Weak Identification in a White Noise Test

Jay Dennis

UNC Chapel Hill, Economics

Current Version

January 29, 2019

Overview

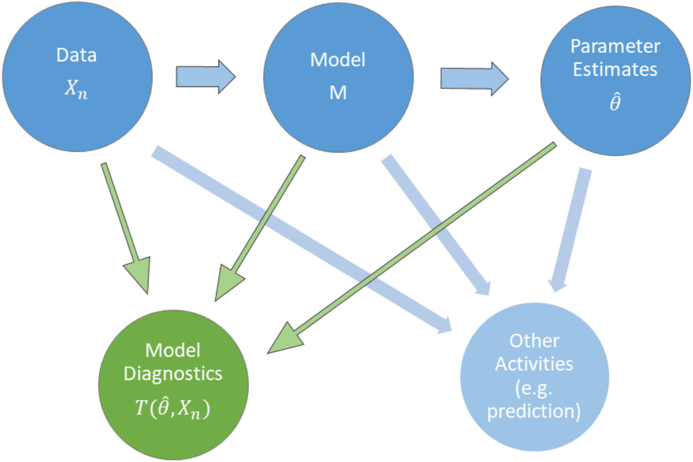
- Parameter identification failure $\Rightarrow$ nonstandard estimator distributions
  - affects testing on the parameters
- Point 1: What happens to *White Noise tests* based on estimated models when parameters are near identification failure?
  - the nonstandardness propagates to our test statistic for serial correlation.
  - Ignoring identification failure can lead to distorted inference:

	Traditional Test	This Test
$(\alpha = 0.05)$	0.096	0.050

Rejection Frequencies, ARMA-GARCH, $T = 100$, Weak Id, Max Correlation Test

- Point 2: What do we do when standard methods fail?

Introduction



Introduction

- This presentation lives in the world of Model Diagnostics
- Focus on White Noise Tests:
  - Does the model capture the dependence in the data?
  - (Data, Model, Parameter Estimates) $\Rightarrow$ Residuals $\hat{\varepsilon}$
  - Examine Correlations: $\hat{\rho}(h) = \frac{1}{n}\sum_{t=h}^{n} \hat{\varepsilon}_{t-h}\hat{\varepsilon}_t / \frac{1}{n}\sum_{t=0}^{n} \hat{\varepsilon}_t^2$
  - Aggregate across many lags:
  - Q-Test: $T(\hat{\theta}, X_n) = \frac{1}{h}\sum_{h=1}^{H} \sqrt{n}\hat{\rho}(h)$
  - Max Corr Test: $T(\hat{\theta}, X_n) = \max_{1\leq h\leq H} |\sqrt{n}\hat{\rho}(h)|$

Introduction

- Traditional methods require 'standard features' in the Data and Model
- Identification Failure violates these standard assumptions
  - $\Rightarrow$ Features are no longer standard
  - $\Rightarrow$ Traditional methods do not work as expected
- Today:
  - How do we determine if there is a problem?
  - What can we do about it?

Related Literature

Outline

Identification Failure?

Consider the ARMA(1,1) model:

$$y_t = \gamma y_{t-1} + \varepsilon_t - \pi\varepsilon_{t-1} \qquad \varepsilon_t \sim iid(0,1)$$

Imagine y_t carries no dependence.

- Typically we assume this means $\gamma = 0$, $\pi = 0$.
- But in general $\gamma = \pi \neq 0 \Rightarrow$ no dependence.
- The general case is called the 'Common Roots Problem.'
- We typically assume this problem away because it prevents identification of γ and π.
- However, there are common situations in which this may be an issue.

Identification Failure?

- Poterba and Summers (1988); Taylor (2005)
  - exploit low levels of dependence in returns?
  - trading rule based on ARMA(1,1) model of returns
- Set $\gamma = \beta + \pi$ to focus on identification of π.

ARMA

- π is not identified when $\beta = 0 \Leftarrow y_t = (\beta + \pi)y_{t-1} + \varepsilon_t - \pi\varepsilon_{t-1}$

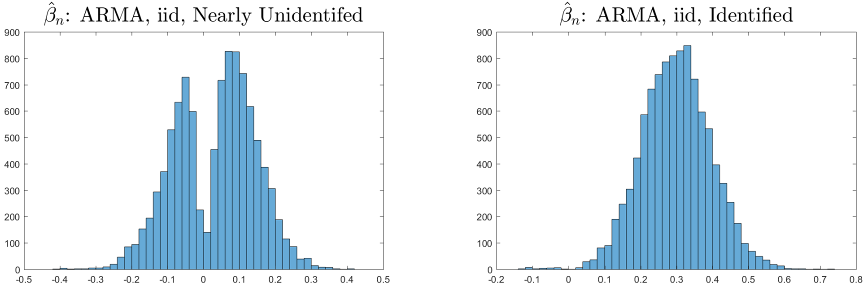
VWRETD, monthly data, ARMA, Unknown Id

	1	2	3	4	5	6
Start Date	31-Jul-1962	31-Jul-1962	31-Oct-1978	31-Jul-1962	31-Jan-1995	31-Oct-1978
End Date	30-Dec-1994	29-Sep-1978	30-Dec-1994	30-Dec-2005	30-Dec-2005	30-Dec-2005
n	389	194	194	521	131	326
$\hat{\beta}_n$	0.075	0.096	0.034	0.070	0.088	0.036
$\hat{se}$	(0.914)	(0.735)	(0.244)	(0.637)	(0.817)	(0.237)

- We don't know if $\beta = 0$ or not.
- β appears to be 'close' to 0
  - **even being close can be problematic.**

Identification Failure for a Parameter

Identification failure leads to non-standard distributions:



- $\Rightarrow$ Non-standard Inference!

ARMA(1,1) $\hat{\beta}_n$, non-id

Identification Failure in our Test

- Imagine we want to use the ARMA model to inform a trading rule (Taylor, 2005).

▸ Other Uses

- We want to know if the ARMA model adequately describes our data.

- Test for serial correlation in the **residuals**:
  - If we detect serial correlation in the residuals, then our model must not be characterizing some dependence.

