Finite-Sample Instrumental Variables Inference using an Asymptotically Pivotal Statistic

Paul A. Bekker* and Frank Kleibergen†
University of Groningen University of Amsterdam

June, 2001

Abstract

The paper considers the K-statistic, Kleibergen's (2000) adaptation of the Anderson-Rubin (AR) statistic in instrumental variables regression. Compared to the AR-statistic this K-statistic shows improved asymptotic efficiency in terms of degrees of freedom in overidentified models and yet it shares, asymptotically, the pivotal property of the AR statistic. That is, asymptotically it has a chi-square distribution whether or not the model is identified. This pivotal property is very relevant for size distortions in finite-sample tests. Whereas Kleibergen (2000) focuses especially on the asymptotic behavior of the statistic, the present paper concentrates on finite-sample properties in a Gaussian framework. In that case the AR statistic has an F-distribution. However, the K-statistic is not exactly pivotal. Its finite-sample distribution is affected by nuisance parameters. Here we consider the two extreme cases, 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. Thus we provide tight bounds for exact confidence sets based on the efficient K-statistic. Asymptotically the two bounds converge, except when there is a large number of redundant instruments.

*Department of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands, E-mail: P.A.Bekker@eco.rug.nl

†Department of Quantitative Economics, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands, E-mail: kleiberg@fee.uva.nl.

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 *et. al.* (1998), and 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.

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 (2000) 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 (2000) 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 does thus 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-test 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 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 briefly the AR statistic. The K-statistic is discussed in Section 3. The computation of its distribution, based on simulations, is described in Section 4. The proof of the tightness of the bounds is given in an Appendix. Section 5 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

$$y = X\beta + \varepsilon, \tag{1}$$

$$X = Z\Pi + V, \tag{2}$$

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 as well as exogenous explanatory variables. The latter are assumed to be columns of Z as well 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

$$\mathrm{AR}(\beta) = \frac{(y - X\beta)' P_Z (y - X\beta)}{(y - X\beta)'(I_n - P_Z)(y - X\beta)}. \tag{3}$$

Minimizing $\mathrm{AR}(\beta)$ over β provides the LIML estimator. However, we are not concerned with estimation but with testing. If the argument in (3) equals the true β, then $y - X\beta = \varepsilon$ and the numerator and denominator of (3) 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 (3), $y - X\beta$ 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 Removing redundant degrees of freedom from the AR Statistic

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

$$V = \varepsilon\lambda' + W, \tag{4}$$

$$\lambda = \frac{\sigma_{V,\varepsilon}}{\sigma_\varepsilon^2}, \tag{5}$$

where W is independent of ε; and $\sigma_{V,\varepsilon}$ 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 (3) by $P_{\tilde{X}}$, with

$$\tilde{X}(\lambda) = P_Z(X - (y - X\beta)\lambda'). \tag{6}$$

For the true value of β, $\tilde{X}(\lambda)$ is stochastic 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 (2000) uses simply a consistent estimator of λ that is stochastic independent of $Z'y$ and $Z'X$ under the null where β is specified. This estimator, $\hat{\lambda}$, is specified as in (4) with $\sigma_{V,\varepsilon}$ and σ_ε^2 replaced by

$$\begin{aligned} \hat{\sigma}_{V,\varepsilon} &= \frac{X'(I_n - P_Z)(y - X\beta)}{n-k}, \\ \hat{\sigma}_\varepsilon^2 &= \frac{(y - X\beta)'(I_n - P_Z)(y - X\beta)}{n-k}. \end{aligned} \tag{7}$$

The resulting K-statistic is given by

$$K(\beta) = \frac{(y - X\beta)' P_{\tilde{X}(\hat{\lambda})}(y - X\beta)}{(y - X\beta)'(I_n - P_Z)(y - X\beta)/(n-k)}. \tag{8}$$

Asymptotically, it is distributed as a chi square with m degrees of freedom. Further discussions are given by Kleibergen (2000).

