Testing parameters in GMM without assuming that they are identified: a comment

Bertille Antoine*and Eric Renault†

November 1, 2007

Abstract:
This note provides a comparative study of the power of two tests proposed within the GMM literature, on the structural parameters, when the identification is not fully ensured: the standard LM test, valid in the framework of Antoine and Renault (2007), and the modified LM test or K-test proposed by Kleibergen (2005). In a more general sense, we compare two approaches with respect to treating identification issues in a GMM context. On one hand, Antoine and Renault (2007) specify identification issues through a partition of the moment conditions, according to the (statistical) information they carry (ie strong or (nearly)-weak). This allows the application of standard test procedures (like LM). On the other hand, as shown by Kleibergen (2005), in the absence of any specification of the identification issues, a modification of the LM test statistic is required. We show that the former specification is not much more involved, while it provides explicit power calculations.

JEL Classification: C12, C32.

Keywords: GMM; (Nearly)-weak identification; LM and modified LM (or K) test; Instrumental variables.

*Simon Fraser University. Email: bertille_ antoine@sfu.ca.
†University of North Carolina at Chapel Hill, CIRANO and CIREQ. Email: renault@email.unc.edu

1 Introduction

In a recent paper published in *Econometrica*, Kleibergen (2005) proposes a GMM-LM based statistic, the K statistic. It uses a modified estimator of the Jacobian, asymptotically uncorrelated with the empirical mean of the moments. This property permits to relax the full rank assumption on the Jacobian and even allows the application of the test in case of weak instruments. In this note, we shed some light on power calculations for the K and LM (or score) test statistics. These calculations are produced for several identification issues, from strong to weak, and for some mixture cases of the former.

Kleibergen (2005) starts with a joint central limit theorem on the moment conditions $\overline{\phi}_T(\theta^0)$ and their associated Jacobian $\left[\partial\overline{\phi}_T(\theta^0)/\partial\theta'\right]$:

Assumption 1 *(Joint CLT from Kleibergen (2005))*

$$\sqrt{T}\begin{bmatrix} \overline{\phi}_T(\theta^0) & - & E\left[\overline{\phi}_T(\theta^0)\right] \\ Vec\left[\frac{\partial\overline{\phi}_T(\theta^0)}{\partial\theta'}\right] & - & Vec\left[J(\theta^0)\right] \end{bmatrix} \quad \textit{with } J(\theta^0) \equiv E\left[\frac{\partial\overline{\phi}_T(\theta^0)}{\partial\theta'}\right]$$

follows an asymptotic gaussian distribution with mean 0 and variance V.

The identification is strong when $J(\theta^0)$ is full-column rank and weak when there exists a deterministic matrix C such that $J(\theta^0) = \frac{C}{\sqrt{T}}$. To test $H_0 : \theta = \theta_0$, Kleibergen proposes the K-test, robust to any case where strong identification fails[1]. It is a modification of the standard GMM score test: instead of computing the LM test as a norm of $\left[\frac{\partial\overline{\phi}_T(\theta_0)}{\partial\theta'}\Omega^{-1}\overline{\phi}_T\right]$, $\left[\frac{\partial\overline{\phi}_T(\theta_0)}{\partial\theta'}\right]$ is replaced by the residual of its regression on the moment conditions. More formally, $\left[\frac{\partial\overline{\phi}_{iT}(\theta_0)}{\partial\theta_j}\right]$ is replaced by

$$\frac{\partial\tilde{\phi}_{iT}(\theta_0)}{\partial\theta_j} = \frac{\partial\overline{\phi}_{iT}(\theta_0)}{\partial\theta_j} - Cov\left[\sqrt{T}\frac{\partial\overline{\phi}_{iT}(\theta_0)}{\partial\theta_j}, \sqrt{T}\overline{\phi}(\theta_0)\right]\left[Var\left(\sqrt{T}\overline{\phi}(\theta_0)\right)\right]^{-1}\overline{\phi}(\theta_0)$$

It has been shown in the literature that this correction generally provides finite sample improvement, without modifying the standard first-order asymptotics[2]: see e.g. Antoine, Bonnal and Renault (2007), Donald and Newey (2000) and Newey and Smith (2004).

[1] The precise identification pattern does not need to be known for the test to be valid and performed. Note also that we distinguish between the true value of the parameter (θ^0) and its value under the null hypothesis (θ_0).

[2] We will show that this is not the case here: the asymptotic behavior of the test statistic may be altered.

We want to investigate the power of the K-test and compare it to the power of the standard score test. For this, we go one step further in the specification of the identification issues. We think that rank deficiencies of the Jacobian must be more tightly related to the moment conditions themselves. More precisely, we use the framework of Antoine and Renault (2007). Everything starts at the moment conditions level: they are partitioned according to the (statistical) information they carry, say strong or nearly-weak. In this framework, the Jacobian naturally inherits a similar pattern, which may explain the asymptotic rank deficiencies. Since the knowledge or the estimation of the degree of weakness of each moment conditions is not required to perform inference, we find that this framework is not much more involved than Kleibergen (2005). Moreover, it helps clarifying power calculations as shown later.

The note is organized as follows. First, we quickly recall the framework of Antoine and Renault (2007). Then, we present the power calculations of the LM and K test statistics against a sequence of local alternatives. We also discuss testing subsets of parameters. Finally, we conclude.

All the proofs are gathered in the appendix.

2 Power against a sequence of local alternatives

2.1 Framework

Antoine and Renault (2007) propose a framework where the moment conditions are partitioned in terms of the (statistical) information they carry. Let us consider here similarly two groups of moment conditions and the associated central limit theorem assumption:

Assumption 2 *(CLT from Antoine and Renault (2007))*

$$\sqrt{T}\left[\begin{array}{c}\bar{\phi}_{1T}(\theta^0)-\rho_1(\theta^0)\\ \bar{\phi}_{2T}(\theta^0)-\frac{\lambda_T}{\sqrt{T}}\rho_2(\theta^0)\end{array}\right]\stackrel{d}{\to}\mathcal{N}(0,S(\theta^0))\quad with\quad 0<<\lambda_T<<\sqrt{T}$$

The first group has k_1 standard moment conditions whereas the second one has k_2 weaker moment conditions. λ_T represents the degree of weakness of the second group of moment conditions, or the speed at which the associated information disappears. This is a convenient way to acknowledge that moment conditions may carry information of heterogeneous quality. Weaker moment conditions contain fragile information that needs to be preserved because it is still relevant for inference. We will see that, with heterogeneous quality of information, the modification of the LM test statistic proposed by Kleibergen may alter the asymptotic behavior of the test statistic. This is in contrast with standard GMM.

