RELATIVE EFFICIENCY WITH EQUIVALENCE CLASSES OF ASYMPTOTIC COVARIANCES

David M. Mandy
and
Carlos Martins-Filho*

University of Missouri and Oregon State University, respectively

February 1998

Abstract. White's [1984] concept of asymptotic variance is shown to allow some ambiguities when used to study asymptotic efficiency. These ambiguities are resolved with some mild conditions on the estimators being studied, because then White's asymptotic variance is an equivalence class in which efficiency conclusions are invariant across members of the class. Among the extant efficiency definitions, the liminf-based definition [White 1994, p. 136] is most informative even though identical conclusions can be obtained under our conditions with earlier definitions, but there are still some notions of efficiency allowed by White's asymptotic variance that can only be detected by weaker efficiency definitions.

JEL Classifications C10, C13, C50.

David M. Mandy
Department of Economics
118 Professional Building
University of Missouri
Columbia, MO 65211 USA
Voice 573-882-1763
FAX 573-882-2697
Internet Mandy @ econ.missouri.edu

Carlos Martins-Filho
Department of Economics
Oregon State University
Ballard Extension Hall 303
Corvallis, OR 97331-3612 USA
Voice 541-737-1476
FAX 541-737-5917
Internet Carlos.Martins @ orst.edu

Key words and phrases. Asymptotic Efficiency.
* We benefitted greatly from particularly insightful comments, and much patience, from an associate editor and an anonymous referee. Naturally they bear no responsibility for any remaining errors.

1. Introduction

Suppose β_n^* and $\tilde{\beta}_n$ are estimators for some parameter vector β and we want to determine their relative asymptotic efficiency. A procedure discussed by White [1994, p. 136] is to first find two sequences of matrices, V_n^* and $\tilde{V}_n$, such that both $V_n^{*-1/2}\sqrt{n}(\beta_n^* - \beta)$ and $\tilde{V}_n^{-1/2}\sqrt{n}(\tilde{\beta}_n - \beta)$ converge in distribution to standard normals. White [1984, p. 66; 1994, p. 91] calls such sequences asymptotic variances, or *avars*. Next these sequences are used to show $\liminf_{n\to\infty} \theta'(\tilde{V}_n - V_n^*)\theta \geq 0 \ \forall\ \theta$, whence we conclude, for example, that β_n^* is asymptotically efficient relative to $\tilde{\beta}_n$. This procedure has an advantage over the traditional (Fisher [1925]) approach of comparing the covariances of the asymptotic distributions, in that V_n^* and $\tilde{V}_n$ need not have limits (White [1982]). But there is a potential disadvantage as well because in all interesting cases there will be other *avar* sequences, say V_n, such that $V_n^{-1/2}\sqrt{n}(\beta_n^* - \beta)$ converges in distribution to a standard normal. Thus, one must inquire whether $\liminf_{n\to\infty} \theta'(\tilde{V}_n - V_n^*)\theta \geq 0 \ \forall\ \theta$ implies $\liminf_{n\to\infty} \theta'(\tilde{V}_n - V_n)\theta \geq 0 \ \forall\ \theta$ for any such V_n, that is, whether statements about asymptotic efficiency are invariant to the *avar* sequences examined. By Fatou's Lemma, this implication holds if $\lim_{n\to\infty}(V_n - V_n^*) = 0$, and the converse holds as well provided the two liminfs are both nonnegative for the same set of $\tilde{V}_n$ sequences (that is, for the same rival estimators $\tilde{\beta}_n$). But without $\lim_{n\to\infty}(V_n - V_n^*) = 0$ anything is possible, including a reversal of the original conclusion, in which case the relative efficiency comparison based on V_n^* would appear definitive when it is in fact ambiguous. Alternatively, V_n^* may yield no conclusion when a conclusion is obtainable using V_n, and the same comments apply to all possible substitutes for V_n^* and $\tilde{V}_n$. Thus, at this level of generality, $\lim_{n\to\infty}(V_n - V_n^*) = 0$ for all relevant sequences V_n^* and V_n is sufficient for meaningful efficiency conclusions. The converse is also of interest, namely, whether two sequences that satisfy this condition are both *avars* given that one of them is.

More formally, we must investigate whether the relation $\mathbf{R}$ defined by

$$\{V_n^*\}_{n=1}^{\infty}\mathbf{R}\{V_n\}_{n=1}^{\infty} \Leftrightarrow \lim_{n\to\infty}(V_n - V_n^*) = 0 \tag{1}$$

has an equivalence class consisting of all sequences that produce the desired asymptotic normal distribution. In this paper we first show this is the appropriate relation for studying relative asymptotic efficiency, in that efficiency conclusions are unambiguous if and only if all pairs of candidate sequences have the relation $\mathbf{R}$. We then show that an *avar* collection is not always an equivalence class with respect to $\mathbf{R}$, but some mild conditions on the underlying random vector are sufficient to ensure that an *avar* collection is indeed an equivalence class with respect to **R**. Hence, when these conditions hold asymptotic efficiency can be studied without ambiguity using the full generality of the *avar* concept.

We demonstrate that the sufficient conditions are equivalent to boundedness of all *avar* sequences and the corresponding sequences of inverses. When *avar* sequences, but not their inverses, are bounded the *avar* collection is a (perhaps proper) subset of an equivalence class with respect to **R**. This is sufficient for the sign of $\liminf_{n\to\infty} \theta'(\tilde{V}_n - V_n^*)\theta$ to be invariant across all V_n^* sequences that are *avar*s of $\sqrt{n}(\beta_n^* - \beta)$, and hence for unambiguous efficiency conclusions by this criterion. However, White [1984, pp. 78-9] provides another definition of efficiency based on the *avar* concept that does not always yield the same efficiency conclusions as this liminf criterion. We show the two definitions are equivalent if *avar* sequences *and* their corresponding sequences of inverses are bounded. Hence, for consistency across definitions *avar* must be an equivalence class with respect to **R**. This rules out both divergences to infinity and approaches to singularity in the collection.

Under these conditions the generalization accomplished by the *avar* concept over the traditional approach is that *avar* accommodates nonconvergent but bounded oscillations in the sequences forming the *avar* class, and their corresponding sequences of inverses. The *avar* concept does not directly deliver assuredly unambiguous efficiency comparisons when there are divergences to infinity or approaches to singularity in the collection. Bounding these sequences is a mild restriction, however, because the random vector can usually be normalized to eliminate such problems before forming the *avar* class, as is customary through multiplication by $\sqrt{n}$. In general, an element of β_n^* can be normalized by any nonstochastic sequence in order to obtain bounded *avars* and inverse *avars*, and the normalizing sequence can differ across elements of β_n^*. Once all estimators under consideration are so normalized, our results show that their relative asymptotic efficiencies can be unambiguously compared irrespective of convergence of the *avar* sequences, the particular sequences examined, or the efficiency definition used. There are two caveats here. First, a multiple correlation between normalized elements of β_n^* might tend to one, in which case some asymptotically redundant element(s) should be dropped before asymptotic efficiency is studied, since this situation would lead to *avar* sequences that approach singularity. Second, because relative efficiency conclusions can be affected if different normalizations are used for β_n^* and $\tilde{\beta}_n$, one should only compare estimators via *avar*s of normalized deviations when corresponding elements are normalized with the same sequence.

Bounds on *avar* sequences and their sequences of inverses are usually (White [1982; 1984, p. 66; 1994, pp. 130-136]) although not always (Bates and White [1993]) imposed, but even when they are imposed they are somewhat unsatisfactory as primitive assumptions because they utilize the *avar* class to restrict itself rather than placing restrictions on the underlying random vector. In contrast, our conditions are imposed directly on the random vector and thus illuminate exactly what is being assumed when bounds are placed on *avar* sequences. To accomplish this we introduce a new order in probability concept, which we call asymptotic linear independence in probability, or *alip*.

Finally, since nonconvergent oscillations are permitted in an *avar* class it is possible to have "one-sided" relative efficiency that is not addressed by either of White's definitions, in the form of either a smaller minimal limiting variance (minimin efficiency) or a smaller maximal limiting variance (minimax efficiency). We introduce new definitions to address these possibilities. Naturally, the new definitions are weaker than White's definitions, so all known *avar* efficiency conclusions automatically hold for minimin and minimax efficiency.