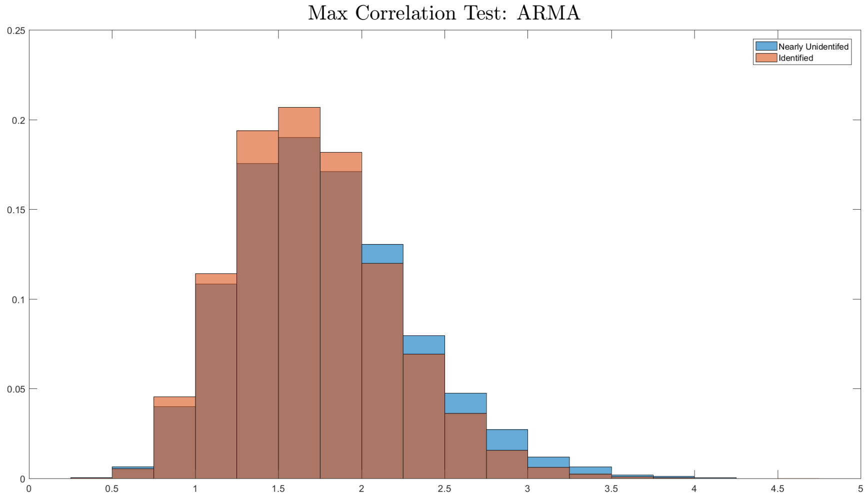
Identification Failure in our Test

- Note: *We are not testing the parameter values!*
- But our test does use the residuals from the estimated model
- We must account for the influence of model estimation on our test statistic:

$$\hat{\mathcal{T}}_n \equiv \hat{\mathcal{T}}_n(\hat{\theta}_n) = \hat{\mathcal{T}}_n(\theta^*) + \underbrace{\sqrt{n}(\hat{\theta}_n - \theta^*)}_{\text{Non-standard}} \hat{\mathcal{D}}_n + o_p(1)$$

- 'Non-standardness' can be felt beyond inference on the parameter values!

Identification Failure in our Test



We show how to correctly conduct inference.

Outline

White Noise Test

- We want to test if $\{\varepsilon_t\}$ is serially correlated:

$$H_0 : \underbrace{\rho(h) = 0 \ \forall h \in \mathbb{N}}_{\text{No Serial Correlation}} \qquad \text{vs.} \qquad H_A : \rho(h) \neq 0 \ \text{ for some } h \in \mathbb{N}$$

- (Data, Model, Parameter Estimates) $\Rightarrow$ Residuals $\hat{\varepsilon}$

ARMA

- We test residuals using the max correlation statistic

$$\hat{\mathcal{T}}_n = \max_{1 \leq h \leq \mathcal{L}_n} |\sqrt{n}\hat{\rho}_n(h)|$$

where $\hat{\rho}(h) = \frac{1}{n}\sum_{t=h}^{n} \hat{\varepsilon}_{t-h}\hat{\varepsilon}_t / \frac{1}{n}\sum_{t=0}^{n} \hat{\varepsilon}_t^2$

Max Tests?
Comments

Critical Values

Recall 2 distributions ▸ Theory

- We don't know which is correct - must bootstrap each individually
  ▸ Bootstrap Procedure
- 2 different critical values - How to combine them?
- ICS - Use data to determine if parameter is identified ▸ ICS Details
  - if so, then use identified cv
  - if not, take the larger of the 2 cvs

Test Properties

- Correct Asymptotic Size (largest rejection probability under the null is α)
  ▸ Theory
- Consistent (Under Alternative, probability of rejecting null $\to 1$) ▸ Theory

This is good, but how does it do in practice?

Outline

Monte Carlo Simulations

Today, I discuss tests based on

- *ICS*: Robust Test with data driven identification category selection
- *S*: Tests that ignore identification failure

for the following tests

- (MC): Max Correlation Test
- (LBQ): Hong's (1996) Standardized Ljung-Box Q test,
- (supLM): Nankervis and Savin's (2010) representation of Andrews and Ploberger's (1996) sup-LM test

Summary of Simulation Results

- Benchmark: Infeasible Tests - We know the true value of the unidentified parameters
- Real World: Feasible Tests - Must evaluate over a grid of nuisance parameters
- Good size and power for infeasible tests
- When truth is near identification failure:
  - Tests ignoring the identification issue are over-sized
  - Robust tests control size well
  - Current issue: Nuisance parameters $\Rightarrow$ low empirical power in feasible tests...

Monte Carlo Simulations - Setup

$$\text{ARMA(1,1)}: \quad y_t = (\beta_n + .5)y_{t-1} + \varepsilon_t - .5\varepsilon_{t-1}$$

with $\beta_n \in \{0, .3/\sqrt{n}, .3\}$. Let $\nu_t \sim$ iid $N(0,1)$.

H_0 (No Serial Correlation)	
iid:	$\varepsilon_t = \nu_t$
GARCH(1,1):	$\varepsilon_t = \sigma_t \nu_t$,
	$\sigma_t^2 = 1 + .3\varepsilon_{t-1}^2 + .6\sigma_{t-1}^2$

H_A (Serial Correlation Present)	
AR(2):	$\varepsilon_t = .5\varepsilon_{t-2} + \nu_t$
MA(10):	$\varepsilon_t = .5\nu_{t-10} + \nu_t$

Simulation Details

ARMA Simulations

Rejection Frequencies: Robust Tests, Identified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.300$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid	GARCH(1,1)	AR(2)	MA(10)	AR(2)	MA(10)
MC ICS	0.0460	0.0300	0.5580	0.7600	0.5120	0.7180
LBQ ICS	0.0400	0.0420	0.5180	0.5420	0.4740	0.4780
sup LM ICS	0.0240	0.0500	0.5040	0.0620	0.4560	0.0460
MC S	0.0520	0.0340			0.5820	0.7640
LBQ S	0.0400	0.0480			0.5360	0.5540
sup LM S	0.0240	0.0540			0.5180	0.0620

Rejection Frequencies: Robust Tests, Nearly Unidentified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.030$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid	GARCH(1,1)	AR(2)	MA(10)	AR(2)	MA(10)
	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 10$	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 10$
MC ICS	0.0520	0.0500	0.8900	0.8120	0.5160	0.1508
LBQ ICS	0.0320	0.0360	0.8540	0.6280	0.4680	0.1111
sup LM ICS	0.0140	0.0380	0.8460	0.0420	0.4540	0.0119
MC S	0.0860	0.0960			0.8980	0.8460
LBQ S	0.0760	0.0700			0.8760	0.7340
sup LM S	0.0880	0.0900			0.8720	0.1100

ARMA Simulations

Rejection Frequencies: Robust Tests, Identified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.300$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid	GARCH(1,1)	AR(2)	MA(10)	AR(2)	MA(10)
MC ICS	0.0460	0.0300	0.5580	0.7600	0.5120	0.7180
LBQ ICS	0.0400	0.0420	0.5180	0.5420	0.4740	0.4780
sup LM ICS	0.0240	0.0500	0.5040	0.0620	0.4560	0.0460
MC S	0.0520	0.0340			0.5820	0.7640
LBQ S	0.0400	0.0480			0.5360	0.5540
sup LM S	0.0240	0.0540			0.5180	0.0620