The Jacobian matrix naturally inherits the above special design:

Assumption 3 *(Assumption 2(iv) and 2(v) from Antoine and Renault (2007))*

$$2(iv) \quad \left[\frac{\partial \rho_1'(\theta^0)}{\partial \theta} \; \frac{\partial \rho_2'(\theta^0)}{\partial \theta}\right] = \underset{T\to\infty}{Plim} \left[\frac{\partial \overline{\phi}_{1T}'(\theta^0)}{\partial \theta} \; \frac{\sqrt{T}}{\lambda_T}\frac{\partial \overline{\phi}_{2T}'(\theta^0)}{\partial \theta}\right]$$

$$2(v) \quad \sqrt{T}\left[\frac{\partial \overline{\phi}_{1T}'(\theta^0)}{\partial \theta} - \frac{\partial \rho_1'(\theta^0)}{\partial \theta}\right] = \mathcal{O}_P(1)$$

2.2 Power of the K-test

We investigate the power of the LM and the K-test. Basically, if H_0, say $\theta = \theta_0$, is false, we would like to know the probability that it will be rejected. Since we work with asymptotic distributions, for any $\theta \neq \theta_0$, the answer is 1 with a consistent test: this does not help the comparison. Hence, instead of looking at an infinite sample, we want to find an approximation for the case of a finite (but reasonably large) sample. The classical solution is to assume that the data-generating process is subject to a Pitman drift. More precisely, the data in a sample of size T are generated by the model element $\theta^{(T)} = \theta_0 + \frac{\gamma}{\delta_T}$ with γ the direction and δ_T the rate of local departure. This device of using a sequence of local alternatives will be the basis of the following discussion of power properties of the LM and K-test. We consider the following sequence of local alternatives:

$$H_0 : \theta = \theta_0 \quad \text{vs} \quad H_{1T} : \theta = \theta^{(T)} \equiv \theta_0 + \frac{\gamma}{\delta_T}$$

where γ is a fixed deterministic p-vector and δ_T a deterministic sequence such that $\delta_T \overset{T}{\to} \infty$. Our approach follows, for instance, Davidson (2000, chapter 12.4)[3]. Note that under the alternative, for each T, the true value of the parameter $\theta^{(T)}$ depends on the sample size.

[3] In this paper, we do not investigate the power properties of specification tests under a sequence of local misspecification alternatives, as done for instance in Newey (1985). Note again the distinction between the

Let us recall first the definitions of the two test statistics, where $\hat{S}_T$ denotes a standard consistent estimator of the long-term covariance matrix $S(\theta_0)$:

Definition 2.1 *To test $H_0 : \theta = \theta_0$, (i) the LM statistic is defined as,*

$$LM(\theta_0) = T\overline{\phi}'_T(\theta_0)\hat{S}_T'^{-1/2} A_T(\theta_0)\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0)$$

with $A_T(\theta_0) = \hat{S}_T^{-1/2} \frac{\partial\overline{\phi}_T(\theta_0)}{\partial\theta'} \left[\frac{\partial\overline{\phi}'_T(\theta_0)}{\partial\theta} \hat{S}_T^{-1} \frac{\partial\overline{\phi}_T(\theta_0)}{\partial\theta'}\right]^{-1} \frac{\partial\overline{\phi}'_T(\theta_0)}{\partial\theta} \hat{S}_T^{-1/2'}$.
(ii) and the K statistic is defined as,

$$K(\theta_0) = T\overline{\phi}'_T(\theta_0)\hat{S}_T'^{-1/2} \tilde{A}_T(\theta_0)\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0)$$

with $\tilde{A}_T(\theta_0) = \hat{S}_T^{-1/2} \frac{\partial\tilde{\phi}_T(\theta_0)}{\partial\theta'} \left[\frac{\partial\tilde{\phi}'_T(\theta_0)}{\partial\theta} \hat{S}_T^{-1} \frac{\partial\tilde{\phi}_T(\theta_0)}{\partial\theta'}\right]^{-1} \frac{\partial\tilde{\phi}'_T(\theta_0)}{\partial\theta} \hat{S}_T^{-1/2'}$.

The above definitions emphasize that the only difference between the two test statistics is their weighting matrices, respectively $A_T(\theta_0)$ and $\tilde{A}_T(\theta_0)$ for LM and K. The main result of this section, theorem 2.5, compares the powers of the above test statistics against sequences of local alternatives (when varying δ_T and γ, respectively the rate and the direction of local departure). To precisely understand when and how the LM and K test statistics behave differently, we study first the asymptotic behavior of the key elements defining them: $\left[\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0)\right]$, the corrected Jacobian and the matrices $A_T(\theta_0)$ and $\tilde{A}_T(\theta_0)$. The following theorems collect these results.

Theorem 2.1 *(Asymptotic behavior of $\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0)$)*
Under H_{1T}:
(i) If $\delta_T = \sqrt{T}$,

$$\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0) \sim \mathcal{N}(0, I) + S^{-1/2}(\theta_0)c \quad \textit{where} \quad c' = \left[\gamma' \frac{\partial\rho'_1(\theta_0)}{\partial\theta} \;\; 0\right]$$

(ii) If $\delta_T = \lambda_T$,
- If $\gamma \in Im\left[\frac{\partial\rho_1(\theta_0)}{\partial\theta'}\right]$, $\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0) = \mathcal{O}_P\left(\frac{\sqrt{T}}{\lambda_T}\right)$
- If $\gamma \in Im\left[\frac{\partial\rho_1(\theta_0)}{\partial\theta'}\right]^{\perp}$, $\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0) \sim \mathcal{N}(0, I) - S^{-1/2}(\theta_0)c \quad$ where $\quad c' = \left[0 \;\; \gamma' \frac{\partial\rho'_2(\theta_0)}{\partial\theta}\right]$

true value of the parameter (θ^0) and its value under the null (θ_0). See also the discussion in Hall (2005, 5.3) about the connection between a rejection of H_0 and a misspecified model. These considerations are beyond the scope of this paper.

Next, we show that the corrected Jacobian does not behave in a standard way with nearly-weak identification.