2. The Equivalence Relation **R** for Studying Relative Asymptotic Efficiency

Let Q^q denote the set of all sequences $\{V_n\}_{n=1}^{\infty}$ of real symmetric positive definite nonstochastic $(q \times q)$ matrices V_n. For such V_n, denote by $V_n^{1/2}$ the unique real symmetric positive definite matrix satisfying $V_n^{1/2}V_n^{1/2} = V_n$. In essence, White's definition of asymptotic covariance is the following.

Definition 1 [White 1984, p. 66; White 1994, p. 91][1]**.** Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of q-dimensional random vectors $(q < \infty)$. The *asymptotic covariance of* $\{x_n\}_{n=1}^{\infty}$, denoted $avar(\{x_n\}_{n=1}^{\infty})$, is

$$avar\left(\{x_n\}_{n=1}^{\infty}\right) \equiv \left\{\{V_n\}_{n=1}^{\infty} \in Q^q : V_n^{-1/2}x_n \xrightarrow{d} z \sim N(0, I_q)\right\}.$$

For brevity, we drop the indexes and write $V_n \in avar(x_n)$ to denote that the sequence $\{V_n\}_{n=1}^{\infty}$ is an element of $avar(\{x_n\}_{n=1}^{\infty})$.

[1] White [1984] actually only requires positive definiteness of V_n for large n. Since only the tail of the sequence is important in the definition, there is no consequential loss of generality from assuming the entire sequence is positive definite.

We say Definition 1 is White's definition "in essence" because neither White [1984] nor White [1994] acknowledge that *avar* is a class. Bates and White [1993] explicitly discuss that the *avar* of their RCASOI class of estimators is a class, but attribute this to the flexibility of RCASOI classes. In fact, in all interesting cases the *avar* of any single random vector is a class because, if there exists one sequence V_n satisfying Definition 1, then $A_n {V_n}^{-1/2} x_n \xrightarrow{d} z$ for any $\{A_n\}_{n=1}^{\infty}$ satisfying $\lim_{n\to\infty} A_n = I_q$. Hence, if $V_n \in avar(x_n)$ then $A_n'^{-1} V_n A_n^{-1} \in avar(x_n)\ \forall\ \{A_n\}_{n=1}^{\infty}$ satisfying $A_n V_n^{-1/2} \in Q^q$ and $\lim_{n\to\infty} A_n = I_q$. In other words, a great profusion of *avar* sequences exists whenever a single *avar* sequence exists.

This observation brings to the forefront the issue of whether conclusions about relative asymptotic efficiency are invariant to which element of an *avar* class is examined. What is needed is an equivalence relation in Q^q whose equivalence classes are precisely those collections of sequences for which conclusions about relative asymptotic efficiency are identical. Given this, if $avar(x_n)$ is an equivalence class of the relation, or even a subset of an equivalence class, then efficiency conclusions are unambiguous. To construct the equivalence relation we first need a formal definition of efficiency. We start with:

Definition 2 [White 1994, p. 136][2]**.** Let β_n^* and $\tilde{\beta}_n$ be consistent estimators of a nonstochastic q-dimensional vector β. Then β_n^* is *asymptotically efficient relative to* $\tilde{\beta}_n$ if there exists $V_n^* \in avar(\sqrt{n}(\beta_n^* - \beta))$ and $\tilde{V}_n \in avar(\sqrt{n}(\tilde{\beta}_n - \beta))$ such that

$$\liminf_{n\to\infty} \theta'(\tilde{V}_n - V_n^*)\theta \geq 0$$

for all $\theta \in \Re^q$ $(\theta \neq 0)$. An estimator is *asymptotically efficient within a class* if it is asymptotically efficient relative to every other estimator in the class.

Notice that one consistent estimator can be asymptotically efficient relative to another only if the *avar*'s of both normalized estimators are nonempty.

According to this definition the relation $\mathbf{R}$ in Q^q that yields unambiguous efficiency conclusions within an equivalence class is $\{V_n^*\}_{n=1}^{\infty}\mathbf{R}\{V_n\}_{n=1}^{\infty}$ if and only if

$$(\forall\ \tilde{V}_n):\ \liminf_{n\to\infty} \theta'(\tilde{V}_n - V_n^*)\theta \geq 0\ \forall\ \theta \Leftrightarrow \liminf_{n\to\infty} \theta'(\tilde{V}_n - V_n)\theta \geq 0\ \forall\ \theta. \tag{2}$$

[2]Definitions 2 and 6 (below) are rephrased from White's original statements to reflect the fact that $avar(x_n)$ is a class. We use the same terminology to accomplish this that Bates and White [1993, Definition 2.5] use for a RCASOI class.

It is trivial to verify that $\mathbf{R}$ is an equivalence relation (satisfying reflexivity, symmetry, and transitivity). More interesting is the fact that this relation is identical to the relation defined in equation (1). That (1) implies (2) is an application of Fatou's Lemma. For the converse, $\tilde{V}_n = V_n^*$ in (2) implies $\liminf_{n\to\infty} \theta'(V_n^* - V_n)\theta \geq 0 \;\forall\, \theta$, while $\tilde{V}_n = V_n$ in (2) implies $\limsup_{n\to\infty} \theta'(V_n^* - V_n)\theta \leq 0 \;\forall\; \theta$. We utilize (1) henceforth since it is more convenient, and henceforth denote the equivalence class of V_n^* with respect to $\mathbf{R}$ by $E_{\mathbf{R}}(V_n^*) \equiv \{V_n \in Q^q\colon\; \lim_{n\to\infty}(V_n^* - V_n) = 0\}$.

3. Sufficient Conditions for *avar*(x_n) to be an Equivalence Class with Respect to $\mathbf{R}$

In general *avar*(x_n) is not an equivalence class with respect to $\mathbf{R}$, so some restrictions are needed in the form of regularity conditions on the underlying random vector x_n. These restrictions include the following notion of asymptotic linear independence in probability.

Definition 3. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of q-dimensional random vectors ($q < \infty$) and $\{y_n\}_{n=1}^{\infty}$ be a sequence of strictly positive nonstochastic real numbers. $\{x_n\}_{n=1}^{\infty}$ is *asymptotically linearly independent in probability of order* $\{y_n\}_{n=1}^{\infty}$ if for every sequence $\{c_n\}_{n=1}^{\infty}$ of real nonstochastic q-dimensional vectors satisfying $\|c_n\| = 1 \;\forall\; n$; there exists a triple (N, ϵ, δ), where N is a natural number, $\epsilon > 0$, and $\delta > 0$; such that

$$n \geq N \Rightarrow P\left(\epsilon < \left|\frac{c_n' x_n}{y_n}\right|\right) > \delta.$$

For brevity, this is denoted $x_n = alip(y_n)$.[3]

It is clear that $x_n = alip(y_n) \Rightarrow x_n \neq o_p(y_n)$. The converse fails because $x_n \neq o_p(y_n)$ still permits subsequences of $\frac{c_n' x_n}{y_n}$ that converge in probability to zero, and also linear combinations of $\frac{x_n}{y_n}$ that converge in probability to zero as long as individual components do not.

Definition 4. A sequence $\{x_n\}_{n=1}^{\infty}$ of q-dimensional random vectors ($q < \infty$) is *avar-regular* if $x_n = O_p(1)$ and $x_n = alip(1)$.

[3] The *alip* concept is similar to Mann and Wald's [1943] notion of ω_p, but these concepts differ in two ways that are important in the present context. First, $alip(y_n)$ is weaker than $\omega_p(y_n)$ in that $\omega_p(y_n)$ requires the probability that a normalized random variable is nonzero approach one, while $alip(y_n)$ merely requires that this probability not approach zero. Second, $alip(y_n)$ is stronger than $\omega_p(y_n)$ in that $alip(y_n)$ places its condition on all bounded nondegenerate linear combinations of a random vector, while $\omega_p(y_n)$ merely places its condition on the individual components of a random vector. In general, $\omega_p(1)$ cannot replace $alip(1)$ in the results below.

Avar-regularity places restrictions on the primitive of the problem, the underlying random vector x_n. Theorem 1 shows this is equivalent to White's [1984, p. 66] approach of bounding the *avar* sequences and the corresponding sequences of inverses.

Theorem 1. *Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of q-dimensional random vectors ($q < \infty$) for which $avar(x_n)$ is nonempty. Then*

(i) $x_n = alip(1)$ if and only if V_n^{-1} is bounded $\forall\ V_n \in avar(x_n)$.