Rejection Frequencies: Robust Tests, Nearly Unidentified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.030$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid	GARCH(1,1)	AR(2)	MA(10)	AR(2)	MA(10)
	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 10$	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 10$
MC ICS	0.0520	0.0500	0.8900	0.8120	0.5160	0.1508
LBQ ICS	0.0320	0.0360	0.8540	0.6280	0.4680	0.1111
sup LM ICS	0.0140	0.0380	0.8460	0.0420	0.4540	0.0119
MC S	0.0860	0.0960			0.8980	0.8460
LBQ S	0.0760	0.0700			0.8760	0.7340
sup LM S	0.0880	0.0900			0.8720	0.1100

ARMA Simulations

Rejection Frequencies: Robust Tests, Identified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.300$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid	GARCH(1,1)	AR(2)	MA(10)	AR(2)	MA(10)
MC ICS	0.0460	0.0300	0.5580	0.7600	0.5120	0.7180
LBQ ICS	0.0400	0.0420	0.5180	0.5420	0.4740	0.4780
sup LM ICS	0.0240	0.0500	0.5040	0.0620	0.4560	0.0460
MC S	0.0520	0.0340			0.5820	0.7640
LBQ S	0.0400	0.0480			0.5360	0.5540
sup LM S	0.0240	0.0540			0.5180	0.0620

Rejection Frequencies: Robust Tests, Nearly Unidentified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.030$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid $\mathcal{L}_n = 5$	GARCH(1,1) $\mathcal{L}_n = 5$	AR(2) $\mathcal{L}_n = 5$	MA(10) $\mathcal{L}_n = 10$	AR(2) $\mathcal{L}_n = 5$	MA(10) $\mathcal{L}_n = 10$
MC ICS	0.0520	0.0500	0.8900	0.8120	0.5160	0.1508
LBQ ICS	0.0320	0.0360	0.8540	0.6280	0.4680	0.1111
sup LM ICS	0.0140	0.0380	0.8460	0.0420	0.4540	0.0119
MC S	0.0860	0.0960			0.8980	0.8460
LBQ S	0.0760	0.0700			0.8760	0.7340
sup LM S	0.0880	0.0900			0.8720	0.1100

ARMA Simulations

Rejection Frequencies: Robust Tests, Identified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.300$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid	GARCH(1,1)	AR(2)	MA(10)	AR(2)	MA(10)
MC ICS	0.0460	0.0300	0.5580	0.7600	0.5120	0.7180
LBQ ICS	0.0400	0.0420	0.5180	0.5420	0.4740	0.4780
sup LM ICS	0.0240	0.0500	0.5040	0.0620	0.4560	0.0460
MC S	0.0520	0.0340			0.5820	0.7640
LBQ S	0.0400	0.0480			0.5360	0.5540
sup LM S	0.0240	0.0540			0.5180	0.0620

Rejection Frequencies: Robust Tests, Nearly Unidentified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.030$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid $\mathcal{L}_n = 5$	GARCH(1,1) $\mathcal{L}_n = 5$	AR(2) $\mathcal{L}_n = 5$	MA(10) $\mathcal{L}_n = 10$	AR(2) $\mathcal{L}_n = 5$	MA(10) $\mathcal{L}_n = 10$
MC ICS	0.0520	0.0500	0.8900	0.8120	0.5160	0.1508
LBQ ICS	0.0320	0.0360	0.8540	0.6280	0.4680	0.1111
sup LM ICS	0.0140	0.0380	0.8460	0.0420	0.4540	0.0119
MC S	0.0860	0.0960			0.8980	0.8460
LBQ S	0.0760	0.0700			0.8760	0.7340
sup LM S	0.0880	0.0900			0.8720	0.1100

ARMA Simulations

Rejection Frequencies: Robust Tests, Identified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.300$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid	GARCH(1,1)	AR(2)	MA(10)	AR(2)	MA(10)
MC ICS	0.0460	0.0300	0.5580	0.7600	0.5120	0.7180
LBQ ICS	0.0400	0.0420	0.5180	0.5420	0.4740	0.4780
sup LM ICS	0.0240	0.0500	0.5040	0.0620	0.4560	0.0460
MC S	0.0520	0.0340			0.5820	0.7640
LBQ S	0.0400	0.0480			0.5360	0.5540
sup LM S	0.0240	0.0540			0.5180	0.0620

Rejection Frequencies: Robust Tests, Nearly Unidentified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.030$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid	GARCH(1,1)	AR(2)	MA(10)	AR(2)	MA(10)
	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 10$	$\mathcal{L}_n = 5$	$\mathcal{L}_n = 10$
MC ICS	0.0520	0.0500	0.8900	0.8120	0.5160	0.1508
LBQ ICS	0.0320	0.0360	0.8540	0.6280	0.4680	0.1111
sup LM ICS	0.0140	0.0380	0.8460	0.0420	0.4540	0.0119
MC S	0.0860	0.0960			0.8980	0.8460
LBQ S	0.0760	0.0700			0.8760	0.7340
sup LM S	0.0880	0.0900			0.8720	0.1100

ARMA Simulations

Rejection Frequencies: Robust Tests, Identified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.300$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid	GARCH(1,1)	AR(2)	MA(10)	AR(2)	MA(10)
MC ICS	0.0460	0.0300	0.5580	0.7600	0.5120	0.7180
LBQ ICS	0.0400	0.0420	0.5180	0.5420	0.4740	0.4780
sup LM ICS	0.0240	0.0500	0.5040	0.0620	0.4560	0.0460
MC S	0.0520	0.0340			0.5820	0.7640
LBQ S	0.0400	0.0480			0.5360	0.5540
sup LM S	0.0240	0.0540			0.5180	0.0620

Rejection Frequencies: Robust Tests, Nearly Unidentified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.030$, $J = 500$

	H_0 True		H_0 False			
	Infeasible		Infeasible		Feasible	
	iid $\mathcal{L}_n = 5$	GARCH(1,1) $\mathcal{L}_n = 5$	AR(2) $\mathcal{L}_n = 5$	MA(10) $\mathcal{L}_n = 10$	AR(2) $\mathcal{L}_n = 5$	MA(10) $\mathcal{L}_n = 10$
MC ICS	0.0520	0.0500	0.8900	0.8120	0.5160	0.1508
LBQ ICS	0.0320	0.0360	0.8540	0.6280	0.4680	0.1111
sup LM ICS	0.0140	0.0380	0.8460	0.0420	0.4540	0.0119
MC S	0.0860	0.0960			0.8980	0.8460
LBQ S	0.0760	0.0700			0.8760	0.7340
sup LM S	0.0880	0.0900			0.8720	0.1100