Theorem 2.2 *(Asymptotic behavior of the corrected Jacobian)*
Under H_{1T}:
(i) If $\delta_T = \sqrt{T}$:

$$\sqrt{T}\frac{\partial \tilde{\phi}_T(\theta_0)}{\partial \theta'} \sim \begin{pmatrix} \sqrt{T}\frac{\partial \rho_1(\theta_0)}{\partial \theta'} \\ \lambda_T \frac{\partial \rho_2(\theta_0)}{\partial \theta'} \end{pmatrix}$$

(ii) If $\delta_T = \lambda_T$ and $\lambda_T^2 >> \sqrt{T}$ (nearly-strong identification):

$$\sqrt{T}\frac{\partial \tilde{\phi}_T(\theta_0)}{\partial \theta'} \sim \begin{pmatrix} \sqrt{T}\frac{\partial \rho_1(\theta_0)}{\partial \theta'} \\ \lambda_T \frac{\partial \rho_2(\theta_0)}{\partial \theta'} \end{pmatrix}$$

(iii) If $\delta_T = \lambda_T$ and $\lambda_T^2 << \sqrt{T}$ (nearly-weak identification):

$$\sqrt{T}\frac{\partial \tilde{\phi}_T(\theta_0)}{\partial \theta'} \sim \begin{pmatrix} \sqrt{T}\frac{\partial \rho_1(\theta_0)}{\partial \theta'} \\ \frac{\sqrt{T}}{\lambda_T} B(\theta_0, \gamma) \end{pmatrix}$$

where $B(\theta_0, \gamma)$ is the $(k_2 \times p)$-matrix with j^{th} column defined as
$B_j = Cov\left(\sqrt{T}\frac{\partial \bar{\phi}_{2T}(\theta_0)}{\partial \theta_j}, \sqrt{T}\bar{\phi}_T(\theta_0)\right) S^{-1}(\theta_0) \begin{pmatrix} \frac{\partial \rho_1(\theta_0)}{\partial \theta'}\gamma \\ 0 \end{pmatrix}$ *for $j = 1, \cdots, p$.*

Theorem 2.3 *(Asymptotic behavior of the matrix $A_T(\theta_0)$)*
$A_T(\theta_0)$ is asymptotically equivalent to the following (full-column rank) projection matrix

$$A(\theta_0) = S^{-1/2}(\theta_0)\frac{\partial \rho(\theta_0)}{\partial \theta'}\left[\frac{\partial \rho'(\theta_0)}{\partial \theta}S^{-1}(\theta_0)\frac{\partial \rho(\theta_0)}{\partial \theta'}\right]^{-1}\frac{\partial \rho'(\theta_0)}{\partial \theta}S^{-1/2\prime}(\theta_0)$$

Next, we show that the weighting matrix $\tilde{A}_T(\theta_0)$ inherits the non-standard behavior of the corrected Jacobian with nearly-weak identification, while it behaves as $A_T(\theta_0)$ with (nearly)-strong identification.

Theorem 2.4 *(Asymptotic behavior of the matrix $\tilde{A}_T(\theta_0)$)*
(i) If $\delta_T = \sqrt{T}$, $\tilde{A}_T(\theta_0)$ is asymptotically equivalent to the projection matrix $A(\theta_0)$.
(ii) If $\delta_T = \lambda_T$ and $\lambda_T^2 >> \sqrt{T}$, $\tilde{A}_T(\theta_0)$ is asymptotically equivalent to the projection matrix

$A(\theta_0)$.

(iii) If $\delta_T = \lambda_T$ *and* $\lambda_T^2 \leq\leq \sqrt{T}$*:*

- *when* N *is full-column rank,* $\tilde{A}_T(\theta_0)$ *is asymptotically equivalent to*

$$\tilde{A}_T(\theta_0) \sim S^{-1/2}(\theta_0) N \left[N' S^{-1}(\theta_0) N\right]^{-1} N' S^{-1/2\prime}(\theta_0)$$

$$\text{with } N \equiv \begin{pmatrix} \frac{\partial \rho_1(\theta_0)}{\partial \theta'} \\ B(\theta_0, \gamma) \end{pmatrix} \quad \text{and } B(\theta_0,\gamma) \text{ has been defined in theorem 2.2}$$

- *when* N *is not full-column rank,* $\tilde{A}_T(\theta_0)$ *is not asymptotically equivalent to a projection matrix of rank* p*.*

The question of interest is to determine when the asymptotic equivalent matrix of $\tilde{A}_T(\theta_0)$ is a projection matrix of rank p and when it is not. In general, we cannot answer this question. There is at least one case where we can conclude that $\tilde{A}_T(\theta_0)$ is not asymptotically equivalent to a projection matrix. This happens when γ is not spanned by the column-space[4] defined by $[\partial \rho_1'(\theta_0)/\partial\theta]$, that is when γ is not identified by the standard group of moment conditions. More formally,

$$\begin{aligned}
\gamma \in Im\left[\frac{\partial \rho_1'(\theta_0)}{\partial \theta}\right]^{\perp} \quad & \Longrightarrow \quad \frac{\partial \rho_1(\theta_0)}{\partial \theta'}\gamma = 0 \\
& \Longrightarrow \quad N = \begin{bmatrix} \frac{\partial \rho_1(\theta_0)}{\partial \theta'} \\ 0 \end{bmatrix} \quad \text{with} \quad Rank\ N < p \\
& \Longrightarrow \quad \left[N' S^{-1}(\theta_0) N\right] \text{ is not invertible} \\
& \Longrightarrow \quad \tilde{A}_T(\theta_0) \text{ is not asymptotically full-column rank, hence not} \\
& \qquad\quad \text{equivalent to a projection matrix of rank } p.
\end{aligned}$$

We now state the main result of this note.

