Assessing the Magnitude of the Concentration Parameter in a Simultaneous Equations Model*

D S Poskitt*  
Department of Econometrics and Business Statistics  
Monash University  
Vic 3800, Australia

C L Skeels  
Department of Economics  
The University of Melbourne  
Victoria 3010, Australia

Don.Poskitt@buseco.monash.edu.au  
Chris.Skeels@unimelb.edu.au

September 20, 2007

Abstract

This paper provides the practitioner with a method of ascertaining when the concentration parameter in a simultaneous equations model is small. We provide some exact distribution theory for a proposed statistic and show that the statistic possesses the minimal desirable characteristics of a test statistic when used to test that the concentration parameter is zero. We also discuss the statistic’s relationship to various other procedures that have appeared in the literature.

* Preliminary and not to be quoted without the authors’ permission.

* Corresponding author. Tel: +61-3-9905 9378, Fax: +61-3-9905 5474
1 Motivation

The concentration parameter is a naturally occurring parameter in the study of simultaneous equations models. When this parameter is close to zero the Fisher information matrix is approximately rank deficient and so the structural form model is weakly identified, a case where standard asymptotic approximations are known to perform poorly. See, *inter alia*, Stock, Wright, and Yogo (2002) and Hahn and Hausman (2003) for surveys of the recent extensive literature on weakly identified models. In contrast to standard asymptotic approximations, however, Poskitt and Skeels (2006) have introduced an alternative approximation to the exact sampling distribution of the two stage least squares estimator that is designed to work well when the concentration parameter is small. Practitioners will therefore need guidance as to when standard asymptotic approximations are likely to break down and use of the Poskitt-Skeels approximation may be deemed appropriate. The results of this note provide practitioners with such a guide.

In what follows we examine the properties of a statistic, designated $A^2$, that can be employed to ascertain the proximity of the concentration parameter to zero. We explore the sampling distribution of $A^2$ under classical assumptions, and demonstrate that $A^2$ provides an admissible test that the concentration parameter is zero, that has a power curve which is monotonically increasing in each of the maximal invariants. In addition we discuss the relationship between $A^2$ and various other statistics that have appeared in the literature. Proofs are assembled in the Appendix.

2 The Model, Notation and Assumptions

Consider the classical structural equation model

$$y = Y\beta + X\gamma + u, \quad u \sim N(0, \sigma_u^2 I_T),$$

with corresponding reduced form

$$[y \quad Y] = [X \quad Z] \begin{bmatrix} \pi_1 & \Pi_1 \\ \pi_2 & \Pi_2 \end{bmatrix} + [v \quad V],$$

where the endogenous variables $y$ and $Y$ are $T \times 1$ and $T \times n$, respectively, the predetermined variables $X$ and $Z$ are $T \times k$ and $T \times \nu$, respectively, with $K = k + \nu$. We shall assume that the $T \times (n + 1)$ matrix $[v \quad V][X \quad Z] \sim N(0, \Omega \otimes I_T)$, so that the rows of $[v \quad V]$ are independent with common $(n + 1) \times (n + 1)$ covariance matrix

$$\Omega = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix},$$

$\omega_{11}$ scalar, where $[v \quad V]$ is partitioned conformably with $[y \quad Y]$. We shall further assume that $[X \quad Z]$ has full column rank almost surely, $\rho([X \quad Z]) = K$, and that the usual compatibility conditions hold, namely $\pi_1 = \Pi_1\beta + \gamma$, $\pi_2 = \Pi_2\beta$ and $u = v - V\beta \sim N(0, \sigma_u^2 I_T)$, with $\sigma_u^2 = [1, -\beta']\Omega[1, -\beta']$.

The two stage least squares estimator for $\beta$ is of the form

$$\hat{\beta} = (Y'PY)^{-1}Y'Py,$$
where $P = R_XZ(Z'R_XZ)^{-1}Z'R_X$, with rank $\rho(P) = \nu$ and $PX = 0$. Now

$$S = [y'Y]P[y'Y][XZ] \sim W_{n+1}(\nu, \Omega, \Delta\Omega^{-1}),$$

(2)

where $\Delta = \Pi_2Z'\Pi_2Z = \Pi_2Z'R_XZ\Pi_2$. In particular, we see that $\Delta$ is a naturally occurring parameter in this model because $\beta$ is a function only of the elements of $S$. It is the proximity of $\Delta$ to zero that is the focus of our attention.