ARMA Simulations

Rejection Frequencies: Robust Tests, Identified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.300$, $J = 500$

	H_0 True Infeasible		H_0 False Infeasible		H_0 False Feasible	
	iid	GARCH(1,1)	AR(2)	MA(10)	AR(2)	MA(10)
MC ICS	0.0460	0.0300	0.5580	0.7600	0.5120	0.7180
LBQ ICS	0.0400	0.0420	0.5180	0.5420	0.4740	0.4780
sup LM ICS	0.0240	0.0500	0.5040	0.0620	0.4560	0.0460
MC S	0.0520	0.0340			0.5820	0.7640
LBQ S	0.0400	0.0480			0.5360	0.5540
sup LM S	0.0240	0.0540			0.5180	0.0620

Rejection Frequencies: Robust Tests, Nearly Unidentified
ARMA, $\alpha = 0.05$, $T = 100$, $\beta = 0.030$, $J = 500$

	H_0 True Infeasible		H_0 False Infeasible		H_0 False Feasible	
	iid $\mathcal{L}_n = 5$	GARCH(1,1) $\mathcal{L}_n = 5$	AR(2) $\mathcal{L}_n = 5$	MA(10) $\mathcal{L}_n = 10$	AR(2) $\mathcal{L}_n = 5$	MA(10) $\mathcal{L}_n = 10$
MC ICS	0.0520	0.0500	0.8900	0.8120	0.5160	0.1508
LBQ ICS	0.0320	0.0360	0.8540	0.6280	0.4680	0.1111
sup LM ICS	0.0140	0.0380	0.8460	0.0420	0.4540	0.0119
MC S	0.0860	0.0960			0.8980	0.8460
LBQ S	0.0760	0.0700			0.8760	0.7340
sup LM S	0.0880	0.0900			0.8720	0.1100

Outline

Empirical Example

- Predictability of stock returns is an active area of research
  - Previous research examines serial correlation in raw return series (Campbell et al., 1997; Nankervis and Savin, 2010)
  - limitation: not appropriate for residuals from an ARMA model
- Exploit low levels of dependence in returns
- trading rule based on ARMA(1,1) model of returns (Taylor, 2005)
- I demonstrate this test as a test of model adequacy

data (Campbell et al., 1997; Nankervis and Savin, 2010):

- CRSP Value-Weighted NYSE/AMEX stock return indices
- Monthly between July 1962 - December 2005

▸ Nankervis & Savin (2010)

Empirical Example

VWRETD, monthly data, ARMA, Unknown Id

	1	2	3	4	5	6
Start Date	31-Jul-1962	31-Jul-1962	31-Oct-1978	31-Jul-1962	31-Jan-1995	31-Oct-1978
End Date	30-Dec-1994	29-Sep-1978	30-Dec-1994	30-Dec-2005	30-Dec-2005	30-Dec-2005
n	389	194	194	521	131	326
$\hat{\beta}_n$	0.075	0.096	0.034	0.070	0.088	0.036
$\hat{se}$	(0.914)	(0.735)	(0.244)	(0.637)	(0.817)	(0.237)

VWRETD, monthly data
Robust CV based Tests, $\mathcal{L}_n = 5$
ARMA, Unknown Id

	1	2	3	4	5	6
Start Date	31-Jul-1962	31-Jul-1962	31-Oct-1978	31-Jul-1962	31-Jan-1995	31-Oct-1978
End Date	30-Dec-1994	29-Sep-1978	30-Dec-1994	30-Dec-2005	30-Dec-2005	30-Dec-2005
n	389	194	194	521	131	326
MC ICS	2.97	1.37	2.47	2.68	1.23	$2.4^{\star\star}$
LBQ ICS	1.52	-0.511	1.12	1.54	-0.945	$1.81^{\star\star}$
sup LM ICS	2.53	1.02	1.07	6.43	2.88	1.83
MC S	$2.97^{\star\star}$	1.37	$2.47^{\star}$	$2.68^{\star\star}$	1.23	$2.4^{\star\star}$
LBQ S	$1.52^{\star}$	-0.511	1.12	$1.54^{\star}$	-0.945	$1.81^{\star\star}$
sup LM S	2.53	1.02	1.07	$6.43^{\star\star}$	2.88	1.83

: Significance levels: * = .1, ** = .05, *** = .01

Conclusion

- Parameter identification failure induces non-standard estimator distributions
- This non-standardness affects the distribution of the white noise test statistic in estimated models
- We can simulate the correct distribution for inference

Thank You!

References I

Andrews, D. W. K. and X. Cheng (2012): "Estimation and Inference with Weak, Semi-Strong, and Strong Identification," *Econometrica*, 80, 2153–2211.

——— (2013): "Maximum likelihood estimation and uniform inference with sporadic identification failure," *Journal of Econometrics*, 173, 36–56.

——— (2014): "GMM estimation and uniform subvector inference with possible identification failure," *Econometic Theory*, 30, 287–333.

Andrews, D. W. K. and W. Ploberger (1996): "Testing for Serial Correlation against an ARMA(1,1) Process," *Journal of the American Statistical Association*, 91, 1331–1342.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997): *The Econometrics of Financial Markets*, Princeton University Press.

Cheng, X. (2015): "Robust inference in nonlinear models with mixed identification strength," *Journal of Econometrics*, 189, 207–228.

Chernozhukov, V., D. Chetverikov, and K. Kato (2013): "Gaussian Approximations and Multiplier Bootstrap for Maxima of Sums of High-Dimensional Random Vectors," *Annals of Statistics*, 41, 2786–2819.

——— (2016): "Central Limit Theorems and Bootstrap in High Dimensions," *Annals of Probability*, forthcoming.

de Haan, L. (1976): "Sample Extremes: An Elementary Introduction," *Statistica Neerlandica*, 30, 161–172.

Fisher, R. A. and L. H. C. Tippet (1928): "Limiting forms of the frequency distribution of the largest or smallest member of a sample," *Proceeding of the Cambridge Philosophical Society*, 24, 180–290.

Francq, C., R. Roy, and J. M. Zakoian (2005): "Diagnostic Checking in ARMA Models with Uncorrelated Errors," *Journal of the American Statistical Association*, 100, 532–544.

Hansen, P. (2005): "A Test for Superior Predictive Ability," *Journal of Business & Economic Statistics*, 23, 365–380.