4 Bounding the Finite-Sample Distribution

Although the asymptotic distribution of (8) is simple, the finite-sample distribution is more complicated. In fact, it depends on nuisance parameters. In this section we will analyse this dependency and show that the distribution can be bounded by distribution functions that do not depend on unknown parameters.

As the denominator of (8) is given by $\hat{\sigma}_\varepsilon^2$, we find that the statistic (8) results from projecting the first column of the following matrix onto the space that is spanned by its last m columns:

$$P_Z\,(y, X) \begin{pmatrix} 1 & 0 \\ -\beta & I_m \end{pmatrix} \begin{pmatrix} \widehat{\sigma}_\varepsilon^{-1} & -\frac{\widehat{\sigma}'_{V,\varepsilon}}{\widehat{\sigma}_\varepsilon^2} \\ 0 & I_m \end{pmatrix} = (0, Z\Pi) + P_Z\,(\varepsilon, V) \begin{pmatrix} \widehat{\sigma}_\varepsilon^{-1} & -\frac{\widehat{\sigma}'_{V,\varepsilon}}{\widehat{\sigma}_\varepsilon^2} \\ 0 & I_m \end{pmatrix}. \tag{9}$$

In order to be able to simulate the distribution of $K(\beta)$, we define a matrix $(\widetilde{\varepsilon}, \widetilde{W})$ as follows

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

where W has been defined in (4) and Σ_W denotes the covariance matrix of its rows. We note that the matrix $(\widetilde{\varepsilon}, \widetilde{W})$ has independent standard normally distributed elements. We define their estimated variance and covariance by

$$\begin{aligned}
\widehat{\sigma}_{\widetilde{W},\widetilde{\varepsilon}} &= \frac{\widetilde{W}'(I_n - P_Z)\widetilde{\varepsilon}}{n-k}, \\
\widehat{\sigma}^2_{\widetilde{\varepsilon}} &= \frac{\widetilde{\varepsilon}'(I_n - P_Z)\widetilde{\varepsilon}}{n-k}.
\end{aligned}$$

The second term in (9) can now be expressed as follows

$$\begin{aligned}
P_Z\left(\varepsilon, V\right)\begin{pmatrix} \widehat{\sigma}_{\varepsilon}^{-1} & -\frac{\widehat{\sigma}'_{V,\varepsilon}}{\widehat{\sigma}^2_{\varepsilon}} \\ 0 & I_m \end{pmatrix} &= P_Z\left(\widetilde{\varepsilon}, \widetilde{W}\right)\begin{pmatrix} \sigma_{\varepsilon} & \frac{\sigma'_{V,\varepsilon}}{\sigma_{\varepsilon}} \\ 0 & \Sigma_W^{1/2} \end{pmatrix}\begin{pmatrix} \widehat{\sigma}_{\varepsilon}^{-1} & -\frac{\widehat{\sigma}'_{V,\varepsilon}}{\widehat{\sigma}^2_{\varepsilon}} \\ 0 & I_m \end{pmatrix} \\
&= P_Z\left(\widetilde{\varepsilon}, \widetilde{W}\right)\begin{pmatrix} \left(\frac{\widehat{\sigma}_{\varepsilon}}{\sigma_{\varepsilon}}\right)^{-1} & \frac{\sigma'_{V,\varepsilon}}{\sigma_{\varepsilon}} - \frac{\sigma_{\varepsilon}\widehat{\sigma}'_{V,\varepsilon}}{\widehat{\sigma}^2_{\varepsilon}} \\ 0 & \Sigma_W^{1/2} \end{pmatrix} \\
&= P_Z\left(\widetilde{\varepsilon}, \widetilde{W}\right)\begin{pmatrix} \widehat{\sigma}_{\widetilde{\varepsilon}}^{-1} & -\frac{\widehat{\sigma}'_{\widetilde{W},\widetilde{\varepsilon}}}{\widehat{\sigma}^2_{\widetilde{\varepsilon}}} \\ 0 & I_m \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \Sigma_W^{1/2} \end{pmatrix},
\end{aligned} \tag{10}$$