### 3 The Test Statistic

Many of the statistics that have been proposed for testing the proximity of $\Delta$ to zero can be interpreted as arising from an examination of the reduced form model for $Y$; see, for example, the partial $R^2$ statistics of Bound, Jaeger, and Baker (1995) and Shea (1997) (see also Godfrey, 1999), and the asymptotic test procedures considered by Cragg and Donald (1993) and Hall, Rudebusch, and Wilcox (1996). In this vane, let us consider the reduced form equation for $Y$ obtained from (1) and its associated assumptions:

$$Y = X\Pi_1 + Z\Pi_2 + \nu. \quad (3)$$

As a first step, partialling out $X$ by premultiplying by $R_X$ yields

$$\tilde{Y} = Z\Pi_2 + \tilde{\nu}, \quad (4)$$

where $[\tilde{Y} \tilde{Z} \tilde{\nu}] = R_X[Y \ Z \ V]$. The likelihood ratio statistic for testing the null hypothesis $H_0^1 : \Pi_2 = 0$ against the alternative that $H_1^1 : \Pi_2 \neq 0$ in (4), or equivalently the reduced form (3), is

$$LR_T = \left[ \frac{\det[Y'R_Z\tilde{Y}]}{\det[Y'Y]} \right]^{T/2}.$$

The statistic $A^2 = [LR_T]^{2/T}$ is a partial version of the vector alienation coefficient introduced by Hotelling (1936), hence our notation. It is easy to show that

$$A^2 = \det[I_n - (Y'R_XY)^{-1/2}(Y'PY)(Y'R_XY)^{-1/2}], \quad (5)$$

from which it follows that $A^2 = 1$ when $\tilde{Y}$ and $\tilde{Z}$ are orthogonal and $A^2 = 0$ if there exists a matrix $D$ of full column rank such that $\tilde{Y} = \tilde{Z}D$. Thus $A^2$ can be viewed as a measure of the perpendicularity between $Y$ and $Z$ having adjusted for the effects of $X$, it represents the proportion of the generalized variance of $\tilde{Y}$ that remains once the regression mean square in the multivariate regression of $\tilde{Y}$ on $\tilde{Z}$ has been accounted for. Noting that $Y'PY$ can be thought of as a sample analogue of $\Pi_2Z'R_XZ\Pi_2$, in as much as $\lim_{T \to \infty} T^{-1}\|Y'PY - \Pi_2Z'R_XZ\Pi_2\| = 0$, suggests that the proportion of the generalized variance of $\tilde{Y}$ explained will be small whenever $\Delta$ is small. This, in turn, suggests that $A^2$ can be employed to test the statistical significance of $\Delta$. Indeed, since we have assumed that $\rho([X \ Z]) = K$, it follows that $Z'Z = Z'R_XZ > 0$ and $\rho(Z) = \rho(X) = \nu$ almost surely. Therefore, testing $H_0^1$ against $H_1^1$ using $LR_T$ is equivalent to testing $H_0^1 : \Delta = 0$ against the alternative $H_1^1 : \Delta \neq 0$ using $A^2$.

---

1We adopt the notational convention $P_A = A(A'A)^{-1}A'$ and $R_A = I_N - P_A$. 
4 Sampling Distribution and Optimality Properties

Lemma 1. Assume that equation (1) and its accompanying assumptions obtain. Then, under $\mathcal{H}_0 : \Delta = 0$, the statistic $A^2$ possesses Wilks' $\Lambda$ distribution, with parameters $n, T - \nu$ and $\nu$, written $A^2 \sim \Lambda(n, T - \nu, \nu)$.

If we adopt an hypothesis testing perspective, then it follows that the set

$$CR\{A^2, \alpha\} = \{A^2 : A^2 < \lambda_\alpha(n, T - \nu, \nu)\},$$

where $\lambda_\alpha(n, T - \nu, \nu)$ denotes the $\alpha \cdot 100\%$ percentile point of the $\Lambda(n, T - \nu, \nu)$ distribution, defines a size $\alpha$ critical region for testing $\mathcal{H}_0$ against $\mathcal{H}_1$. Equally, we might use p-values of the form

$$p = \mathbb{P}(\Lambda(n, T - \nu, \nu) < A^2)$$

to provide a probability scale indicative of the magnitude of $\Delta$. We imagine that in many situations such a scale may be of greater practical relevance than is a test that $\Delta$ is identically equal to zero.