Hill, J. (2017): "Supplemental Material II for "A Smoothed P-Value Test When There is a Nuisance Parameter under the Alternative." STAR Model Misspecification Test," *UNC Working Paper*.

Hill, J. and J. Dennis (2018): "Testing Many Zero Restrictions Where a Subset May Lie on the Boundary," *UNC Working Paper*.

Hill, J. and K. Motegi (2017): "A Max-Correlation White Noise Test for Weakly Dependent Time Series," *UNC Working Paper*.

Hong, Y. (1996): "Consistent Testing for Serial Correlation of Unknown Form," *Econometrica*, 64, 837–864.

Luukkonen, R., P. Saikkonen, and T. Terasvirta (1988): "Testing linearity against smooth transition autoregressive models," *Biometrika*, 75, 491–499.

References II

NANKERVIS, J. C. AND N. E. SAVIN (2010): "Testing for Serial Correlation: Generalized Andrews-Ploberger Tests," *Journal of Business and Economic Statistics*, 28, 246–255.

——— (2012): "Testing for Uncorrelated Errors in ARMA Models: Non-Standard Andrews-Ploberger Tests," *Econometrics Journal*, 15, 516–534.

POTERBA, J. M. AND L. H. SUMMERS (1988): "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics*, 22, 27–59.

ROMANO, J. AND L. THOMBS (1996): "Inference for autocorrelations under Weak Assumptions," *Journal of the American Statistical Association*, 91, 590–600.

SHAO, X. (2010): "The Dependent Wild Bootstrap," *Journal of the American Statistical Association*, 105, 218–235.

——— (2011): "A Bootstrap-Assisted Spectral Test of White Noise under Unknown Dependence," *Journal of Econometrics*, 162, 213–224.

SKALIN, J. AND T. TERÄSVIRTA (2001): "Modeling asymmetries and moving equilibria in unemployment rates," *Macroeconomic Dynamics*.

TAYLOR, S. (2005): *Asset Price Dynamics, Volatility, and Prediction*, Princeton, NJ: Princeton University Press.

TERASVIRTA, T. (1994): "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association*, 89, 208–218.

TERÄSVIRTA, T. AND H. M. ANDERSON (1992): "Characterizing nonlinearities in business cycles using smooth transition autoregressive models," *Journal of Applied Econometrics*, 7, S119–S136.

VAN DIJK, D., T. TERÄSVIRTA, AND P. H. FRANSES (2002): "Smooth Transition Autoregressive Models - A Survey of Recent Developments," *Econometric Reviews*, 21, 1–47.

XIAO, H. AND W. B. WU (2014): "Portmanteau Test and Simultaneous Inference for Serial Covariances," *Statistica Sinica*, 24, 577–599.

ZHANG, D. AND W. B. WU (2017): "Gaussian Approximation for High Dimensional Time Series," *Annals of Statistics Forthcoming*.

ZHANG, X. AND G. CHENG (2017): "Gaussian Approximation for High Dimensional Vector Under Physical Dependence," *Bernoulli Forthcoming*.

Appendix Slides follow

Appendix Slides follow

Related Literature

We build on ideas from many research areas:

- White noise tests:
  (Hong, 1996; Andrews and Ploberger, 1996; Shao, 2011; Nankervis and Savin, 2012; Xiao and Wu, 2014; Hill and Motegi, 2017)
- Weak Identification/Parametric Identification Failure:
  (Andrews and Cheng, 2012, 2013, 2014; Cheng, 2015),
- Max Tests:
  (Chernozhukov et al., 2013, 2016; Zhang and Cheng, 2017; Zhang and Wu, 2017; Hill and Motegi, 2017; Hill and Dennis, 2018; Hansen, 2005)
- The Dependent Wild Bootstrap:
  (Shao, 2010, 2011).
- Testing for serial correlation when there is weak identification in the model estimation step is new.

Return

Empirical Example: Nankervis and Savin (2010)

Nankervis and Savin (2010) CRSP EWRETD, monthly

	1	2	3	4	5	6
Start Date	31-Jul-1962	31-Jul-1962	31-Oct-1978	31-Jul-1962	31-Jan-1995	31-Oct-1978
End Date	30-Dec-1994	29-Sep-1978	30-Dec-1994	30-Dec-2005	30-Dec-2005	30-Dec-2005
sup LM	9.4**	5.0	3.5	13.3**	5.3	11.3**
Exp-LM_0	6.3**	3.6*	2.3	9.0**	3.3	6.9**
Exp-LM_∞	3.8**	2.0	1.3	5.6**	2.0*	4.3**
BP5	11.0	5.1	7.4	16.1**	7.5	15.3**

- Tests *directly on the unfiltered returns* suggests that there may be some dependence that can be modeled here
- Let's model these series with an ARMA(1,1) and run diagnostics

Return to Empirical Example

Identification Failure - Economic Models

- Many commonly used models in Economics can have parameter identification failure.
- Today, we focus on the ▸ ARMA(1,1) model
  - parsimonious representations of many different time series(Andrews and Ploberger, 1996).
  - mean-reverting financial time series (Poterba and Summers, 1988),
  - price-trend models (Taylor, 2005).
- We also consider the ▸ LSTAR model (Terasvirta, 1994).
  - business cycle asymmetry with recession/expansion regimes (Teräsvirta and Anderson, 1992; Skalin and Teräsvirta, 2001).
  - Rich class of specifications by changing the transition function van Dijk et al. (2002)
- *Remember, we are testing the residuals for serial correlation*

▸ Return

Additive Non-linear Models

- Additive Non-linear Models

$$y_t = \beta g(X_t, \pi) + \zeta X_t + \varepsilon_t$$

$$\varepsilon_t(\theta) = y_t - \beta g(X_t, \pi) - \zeta X_t$$

- g is a smooth nonlinear function
- Estimate with Least Squares:

$$Q_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t(\theta)^2 / 2$$

- examples are Smooth Transition Auto-Regressive models (STAR) (Terasvirta, 1994)

Return to 1st Slide   Go back to WN Test Recap

ARMA Model

$$y_t = (\beta + \pi) y_{t-1} + \varepsilon_t - \pi \varepsilon_{t-1}$$

$$\varepsilon_t(\theta) = y_t - \beta \sum_{j=0}^{\infty} \pi^j y_{t-j-1}$$

- Estimate with QML:

$$Q_n(\theta) = \frac{1}{2} log \zeta + \frac{1}{2\zeta} \frac{1}{n} \sum_{t=1}^{n} \left(y_t - \beta \sum_{j=0}^{t-1} \pi^j y_{t-j-1} \right)^2$$