(ii) $x_n = O_p(1)$ if and only if V_n is bounded $\forall\ V_n \in avar(x_n)$.

Hence x_n is avar-regular if and only if both V_n and V_n^{-1} are bounded for every $V_n \in avar(x_n)$.

Proof. All proofs are in the Appendix.

The main result of this section is that *avar*-regularity is sufficient to ensure $avar(x_n)$ is an equivalence class with respect to $\mathbf{R}$. Bates and White [1993, Theorem 2.3] investigate whether $V_n \in avar(x_n) \Rightarrow avar(x_n) \subseteq E_{\mathcal{R}}(V_n)$, using the slightly different relation $\mathcal{R}$ in Q^q defined by

$$\{V_n\}_{n=1}^{\infty} \mathcal{R} \{\tilde{V}_n\}_{n=1}^{\infty} \Leftrightarrow \lim_{n\to\infty} V_n^{-1/2} \tilde{V}_n V_n^{-1/2} = I_q. \tag{3}$$

As with $\mathbf{R}$, it is straightforward to verify that $\mathcal{R}$ is an equivalence relation. However, $\mathcal{R}$ does not have the same equivalence classes as $\mathbf{R}$ unless attention is restricted to sequences that are bounded and have bounded inverses. Hence, from equation (2), $\mathcal{R}$ is not always the appropriate relation for studying relative asymptotic efficiency. But Theorem 2 below shows that *avar*-regularity, which by Theorem 1 bounds the candidate sequences and their inverses, implies $avar(x_n)$ is an equivalence class with respect to both $\mathcal{R}$ and $\mathbf{R}$. The proof of Theorem 2 relies on the following preliminary results.

Lemma 1. *Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of q-dimensional random vectors ($q < \infty$). If $x_n \xrightarrow{d} z \sim (0, \Sigma)$, where Σ is positive definite, then $x_n = alip(1)$.*

Lemma 2. *Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of q-dimensional random vectors ($q < \infty$). If $V_n, \tilde{V}_n \in avar(x_n)$ then $\tilde{V}_n^{1/2} V_n^{-1/2}$ is bounded.*

Theorem 2. *Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of q-dimensional random vectors ($q < \infty$), $\mathbf{R}$ be defined by equation (1), $\mathcal{R}$ be defined by equation (3), and $V_n \in avar(x_n)$. Then $avar(x_n) = E_{\mathcal{R}}(V_n)$, and* *(i)* $x_n = alip(1) \Rightarrow E_{\mathcal{R}}(V_n) \supseteq E_{\mathbf{R}}(V_n)$.

(ii) $x_n = O_p(1) \Rightarrow E_{\mathcal{R}}(V_n) \subseteq E_{\mathbf{R}}(V_n)$.

Hence avar-regularity implies that avar(x_n) *is an equivalence class with respect to both* $\mathbf{R}$ *and* $\mathcal{R}$.

Theorem 2 shows that, for an *avar*-regular random vector x_n and an *avar* sequence $V_n \in avar(x_n)$, $avar(x_n)$ is precisely the set of sequences $\{\tilde{V}_n\}_{n=1}^{\infty}$ of real symmetric positive definite matrices such that $\lim_{n\to\infty}(V_n - \tilde{V}_n) = 0_q$. And from equation (2), this is precisely the set of sequences that yield unambiguous efficiency conclusions when Definition 2 is used to define relative efficiency.

It is worth remarking on the role of normality in obtaining this conclusion. None of the proofs given in the Appendix rely on normality of the limiting random vector z, except for the use of the normal characteristic function in establishing $avar(x_n) = E_{\mathcal{R}}(V_n)$ in Theorem 2. Moreover, a proof of $x_n = alip(1) \Rightarrow avar(x_n) \supseteq E_{\mathbf{R}}(V_n)$ that does not rely on normality is available from the authors on request. Hence most results given here hold even if Definition 1 is relaxed to permit convergence to any common random vector z that has zero mean and identity covariance, as in the RCASOI class of estimators discussed by Bates and White [1993]. However, for the purpose of obtaining definitive efficiency comparisons within the *avar* class, $avar(x_n) \subseteq E_{\mathbf{R}}(V_n)$ is the crucial property. This relies on both normality and $x_n = O_p(1)$ in our proof, and Example 2 below shows that $x_n = O_p(1)$ cannot be discarded. Whether normality is necessary for $avar(x_n) \subseteq E_{\mathbf{R}}(V_n)$ or $avar(x_n) \subseteq E_{\mathcal{R}}(V_n)$ is an open question, as we have no counter-examples to these when the limit distribution is non-normal. Bates and White [1993, p. 648] propose a proof of $avar(x_n) \subseteq E_{\mathcal{R}}(V_n)$ that does not rely on normality, but an important and questionable step therein is that $\lim_{n\to\infty} V_n^{-1/2}\tilde{V}_n^{1/2} = I_q$ is implied by both y_n and $(V_n^{-1/2}\tilde{V}_n^{1/2})y_n$ converging in distribution to the same (potentially non-normal) $(0, I_q)$ random vector, for some sequence of $(0, I_q)$ random vectors y_n. This is not obvious. Indeed, our proof resorts to an application of the Mean Value Theorem on the normal characteristic function to make the transformation from convergence in distribution to convergence of sequences of matrix products, and it is not clear how this step might be accomplished for an arbitrary characteristic function.

The following example shows that $x_n = alip(1)$ is needed in Theorems 1 and 2 and cannot be replaced by $x_n \neq o_p(1)$ (or by $x_n = \omega_p(1)$).

Example 1: The role of $x_n = alip(1)$ in making $avar(x_n)$ an equivalence class with respect to

R. Let

$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} \end{bmatrix}\right), \quad x_n = \begin{bmatrix} z_1 - \frac{z_2}{n} \\ z_1 + \frac{z_2}{n} \end{bmatrix}, \quad V_n^{*-1/2} = \begin{bmatrix} n+1 & -n \\ -n & n \end{bmatrix}, \text{ and}$$

$$V_n^{-1/2} = \begin{bmatrix} n^2+1 & -n^2 \\ -n^2 & n^2 \end{bmatrix}.$$

Then

$$V_n^{*-1/2} x_n = \begin{bmatrix} z_1 - 2z_2 - \frac{z_2}{n} \\ 2z_2 \end{bmatrix} \xrightarrow{p} \begin{bmatrix} z_1 - 2z_2 \\ 2z_2 \end{bmatrix} \sim N(0, I_2),$$

so $V_n^* \in avar(x_n)$, while

$$V_n^{-1/2} x_n = \begin{bmatrix} -2nz_2 + z_1 - \frac{z_2}{n} \\ 2nz_2 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} n^2 - 2n + 3 + \frac{1}{4n^2} - \frac{1}{n} & n - n^2 - \frac{1}{4} \\ n - n^2 - \frac{1}{4} & n^2 \end{bmatrix}\right),$$

so $V_n \notin avar(x_n)$. But

$$V_n^* = \begin{bmatrix} 2 & 2 + \frac{1}{n} \\ 2 + \frac{1}{n} & 1 + (\frac{n+1}{n})^2 \end{bmatrix} \to \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \quad \text{and} \quad V_n = \begin{bmatrix} 2 & 2 + \frac{1}{n^2} \\ 2 + \frac{1}{n^2} & 1 + \left(\frac{n^2+1}{n^2}\right)^2 \end{bmatrix} \to \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix},$$

so $\lim_{n\to\infty}(V_n - V_n^*) = 0_2$, while

$$V_n^{-1/2} V_n^* V_n^{-1/2} = \begin{bmatrix} n^2 - 2n + 2 & n(1-n) \\ n(1-n) & n^2 \end{bmatrix} \to \begin{bmatrix} \infty & -\infty \\ -\infty & \infty \end{bmatrix}.$$

Note that $x_n \neq o_p(1)$ here, but $x_n \neq alip(1)$ since $c_n' = [\,-2^{-1/2} \quad 2^{-1/2}\,]\ \forall\ n$ in Definition 3 yields $c_n' x_n = \frac{z_2\sqrt{2}}{n} \xrightarrow{p} 0$. Hence when $x_n \neq alip(1)$ elements of $avar(x_n)$ can have unbounded inverses, and $E_{\mathcal{R}}(V_n^*)$ and $avar(x_n)$ can both be smaller than $E_{\mathbf{R}}(V_n^*)$, even when $x_n \neq o_p(1)$. Thus *alip* is strictly stronger than o_p. □

The next example shows that $x_n = O_p(1)$ is needed in Theorems 1 and 2 as well.