where the last equality follows from

$$\begin{aligned}
\widehat{\sigma}_{\widetilde{W},\widetilde{\varepsilon}} &= \frac{\Sigma_W^{-1/2}}{\sigma_{\varepsilon}}\widehat{\sigma}_{W,\varepsilon} \\
&= \frac{\Sigma_W^{-1/2}}{\sigma_{\varepsilon}}\left(\hat{\sigma}_{V,\varepsilon} - \frac{\hat{\sigma}^2_{\varepsilon}\sigma_{V,\varepsilon}}{\sigma^2_{\varepsilon}}\right), \\
\hat{\sigma}^2_{\widetilde{\varepsilon}} &= \frac{\hat{\sigma}^2_{\varepsilon}}{\sigma^2_{\varepsilon}}.
\end{aligned}$$

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

$$K(\beta) = \frac{\widetilde{\varepsilon}' P_{\widetilde{X}} \widetilde{\varepsilon}}{\hat{\sigma}^2_{\widetilde{\varepsilon}}} \quad (11)$$

$$\widetilde{X} = P_Z \left(\widetilde{W} - \widetilde{\varepsilon} \frac{\widehat{\sigma}'_{\widetilde{W},\widetilde{\varepsilon}}}{\widehat{\sigma}^2_{\widetilde{\varepsilon}}} \right) + Z\Pi\Sigma_W^{-1/2}. \quad (12)$$

If the matrix $\Pi\Sigma^{-1/2}$ were known, the distribution of $K(\beta)$ could be simulated, since $(\widetilde{\varepsilon}, \widetilde{W})$ can be simulated, as they have a standard normal distribution. However, the matrix $\Pi\Sigma_W^{-1/2}$ is not known. Still we can distinguish between two extreme cases.

One extreme case is given by total underidentification: $\Pi = 0$. In this case the distribution can be simulated easily. The distribution depends only on the dimension parameters k, m and n.

Another extreme case is when $\Pi = \theta\Pi^\star$, where $\Pi^\star$ is fixed and the scalar θ grows to infinity, $\theta \to \infty$. In the latter case $P_{\widetilde{X}}$ converges in probability to $P_{Z\Pi}$ and so the distribution of $K(\beta)/m$ converges to $F_{m,n-k}$. These two extreme cases in which we can construct the exact distribution of $K(\beta)/m$ can be shown to give upper and lower bounds on the exact finite sample distribution of $K(\beta)/m$. This results since as $|\theta|$ increases, then $K(\beta)$ decreases stochastically, i.e. the distribution function moves to the left. Intuitively, this can be understood by noticing that $\widetilde{\varepsilon}$ is correlated with the variables in $\widetilde{X}$. A more formal justification is given in the Appendix.

The above 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 $\widehat{\sigma}_{\widehat{W},\widehat{\varepsilon}}$ is small, in probability, and the correlation between $\widetilde{\varepsilon}$ and $\widetilde{X}$ 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 (2000).

Would this limit distibution also provide an accurate approximation if there are many redundant instruments? In order 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 $\Pi = 0$, we can use the following limits, which can be easily verified.

$$\widehat{\sigma}^2_{\widetilde{\varepsilon}} \xrightarrow{p} 1,$$

$$\widehat{\sigma}_{\widetilde{W},\widetilde{\varepsilon}} \xrightarrow{p} 0,$$

$$\widetilde{W}' P_Z \widetilde{W}/k \xrightarrow{p} I_m,$$

$$\widetilde{\varepsilon}' P_Z \widetilde{\varepsilon}/k \xrightarrow{p} 1,$$

$$\widetilde{W}' P_Z \widetilde{\varepsilon}/k \xrightarrow{p} 0,$$

so

$$\widetilde{X}' \widetilde{X}/k \xrightarrow{p} I_m. \tag{13}$$

Furthermore,

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

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

so

$$\widetilde{X}' \widetilde{\varepsilon}/(k^{1/2}) \overset{A}{\sim} \mathcal{N}(0, (1-\alpha)^{-1} I_m). \tag{14}$$

Consequently, we find for the upper bound, where $\Pi = 0$,

$$(1 - k/n) K(\beta) \overset{A}{\sim} \chi^2_m, \tag{15}$$

which holds whether or not $\alpha > 0$. In particular if $\alpha > 0$ we 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.