Return to Identification

Return to WN Test

Max Tests

$$\hat{\mathcal{T}}_n = \max_{1 \le h \le \mathcal{L}_n} |\sqrt{n}\hat{\rho}_n(h)|$$

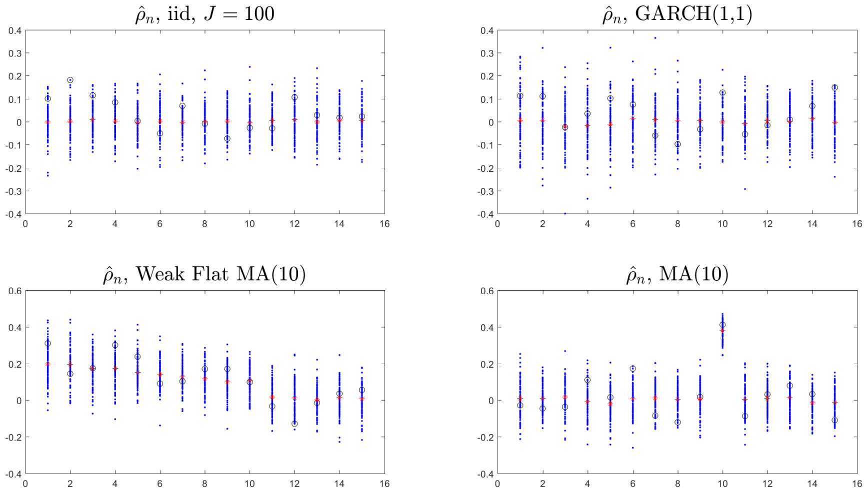
- do not require inversion of a large covariance matrix
- Utilize the most informative of a sequence of estimators
- trade-off: ignore information from everything that is not the maximum

▸ Example ▸ Example 2

- Inference usually relies on calculation of asymptotic distribution. (de Haan, 1976; Xiao and Wu, 2014)
- But EVT requires conditions that might be too restrictive (Hill and Dennis, 2018; Hill and Motegi, 2017)
  - (e.g. dependence properties, non-standard distributions)
- We will side-step EVT

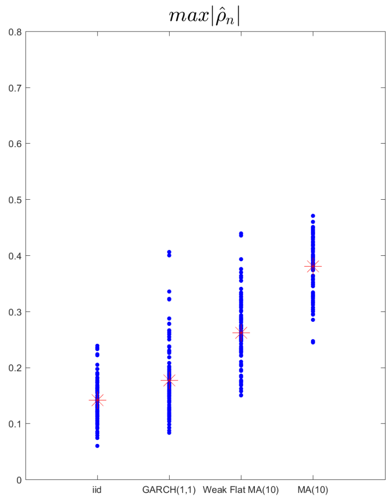
▸ Return to White Noise Test

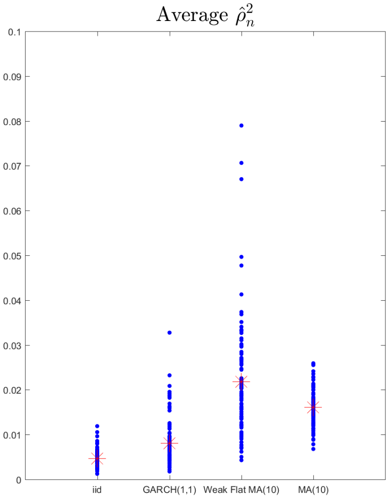
Max Tests - Correlation Examples



Combining Correlations
Go back

Max Tests - Max Correlation Examples





Go back

Bootstrap

Bootstrapping this type of test is relatively new (Chernozhukov et al., 2013, 2016; Zhang and Cheng, 2017; Zhang and Wu, 2017; Hill and Motegi, 2017; Hill and Dennis, 2018)

We must account for

- Dependence (uncorrelated, dependent errors)
- Non-standard distributions (weak identification)
- Unobserved random variables (residuals)

Return to Limiting Distribution of Test Stat

Appendix - Identification Robust Max Correlation Test

Appropriate for models with

- potentially unidentified parameters
- known sources of identification failure (Andrews and Cheng, 2012)

White Noise Test

- Only requires uncorrelatedness under the null (Romano and Thombs, 1996; Francq et al., 2005; Nankervis and Savin, 2010)
- Appropriate for residuals (estimated models)

Based on the maximum sample serial correlation (de Haan, 1976; Xiao and Wu, 2014; Hill and Motegi, 2017)

Note: Our inference procedure does *not* require the max correlation test.

▸ Return to Max Test

Appendix - White Noise Test - The Test Statistic

Theorem ($\mathcal{T}$ Limit Law)

Let some assumptions and H_0 hold. Then for some non-unique sequence of positive integers $\{\mathcal{L}_n\}$ with $\mathcal{L}_n \to \infty$ and $\mathcal{L}_n = o(n)$

a. *Under weak identification,* $\left| \hat{\mathcal{T}}_n - \max_{1 \leq h \leq \mathcal{L}_n} |\mathcal{Z}^{\psi}(h, \pi^*(b, \pi_0))| \right| \xrightarrow{p} 0.$
b. *Under strong identification,* $\left| \hat{\mathcal{T}}_n - \max_{1 \leq h \leq \mathcal{L}_n} |\mathcal{Z}^{\theta}(h)| \right| \xrightarrow{p} 0.$

Limiting Distributions

- Take away: Distributions are different $\Rightarrow$ possibility of distorted inference when ignoring possibility of identification failure
  - Under Strong Id, the limit is standard
  - Under Weak Id, $\hat{\pi}_n$ is not consistent, and the limiting distribution is complicated!
- We will bootstrap these distributions...

Bootstapping a Max Statistic?

Return to CV slide

Appendix - Bootstrap

- First order expansion
  - Accounts for the influence of model estimation ($\hat{\theta}_n$)
  - Differs depending on identification scenario
- Two limiting distributions:
- Strong Identification
  - expand about true parameter $\theta_n = (\beta_n, \zeta, \pi)$
  - the limit is standard
- Identification Failure
  - $\hat{\pi}_n$ is not consistent ($\hat{\pi}_n \xrightarrow{d} \pi^*$): additional source of randomness to replicate
  - expand about point of identification failure: $\psi_{0,n} = (0, \zeta)$: Two bias terms

Continue to Bootstrap Detail

Return to CV slide

Appendix - Bootstrap

- Dependent Wild Bootstrap (Shao, 2010, 2011).
- block size: k_n
- $\tilde{z}_1, \ldots, \tilde{z}_{n/k_n} \sim$ iid $N(0,1)$ random variables

$\tilde{z}_1$			$\tilde{z}_2$			$\ldots$	$\tilde{z}_{n/k_n}$		
$\tilde{z}_1$	$\tilde{z}_1$	$\tilde{z}_1$	$\tilde{z}_2$	$\tilde{z}_2$	$\tilde{z}_2$	$\ldots$	$\tilde{z}_{n/k_n}$	$\tilde{z}_{n/k_n}$	$\tilde{z}_{n/k_n}$
z_1	z_2	z_3	z_4	z_5	z_6	$\ldots$	z_{n-2}	z_{n-1}	z_n

- $\{z_t\}$ is a sequence of Gaussian multipliers

Formally...