Example 2: The role of $x_n = O_p(1)$ in making $avar(x_n)$ an equivalence class with respect to R. Let $z \sim N(0,1)$, $x_n = nz$, $V_n^* = n^2$, and $V_n = (n+1)^2$. Clearly $V_n^*, V_n \in avar(x_n)$ and $V_n^* \mathcal{R} V_n$. But $V_n - V_n^* = (n+1)^2 - n^2 = 2n + 1 \to \infty$, so V_n^* and V_n are not in the same equivalence class with respect to $\mathbf{R}$. The problem is $x_n \neq O_p(1)$, in which case elements of $avar(x_n)$ can be unbounded, and $E_{\mathcal{R}}(V_n)$ and $avar(x_n)$ can both be larger than $E_{\mathbf{R}}(V_n)$. □

4. Other Definitions of Asymptotic Efficiency for *avar* Equivalence Classes

White offers another definition of relative asymptotic efficiency based on the *avar* concept. This is the definition used by Bates and White as well.

Definition 6 [White 1984, pp. 78-79]. Let β_n^* and $\tilde{\beta}_n$ be consistent estimators of a nonstochastic vector β. Then β_n^* is *asymptotically efficient relative to* $\tilde{\beta}_n$ if there exists $V_n^* \in avar(\sqrt{n}(\beta_n^* - \beta))$ and $\tilde{V}_n \in avar(\sqrt{n}(\tilde{\beta}_n - \beta))$, and an integer N, such that $\tilde{V}_n - V_n^*$ is positive semidefinite for all $n \geq N$. An estimator is *asymptotically efficient within a class* if it is asymptotically efficient relative to every other estimator in the class.

It is clear that if β_n^* is asymptotically efficient relative to $\tilde{\beta}_n$ according to Definition 6 then β_n^* is asymptotically efficient relative to $\tilde{\beta}_n$ according to Definition 2 as well. So, those estimators identified as efficient using Definition 6 (for example, those studied by White [1984 and 1994]) are *unambiguously* efficient by Definition 2, provided we impose *avar*-regularity (or at least $O_p(1)$) to resolve ambiguity when using Definition 2. More interesting is that the converse holds as well, as shown by Theorem 3, *if the normalized deviation of* $\tilde{\beta}_n$ *is avar-regular*, but not without *avar*-regularity, as shown by Examples 3 and 4 below.

Theorem 3. *Let β_n^* and $\tilde{\beta}_n$ be consistent estimators of a nonstochastic q-dimensional vector β and assume $\sqrt{n}(\tilde{\beta}_n - \beta)$ is avar-regular. If β_n^* is asymptotically efficient relative to $\tilde{\beta}_n$ according to Definition 2 then β_n^* is asymptotically efficient relative to $\tilde{\beta}_n$ according to Definition 6.*

Example 3: The role of $x_n = alip(1)$ in Theorem 3. Let $z \sim N(0,1)$, $\sqrt{n}(\beta_n^* - \beta) = \frac{2z}{n}$, and $\sqrt{n}(\tilde{\beta}_n - \beta) = \frac{z}{n}$. Since $\frac{4}{n^2} \in avar(\sqrt{n}(\beta_n^* - \beta))$, $\frac{1}{n^2} \in avar(\sqrt{n}(\tilde{\beta}_n - \beta))$, and $\liminf_{n\to\infty}(\frac{1}{n^2} - \frac{4}{n^2}) = 0$; β_n^* is efficient relative to $\tilde{\beta}_n$ according to Definition 2. For arbitrary $V_n^* \in avar(\sqrt{n}(\beta_n^* - \beta))$ we have

$$V_n^{*-1/2}\frac{2z}{n} \sim N\left(0, V_n^{*-1}\frac{4}{n^2}\right) \xrightarrow{d} N(0,1),$$

so $V_n^* \frac{n^2}{4} \to 1$. Similarly, for arbitrary $\tilde{V}_n \in avar(\sqrt{n}(\tilde{\beta}_n - \beta))$ we have $\tilde{V}_n n^2 \to 1$. So, there exists N such that

$$n \geq N \Rightarrow \begin{cases} V_n^* > \frac{2}{n^2} \\ \tilde{V}_n < \frac{3}{2n^2}, \end{cases}$$

implying $\tilde{V}_n - V_n^* < -\frac{1}{2n^2}$ for $n \geq N$. Hence β_n^* is not efficient relative to $\tilde{\beta}_n$ according to Definition 6, due to the fact that $\sqrt{n}(\tilde{\beta}_n - \beta) \neq alip(1)$. □

Example 4: The role of $x_n = O_p(1)$ in Theorem 3. The condition $\sqrt{n}(\tilde{\beta}_n - \beta) = O_p(1)$ is only used in the proof of Theorem 3 to establish equicontinuity of the sequence of quadratic forms $f_n(\theta) = \theta'(\tilde{V}_n - V_n^*)\theta$ on the unit sphere. Hence we construct an example in which a sequence of quadratic forms is not equicontinuous on the unit sphere. Let $z \sim N(0, I_2)$, $\sqrt{n}(\beta_n^* - \beta) = V_n^{*1/2} z$, and $\sqrt{n}(\tilde{\beta}_n - \beta) = \tilde{V}_n^{1/2} z$; where

$$V_n^* = \begin{bmatrix} \csc^2 \phi_n & 0 \\ 0 & 1 \end{bmatrix}, \quad \tilde{V}_n = \Phi_n + V_n^*;$$

and ϕ_n and Φ_n are chosen to produce the desired quadratic form. Choosing $\phi_n \in (0, \frac{\pi}{4})$ such that $\phi_n \to 0$ and

$$\Phi_n = \frac{1}{\cos^4 \phi_n - \sin^4 \phi_n} \begin{bmatrix} (\cos^2 \phi_n - \sin^2 \phi_n)(\cot^2 \phi_n - \sin^2 \phi_n) & \sin \phi_n \cos \phi_n (\cot^2 \phi_n - 2\sin^2 \phi_n) \\ \sin \phi_n \cos \phi_n (\cot^2 \phi_n - 2\sin^2 \phi_n) & (\cos^2 \phi_n - \sin^2 \phi_n) \sin^2 \phi_n \end{bmatrix}$$

makes $f_n(\theta) = \theta' \Phi_n \theta$ a saddle (indefinite) rotated ϕ_n from standard position that collapses around the rotated θ_2 axis as $n \to \infty$ but satisfies $f_n(\theta_n) = -1 \ \forall \ n$, where $\theta_n' = (\, -\sin \phi_n \quad \cos \phi_n \,)$. Note that

$$\Phi_n \to \begin{bmatrix} \infty & \infty \\ \infty & 0 \end{bmatrix} \quad \text{and} \quad V_n^* \to \begin{bmatrix} \infty & 0 \\ 0 & 1 \end{bmatrix},$$

so by Theorem 1(ii) we have $\sqrt{n}(\tilde{\beta}_n - \beta) \neq O_p(1)$. It can be shown, however, that $\sqrt{n}(\tilde{\beta}_n - \beta) = \mathit{alip}(1)$ and that $\tilde{V}_n$ is positive definite for n large. Because the $(1,1)$ element of Φ_n is $O(\csc^2 \phi_n)$ while the off-diagonal elements are $O(\csc \phi_n)$, we have

$$\liminf_{n \to \infty} f_n(\theta) = \begin{cases} \infty & \text{if } \theta_1 \neq 0 \\ 0 & \text{if } \theta_1 = 0, \end{cases}$$

so β_n^* is efficient relative to $\tilde{\beta}_n$ according to Definition 2. To show that β_n^* is not efficient relative to $\tilde{\beta}_n$ according to Definition 6, let W_n^* and $\tilde{W}_n$ be arbitrary elements of $\mathit{avar}(\sqrt{n}(\beta_n^* - \beta))$ and $\mathit{avar}(\sqrt{n}(\tilde{\beta}_n - \beta))$, respectively. Then, by the same arguments used in Example 3 we have $V_n^{*-1/2} W_n^* V_n^{*-1/2}$ and $\tilde{V}_n^{-1/2} \tilde{W}_n \tilde{V}_n^{-1/2}$ both converging to I_2. Now write

$$\begin{aligned} \theta_n'(\tilde{W}_n - W_n^*)\theta_n &= \theta_n' \tilde{V}_n^{1/2} [\tilde{V}_n^{-1/2} \tilde{W}_n \tilde{V}_n^{-1/2} - I_2] \tilde{V}_n^{1/2} \theta_n + f_n(\theta_n) \\ &\quad + \theta_n' V_n^{*1/2} [I_2 - V_n^{*-1/2} W_n^* V_n^{*-1/2}] V_n^{*1/2} \theta_n. \end{aligned} \tag{4}$$

