An Alternative Sense of Asymptotic Efficiency

Ulrich K. Müller
Princeton University

June 20, 2008
Introduction

Asymptotically efficient tests for nonstandard problems


**Standard Asymptotic Efficiency I**

1. Consider first canonical parametric version of model, typically with Gaussian i.i.d. disturbances

   Unit Root Test: \( y_t = \rho y_{t-1} + \varepsilon_t, \ y_0 = 0, \ \varepsilon_t \sim \text{i.i.d.} \mathcal{N}(0, 1) \)

   \( H_0 : \rho = 1 \quad \text{against} \quad H_1 : \rho = \rho_1 < 1 \)

2. Derive small sample efficient test in that canonical model

   \( \text{LR}_T = \exp[-\frac{1}{2}(1 - \rho_1)(y_T^2 - \sum(\Delta y_t)^2) - \frac{1}{2}(1 - \rho_1)^2 \sum y_{t-1}^2] \)

3. Take limits of small sample efficient test against local alternatives

   \( \rho = \rho_T = 1 - \theta / T \)

   \( H_0 : \theta = 0 \quad \text{against} \quad H_1 : \theta = \theta_1 > 0 \)

   \( \text{LR}_T \sim \exp[-\frac{1}{2}\theta_1(J_{\theta}(1)^2 - 1) - \frac{1}{2}\theta_1^2 \int J_{\theta}(s)^2 ds] \)

   \( J_{\theta}(s) = \int_0^s e^{-\theta(s-r)}dW(r) \)
Standard Asymptotic Efficiency II

4. Construct robustified version of test statistic to obtain correct coverage also for non-canonical versions of the model

\[ \hat{\text{LR}}_T = \exp\left[-\frac{1}{2}\theta_1(\hat{J}_T(1)^2 - 1) - \frac{1}{2}\theta_1^2 \int \hat{J}_T(s)^2 ds\right] \]
\[ \hat{J}_T(\cdot) = T^{-1/2}\hat{\omega}_T^{-1}y[.T] \rightsquigarrow J_\theta(\cdot) \]

End-product is test that

i. is asymptotically efficient in canonical model

ii. has same asymptotic rejection probabilities for all models where robustified version yields same weak limits

whenever \( \hat{\omega}_T \) and \( \varepsilon_t \) are such that \( T^{-1/2}\hat{\omega}_T^{-1}y[.T] \rightsquigarrow J_\theta(\cdot) \)

But: There could exist a test that is also efficient in the canonical model, with higher asymptotic power against some non-canonical model
Semiparametric Efficiency

- Model has infinitely dimensional nuisance parameter, such as distribution of disturbance

- Well developed only for standard problem with locally asymptotic normal likelihood ratios

- Jansson (2007)
  - Unit root test for model $y_t = \rho y_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim iidF(0, 1)$, and $F$ is nuisance parameter
  - Semiparametric power envelope with $F$ known
  - Adaption only possible for symmetric $F$. Otherwise, still asymptotic power gains over $\widehat{LR}_T$

  $\Rightarrow \widehat{LR}_T$ test asymptotically inadmissible in this set-up
New Sense of Asymptotic Efficiency

1. For test of $H_0 : \theta = \theta_0$ against $H_1 : \theta = \theta_1$ on double-array data $Y_T \in \mathbb{R}^{nT}$, consider typical weak convergence $X_T = h(Y_T) \sim X \sim P(\theta)$

   $$X_T = \hat{J}_T(\cdot) = T^{-1/2} \hat{\omega}_{T, \cdot}^{-1} y_{T, \cdot} \sim X = J_\theta(\cdot)$$

2. Derive best test in limiting problem with $X$ observed

   $$L(X) = \exp\left[-\frac{1}{2} \theta_1 (X(1)^2 - 1) - \frac{1}{2} \theta_1^2 \int X(s)^2 ds\right]$$ is RN-derivative by Neyman-Pearson: reject for large values of $L(X)$

3. Robustness requirement: Tests in original problem must have correct asymptotic rejection probability whenever $X_T \sim X \sim P(\theta_0)$

   Robust unit root tests do not overreject whenever $\hat{J}_T(\cdot) \sim W(\cdot)$

   further step in progression to weaker assumptions about disturbances

Stock (1994), White (2001), Breitung (2002), Davidson (2007) define $I(1)$ property in terms of $T^{-1/2} y_{T, \cdot} \sim \omega W(\cdot)$
Main Result