Theorem 2.5 *(Power of* LM *and* K *test statistics)*
(i) With strong identification (only standard moment conditions), $LM(\theta_0)$ *and* $K(\theta_0)$ *are*

[4]For any $(n \times m)$-matrix M, $Im[M]$ represents the subspace of $\mathbb{R}^n$ generated by the column vectors of M. It is also referred to as $Col[M]$ and $Range[M]$.

asymptotically equivalent. They have the following power against local alternatives H_{1T} at rate $\delta_T = \sqrt{T}$:

$$K(\theta_0) \sim LM(\theta_0) \xrightarrow{d} \chi^2_p(\mu) \qquad \text{(under } H_{1T}\text{)}$$

$$\text{with} \quad \mu = \gamma' \frac{\partial \rho'(\theta_0)}{\partial \theta} S^{-1}(\theta_0) \frac{\partial \rho(\theta_0)}{\partial \theta'} \gamma \quad \text{and} \quad \frac{\partial \rho(\theta_0)}{\partial \theta'} \gamma \neq 0 \;\; \forall \, \gamma$$

(ii) With nearly-strong identification ($\lambda_T^2 >> \sqrt{T}$), $LM(\theta_0)$ and $K(\theta_0)$ are asymptotically equivalent. They have the following power against local alternatives H_{1T}:

- when $\delta_T = \sqrt{T}$ and $\gamma \in Im\,[\partial \rho_1'(\theta_0)/\partial \theta]$:

$$K(\theta_0) \sim LM(\theta_0) \xrightarrow{d} \chi^2_p(\mu) \qquad \text{(under } H_{1T}\text{)}$$

$$\text{with} \quad \mu = [\gamma' \frac{\partial \rho_1'(\theta_0)}{\partial \theta} \;\; 0] S^{-1}(\theta_0) \begin{bmatrix} \partial \rho_1(\theta_0)/\partial \theta' \gamma \\ 0 \end{bmatrix} \quad \text{and} \quad \frac{\partial \rho_1(\theta_0)}{\partial \theta'} \gamma \neq 0$$

- when $\delta_T = \sqrt{T}$ and $\gamma \in Im\,[\partial \rho_1'(\theta_0)/\partial \theta]^{\perp}$, the power is equal to the size:

$$K(\theta_0) \sim LM(\theta_0) \xrightarrow{d} \chi^2_p \qquad \text{(under } H_{1T}\text{)}$$

- when $\delta_T = \lambda_T$ and $\gamma \in Im\,[\partial \rho_1'(\theta_0)/\partial \theta]$:

$$K(\theta_0) \sim LM(\theta_0) \sim \mathcal{O}_P\left(\frac{T}{\lambda_T^2}\right) \quad \text{and} \quad K(\theta_0) > 0 \qquad \text{(under } H_{1T}\text{)}$$

- when $\delta_T = \lambda_T$ and $\gamma \in Im\,[\partial \rho_1'(\theta_0)/\partial \theta]^{\perp}$:

$$K(\theta_0) \sim LM(\theta_0) \xrightarrow{d} \chi^2_p(\mu) \qquad \text{(under } H_{1T}\text{)}$$

$$\text{with} \quad \mu = [0 \;\; \gamma' \frac{\partial \rho_2'(\theta_0)}{\partial \theta}] S^{-1}(\theta_0) \begin{bmatrix} 0 \\ \partial \rho_2(\theta_0)/\partial \theta' \gamma \end{bmatrix} \quad \text{and} \quad \frac{\partial \rho_2(\theta_0)}{\partial \theta'} \gamma \neq 0$$

(iii) With nearly-weak identification ($\lambda_T^2 \leq\leq \sqrt{T}$), $LM(\theta_0)$ and $K(\theta_0)$ are not asymptotically equivalent. $LM(\theta_0)$ has the same asymptotic behavior and power as in case (ii).

- when $\delta_T = \sqrt{T}$: same as the similar case in (ii)

- when $\delta_T = \lambda_T$ and $\gamma \in Im\,[\partial \rho_1'(\theta_0)/\partial \theta]$: we cannot conclude about the asymptotic behavior of $\tilde{A}_T(\theta_0)$

- when $\delta_T = \lambda_T$ and $\gamma \in Im\,[\partial \rho_1'(\theta_0)/\partial \theta]^{\perp}$: $\tilde{A}_T(\theta_0)$ is not asymptotically full-column rank, hence not asymptotically equivalent to a projection matrix of rank p.

With strong identification, both tests have power against every direction of the local alternatives at the standard rate $\sqrt{T}$. When strong identification fails, this is not the case anymore. In particular, the framework of Antoine and Renault (2007) allows us to find standard directions (at rate $\sqrt{T}$) against which the tests only have power equal to the size. We can also see that the tests have some power against slower alternatives (at rate λ_T): power that may depend (again) on the direction of departure from the null hypothesis. Such a power study is only possible because we decided to go one step further in the specification of the identification issues.

3 Testing hypotheses on subvectors

So far, we have focused on testing jointly the entire vector of the structural parameters θ. We might also be interested in testing a subset of these parameters, say $H_0^* : \beta = \beta_0$ when $\theta = (\alpha' \ \beta')'$. To do so, Kleibergen (2005) needs an additional assumption ensuring the full rank of the partial expected Jacobian with respect to the free parameters:

Assumption 4 *(Full rank of the partial expected Jacobian from Kleibergen (2005))*

$$\lim_{T\to\infty} E\left[\frac{\partial\bar{\phi}_T(\theta)}{\partial\alpha'}\right] \quad \textit{is a continuous function of } \theta \textit{ and has full rank at } \theta_0 = (\alpha_0' \ \beta_0')'.$$

Checking the validity of the above assumption involves several difficulties. First, of course, θ_0 is partially unknown under H_0^*. But, more generally, as mentioned p1111 in Kleibergen (2005), "it is not always straightforward to determine the parameters for which the assumption is satisfied".

However, in the framework of Antoine and Renault (2007), we do not meet such difficulties. After the convenient reparametrization, assumption 4 naturally holds for each (new) subvector identified by a group of moment conditions. Post-multiplying the initial Jacobian matrix by the matrix of reparametrization allows us to reinterpret the new Jacobian matrix as in assumption 4:

$$E\left[\frac{\partial\bar{\phi}_T(\theta)}{\partial\theta'}\right] R^0 = E\left[\frac{\partial\bar{\phi}_T(\theta)}{\partial\eta_1'} \ \frac{\partial\bar{\phi}_T(\theta)}{\partial\eta_2'}\right] \equiv \left[\frac{\partial\rho(\theta)}{\partial\eta_1'} \ \frac{\partial\rho(\theta)}{\partial\eta_2'}\right]$$

and each submatrix $[\partial\rho(\theta_0)/\partial\eta_i']$ has full column rank. In other words, testing the entire subvector η_1 or the entire subvector η_2 with the standard LM procedure works without any additional hypothesis.

Finally, note that, in general, an additional assumption is required when testing linear combinations of the structural parameters. This is to avoid perverse asymptotic correlations happening because of the multiplicity of rates of convergence. See also section 4 in Antoine and Renault (2007) for further details on Wald testing any transformation of the parameters.

4 Conclusion

In this note, we have performed a comparative power study between the standard GMM-LM test and its correction proposed by Kleibergen (2005).