Since $\theta_n' V_n^* \theta_n = 1 + \cos^2 \phi_n$ and $\theta_n' \tilde{V}_n \theta_n = f_n(\theta_n) + \theta_n' V_n^* \theta_n = \cos^2 \phi_n$, both $\theta_n' V_n^{*1/2}$ and $\theta_n' \tilde{V}_n^{1/2}$ are bounded. Hence the first and last terms of (4) approach zero, leaving only $f_n(\theta_n) = -1$ for n large, so β_n^* is not efficient relative to $\tilde{\beta}_n$ according to Definition 6. □

We have seen that if two estimators are *avar*-regular then Definition 2 can be used to make an unambiguous efficiency comparison despite the fact that their *avar*'s are classes. In this case, any elements of the classes are representative and so the researcher can just select a convenient element for each estimator to make the efficiency comparison. This is not true of Definition 6, even though efficiency *conclusions* by the two definitions are equivalent under *avar*-regularity. Definition 2, being a limit-based definition, has the advantage that the *avar* comparisons it calls for always yield the same answer within an equivalence class with respect to $\mathbf{R}$, as noted in equation (2) above, and Theorem 2 shows that such an equivalence class is exactly the set of sequences under consideration.

In contrast, the *avar* comparisons called for by Definition 6 are not limit-based and therefore do not always yield the same answer within an equivalence class with respect to $\mathbf{R}$. This can happen even when *avar* sequences have traditional Fisherian limits if the estimators are equally efficient, and does not violate Theorem 3 because Theorem 3 only promises *one pair* of *avar* sequences demonstrating the efficiency conclusion that prevails in the limit (i.e., in Definition 2). This ambiguity is noted by Bates and White [1993, p. 639], who define the concept of a *canonical avar sequence* as a solution. No matter how it is solved, the problem arises only because Definition 6 is not limit-based. Hence, Theorem 3 shows that the problem can also be solved with no change in our concept of asymptotic efficiency by relying on Definition 2 rather than Definition 6, provided we impose *avar*-regularity. Even though only $x_n = O_p(1)$ is really needed to get unambiguous efficiency conclusions from Definition 2, if we do not impose $x_n = alip(1)$ as well then the use of Definition 2 in lieu of Definition 6 comes at the price of a slightly weaker efficiency concept, in that one might conclude an estimator is efficient that would not be found efficient by Definition 6. Put another way, although the two definitions may differ when $x_n \neq alip(1)$, either definition can be used to unambiguously study efficiency when $x_n \neq alip(1)$, as in the RCANI class of Bates and White, but $alip(1)$ is relaxed at the expense of either a slightly weaker efficiency concept (if Definition 2 is used) or of having to find canonical *avar* sequences to resolve ambiguity (if Definition 6 is used). With $alip(1)$ in place these problems do not arise, since use of Definition 2 is then unambiguous and equivalent to Definition 6, but of course $alip(1)$ is itself restrictive. Although the fact that *avar* is a class is not acknowledged in White [1984 and 1994], and hence the potential ambiguity of the efficiency definition used (Definition 6) is also not mentioned, the efficiency proofs there rely on the use of canonical *avar* sequences.

In some situations Definitions 2 and 6 are both too strong to detect some potentially informative relative efficiencies, because they rely on liminf's of differences, which are not the same as differences of liminf's. In these situations the following may prove useful.

Definition 7. Let β_n^* and $\tilde{\beta}_n$ be consistent estimators of a nonstochastic q-dimensional vector β and suppose $\sqrt{n}(\beta_n^* - \beta)$ and $\sqrt{n}(\tilde{\beta} - \beta)$ are both *avar*-regular. Further assume $avar(\sqrt{n}(\beta_n^* - \beta))$ and $avar(\sqrt{n}(\tilde{\beta}_n - \beta))$ are both nonempty. We say β_n^* possesses *minimin asymptotic efficiency relative to* $\tilde{\beta}_n$ if

$$\liminf_{n\to\infty} \theta' \tilde{V}_n \theta \geq \liminf_{n\to\infty} \theta' V_n^* \theta \text{ for all } \tilde{V}_n \in avar(\sqrt{n}(\tilde{\beta}_n - \beta)),\ V_n^* \in avar(\sqrt{n}(\beta_n^* - \beta)), \text{ and } \theta \in \Re^q\ (\theta \neq 0).$$

We say β_n^* possesses *minimax asymptotic efficiency relative to* $\tilde{\beta}_n$ if

$$\limsup_{n\to\infty} \theta' \tilde{V}_n \theta \geq \limsup_{n\to\infty} \theta' V_n^* \theta \text{ for all } \tilde{V}_n \in avar(\sqrt{n}(\tilde{\beta}_n - \beta)),\ V_n^* \in avar(\sqrt{n}(\beta_n^* - \beta)), \text{ and } \theta \in \Re^q\ (\theta \neq 0).$$

Minimin efficiency focuses on best asymptotic performance, by which is meant smallest limiting *avar*'s, while minimax focuses on worst asymptotic performance, by which is meant largest limiting *avar*'s. Under *avar*-regularity, efficiency by either Definition 2 or 6 implies both minimin and minimax efficiency. Examples can be constructed using oscillations in which minimin and minimax efficiency of *avar*-regular sequences hold but Definitions 2 and 6 do not, even though the estimators being compared are conceptually no different from estimators that can be compared with Definitions 2 and 6. Because of Theorem 2, it is equivalent to only require the defining minimin and minimax inequalities to hold for one pair of *avars*.

Note finally that if one *avar* sequence for each of two *avar*-regular estimators has a positive definite limit then the Fisher definition of asymptotic efficiency is applicable to these limiting covariance matrices. In this case all efficiency definitions involve these same limits and are therefore equivalent. That is, if there exist positive definite matrices V^* and $\tilde{V}$ such that $\sqrt{n}(\beta_n^* - \beta) \xrightarrow{d} N(0, V^*)$ and $\sqrt{n}(\tilde{\beta}_n - \beta) \xrightarrow{d} N(0, \tilde{V})$ then β_n^* is asymptotically efficient relative to $\tilde{\beta}_n$ according to any of the definitions discussed herein if and only if $\tilde{V} - V^*$ is positive semidefinite.

Appendix

The proof of Theorem 1 uses the following fact.

Fact. *Let x be a random variable with $E(x) = \mu$ and $V(x) = \sigma^2 < \infty$. Then $P(\sigma \leq |x - \mu|) > 0$.*

Proof. If $P(\sigma > |x-\mu|) = 1$ then $\sigma > 0$, in which case $\sigma^2 = \int_{|x-\mu|<\sigma}(x-\mu)^2 dF(x) < \int_{|x-\mu|<\sigma}\sigma^2 dF(x) = \sigma^2$, a contradiction. □

Proof of Theorem 1. Fix $V_n \in avar(x_n)$, let $(\, e_{n1} \quad \dots \quad e_{nq}\,)$ be an orthonormal linearly independent set of eigenvectors for V_n, and $(\, \lambda_{n1} \quad \dots \quad \lambda_{nq}\,)$ be the corresponding real strictly positive eigenvalues (a full set of real strictly positive eigenvalues and orthonormal linearly independent eigenvectors exists since V_n is symmetric and positive definite).