Appendix - Bootstrap Overview

Weak Identification:

1. Simulate a random draw, $\pi^*_{(bs)}(b, \pi_0)$, from the distribution $\pi^*(b, \pi_0)$ using the z_t
2. Use $\pi^*_{(bs)}(b, \pi_0)$ to construct the components of our test statistic under weak identification, which are functions of π.
3. Use the draws z_t to construct the wild bootstrap version of the test statistic.
4. deal with nuisance parameters b and π_0

Go to the Bootstrap Algorithm for Weak Id

Strong Identification:

1. Construct the components of our test statistic using $\hat{\theta}_n$.
2. Use the draws z_t to construct the wild bootstrap version of the test statistic.

Go to the Bootstrap Algorithm for Strong Id

Appendix - Critical Value Computation Overview

1. Repeat the above procedures M times for each identification category.
2. Order the resulting test statistics within each category.
3. α-level critical values are the statistics in $[(1-\alpha)\cdot M]$th ordered positions
4. The critical value under weak identification depends on nuisance parameters. Sup over these nuisance parameters.

Go To Critical Value Computation Algorithm

Return to CV slide

Appendix - Critical Value Computation

Theorem

Under weak identification, let $k = w$, and under (semi-) strong identification, let $k = s$.
Let the number of bootstrap samples $M_n \to \infty$.
There is a non-unique sequence of positive integers $\{\mathcal{L}_n\}$ with $\mathcal{L}_n \to \infty$ and $\mathcal{L}_n = o(n)$ such that $\hat{c}^{(k)}_{1-\alpha,n} \xrightarrow{p} c^{(k)}_{1-\alpha}$.
Moreover, under the alternative hypothesis, $P(\hat{T}_n > \hat{c}^{(k)}_{1-\alpha,n}) \to 1$ for $k = w$ and $k = s$.

- Critical values are consistent
- Tests based on either critical value are consistent under the alternative
- But this implies 2 different tests...

Return to CV slide

Appendix - Putting the Critical Values Together

Robust Critical Values:

- Least Favorable (LF)
  - always take the larger of the critical values
  - $c^{(LF)}_{1-\alpha} = \max\{c^{(w)}_{1-\alpha}, c^{(s)}_{1-\alpha}\}$
- Identification-Category Selection (ICS)
  - data driven pre-test for the id category
  - Step 1: Use data to determine if $b = \lim_n \sqrt{n}\beta_n$ is finite
  - Step 2:
    - if we believe b is finite, use LF cv
    - otherwise, use (semi-) Strong identification cv

Go To ICS Details

Decision Rule: Reject the null hypothesis when $\hat{T}_n > c^{(\cdot)}_{1-\alpha}$.

Return to CV slide

Appendix - Putting the Critical Values Together

For any critical value, $c_{1-\alpha,n}$, the asymptotic size of the test is the maximum rejection probability over distributions consistent with the null hypothesis:

$$AsySz = \limsup_{n\to\infty} \sup_{\gamma\in\Gamma^*} P_\gamma(\mathcal{T}_n > c_{1-\alpha,n} |\ H_0).$$

Theorem (Andrews and Cheng (2012))

Under Andrews and Cheng's (2012) assumptions and H_0, the tests based on LF and ICS critical values $c^{(\cdot)}_{1-\alpha,n}$ satisfy $AsySz = \alpha$.

Return to CV slide

Monte Carlo Simulations

- $J = 500$ samples of size
- $n \in \{100, 250, 500, 1000\}$

An assortment of lag lengths:

- $\mathcal{L}_n \in \{5, [n^{1/3}], [\sqrt{n}/(\ln(n)/4)], [\sqrt{n}/(\ln(n)/5)], [\sqrt{n}], [.5n/\ln(n)]\}$.
- Additionally, we use $\mathcal{L}_n = [n/\ln(n)]$ when $n = 500, 1000$ leading to lag lengths 80 and 144, respectively.

For the Bootstrap, we use

- $M = 500$ bootstrap samples
- DWB block size: $k_n = [\sqrt{n}]$.

Return to Simulation Setup

STAR Simulations - Observations - Size/Power

- Size (H_0 true):
  - We see empirical size shrinkage for MC and LBQ across all specifications as $\mathcal{L}_n$ increases. Size Shrinkage
  - Less of an issue for MC
  - At lower $\mathcal{L}_n$, sizes of ICS based statistics are close to nominal for iid errors, and conservative for GARCH errors
- Power (H_0 false):
  - ICS based tests are comparable to S based counterparts
  - MC has comparable power to LBQ and sometimes smaller power than sup LM and CvM for H_A with close correlation. AR(2) Power MA(1) Power
  - MC dominates for H_A with distant correlation MA(10) Power
- Other Observations:
  - Under Strong Id, ICS is comparable to S MC ICS vs. MC S
  - Size distortions of S only sometimes noticeable ICS vs. S
  - Under weak and non-id, ICS does not always appear to dominate S. ICS vs. S 2

Return to ARMA sims

ARMA Simulations

Rejection Frequencies: Infeasible CV based Tests
ARMA, Weak Id, $\alpha = 0.05$
$T = 500$, $\beta = 0.013$, $J = 500$

	H_0 True		H_0 False			
	iid, $\mathcal{L}_n = 5$	GARCH(1,1), $\mathcal{L}_n = 5$	AR(2), $\mathcal{L}_n = 5$	MA(10), $\mathcal{L}_n = 22$	MA(21), $\mathcal{L}_n = 80$	MA(
MC ICS	0.0280	0.0460	1.0000	1.0000	1.0000	
LBQ ICS	0.0260	0.0300	1.0000	1.0000	0.9680	
sup LM ICS	0.0080	0.0280	1.0000	0.0540	0.0400	
CvM ICS	0.0020	0.0120	1.0000	0.0300	0.0200	
MC S	0.0460	0.0820	1.0000	1.0000	1.0000	
LBQ S	0.0540	0.0660	1.0000	1.0000	0.9820	
sup LM S	0.0440	0.0720	1.0000	0.1260	0.0900	
CvM S	0.0320	0.0420	1.0000	0.3260	0.1460	