We have already observed that the magnitude $A^2$ can be interpreted as a measure of the perpendicularity, or lack of correlation, between $\mathbf{Y}$ and $\mathbf{Z}$. To formalize this idea let us assume, in a natural extension of equation (1), that $\mathbf{Z} = \mathbf{X}\Pi_3 + \mathbf{U}_2$ and

$$[\mathbf{Y} \mathbf{Z}] = \mathbf{X}[(\Pi_1 + \Pi_3\Pi_2) \Pi_3] + [\mathbf{U}_1 \mathbf{U}_2]^\dagger, \quad (6)$$

where the conditional distribution of $\mathbf{U} = [\mathbf{U}_1 \mathbf{U}_2] = [(\mathbf{V} + \mathbf{U}_2\Pi_2) \mathbf{U}_2]$ given $\mathbf{X}$ is Gaussian with mean zero and variance-covariance $\Sigma \otimes I_T, \mathbf{U}\mathbf{X} \sim N(0, \Sigma \otimes I_T)$. If we regard (6) as a specification for the joint distribution of $[\mathbf{Y} \mathbf{Z}]$ conditional on $\mathbf{X}$ then we can contemplate testing that the instruments are orthogonal to the endogenous regressors by testing $\mathcal{H}_0' : \Sigma_{12} = 0$ against the alternative that $\mathcal{H}_1' : \Sigma_{12} \neq 0$, where $\Sigma_{12}$ denotes the covariance between $\mathbf{U}_1$ and $\mathbf{U}_2$. The likelihood ratio statistic for testing $\mathcal{H}_0'$ against $\mathcal{H}_1'$ is, again, $LR_T$. This, of course, is to be expected because if $\mathcal{H}_0'$ is true in (6) then $\mathbf{Z}$ does not appear in the conditional mean of $\mathbf{Y}$ given $\mathbf{Z}$, implying that $\Pi_2 = 0$, so that $\mathcal{H}_1'$ and $\mathcal{H}_0$ are also true. Given that for the purposes of testing $A^2$ is equivalent to $LR_T$, we can anticipate that $A^2$ will inherit any desirable properties of the likelihood ratio test. More specifically, we have the following result.

Theorem 1. Assume that equation (6) and its accompanying assumptions hold, and let $\rho_1^2 \geq \ldots \geq \rho_n^2$, denote the (population) canonical correlations, the characteristic roots of $\Sigma_{11}^{-1/2}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11}^{-1/2}$. Then the statistic $A^2$ yields an admissible test, invariant under the group of linear transformations $(\mathbf{Y}, \mathbf{Z}) \leftrightarrow (\mathbf{Y}G, \mathbf{ZH})$ where $\mathbf{G}$ and $\mathbf{H}$ are arbitrary non-singular matrices of dimension $n \times n$ and $\nu \times \nu$, respectively. Moreover, the power function of $A^2$ is monotonically increasing in each of the maximal invariants $\rho_i^2$, $i = 1, \ldots, n$.

Theorem 1 indicates: First, that within the class of tests invariant under the group of transformations $(\mathbf{Y}, \mathbf{Z}) \leftrightarrow (\mathbf{Y}G, \mathbf{ZH})$ there is no test which dominates $A^2$ in the sense of having better power for at least one point in the parameter space

---

For a survey of some computational aspects of Wilks’-$\Lambda$ distribution the interested reader is referred to Poskitt and Skeels (2004, Section 4.2).
Assessing the Concentration Parameter

and no less power elsewhere. Second, that the power function of \( A^2 \) depends upon all of the maximal invariants and will, therefore, be sensitive to any departures from the null. These are not especially strong properties but, as we shall see, even these are not possessed by other test procedures that have been proposed in the literature.

5 Discussion and Alternative Procedures

In this section we compare \( A^2 \) with three other procedures that might be employed in this context. The three procedures considered are (i) a generalization of the partial \( R^2 \) of Bound et al. (1995) to multivariate situations, (ii) a variant of the Cragg and Donald (1993) procedure, and (iii) the Roy (1957) maximum root test.