5 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 a few values of k and n. Figures 1 to 7 show the lower and upper bounds on the exact distribution function of $\frac{K(\beta)}{m}$ (8). The figures show a considerable difference between the lower and upper bound, when k/n is large. 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 $\frac{\chi^2_m}{m}$, based on (15), we used the $F_{m,n-k}$ distribution, which is more accurate in small samples and more convenient as well, since it is also used for the lower bound. The figures show 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 $\frac{K(\beta)}{m}$.

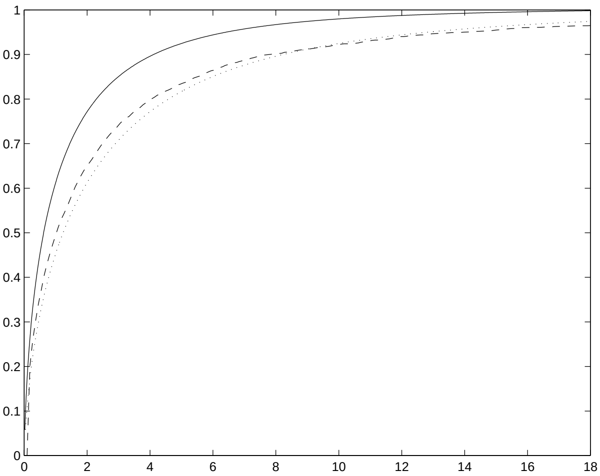

Figure 1: Lower bound (—) and upper bound (- -) and limiting approximation of upper bound (...) of the distribution function of $K(\beta)/m$ for $n = 10$, $k = 5$ and $m = 1$.

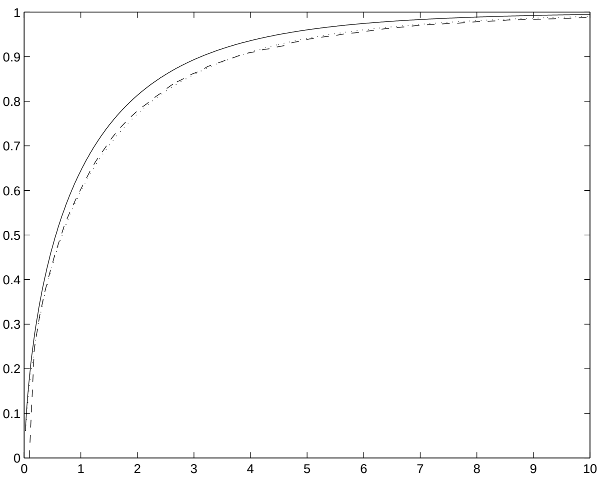

Figure 2: Lower bound (—) and upper bound (- -) and limiting approximation of upper bound (...) of the distribution function of $K(\beta)/m$ for $n = 25$, $k = 5$ and $m = 1$.

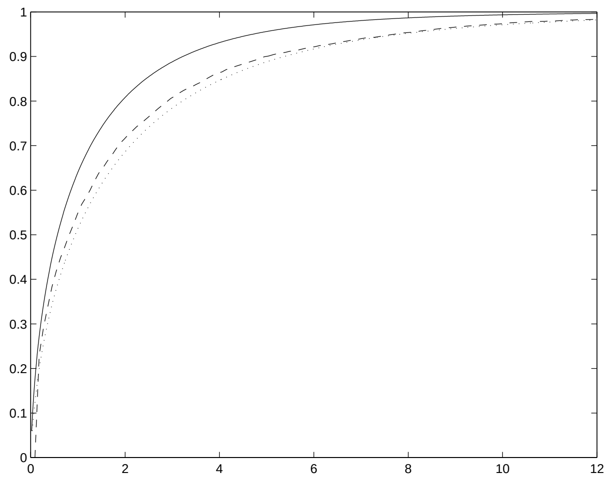

Figure 3: Lower bound (—) and upper bound (- -) and limiting approximation of upper bound (...) of the distribution function of $K(\beta)/m$ for $n = 25$, $k = 10$ and $m = 1$.