4. Asymptotically best robust test is optimal test in limiting problem, evaluated at sample analogues

   Rejecting for large values of
   
   \[ L(X_T) = \widehat{LR}_T = \exp\left[ -\frac{1}{2} \theta_1 \hat{J}_T((1)^2 - 1) - \frac{1}{2} \theta_1^2 \int \hat{J}_T(s)^2 ds \right] \]

   identical to ERS

\[ \Rightarrow \] For any test that has higher asymptotic power than best robust test, there exists model with \( X_T \sim P(\theta_0) \) where test overrejects asymptotically

   there exists \( T^{-1/2} \hat{\omega}_T^{-1} y_T, [\cdot, T] \sim W(\cdot) \) for which Jansson’s (2007) test

   has asymptotic rejection probability larger than nominal level
Generalizations

• Generic result whenever null and alternative hypotheses induce weak convergences. Very weak regularity conditions beyond a.e. continuity of best limiting test.

• General set-up allows for
  – consistently estimable parameters
  – additional restrictions on tests: unbiasedness, (conditional) similarity, invariance
Discussion

• Complete class of tests in limiting problem that are continuous a.e., evaluated at sample analogues, form an "asymptotically essentially complete class of robust tests".

• Appeal of efficiency property of depends on appropriateness of robustness constraint
  – Weak convergences as regularity condition. Much more natural in time series context.
  – Conservative. How sure are we about conventional primitive conditions, such as mixing?
  – Quality of small sample approximation.

• Alternative approach to construction of reasonable tests for complicated models.
Applications

- Broader sense of asymptotic efficiency of tests mentioned in introductory slide

- Precise sense of optimality of Sowell’s (1996) GMM parameter stability tests

- Application to Müller and Watson (2007, 2008) and Ibragimov and Müller (2007) that take weak convergence as a starting point
Heuristic Proof I

- Reconsider unit root testing problem, where

\[ X_T = \hat{J}_T(\cdot) = T^{-1/2}\hat{\omega}_T^{-1}y_{T,\cdot}T \sim X = J_\theta(\cdot) \]

and in limiting problem with \( X \) observed, Neyman-Pearson test of \( H_0 : \theta = 0 \) against \( H_1 : \theta = \theta_1 > 0 \) rejects for large values of

\[ L(X) = \exp\left[-\frac{1}{2}\theta_1(X(1)^2 - 1) - \frac{1}{2}\theta_1^2 \int X(s)^2 ds \right] \]

- Idea of proof: For any \( X_T \sim Q_T \sim J_{\theta_1}(\cdot) \), one can construct \( X_T \sim P_T \sim J_0(\cdot) \) such that the best small sample test of

\[ H_{T,0} : X_T \sim P_T \quad \text{against} \quad H_{T,1} : X_T \sim Q_T \]

rejects for \( L(X_T) \). Robust tests must control asymptotic size under \( H_{0,T} \), and no test can have a better asymptotic level and power trade-off than a sequence of small sample optimal tests.
Heuristic Proof II

- For simplicity, pretend that $\frac{1}{L}$ is bounded and continuous.

- Let $Q_T$ be any given probability measure of $X_T \sim J_{\theta_1}(\cdot) \sim Q$. Construct measure $P_T$ as

$$
\int_A dP_T = \kappa_T^{-1} \int_A \frac{1}{L} dQ_T \quad \text{for all measurable } A \subset \mathbb{R}^T
$$

$$
\kappa_T = \int \frac{1}{L} dQ_T
$$

where $L : D_{[0,1]} \mapsto \mathbb{R}$ is Radon-Nikodym derivative of $Q$ with respect to $P$ ($L = dQ/dP$).