We have shown that this correction does have asymptotic consequences, especially with heterogeneous identification patterns. Hence, we recommend carefulness, especially when instruments of heterogeneous quality are used. Moreover, we also recommend using the framework of Antoine and Renault (2007). As shown in this note, it is not much more involved in terms of specifying the identification issues. In addition, not only it enables the use of (valid) standard test procedures (like GMM-LM and Wald), but also it helps identify the directions against which the tests have power.

In terms of testing hypothesis on subvectors, the superiority of the framework of Antoine and Renault (2007) is clear. The reparametrization (see section 3 in Antoine and Renault (2007)) precisely identifies the directions in the parameter space for which the standard GMM-LM test can be performed. In particular, no additional assumption on the free (remaining) parameters is required as for the K-test of Kleibergen (2005). More generally, this framework also deals with (nonlinear) transformations of the structural parameters. This is beyond the scope of Kleibergen (2005) (see section 4 in Antoine and Renault (2007)).

References

[1] B. Antoine, H. Bonnal, and E. Renault, *On the Efficient Use of the Informational Content of Estimating Equations: Implied Probabilities and Euclidean Empirical Likelihood*, Journal of Econometrics **138** (2007), 461–487.

[2] B. Antoine and E. Renault, *Efficient GMM with Nearly-weak Identification*, Working Paper (2007).

[3] J. Davidson, *Econometric Theory*, Blackwell Publisher, Oxford, 2000.

[4] S.G. Donald and W.K. Newey, *A Jackknife interpretation of the Continuous Updating Estimator*, Economics Letter **67** (2000), 239–243.

[5] C. Gouriéroux and A. Monfort, *Statistics and Econometric Models*, Cambridge University Press, Cambridge, 1995.

[6] A.R. Hall, *Generalized Method of Moments*, Advanced Texts in Econometrics, Oxford University Press, 2005.

[7] F. Kleibergen, *Testing parameters in GMM without assuming that they are identified*, Econometrica **73** (2005), 1103–1123.

[8] W.K. Newey, *Generalized Method of Moments Specification Testing*, Journal of Econometrics **29** (1985), 229–256.

[9] W.K. Newey and R.J. Smith, *Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators*, Econometrica **72** (2004), 219–255.

Appendix

Proof of Theorem 2.1 *(Behavior of* $\sqrt{T}\hat{S}_T^{-1/2}(\theta_0)\overline{\phi}_T(\theta_0)$*)*:
The application of the mean-value theorem gives:

$$\overline{\phi}_T(\theta_0) = \overline{\phi}_T(\theta^{(T)} - \gamma/\delta_T) = \overline{\phi}_T(\theta^{(T)}) + \frac{\partial \overline{\phi}_T^*}{\partial \theta'}(\theta_0 - \theta^{(T)}) = \overline{\phi}_T(\theta^{(T)}) - \frac{\partial \overline{\phi}_T^*}{\partial \theta'}\frac{\gamma}{\delta_T}$$

with $\left[\frac{\partial \overline{\phi}_T^*}{\partial \theta'}\right]$ the Jacobian matrix evaluated at a vector with each component between θ_0 and $\theta^{(T)}$. In addition, we have:

$$\begin{aligned} Var[\sqrt{T}\overline{\phi}_T(\theta_0)] &= Var\left(\sqrt{T}\overline{\phi}(\theta^{(T)}) - \sqrt{T}\frac{\partial \overline{\phi}_T^*}{\partial \theta'}\frac{\gamma}{\delta_T}\right) \\ &= Var[\sqrt{T}\overline{\phi}_T(\theta^{(T)})] + Var\left(\sqrt{T}\frac{\partial \overline{\phi}_T^*}{\partial \theta'}\frac{\gamma}{\delta_T}\right) \\ &\quad -2Cov\left(\sqrt{T}\overline{\phi}_T(\theta^{(T)}), \sqrt{T}\frac{\partial \overline{\phi}_T^*}{\partial \theta'}\frac{\gamma}{\delta_T}\right) \\ &\sim Var[\sqrt{T}\overline{\phi}_T(\theta^{(T)})] \end{aligned}$$

Finally,

$$\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0) = \sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta^{(T)}) - \sqrt{T}\hat{S}_T^{-1/2}\frac{\partial \overline{\phi}_T^*}{\partial \theta'}\frac{\gamma}{\delta_T}$$

We can also deduce that under H_{1T}[5]:

$$RHS(1) \sim \mathcal{N}(0, I) \quad \text{and} \quad RHS(2) \sim S^{-1/2}(\theta_0)\begin{pmatrix} \frac{\sqrt{T}}{\delta_T}\frac{\partial \rho_1(\theta_0)}{\partial \theta'}\gamma \\ \frac{\lambda_T}{\delta_T}\frac{\partial \rho_2(\theta_0)}{\partial \theta'}\gamma \end{pmatrix}$$

(i) $\delta_T = \sqrt{T}$:

$$\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0) \quad \sim \quad \mathcal{N}(0, I) - S^{-1/2}(\theta_0)\begin{pmatrix} \frac{\partial \rho_1(\theta_0)}{\partial \theta'}\gamma \\ 0 \end{pmatrix}$$

(ii) $\delta_T = \lambda_T$:

$$RHS(2) \sim S^{-1/2}(\theta_0)\begin{pmatrix} \frac{\sqrt{T}}{\lambda_T}\frac{\partial \rho_1(\theta_0)}{\partial \theta'}\gamma \\ \frac{\partial \rho_2(\theta_0)}{\partial \theta'}\gamma \end{pmatrix}$$

[5] Note that the result for $RHS(1)$ is a little more involved because we now deal with an element of a triangular array. See Davidson (2000) p298 for a similar discussion and appropriate regularity conditions.