(i) First assume $x_n = alip(1)$ and suppose the elements of V_n^{-1} are not bounded. Then there is an unbounded inverse eigenvalue sequence λ_{ni}^{-1}, in which case there is a subsequence $\{\lambda_{k_n i}\}_{n=1}^{\infty}$ such that $\lim_{n\to\infty}\lambda_{k_n i} = 0$. Since $(\lambda_{ni}^{1/2}, e_{ni})$ is an (eigenvalue,eigenvector) pair for $V_n^{1/2}$, this implies $\lim_{n\to\infty} V_{k_n}^{1/2} e_{k_n i} = \lim_{n\to\infty}\lambda_{k_n i}^{1/2} e_{k_n i}$ $= 0$ (using orthonormality of the eigenvectors to bound $e_{k_n i}$). That is, $e'_{k_n i}V_{k_n}^{1/2} = o(1)$. But then $e'_{k_n i}x_{k_n} =$ $e'_{k_n i}V_{k_n}^{1/2}V_{k_n}^{-1/2}x_{k_n} = o(1)O_p(1) = o_p(1)$, which contradicts $x_{k_n} = alip(1)$. For the converse, assume V_n^{-1} is bounded and suppose $x_n \neq alip(1)$. Since $(\, \lambda_{n1}^{-1} \quad \dots \quad \lambda_{nq}^{-1}\,)$ are the eigenvalues for V_n^{-1}, λ_{ni}^{-1} is bounded for $i = 1, \dots, q$. That is, there exists $M \in (0,\infty)$ such that $\lambda_{ni}^{-1} < M$, implying

$$\lambda_{ni} > M^{-1}, \text{ for } i = 1, \dots, q;\ \forall\, n.$$

Also, since $x_n \neq alip(1)$ there exists a sequence $\{c_n\}_{n=1}^{\infty}$ on the unit ball in $\Re^q$ with the following property:

$$\text{To each } n \text{ there corresponds } k_n \geq n \text{ such that } P\left(\frac{M^{-1/2}}{2} \leq |c'_{k_n}x_{k_n}|\right) \leq \frac{1}{n}.$$

That is,

$$\lim_{n\to\infty} P\left(\frac{M^{-1/2}}{2} \leq |c'_{k_n}x_{k_n}|\right) = 0.$$

Since $(\, e_{n1} \quad \dots \quad e_{nq}\,)$ is an orthonormal basis for $\Re^q$ we may write c_n as

$$c_n = \sum_{i=1}^{q}(c'_n e_{ni})e_{ni},$$

so $\sum_{i=1}^{q}(c'_n e_{ni})^2 = c'_n c_n = 1\ \forall\, n$, since c_n is on the unit ball. Thus, by definition of e_{ni} and λ_{ni}, and orthonormality of $(\, e_{n1} \quad \dots \quad e_{nq}\,)$,

$$\|c'_n V_n^{1/2}\| = (c'_n V_n c_n)^{1/2} = \left(\sum_{i=1}^{q}(c'_n e_{ni})^2\lambda_{ni}\right)^{1/2} > \left(M^{-1}\sum_{i=1}^{q}(c'_n e_{ni})^2\right)^{1/2} = M^{-1/2}\ \forall\, n.$$

Now write

$$c'_{k_n} x_{k_n} = \|c'_{k_n} V_{k_n}^{1/2}\| a'_{k_n} V_{k_n}^{-1/2} x_{k_n}, \quad \text{where } a'_{k_n} \equiv \frac{c'_{k_n} V_{k_n}^{1/2}}{\|c'_{k_n} V_{k_n}^{1/2}\|},$$

so that $\|a_{k_n}\| = 1 \,\forall\, n$. Then a_{k_n} has a convergent subsequence $a_{\ell_{k_n}} \to a_0$, where $\|a_0\| = 1$, so $a'_{\ell_{k_n}} V_{\ell_{k_n}}^{-1/2} x_{\ell_{k_n}} \xrightarrow{d} a'_0 z \sim (0, a'_0 a_0) = (0,1)$.[4] By the Fact, $P(1 \le |a'_0 z|) > 0$. Moreover, since all distribution functions are continuous off a countable set, there exists $\epsilon \in [\frac{1}{2}, 1]$ at which the distribution function of $|a'_0 z|$ is continuous. Then, by convergence in distribution there exists N such that

$$\begin{aligned} n \ge N \Rightarrow P\left(\frac{M^{-1/2}}{2} \le |c'_{\ell_{k_n}} x_{\ell_{k_n}}|\right) &= P\left(\frac{M^{-1/2}}{2} \le \|c'_{\ell_{k_n}} V_{\ell_{k_n}}^{1/2}\| |a'_{\ell_{k_n}} V_{\ell_{k_n}}^{-1/2} x_{\ell_{k_n}}|\right) \\ &\ge P\left(\frac{1}{2} \le |a'_{\ell_{k_n}} V_{\ell_{k_n}}^{-1/2} x_{\ell_{k_n}}|\right) \\ &\ge P(\epsilon \le |a'_{\ell_{k_n}} V_{\ell_{k_n}}^{-1/2} x_{\ell_{k_n}}|) \\ &> \frac{P(\epsilon \le |a'_0 z|)}{2} \\ &\ge \frac{P(1 \le |a'_0 z|)}{2} > 0, \end{aligned}$$

a contradiction.

(ii) First assume $x_n = O_p(1)$ and suppose the elements of V_n are not bounded. Then, as in (i), there is an eigenvalue subsequence $\{\lambda_{k_n i}\}_{n=1}^{\infty}$ such that $\lim_{n\to\infty} V_{k_n}^{-1/2} e_{k_n i} = \lim_{n\to\infty} \lambda_{k_n i}^{-1/2} e_{k_n i} = 0$. That is, $e'_{k_n i} V_{k_n}^{-1/2} = o(1)$. Since $e_{k_n i}$ is on the unit ball in $\Re^q$ there is a convergent subsequence $e_{\ell_{k_n} i} \to e_0$, where $\|e_0\| = 1$, so $e'_{\ell_{k_n} i} V_{\ell_{k_n}}^{-1/2} x_{\ell_{k_n}} \xrightarrow{d} e'_0 z \sim (0, e'_0 e_0) = (0,1)$, implying $e'_{\ell_{k_n} i} V_{\ell_{k_n}}^{-1/2} x_{\ell_{k_n}} \ne o_p(1)$ by the Fact. But since $x_n = O_p(1)$ we have $e'_{\ell_{k_n} i} V_{\ell_{k_n}}^{-1/2} x_{\ell_{k_n}} = o(1) O_p(1) = o_p(1)$, a contradiction. For the converse, just note that $V_n^{1/2}$ is bounded whenever V_n is bounded, so $x_n = V_n^{1/2} V_n^{-1/2} x_n = O(1) O_p(1) = O_p(1)$. □

Proof of Lemma 1. Since $c'\Sigma c$ is a continuous function of c, there exists $\bar{c} \in \partial B_1(0)$ (the boundary of the unit ball in $\Re^q$) such that

$$0 < \bar{c}'\Sigma\bar{c} \le c'\Sigma c \;\forall\; c \in \partial B_1(0).$$

Now suppose $x_n \ne$ *alip*(1). Then as in Theorem 1(i) there exists a sequence $\{c_n\}_{n=1}^{\infty}$ on the unit ball with the following property:

$$\text{To each } n \text{ there corresponds } k_n \ge n \text{ such that } P\left(\frac{\bar{c}'\Sigma\bar{c}}{2} \le |c'_{k_n} x_{k_n}|\right) \le \frac{1}{n}.$$

[4]In Definition 1, z is a standard normal random variable. Theorem 1 actually holds for any limiting random vector z with 0 mean and identity covariance, irrespective of normality. To demonstrate this, we avoid use of normality in the present proof.

