2}\hat{\sigma}_{\widetilde{W},\tilde{\varepsilon}} \xrightarrow{A} \mathcal{N}(0, I_m),$$

so

$$\widetilde{X}'\tilde{\varepsilon}/(k^{1/2}) \xrightarrow{A} \mathcal{N}(0, (1-\alpha)^{-1}I_m).$$

Consequently,

$$(1-\alpha)K(\beta) \xrightarrow{A} \chi_m^2. \qquad \blacksquare$$

https://doi.org/10.1017/S0266466603195023 Published online by Cambridge University Press