In order to compare \( A^2 \) with a generalized partial \( R^2 \) first observe that one possible generalization of the univariate partial \( R^2 \) of Bound et al. (1995) is a partial version of Hotelling’s coefficient of vector correlation

\[
R^2 = \frac{\det[\tilde{Y}'\tilde{P}\tilde{Z}\tilde{Y}]}{\det[\tilde{Y}'\tilde{Y}]} = \prod_{i=1}^{n} r_i^2,
\]

where \( r_1^2 \geq \ldots \geq r_n^2 \) lists in descending order the partial canonical correlations between \( Y \) and \( Z \) having adjusted for the effects of \( X \). Similarly, from (5),

\[
A^2 = \prod_{i=1}^{n} (1 - r_i^2).
\]

Following our earlier developments, it is straight-forward to show that, when \( \Delta = 0 \), \( R^2 \sim \Lambda(n, \nu, T - \nu) \). It is important to recognize, however, that in general \( R^2 \neq 1 - A^2 \) and so probability calculations based on \( A^2 \) and \( R^2 \) will not be identical. This raises the question of which measure is most appropriate for our current needs.

From expressions (7) and (8) it follows that it is only necessary for the largest (smallest) partial canonical correlation to deviate substantially from zero (one) for \( A^2 (R^2) \) to deviate significantly from unity. Thus, whereas \( A^2 \) will be sensitive to departures from orthogonality \( R^2 \) is designed to detect exact correlation. Now recall that the use of Wilks’-\( \Lambda \) distribution is contingent on \( \Delta \) being equal to zero, which we have already observed is equivalent to the hypothesis that \( \tilde{Y} \) and \( \tilde{Z} \) are uncorrelated, and hence that \( \rho_1^2 = \ldots = \rho_n^2 = 0 \). Therefore, \( A^2 \) is more in accord with the basic assumption underlying the application of Wilks’-\( \Lambda \) distribution than is \( R^2 \). Hence, \( A^2 \) appears to be more suited to our purpose.

Another statistic seemingly related to \( A^2 \) is that version of the Cragg and Donald (1993) procedure for testing the rank of \( \Pi_2 \) that is “concerned with whether \( X_2(\tilde{Z}) \) can serve as instruments for \( Y_2(\tilde{Y}) \) in the sense that there is enough correlation” is given by (in the notation of this paper) the smallest eigenvalue of \( \tilde{Y}'\tilde{P}\tilde{Z}\tilde{Y} \) in the metric of \( \tilde{Y}'\tilde{R}_Z\tilde{Y} \); see hypothesis \( H_0^f \) and Theorem 3 of Cragg and Donald (1993). Noting that

\[
det[\tilde{Y}'\tilde{P}\tilde{Z}\tilde{Y} - \lambda \tilde{Y}'\tilde{R}_Z\tilde{Y}] = \\
det[(1 + \lambda)(\tilde{Y}'\tilde{Y})^{-1/2}\tilde{Y}'\tilde{P}\tilde{Z}\tilde{Y} (\tilde{Y}'\tilde{Y})^{-1/2} - \frac{\lambda}{1 + \lambda} I_n],
\]

where \( \lambda \) is a positive eigenvalue of \( \tilde{R}_Z \) and \( I_n \) is the \( n \times n \) identity matrix.
we can conclude that $\lambda/(1 + \lambda) = r^2$ and hence that this version of Cragg and Donald’s statistic is equivalent to testing the significance of the smallest canonical correlation. Hall et al. (1996) have also advocated using the smallest canonical correlations between $\tilde{Y}$ and $\tilde{Z}$ to assess the relevance of the instruments for the estimation of $\beta$; see also Bowden and Turkington (1984, §2.3).

If we are interested in looking for evidence that $\Delta$ is small then our previous discussion suggests that we should examine those linear combinations $\tilde{Y}$ and $\tilde{Z}$ that yield evidence in favour of the hypothesis that $\tilde{Y}$ and $\tilde{Z}$ are uncorrelated. The Union-Intersection principle of Roy (1957) indicates that this ultimately leads to a procedure akin to those considered by Cragg and Donald (1993) and Hall et al. (1996), except that we should examine the size of the largest rather than the smallest canonical correlation. It is of interest to note that Theorem 8.10.4 of Anderson (2003) implies that whereas Roy’s maximum root test with an acceptance region of the form $\{r_1^2 : r_n^2 \leq \kappa_\alpha\}$ is admissible, in the sense that it cannot be improved upon by reducing the probability of Type I and/or Type II errors, the minimum root test with acceptance region $\{r_n^2 : r_n^2 \leq \kappa_\alpha\}$ is not admissible.