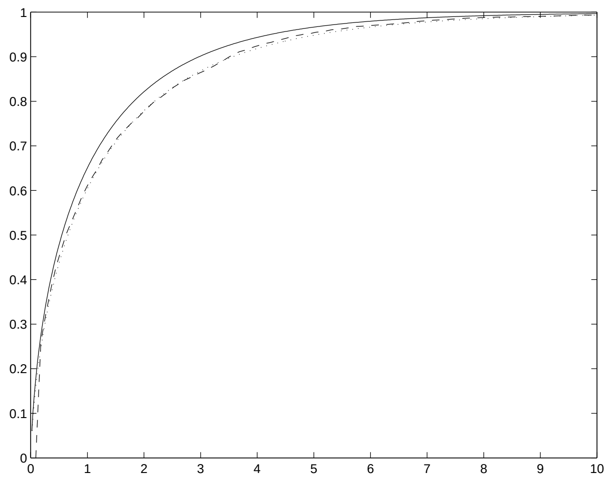

Figure 4: Lower bound (—) and upper bound (- -) and limiting approximation of upper bound (...) of the distribution function of $K(\beta)/m$ for $n = 50$, $k = 10$ and $m = 1$.

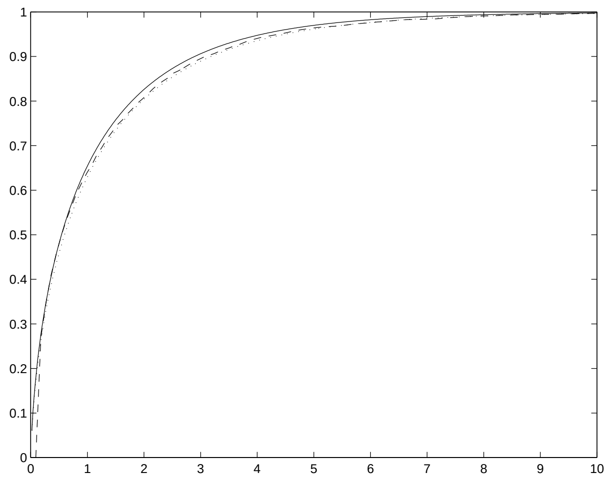

Figure 5: Lower bound (—) and upper bound (- -) and limiting approximation of upper bound (...) of the distribution function of $K(\beta)/m$ for $n = 100$, $k = 10$ and $m = 1$.

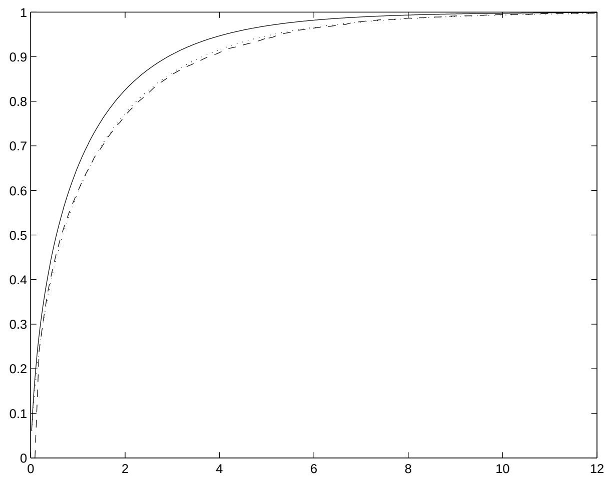

Figure 6: Lower bound (—) and upper bound (- -) and limiting approximation of upper bound (...) of the distribution function of $K(\beta)/m$ for $n = 100$, $k = 25$ and $m = 1$.