- If $\gamma \in Im\left[\frac{\partial \rho_1(\theta_0)}{\partial \theta'}\right]$, $\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0) = \mathcal{O}_p(1) + \mathcal{O}_p\left(\frac{\sqrt{T}}{\lambda_T}\right) = \mathcal{O}_p\left(\frac{\sqrt{T}}{\lambda_T}\right)$
- If $\gamma \in Im\left[\frac{\partial \rho_1(\theta_0)}{\partial \theta'}\right]^{\perp}$, $\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0) \sim \mathcal{N}(0, I) - S^{-1/2}(\theta_0)\begin{pmatrix} 0 \\ \frac{\partial \rho_2(\theta_0)}{\partial \theta'}\gamma \end{pmatrix}$ ■

Proof of Theorem 2.2 *(Behavior of the corrected Jacobian)*:
At the beginning of this proof, we treat each component of the moment conditions separately: therefore, the index $i = 1, \cdots, K$ refers to the component and not to the group of moment conditions as in the main text. Recall first the definition of the corrected Jacobian:

$$\sqrt{T}\frac{\partial \tilde{\phi}_{iT}(\theta_0)}{\partial \theta_j} = \sqrt{T}\frac{\partial \overline{\phi}_{iT}(\theta_0)}{\partial \theta_j} - B_{ij}\sqrt{T}\overline{\phi}_T(\theta_0)$$

where $B_{ij} = Cov\left(\sqrt{T}\frac{\partial \overline{\phi}_{iT}(\theta_0)}{\partial \theta_j}, \sqrt{T}\overline{\phi}_T(\theta_0)\right) S^{-1}(\theta_0)$ for $i = 1, \cdots, K$ and $j = 1, \cdots, p$.
The application of the mean-value theorem gives:

$$\frac{\partial \overline{\phi}_{iT}(\theta_0)}{\partial \theta_j} = \frac{\partial \overline{\phi}_{iT}(\theta^{(T)})}{\partial \theta_j} - \frac{\partial}{\partial \theta'}\left[\frac{\partial \overline{\phi}_{iT}}{\partial \theta_j}\right]^{**}\frac{\gamma}{\delta_T}$$

where $[.]^{**}$ denotes the Hessian evaluated at a vector whose components are between θ_0 and $\theta^{(T)}$. Recall also (see proof of theorem 2.1):

$$\sqrt{T}\overline{\phi}_T(\theta_0) \quad = \quad \sqrt{T}\overline{\phi}_T(\theta^{(T)}) - \frac{\partial \overline{\phi}_T^*}{\partial \theta'}\sqrt{T}\frac{\gamma}{\delta_T}$$

We deduce:

$$\sqrt{T}\frac{\partial \tilde{\phi}_{iT}(\theta_0)}{\partial \theta_j} \quad = \quad \sqrt{T}\frac{\partial \overline{\phi}_{iT}(\theta^{(T)})}{\partial \theta_j} - \frac{\partial}{\partial \theta'}\left[\frac{\partial \overline{\phi}_{iT}}{\partial \theta_j}\right]^{**}\sqrt{T}\frac{\gamma}{\delta_T} - B_{ij}\sqrt{T}\overline{\phi}_T(\theta^{(T)}) + B_{ij}\frac{\partial \overline{\phi}_T^*}{\partial \theta'}\sqrt{T}\frac{\gamma}{\delta_T}$$

Define first the block-diagonal matrix Λ_T,
$\Lambda_T = \begin{pmatrix} Id_{k_1} & 0 \\ 0 & \frac{\lambda_T}{\sqrt{T}}Id_{k_2} \end{pmatrix}$ and μ_{iT} the rate of convergence associated to the i^{th} component, $\mu_{iT} = \sqrt{T}$ for $1 \leq i \leq k_1$ and $\mu_{iT} = \lambda_T$ for $k_1 + 1 \leq i \leq K$.
- with assumption 2(iv) from Antoine and Renault (2007) (about the Plim of the well-scaled Jacobian) we get: $RHS(1) \sim \mu_{iT}\frac{\partial \rho_i(\theta^{(T)})}{\partial \theta_j}$.
- with assumption 3(ii) from Antoine and Renault (2007)p (about the Plim of the well-scaled Hessian) we get: $RHS(2) \sim \mu_{iT}H_i\frac{\gamma}{\delta_T}$ for some fixed matrix H_i. This is dominated by $RHS(1)$.
- $RHS(3) \sim B_{ij}\sqrt{T}\Lambda_T\rho(\theta^{(T)})$ and $\rho(\theta^{(T)}) = 0$ under H_{1T}.
- $RHS(4) \sim B_{ij}\Lambda_T\frac{\partial \rho(\theta^{(T)})}{\partial \theta'}\frac{\gamma}{\delta_T}$.

Finally,

$$\sqrt{T}\frac{\partial \tilde{\phi}_{iT}(\theta_0)}{\partial \theta_j} \quad \sim \quad \mu_{iT}\frac{\partial \rho_i(\theta^{(T)})}{\partial \theta_j} + B_{ij}\left[\begin{array}{c} \sqrt{T}\frac{\partial \rho_1(\theta^{(T)})}{\partial \theta'} \\ \lambda_T \frac{\partial \rho_2(\theta^{(T)})}{\partial \theta'} \end{array}\right]\frac{\gamma}{\delta_T}$$

• Study of the terms of the RHS:

- when $\mu_{iT} = \sqrt{T}$ (ie the i-th component is strong):

$$\sqrt{T}\frac{\partial \tilde{\phi}_{iT}(\theta_0)}{\partial \theta_j} = \sqrt{T}\frac{\partial \rho_i(\theta^{(T)})}{\partial \theta_j} + B_{ij}\left[\begin{array}{c} \sqrt{T}\frac{\partial \rho_1(\theta^{(T)})}{\partial \theta'} \\ \lambda_T \frac{\partial \rho_2(\theta^{(T)})}{\partial \theta'} \end{array}\right]\frac{\gamma}{\delta_T} \quad \Rightarrow \quad \frac{\partial \tilde{\phi}_{iT}(\theta_0)}{\partial \theta_j} \sim \frac{\partial \rho_i(\theta_0)}{\partial \theta_j}$$

- when $\mu_{iT} = \lambda_T$ (ie the i-th component is nearly-weak):

$$\sqrt{T}\frac{\partial \tilde{\phi}_{iT}(\theta_0)}{\partial \theta_j} \sim \lambda_T\frac{\partial \rho_i(\theta_0)}{\partial \theta_j} + B_{ij}\left[\begin{array}{c} \sqrt{T}\frac{\partial \rho_1(\theta^{(T)})}{\partial \theta'} \\ \lambda_T \frac{\partial \rho_2(\theta^{(T)})}{\partial \theta'} \end{array}\right]\frac{\gamma}{\delta_T}$$

* if $\delta_T = \sqrt{T}$:

$$\sqrt{T}\frac{\partial \tilde{\phi}_{iT}(\theta_0)}{\partial \theta_j} \sim \lambda_T\frac{\partial \rho_i(\theta_0)}{\partial \theta_j}$$

* if $\delta_T = \lambda_T$ and $\lambda_T >> \sqrt{T}/\lambda_T$:

$$\sqrt{T}\frac{\partial \tilde{\phi}_{iT}(\theta_0)}{\partial \theta_j} \sim \lambda_T\frac{\partial \rho_i(\theta_0)}{\partial \theta_j}$$

* if $\delta_T = \lambda_T$ and $\lambda_T << \sqrt{T}/\lambda_T$:

$$\sqrt{T}\frac{\partial \tilde{\phi}_{iT}(\theta_0)}{\partial \theta_j} \sim B_{ij}\left[\begin{array}{c} \frac{\sqrt{T}}{\lambda_T}\frac{\partial \rho_1(\theta_0)}{\partial \theta'} \\ \frac{\partial \rho_2(\theta_0)}{\partial \theta'} \end{array}\right]\gamma$$

• Extension to treat all components simultaneously: we are back to our regular formalism where the indexes 1 and 2 refer to the groups of moment conditions.