That is, $\lim_{n\to\infty} P\left(\frac{\bar{c}'\Sigma\bar{c}}{2} \leq |c'_{k_n} x_{k_n}|\right) = 0$. Since c_{k_n} is on the unit ball there exists a convergent subsequence $c_{\ell_{k_n}} \to c_0 \in \partial B_1(0)$. Hence $c'_{\ell_{k_n}} x_{\ell_{k_n}} \xrightarrow{d} c'_0 z \sim (0, c'_0\Sigma c_0)$. By the Fact, $P(c'_0\Sigma c_0 \leq |c'_0 z|) > 0$. Moreover, since $\frac{\bar{c}'\Sigma\bar{c}}{2} < \bar{c}'\Sigma\bar{c} \leq c'_0\Sigma c_0$, and since all distribution functions are continuous off a countable set, there exists $\epsilon \in [\bar{c}'\Sigma\bar{c}/2, c'_0\Sigma c_0]$ at which the distribution function of $|c'_0 z|$ is continuous. Thus, by convergence in distribution there exists N such that

$$n \geq N \Rightarrow P\left(\frac{\bar{c}'\Sigma\bar{c}}{2} \leq |c'_{\ell_{k_n}} x_{\ell_{k_n}}|\right) \geq P(\epsilon \leq |c'_{\ell_{k_n}} x_{\ell_{k_n}}|) > \frac{P(\epsilon \leq |c'_0 z|)}{2} \geq \frac{P(c'_0\Sigma c_0 \leq |c'_0 z|)}{2} > 0,$$

a contradiction. □

Proof of Lemma 2. Suppose not. Then $(\tilde{V}_n^{1/2} V_n^{-1/2})'(\tilde{V}_n^{1/2} V_n^{-1/2})$ is an unbounded, symmetric, positive definite matrix. Hence it can be represented as $E_n\Lambda_n E'_n$, where E_n is orthogonal. As in Theorem 1, by unboundedness there exists an eigenvalue subsequence $\lambda_{k_n i}^{-1} \to 0$. Denote by f_i the unit vector in direction i, and let

$$c'_n \equiv \frac{f'_i E'_n V_n^{1/2} \tilde{V}_n^{-1/2}}{\|f'_i E'_n V_n^{1/2} \tilde{V}_n^{-1/2}\|},$$

so that $\|c_n\| = 1 \;\forall\; n$. Note that

$$\begin{aligned}
\|f'_i E'_n V_n^{1/2} \tilde{V}_n^{-1/2}\| &= (f'_i E'_n (V_n^{1/2} \tilde{V}_n^{-1} V_n^{1/2}) E_n f_i)^{1/2} \\
&= (f'_i E'_n E_n \Lambda_n^{-1} E'_n E_n f_i)^{1/2} \\
&= (f'_i \Lambda_n^{-1} f_i)^{1/2} = \lambda_{ni}^{-1/2}.
\end{aligned}$$

Hence

$$\begin{aligned}
c'_n \tilde{V}_n^{-1/2} x_n &= c'_n \tilde{V}_n^{-1/2} V_n^{1/2} V_n^{-1/2} x_n \\
&= \lambda_{ni}^{1/2} [f'_i E'_n V_n^{1/2} \tilde{V}_n^{-1/2} \tilde{V}_n^{-1/2} V_n^{1/2}] V_n^{-1/2} x_n \\
&= \lambda_{ni}^{1/2} [f'_i E'_n E_n \Lambda_n^{-1} E'_n] V_n^{-1/2} x_n \\
&= \lambda_{ni}^{1/2} [f'_i \Lambda_n^{-1} E'_n] V_n^{-1/2} x_n \\
&= \lambda_{ni}^{-1/2} e'_{ni} V_n^{-1/2} x_n,
\end{aligned}$$

where e_{ni} is column i of E_n. Since $e_{k_n i}$ is on the unit ball in $\Re^q$, there is a convergent subsequence $e_{\ell_{k_n} i} \to e_0 \in \partial B_1(0)$, so $e'_{\ell_{k_n} i} V_{\ell_{k_n}}^{-1/2} x_{\ell_{k_n}} \xrightarrow{d} e'_0 z \sim (0, e'_0 e_0) = (0, 1)$.[5] Hence $\lambda_{\ell_{k_n} i}^{-1/2} e'_{\ell_{k_n} i} V_{\ell_{k_n}}^{-1/2} x_{\ell_{k_n}} \xrightarrow{p} 0$. But, by Lemma 1, $\tilde{V}_{\ell_{k_n}}^{-1/2} x_{\ell_{k_n}} = \mathit{alip}(1)$, a contradiction. □

[5] As in Theorem 1, normality of z is not needed here.

Proof of Theorem 2. Suppose first that $\tilde{V}_n \in E_{\mathcal{R}}(V_n)$. Then $\lim_{n\to\infty} V_n^{1/2}\tilde{V}_n^{-1}V_n^{1/2} = I_q$, so $V_n^{1/2}\tilde{V}_n^{-1/2} = O(1)$. Denote the characteristic function of a random vector y by f_y. By the continuity theorem [Lukacs 1970, p. 48], $f_{V_n^{-1/2}x_n}(\theta) \to f_z(\theta)$ pointwise. Hence, by White [1984, p. 66] Lemma 4.23,

$$f_{\tilde{V}_n^{-1/2}x_n}(\theta) - f_z(V_n^{1/2}\tilde{V}_n^{-1/2}\theta) = f_{(\tilde{V}_n^{-1/2}V_n^{1/2})V_n^{-1/2}x_n}(\theta) - f_z(V_n^{1/2}\tilde{V}_n^{-1/2}\theta) \to 0 \text{ pointwise in } \theta.$$

Recalling that $f_z(\theta) = exp\left(-\frac{\theta'\theta}{2}\right)$,

$$f_z(V_n^{1/2}\tilde{V}_n^{-1/2}\theta) = exp\left(-\frac{\theta'\tilde{V}_n^{-1/2}V_n\tilde{V}_n^{-1/2}\theta}{2}\right) \to exp\left(-\frac{\theta'\theta}{2}\right),$$

so $f_{\tilde{V}_n^{-1/2}x_n}(\theta) \to f_z(\theta)$. Applying the continuity theorem again shows that $\tilde{V}_n \in avar(x_n)$.

Now consider the converse by supposing $\tilde{V}_n \in avar(x_n)$. By Lemma 2, $M \geq 1$ can denote a common bound for all elements of $\tilde{V}_n^{1/2}V_n^{-1/2}$. Fix $\epsilon > 0$ and set $r = q^2M$. Recalling that $f_{V_n^{-1/2}x_n}(\theta) = f_{x_n}(V_n^{-1/2}\theta)$, by the second continuity theorem [Lukacs 1970, p. 53] there exists N_ϵ such that

$$n \geq N_\epsilon \Rightarrow \begin{cases} \left|f_{x_n}(V_n^{-\frac{1}{2}}\theta) - exp\left(-\frac{\theta'\theta}{2}\right)\right| & < \frac{\epsilon}{4}exp\left(-\frac{r^2}{2}\right) \ \forall\, \theta \in \overline{B_r(0)} \\ \left|f_{x_n}(\tilde{V}_n^{-\frac{1}{2}}t) - exp\left(-\frac{t't}{2}\right)\right| & < \frac{\epsilon}{4}exp\left(-\frac{r^2}{2}\right) \ \forall\, t \in \overline{B_r(0)}, \end{cases}$$

where $\overline{B_r(0)}$ is the closed ball of radius r about 0 in $\Re^q$. Since $\tilde{V}_n^{\frac{1}{2}}V_n^{-\frac{1}{2}}\theta \in \overline{B_r(0)}$ for every n and $\theta \in \partial B_1(0)$, setting $t = \tilde{V}_n^{\frac{1}{2}}V_n^{-\frac{1}{2}}\theta$ yields

$$n \geq N_\epsilon \Rightarrow \begin{cases} \left|f_{x_n}(V_n^{-\frac{1}{2}}\theta) - exp\left(-\frac{\theta'\theta}{2}\right)\right| & < \frac{\epsilon}{4}exp\left(-\frac{r^2}{2}\right) \\ \left|f_{x_n}(V_n^{-\frac{1}{2}}\theta) - exp\left(-\frac{\theta'V_n^{-\frac{1}{2}}\tilde{V}_nV_n^{-\frac{1}{2}}\theta}{2}\right)\right| & < \frac{\epsilon}{4}exp\left(-\frac{r^2}{2}\right) \end{cases}$$

for every $\theta \in \partial B_1(0)$. Hence

$$n \geq N_\epsilon \Rightarrow \left|exp\left(-\frac{\theta'V_n^{-\frac{1}{2}}\tilde{V}_nV_n^{-\frac{1}{2}}\theta}{2}\right) - exp\left(-\frac{\theta'\theta}{2}\right)\right| < \frac{\epsilon}{2}exp\left(-\frac{r^2}{2}\right) \ \forall\, \theta \in \partial B_1(0).$$