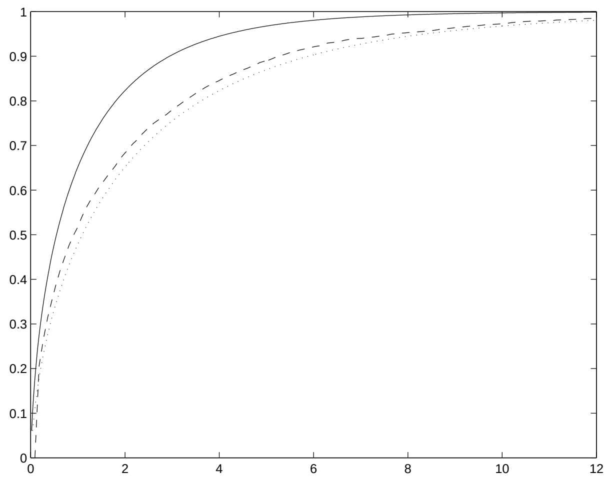

Figure 7: Lower bound (—) and upper bound (- -) and limiting approximation of upper bound (...) of the distribution function of $K(\beta)/m$ for $n = 100$, $k = 50$ and $m = 1$.

Appendix

In order to prove the assertion in Section 4 that $K(\beta)$ is stochastically decreasing in $|\theta|$, where $\Pi = \theta\Pi^\star$, we first condition on $(I_n - P_Z)(\widetilde{\varepsilon}, \widetilde{W})$. This does not affect the distribution of $P_Z(\widetilde{\varepsilon}, \widetilde{W})$. Let

$$
\begin{aligned}
H &= (Z'Z)^{1/2}\Pi^\star\Sigma_W^{-1/2}, \\
u &= (Z'Z)^{-1/2}Z'\widetilde{\varepsilon}, \\
U &= (Z'Z)^{-1/2}Z'(\widetilde{W} - \widetilde{\varepsilon}\widetilde{\lambda}'), \\
\widetilde{\lambda} &= \frac{\widehat{\sigma}_{\widetilde{W},\widetilde{\varepsilon}}}{\widehat{\sigma}^2_{\widetilde{\varepsilon}}}.
\end{aligned}
$$

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

$$
\begin{aligned}
\widetilde{\varepsilon}P_{\widetilde{X}}\widetilde{\varepsilon} &= (U\gamma + \nu)'P_L(U\gamma + \nu), \\
L &= H\theta + U.
\end{aligned}
$$

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: $(\gamma'U'P_LU\gamma)^{1/2}$. Furthermore let $\xi = Q'(L'L)^{-1/2}L'\nu$, whose elements are independently normally distributed, then

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

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 non-central χ^2 distribution that decreases as the non-centrality parameter decreases. So we only have to prove that the non-random function

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

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

References

[1] Anderson, T.W. and H. Rubin. Estimators of the Parameters of a Single Equation in a Complete Set of Stochastic Equations. *The Annals of Mathematical Statistics*, **21**:570–582, (1949).

[2] Bekker, P.A. Alternative Approximations to the Distributions of Instrumental Variable Estimators. *Econometrica*, **62**:657–681, 1994.

[3] Hausman, J. and J. Hahn. A New Specification Test for the Validity of Instrumental Variables. *Econometrica*, forthcoming.

[4] Kleibergen, F. Pivotal Statistics for testing Structural Parameters in Instrumental Variables Regression. Tinbergen Institute Discussion Paper TI 2000-055/4, 2000.

[5] Nelson, C.R. and R. Startz. Some Further Results on the Exact Small Sample Properties of the Instrumental Variables Estimator. *Econometrica*, **58**:967–976, 1990.

[6] Staiger, D. and J.H. Stock. Instrumental variables regression with weak instruments. *Econometrica*, **65**:557–586, 1997.

[7] Zivot, E., R. Startz, and C. R. Nelson. Valid Confidence Intervals and Inference in the Presence of Weak Instruments. *International Economic Review*, **39**:1119–1144, 1998.