- Note that $\kappa_T = \int \frac{1}{L} dQ_T \to \int \frac{1}{L} dQ = 1$. Furthermore, for any bounded and continuous function $\vartheta : D_{[0,1]} \mapsto \mathbb{R}$, the distribution $P_T$ satisfies

$$
\int \vartheta dP_T = \kappa_T^{-1} \int \frac{\vartheta}{L} dQ_T \to \int \frac{\vartheta}{L} dQ = \int \vartheta dP
$$

so that by the CMT, $P_T \sim J_0(\cdot) \sim P$. 
Weak IV Regression I

- Structural and reduced form equation

\[
y_{1,t} = y_{2,t}\beta + u_{t,1} \\
y_{2,t} = z't\pi + v_{t,2}
\]

- Reduced form

\[
y_{1,t} = z't\pi\beta + v_{t,1}
\]

- AMS consider small sample efficient tests of

\[ H_0 : \beta = \beta_0 \]

for nonstochastic \(z_t\) and \(v_t = (v_{1,t}, v_{2,t})' \sim i.i.d.\mathcal{N}(0, \Omega)\) with \(\Omega\) known.

By sufficiency, tests may be restricted to functions of

\[
\sum_{t=1}^{T} \begin{pmatrix} z_ty_{1,t} \\ z_ty_{2,t} \end{pmatrix} \sim \mathcal{N} \left( \begin{pmatrix} S_z\pi\beta \\ S_z\pi \end{pmatrix}, \Omega \otimes S_z \right), \quad S_z = \sum_{t=1}^{T} z_tz_t'
\]
Weak IV Regression II

• AMS derive WAP maximizing similar tests that are invariant to the group of transformations

\[ \{z_t\}_{t=1}^T \rightarrow \{Oz_t\}_{t=1}^T \text{ for any orthogonal matrix } O. \]

• AMS then consider Staiger and Stock (1997) weak instrument asymptotics, where \( \pi = T^{-1/2}C \) for some fixed \( C \)

• AMS derive test that

1. maximizes WAP among all asymptotically invariant and asymptotically similar tests when \( v_t \sim i.i.d. \mathcal{N}(0, \Omega) \) independent of \( \{z_t\}_{t=1}^T \)

2. yields correct asymptotic null rejection probability under much broader conditions, including heteroskedastic and autocorrelated \( v_t \)
Weak IV Regression III

- Typical weak convergence under weak IV asymptotics

\[
\hat{D}_z = T^{-1} \sum_{t=1}^{T} z_t z_t' \xrightarrow{p} D_z \\
\hat{\Sigma} \xrightarrow{p} \Sigma \\
X_T = T^{-1/2} \sum_{t=1}^{T} \begin{pmatrix} z_t y_{1,t} \\ z_t y_{2,t} \end{pmatrix} \sim X \sim \mathcal{N} \left( \begin{pmatrix} D_z C \beta \\ D_z C \end{pmatrix}, \Sigma \right)
\]

- In just-identified case ($D_z$ a scalar), rely on AMS small sample result for uniformly most powerful unbiased test in limiting problem (actually Moreira (2001))

\[
\Rightarrow \text{reject for large values of Anderson-Rubin statistic } (b_0' X_T)^2 / b_0' \hat{\Sigma} b_0,
\]

where $b_0 = (1, -\beta_0)'$

\[
\Rightarrow \text{by results here, uniformly most asymptotically powerful unbiased robust test}
\]

- In overidentified case, no general known solution for best limiting test (AMS results apply only when $\Sigma = \Omega \otimes D_z$)
Weak IV Regression IV

- Robustness constraint is large, since weak convergence

\[ T^{-1/2} \sum_{t=1}^{T} \begin{pmatrix} z_{ty1,t} \\ z_{ty2,t} \end{pmatrix} \rightsquigarrow \mathcal{N} \left( \begin{pmatrix} D_z C \beta \\ D_z C \end{pmatrix}, \Sigma \right) \]

can hold for many 'weird' data generating processes.

- Assume in addition the weak convergence

\[ T^{-1/2} \sum_{t=1}^{\lfloor T \rfloor} \begin{pmatrix} z_{ty1,t} \\ z_{ty2,t} \end{pmatrix} \rightsquigarrow G(\cdot) \]

\[ G(s) = s \begin{pmatrix} D_z C \beta \\ D_z C \end{pmatrix} + \Sigma^{1/2} W(s) \]

Since \( G(1) \) sufficient for \((C, \beta)\), does not change best test in limiting problem.
Conclusion

• Alternative sense of asymptotic efficiency
  1. for numerous efficient tests in canonical models
  2. for nonstandard methods that start with weak convergence assumption

• Asymptotic efficient tests in this sense simply best tests of limiting problem, evaluated at sample analogues

• Stringent robustness constraint
  ⇒ Most natural in time series context