Roy’s maximum root test is invariant under the group of non-singular linear transformations prompting comparison with $CR\{A^2, \alpha\}$, which is also invariant and admissible under this group of transformations. Simulation results presented in Schatzoff (1966), based on an experimental design structured in terms of the maximal invariants, indicate that although Roy’s maximum root test will have good power in alternative directions where $\rho_1^2 > 0$ and $\rho_2^2 = \ldots = \rho_n^2 = 0$, its performance will be inferior to that of $A^2$ in more general directions. Indeed, Schatzoff goes so far as to suggest that “…the largest root should not be used except to test specifically against an alternative of rank one” (Schatzoff, 1966, p.429). He concludes that the test based on Hotelling’s alienation coefficient is to be preferred. An obvious advantage of using $A^2$ as a basis for calibrating the magnitude of the concentration parameter is that it does not focus on a particular canonical correlation but summarizes the simultaneous impact of all $\rho_i^2$, $i = 1, \ldots, n$, suggesting that $A^2$ will be sensitive to deviations of $\Delta$ from zero in all possible directions.

References


Assessing the Concentration Parameter


Appendix: Proofs

Proof of Lemma 1: Let $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_n)' \neq \mathbf{0}$ be an arbitrary constant vector. Post-multiplying equation (4) by $\boldsymbol{\tau}$ yields

$$\tilde{Y}\boldsymbol{\tau} = \tilde{Z}\phi + \eta,$$

where $\phi = \Pi_2\tau$, $\eta = R_XV\tau$, and $\eta|X, Z \sim N(0, \sigma^2R_X)$, where $\sigma^2 = \tau'\Omega_{22}\tau$.

The operators $P_Z$ and $R_Z$ are idempotent with ranks $\rho$ and $T - \rho$, respectively, where $\rho(Z) = \rho(Z) = \nu$, the number of instruments used in addition to $X$, supposing that $X$ is employed as its own instrument. We can therefore conclude that, for given $Z$ and $X$, the quadratic forms $\tau'\tilde{Y}'P_Z\tilde{Y}\tau$ and $\tau'\tilde{Y}'R_Z\tilde{Y}\tau$ are independently distributed as $\sigma^2 \cdot \chi^2(\nu, \mu)$, $\mu = \phi'\tilde{Z}'\tilde{Z}\phi/\sigma^2$, and $\sigma^2 \cdot \chi^2(T - \nu)$ random variables, respectively. Since $\tau$ is arbitrary we have from Rao (1973, §8b.2 (ii) & (iii)) that the matrices $\tilde{Y}'P_Z\tilde{Y}$ and $\tilde{Y}'R_Z\tilde{Y}$ will have independent Wishart distributions:

$$\tilde{Y}'P_Z\tilde{Y} \sim W_n(\nu, \Omega_{22}, \Gamma_{22}) \quad \text{and} \quad \tilde{Y}'R_Z\tilde{Y} \sim W_n(T - \nu, \Omega_{22}).$$
When $\Delta = 0$ the non-centrality parameter $\mu = \tau' \Delta / \sigma^2 = 0$ and both $\tilde{Y}' P_{\tilde{Z}} \tilde{Y}$ and $\tilde{Y}' R_{\tilde{Z}} \tilde{Y}$ will have central Wishart distributions. Writing $A^2$ as the ratio of $\text{det}(\tilde{Y}' R_{\tilde{Z}} \tilde{Y})$ to $\text{det}(\tilde{Y}' (R_{\tilde{Z}} + P_{\tilde{Z}}) \tilde{Y})$ it follows that when $\Delta = 0$ the statistic $A^2$ will possess Wilks’-A distribution (Wilks, 1962, §18.5.1), as required.