The above calculations lead to:

$$\sqrt{T}\frac{\partial \tilde{\phi}_T}{\partial \theta'} \quad \sim \quad \left(\begin{array}{c} \sqrt{T}\frac{\partial \rho_1(\theta_0)}{\partial \theta'} \\ \lambda_T \frac{\partial \rho_2(\theta_0)}{\partial \theta'} + B_T(\theta_0, \gamma, \delta_T) \end{array}\right)$$

where $B_T(.)$ is the $(k_2 \times p)$-matrix with j^{th} column defined as
$B_{Tj} = Cov(\sqrt{T}\frac{\partial \bar{\phi}_{2T}(\theta_0)}{\partial \theta_i}, \sqrt{T}\phi_T(\theta_0))S^{-1}(\theta_0) \begin{pmatrix} \frac{\sqrt{T}}{\delta_T}\frac{\partial \rho_1(\theta_0)}{\partial \theta'}\gamma \\ \frac{\lambda_T}{\delta_T}\frac{\partial \rho_2(\theta_0)}{\partial \theta'}\gamma \end{pmatrix}$ for $j = 1, \cdots, p$.
To conclude:
(i) $\delta_T = \sqrt{T}$ or (ii) $\delta_T = \lambda_T$ and $\lambda_T^2 >> \sqrt{T}$:

$$\sqrt{T}\frac{\partial \tilde{\phi}_T(\theta_0)}{\partial \theta'} \sim \begin{pmatrix} \sqrt{T}\frac{\partial \rho_1(\theta_0)}{\partial \theta'} \\ \lambda_T\frac{\partial \rho_2(\theta_0)}{\partial \theta'} \end{pmatrix}$$

(iii) $\delta_T = \lambda_T$ and $\lambda_T^2 << \sqrt{T}$:

$$\sqrt{T}\frac{\partial \tilde{\phi}_T(\theta_0)}{\partial \theta'} \sim \begin{pmatrix} \sqrt{T}\frac{\partial \rho_1(\theta_0)}{\partial \theta'} \\ \frac{\sqrt{T}}{\lambda_T}B(\theta_0, \gamma) \end{pmatrix}$$

■

Proof of Theorem 2.3 *(Asymptotic behavior of the matrix $A_T(\theta_0)$)*:
Recall the mean-value theorem on the Jacobian:

$$\sqrt{T}\frac{\partial \bar{\phi}_{iT}(\theta_0)}{\partial \theta_j} = \sqrt{T}\frac{\partial \bar{\phi}_T(\theta^{(T)})}{\partial \theta_j} - \sqrt{T}\frac{\partial}{\partial \theta'}\left[\frac{\partial \bar{\phi}_{iT}}{\partial \theta_j}\right]^{**}\frac{\gamma}{\delta_T} \sim \mu_{iT}\frac{\partial \rho_i(\theta^{(T)})}{\partial \theta_j}$$

Λ_T, as defined in the proof of theorem 2.2, is invertible for any sample size T. We deduce:

$$\begin{aligned} A_T(\theta_0) &= \hat{S}_T^{-1/2}\frac{\partial \bar{\phi}_T(\theta_0)}{\partial \theta'}\Lambda_T\left[\Lambda_T\frac{\partial \bar{\phi}_T'(\theta_0)}{\partial \theta}\hat{S}_T^{-1}\frac{\partial \bar{\phi}_T(\theta_0)}{\partial \theta'}\Lambda_T\right]^{-1}\Lambda_T\frac{\partial \bar{\phi}_T'(\theta_0)}{\partial \theta}\hat{S}_T^{-1/2\prime} \\ &\sim S^{-1/2}(\theta_0)\frac{\partial \rho(\theta_0)}{\partial \theta'}\left[\frac{\partial \rho'(\theta_0)}{\partial \theta}S^{-1}(\theta_0)\frac{\partial \rho(\theta_0)}{\partial \theta'}\right]^{-1}\frac{\partial \rho'(\theta_0)}{\partial \theta}S^{-1/2\prime}(\theta_0) \end{aligned}$$

which is a projection matrix because by assumption $[\partial \rho(\theta_0)/\partial \theta']$ is full column-rank. ■

Proof of Theorem 2.4 *(Asymptotic behavior of matrix $\tilde{A}_T(\theta_0)$)*:
(i) $\delta_T = \sqrt{T}$:

$$\begin{aligned} \tilde{A}_T(\theta_0) &= \hat{S}_T^{-1/2}\frac{\partial \tilde{\phi}_T(\theta_0)}{\partial \theta'}\left(\frac{\partial \tilde{\phi}_T'(\theta_0)}{\partial \theta}\hat{S}_T^{-1}\frac{\partial \tilde{\phi}_T(\theta_0)}{\partial \theta'}\right)^{-1}\frac{\partial \tilde{\phi}_T'(\theta_0)}{\partial \theta}\hat{S}_T^{-1/2\prime} \\ &= \hat{S}_T^{-1/2}\frac{\partial \tilde{\phi}_T(\theta_0)}{\partial \theta'}\Lambda_T\left(\Lambda_T\frac{\partial \tilde{\phi}_T'(\theta_0)}{\partial \theta}\hat{S}_T^{-1}\frac{\partial \tilde{\phi}_T(\theta_0)}{\partial \theta'}\Lambda_T\right)^{-1}\Lambda_T\frac{\partial \tilde{\phi}_T'(\theta_0)}{\partial \theta}\hat{S}_T^{-1/2\prime} \\ &\sim S^{-1/2}(\theta_0)\frac{\partial \rho(\theta_0)}{\partial \theta'}\left(\frac{\partial \rho'(\theta_0)}{\partial \theta}S^{-1}(\theta_0)\frac{\partial \rho(\theta_0)}{\partial \theta'}\right)^{-1}\frac{\partial \rho'(\theta_0)}{\partial \theta}S^{-1/2\prime}(\theta_0) \end{aligned}$$

where Λ_T is the invertible matrix defined in the proof of theorem 2.3. The last matrix is $A(\theta_0)$, a projection matrix of rank p because $S^{-1/2}(\theta_0)\partial\rho/\partial\theta'$ is full-column rank by assumption.