By the Mean Value Theorem,

$$\left|exp\left(-\frac{\theta'V_n^{-\frac{1}{2}}\tilde{V}_nV_n^{-\frac{1}{2}}\theta}{2}\right) - exp\left(-\frac{\theta'\theta}{2}\right)\right| = \left|-\frac{1}{2}exp\left(-\frac{c}{2}\right)\right| \left|\theta'V_n^{-\frac{1}{2}}\tilde{V}_nV_n^{-\frac{1}{2}}\theta - \theta'\theta\right|$$

for some c between $\theta'V_n^{-\frac{1}{2}}\tilde{V}_nV_n^{-\frac{1}{2}}\theta$ and $\theta'\theta$. Thus

$$n \geq N_\epsilon \Rightarrow \left|\theta'V_n^{-\frac{1}{2}}\tilde{V}_nV_n^{-\frac{1}{2}}\theta - \theta'\theta\right| < \epsilon exp\left(\frac{c - r^2}{2}\right)$$

for every $\theta \in \partial B_1(0)$ and the corresponding c (which also depends on n). Since $\theta' V_n^{-\frac{1}{2}} \tilde{V}_n V_n^{-\frac{1}{2}} \theta \le r^2$ for every n and $\theta \in \partial B_1(0)$, and $\theta'\theta = 1 \le r^2$ for such θ, we have $c \le r^2$. Thus $exp\left(\frac{c-r^2}{2}\right) \le 1$ for every n and $\theta \in \partial B_1(0)$. That is,

$$n \ge N_\epsilon \Rightarrow \left|\theta' V_n^{-\frac{1}{2}} \tilde{V}_n V_n^{-\frac{1}{2}} \theta - \theta'\theta\right| < \epsilon \;\forall\; \theta \in \partial B_1(0).$$

Since θ can be any vector on the unit ball, this implies $\lim_{n\to\infty} V_n^{-1/2}\tilde{V}_n V_n^{-1/2} = I_q$, or $\tilde{V}_n \in E_{\mathcal{R}}(V_n)$.

Next consider (i). Since V_n^{-1} is bounded by Theorem 1(i) and $V_n^{-1/2}\tilde{V}_n V_n^{-1/2} - I_q = V_n^{-1/2}[\tilde{V}_n - V_n]V_n^{-1/2}$, $\lim_{n\to\infty}[V_n^{-1/2}\tilde{V}_n V_n^{-1/2} - I_q] = 0$ when $\lim_{n\to\infty}(\tilde{V}_n - V_n) = 0$.

Finally, consider (ii). Since V_n is bounded by Theorem 1(ii) and $\tilde{V}_n - V_n = V_n^{1/2}[V_n^{-1/2}\tilde{V}_n V_n^{-1/2} - I_q]V_n^{1/2}$, $\lim_{n\to\infty}(\tilde{V}_n - V_n) = 0$ when $\lim_{n\to\infty}[V_n^{-1/2}\tilde{V}_n V_n^{-1/2} - I_q] = 0$. □

Proof of Theorem 3. Select $V_n^* \in avar(\sqrt{n}(\beta_n^* - \beta))$ and $\tilde{V}_n \in avar(\sqrt{n}(\tilde{\beta}_n - \beta))$ such that $f(\theta) \equiv \liminf_{n\to\infty} f_n(\theta) = \lim_{n\to\infty} g_n(\theta) \ge 0 \;\forall\; \theta \in \Re^q$ $(\theta \ne 0)$, where $f_n(\theta) \equiv \theta'(\tilde{V}_n - V_n^*)\theta$ and $g_n(\theta) \equiv \inf\{f_m(\theta)\colon m \ge n\}$. We first establish uniform equicontinuity of the sequence $f_n(\theta)$ on the boundary of the unit ball in $\Re^q$. By Theorem 1(ii), $\tilde{V}_n$ is bounded. It is straightforward to use this along with $f(\theta) \ge 0 \;\forall\; \theta$ to conclude that V_n^* is bounded as well, so denote by M a common bound for all elements of $\tilde{V}_n$ and V_n^*. For any $\theta, \theta_0 \in \partial B_1(0)$ we have

$$|f_n(\theta) - f_n(\theta_0)| \le \|\theta - \theta_0\| \left[\|(\tilde{V}_n - V_n^*)\theta\| + \|\theta_0'(\tilde{V}_n - V_n^*)\|\right] \le \|\theta - \theta_0\| 4Mq.$$

Hence for every $\epsilon > 0$ there exists $\delta = \frac{\epsilon}{4Mq}$ such that $\|\theta - \theta_0\| < \delta \Rightarrow |f_n(\theta) - f_n(\theta_0)| < \epsilon \;\forall\; n$, or $f_n(\theta)$ is uniformly equicontinuous on $\partial B_1(0)$. Now, by definition of $g_n(\theta)$ there exists a subsequence $f_{i_n}(\theta)$ such that $g_n(\theta) + \frac{1}{n} > f_{i_n}(\theta) \ge g_n(\theta)$. Thus $f_{i_n}(\theta)$ converges pointwise to $f(\theta)$, so by equicontinuity this convergence is actually uniform on $\partial B_1(0)$. This implies $\lim_{n\to\infty} f_{i_n}(\theta) - \frac{1}{n} = f(\theta)$ uniformly on $\partial B_1(0)$. Hence, since $f_n(\theta) \ge g_n(\theta)$ by definition of $g_n(\theta)$, for every natural number k there exists a natural number N_k such that

$$n \ge N_k \Rightarrow f_n(\theta) \ge g_n(\theta) > f_{i_n}(\theta) - \frac{1}{n} > f(\theta) - \frac{1}{k} \;\forall\; \theta \in \partial B_1(0).$$

We may summarize this by saying that the convergence to the limit inferior is uniform on $\partial B_1(0)$ in the sense that for every k there exists N_k such that $n \ge N_k \Rightarrow f_n(\theta) + \frac{1}{k} > 0 \;\forall\; \theta \in \partial B_1(0)$, and without loss of generality we may assume $1 < N_1 < N_2 < \cdots$. Let k_n be the largest natural number k satisfying $N_k \le n$ for $n \ge N_1$

and note that $n \geq N_{k_n}$ $\forall$ $n \geq N_1$ and $\lim_{n\to\infty} \frac{1}{k_n} = 0$. Thus $\tilde{V}_n + \frac{1}{k_n} I_q \in avar(\sqrt{n}(\tilde{\beta}_n - \beta))$ by Theorem 2(i), so that β_n^* is asymptotically efficient relative to $\tilde{\beta}_n$ according to Definition 6 if $\left(\tilde{V}_n + \frac{1}{k_n} I_q\right) - V_n^*$ is positive semidefinite for n large. For any $n \geq N_1$ and $\theta \neq 0$ we have

$$\begin{aligned}\left(\frac{\theta'}{\|\theta\|}\right)\left[\left(\tilde{V}_n + \frac{1}{k_n} I_q\right) - V_n^*\right]\left(\frac{\theta}{\|\theta\|}\right) &= \left(\frac{\theta'}{\|\theta\|}\right)\left[\tilde{V}_n - V_n^*\right]\left(\frac{\theta}{\|\theta\|}\right) + \frac{\theta'\theta}{k_n\|\theta\|^2} \\ &= f_n\left(\frac{\theta}{\|\theta\|}\right) + \frac{1}{k_n} > 0,\end{aligned}$$

so $\theta'\left[\left(\tilde{V}_n + \frac{1}{k_n} I_q\right) - V_n^*\right]\theta > 0$ $\forall$ $n \geq N_1$ and $\forall$ $\theta \neq 0$. That is, $\left(\tilde{V}_n + \frac{1}{k_n} I_q\right) - V_n^*$ is positive definite for $n \geq N_1$. □

References

Bates, C. E. and H. White, *Determination of Estimators with Minimum Asymptotic Covariance Matrices*, Econometric Theory **9** (1993), 633-648.

Fisher, R. A., *Theory of Statistical Estimation*, Proceedings of the Cambridge Philosophical Society **22** (1925), 700-725.

Lukacs, E., *Characteristic Functions*, Hafner, New York, 1970.

Mann, H. B. and A. Wald, *On Stochastic Limit and Order Relationships*, Annals of Mathematical Statistics **14** (1943), 217-226.

White, H., *Instrumental Variables Regression with Independent Observations*, Econometrica **50** (1982), 483-499.

———, *Asymptotic Theory for Econometricians*, Academic Press (Harcourt Brace Jovanovich), Orlando, 1984.

———, *Estimation, Inference and Specification Analysis*, Cambridge University Press (Econometric Society Monograph 22), Cambridge, 1994.