**Proof of Theorem 1:** First note that the problem of testing the hypothesis that $\Sigma_{12} = 0$ or, equivalently, $\rho_i^2 = 0$, $i = 1, \ldots, n$, is invariant under the group of non-singular linear transformations. It is well known that the canonical correlations are the maximal invariants under this group of transformations and so, from (8), we see that $A^2$ is an invariant test statistic. Admissibility follows by writing the acceptance region of the test as

$$AR\{A^2, \alpha\} = \left\{ A^2 : \prod_{i=1}^{n} (1 - r_i^2)^{-1} \leq \kappa_\alpha \right\} = \left\{ A^2 : \prod_{i=1}^{n} (1 + \lambda_i) \leq \kappa_\alpha \right\}, \quad (A.2)$$

where the $\lambda_i = r_i^2/(1 - r_i^2)$ coincide with the non-zero characteristic roots of $P_{\tilde{Z}} \tilde{Y} (\tilde{Y}' R_{\tilde{Z}} \tilde{Y})^{-1} \tilde{Y}' P_{\tilde{Z}}$, and applying Corollary 8.10.2 of Anderson (2003).

Since $R_X$ is idempotent of rank $T - k$ there exists a $T \times (T - k)$ column orthonormal matrix $Q_X$, $Q_X Q_X = I_{T - k}$, such that $R_X = Q_X Q_X'$ and $Q_X [Y \ Z]' = Q_X [U_1 \ U_2]$. There also exists two non-singular matrices $A$ and $G$ that map $U_1$ and $U_2$, respectively, to the canonical variates so that $[V_1 \ V_2] = Q_X [U_1 A \ U_2 G]$ is distributed

$$N\left( [0 \ 0], \begin{bmatrix} I_n & [A \ 0] \end{bmatrix} \otimes I_{T - k} \right),$$

where $A = \text{diag}[\rho_1, \ldots, \rho_n]$. Given the instruments, the conditional distribution of $V_1$ is $N(V_2 | A 0', (I_n - A^2) \otimes I_{T - k})$ and $W_1 = V_1 (I_n - A^2)^{-1/2} / \sim N(V_2 M', I_n \otimes I_{T - k})$, where $M = [\text{diag}[\delta_1, \ldots, \delta_n] \ 0]$, $\delta_i = \rho_i/(1 - \rho_i^2)^{1/2}$.

Define $O$ as the $(T - k) \times (T - k)$ orthogonal matrix

$$O = \begin{bmatrix} (V_2^2)^{-1/2} V_2' \\ (W_1^2 R_{V_2 W_1})^{-1/2} W_1' R_{V_2} \\ O_3 \end{bmatrix}$$

where $O_3$ is a $(T - k - \nu - n) \times (T - k)$ matrix that makes $O$ orthogonal. Then

$$OW_1 = \begin{bmatrix} (V_2^2)^{-1/2} V_2' W_1 \\ (W_1^2 R_{V_2 W_1})^{1/2} \\ O_3 W_1 \end{bmatrix} = \begin{bmatrix} W_{11} \\ W_{12} \\ 0 \end{bmatrix}$$

and $W_{11} \sim N((V_2^2)^{-1/2} M', I_n \otimes I_n)$ is distributed independently of $W_{12} \sim N(0, I_n \otimes I_n)$. Moreover, by construction, the $\lambda_i$, $i = 1, \ldots, n$, of expression (A.2) are the non-zero characteristic roots of $W_{11}(W_{12}' W_{12})^{-1} W_{11}'$. Applying the same argument as that used by Anderson (2003, pp. 368–369) it follows that $AR\{A^2, \alpha\}$ is convex in each row of $W_{11}$ given $W_{12}$ and the other rows of $W_{11}$ and hence, by Theorem 8.10.6 of Anderson (2003), the conditional power of $A^2$, given the instruments, is monotonically increasing in the characteristic roots of $MV_2^2 V_2 M'$. But the characteristic roots of $MV_2^2 V_2 M'$ are all monotonically increasing in $\rho_i$, $i = 1, \ldots, n$, by Lemma 9.10.2 of Anderson (2003). Taking the unconditional power, recognizing that the marginal distributions of $W_{12}' W_{12} \sim \mathcal{W}_n(n, I_n)$ and
\[ V'_2 V_2 \sim W_\nu (T - k, I_\nu) \] do not depend on the \( \rho_i, i = 1, \ldots, n \), gives us the result that for all possible sets of the instruments the power of \( \mathcal{A}^2 \) is monotonically increasing in each \( \rho_i, i = 1, \ldots, n \). \qed