(ii) $\delta_T = \lambda_T$ and $\lambda_T^2 >> \sqrt{T}$: similar to (i).

(iii) $\delta_T = \lambda_T$ and $\lambda_T^2 << \sqrt{T}$, we proceed as in the case (i). From theorem 2.2(iii), we have:

$$\frac{\partial\tilde{\phi}_T(\theta_0)}{\partial\theta'}\begin{bmatrix} Id_{s1} & 0 \\ 0 & \lambda_T Id_{p-s1}\end{bmatrix} \sim \begin{pmatrix} \frac{\partial\rho_1(\theta_0)}{\partial\theta'} \\ B(\theta_0,\gamma)\end{pmatrix} \equiv N$$

- When N is full-column rank, $\left[N'S^{-1}(\theta_0)N\right]$ is full rank, hence invertible. We have:

$$\tilde{A}_T(\theta_0) \sim S^{-1/2}(\theta_0)N\left[N'S^{-1}(\theta_0)N\right]^{-1}N'S^{-1/2\prime}(\theta_0)$$

- When N is not full-column rank, $\left[N'S^{-1}(\theta_0)N\right]$ is not invertible. Then, $\tilde{A}_T(\theta_0)$ is not asymptotically equivalent to a projection matrix. ■

Proof of Theorem 2.5 *(Power of LM and K test statistics)*:
(i) From theorems 2.3 and 2.4(i): $A_T(\theta_0) \sim \tilde{A}_T(\theta_0) \sim A(\theta_0)$

$$\Rightarrow\ LM(\theta_0) \sim K(\theta_0) \sim T\overline{\phi}_T'(\theta_0)\hat{S}_T^{-1/2\prime}A(\theta_0)\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0)$$

From theorem 2.1(i): $\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0) \sim \mathcal{N}(0,I) - S^{-1/2}(\theta_0)c$ where $c' = \left[\gamma'\frac{\partial\rho'(\theta_0)}{\partial\theta}\right]$. We get (after applying Corollary B.3 from Gouriéroux and Monfort (1995)):

$$\mu = c'S^{-1/2\prime}(\theta_0)A(\theta_0)S^{-1/2}(\theta_0)c = \frac{1}{2}\gamma'\frac{\partial\rho'(\theta_0)}{\partial\theta}S^{-1}(\theta_0)\frac{\partial\rho(\theta_0)}{\partial\theta'}\gamma$$

(ii) From theorems 2.3 and 2.4(i) and (ii): $A(\theta_0) \sim \tilde{A}_T(\theta_0) \sim A(\theta_0)\ \Rightarrow\ K(\theta_0) \sim LM(\theta_0)$.
- if $\delta_T = \sqrt{T}$, from theorem 2.1(i): $\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0) \sim \mathcal{N}(0,I) - S^{-1/2}(\theta_0)c$ where $c' = \left[\gamma'\frac{\partial\rho_1'(\theta_0)}{\partial\theta}\ 0\right]$. Then the calculation is similar to (i).
Note also that: $\gamma \in Im\left[\frac{\partial\rho_1'(\theta_0)}{\partial\theta}\right]\ \Rightarrow\ \frac{\partial\rho_1(\theta_0)}{\partial\theta'}\gamma \neq 0$; and $\gamma \in Im\left[\frac{\partial\rho_1'(\theta_0)}{\partial\theta}\right]^{\perp}\ \Rightarrow\ \frac{\partial\rho_1(\theta_0)}{\partial\theta'}\gamma = 0$.
- if $\delta_T = \lambda_T$ and $\gamma \in Im\left[\frac{\partial\rho_1'(\theta_0)}{\partial\theta}\right]$, from theorem 2.1(ii):
$\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0) = \mathcal{O}_P\left(\frac{\sqrt{T}}{\lambda_T}\right)\ \Rightarrow\ LM(\theta_0) = \mathcal{O}_P\left(\frac{T}{\lambda_T^2}\right)$.
Also, $LM(\theta_0) > 0$ as a quadratic form with a positive definite weighting matrix $A(\theta_0)$
- if $\delta_T = \lambda_T$ and $\gamma \in Im\left[\frac{\partial\rho_1'(\theta_0)}{\partial\theta}\right]^{\perp}$, from theorem 2.1(ii):
$\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0) = \mathcal{N}(0,I) - S^{-1/2}(\theta_0)c$ where $c' = \left[0\ \gamma'\frac{\partial\rho_2'(\theta_0)}{\partial\theta}\right]$. Then the calculation is similar to (i).

(iii) - From theorems 2.1 and 2.3, there is no distinction in the asymptotic behavior of $\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0)$ and $A_T(\theta_0)$ in cases where $\lambda_T^2 << \sqrt{T}$ or $\lambda_T^2 >> \sqrt{T}$. Hence, $LM(\theta_0)$ behaves similarly in cases (ii) and (iii).
- if $\delta_T = \sqrt{T}$, from theorems 2.1(i), 2.2(i) and 2.4(i), there is no distinction in the asymptotic behavior of $\sqrt{T}\hat{S}_T^{-1/2}\overline{\phi}_T(\theta_0)$ and $\tilde{A}_T(\theta_0)$ in cases where $\lambda_T^2 << \sqrt{T}$ or $\lambda_T^2 >> \sqrt{T}$. Hence, $K(\theta_0)$ behaves similarly as in the similar case in (ii).
- if $\delta_T = \lambda_T$ and $\gamma \in Im\left[\frac{\partial\rho_1'(\theta_0)}{\partial\theta}\right]$, from theorem 2.4(iii) the asymptotic behavior of $\tilde{A}_T(\theta_0)$ is not clear. So we cannot conclude.
- if $\delta_T = \lambda_T$ and $\gamma \in Im\left[\frac{\partial\rho_1'(\theta_0)}{\partial\theta}\right]^{\perp}$, from theorem 2.4(iii) and the comment following it, $\tilde{A}_T(\theta_0)$ is not asymptotically full-column rank, hence not asymptotically equivalent to a projection matrix of rank p. $\blacksquare$