Rejection Frequencies: Infeasible CV based Tests
ARMA, Strong Id, $\alpha = 0.05$
$T = 500$, $\beta = 0.300$, $J = 500$

	H_0 True		H_0 False			
	iid, $\mathcal{L}_n = 5$	GARCH(1,1), $\mathcal{L}_n = 5$	AR(2), $\mathcal{L}_n = 5$	MA(10), $\mathcal{L}_n = 22$	MA(21), $\mathcal{L}_n = 80$	MA(
MC ICS	0.0340	0.0280	0.9560	1.0000	1.0000	
LBQ ICS	0.0340	0.0160	0.9520	1.0000	0.9640	
sup LM ICS	0.0200	0.0100	0.9540	0.0540	0.0240	
CvM ICS	0.0060	0.0100	0.8520	0.0780	0.0320	
MC S	0.0340	0.0280	0.9560	1.0000	1.0000	
LBQ S	0.0340	0.0160	0.9520	1.0000	0.9640	
sup LM S	0.0200	0.0100	0.9540	0.0540	0.0240	
CvM S	0.0060	0.0100	0.8520	0.0780	0.0320	

Return to ARMA Sims

VWRETD, annual data
Robust CV based Tests, $\mathcal{L}_n = .5 * n/\log(n)$
ARMA, Unknown Id

	1	2	3	4	5	6	7
Start Date	31-Dec-1962	31-Dec-1962	29-Dec-1978	31-Dec-1962	29-Dec-1995	29-Dec-1978	30-Dec-19
End Date	30-Dec-1994	29-Dec-1978	30-Dec-1994	30-Dec-2005	30-Dec-2005	30-Dec-2005	30-Dec-20
n	32	16	16	43	10	27	17
MC ICS	2.32	0.986	1.04	1.79	0.561	0.451	1.16
LBQ ICS	$2.34^{\star}$	-0.468	-0.452	0.524	-1.58	-1.29	-0.785
sup LM ICS	0.63	0.141	1.08	0.364	0.384	0.376	2.41
CvM ICS	$5.56e-06$	$1.48e-05$	$7.01e-05$	$1.27e-07$	$7.29e-10$	$7.7e-08$	$2.99e-08$
MC S	2.32	$0.986^{\star\star\star}$	1.04	1.79	0.561	0.451	1.16
LBQ S	$2.34^{\star}$	$-0.468^{\star\star\star}$	-0.452	0.524	-1.58	-1.29	-0.785
sup LM S	0.63	0.141	1.08	0.364	0.384	0.376	2.41
CvM S	$5.56e-06$	$1.48e-05$	$7.01e-05$	$1.27e-07$	$7.29e-10$	$7.7e-08$	$2.99e-08$

Return to Empirical Example

EWRETD, annual data
Robust CV based Tests, $\mathcal{L}_n = .5 * n/\log(n)$
ARMA, Unknown Id

	1	2	3	4	5	6	7
Start Date	31-Dec-1962	31-Dec-1962	29-Dec-1978	31-Dec-1962	29-Dec-1995	29-Dec-1978	30-Dec-198
End Date	30-Dec-1994	29-Dec-1978	30-Dec-1994	30-Dec-2005	30-Dec-2005	30-Dec-2005	30-Dec-200
n	32	16	16	43	10	27	17
MC ICS	1.86	0.313	0.689	2.54	0.794	0.974	1.2
LBQ ICS	0.413	-0.917	-0.588	1.1	-1.43	-0.772	$-0.748^{\star}$
sup LM ICS	2.96	0.0805	0.546	4.61	0.368	0.81	$4.15^{\star\star\star}$
CvM ICS	$5.53e-07$	$4.95e-06$	0.000114	$6.75e-08$	$2.84e-05$	$5.65e-06$	$2.47e-0$
MC S	$1.86^{\star\star\star}$	0.313	0.689	$2.54^{\star\star\star}$	0.794	0.974	1.2
LBQ S	$0.413^{\star\star}$	-0.917	-0.588	$1.1^{\star\star\star}$	-1.43	-0.772	$-0.748^{\star}$
sup LM S	$2.96^{\star\star}$	0.0805	0.546	$4.61^{\star\star}$	0.368	0.81	$4.15^{\star\star\star}$
CvM S	$5.53e-07$	$4.95e-06$	0.000114	$6.75e-08$	$2.84e-05^{\star}$	$5.65e-06$	$2.47e-0$

Return to Empirical Example

EWRETD, monthly data
Robust CV based Tests, $\mathcal{L}_n = .5 * n/\log(n)$
ARMA, Unknown Id

	1	2	3	4	5	6	7
Start Date	31-Jul-1962	31-Jul-1962	31-Oct-1978	31-Jul-1962	31-Jan-1995	31-Oct-1978	29-Jan-198
End Date	30-Dec-1994	29-Sep-1978	30-Dec-1994	30-Dec-2005	30-Dec-2005	30-Dec-2005	30-Dec-200
n	389	194	194	521	131	326	215
MC ICS	$3.16^{\star}$	2.77	2.08	3.32	1.62	2.73	1.59
LBQ ICS	0.542	0.522	-1.03	1.42	-0.801	0.427	-1.34
sup LM ICS	0.557	0.823	0.668	0.815	0.948	0.911	0.0674
CvM ICS	$1.71e-11$	$1.58e-10$	$3.81e-09$	$4.32e-12$	$7.93e-11$	$6.9e-10$	$1.98e-1$
MC S	$3.16^{\star}$	2.77	2.08	3.32	1.62	2.73	1.59
LBQ S	0.542	0.522	-1.03	1.42	-0.801	0.427	-1.34
sup LM S	0.557	0.823	0.668	0.815	0.948	0.911	0.0674
CvM S	$1.71e-11$	$1.58e-10$	$3.81e-09$	$4.32e-12$	$7.93e-11$	$6.9e-10$	$1.98e-1$

Return to Empirical Example

Bootstrap

- Dependent Wild Bootstrap (Shao, 2010, 2011).
- First, draw standard normal random variables with perfect dependence within blocks and independence across blocks:

Formally,

- Select a block size k_n s.t. $1 \leq k_n \leq n$, $k_n \to \infty$, and $k_n/n \to 0$.
- Define blocks by $\mathbb{B}_s = \{(s-1)k_n + 1, \ldots, sk_n\}$ for $s = 1, \ldots, n/k_n$.
- Generate iid $N(0,1)$ random variables $\{\tilde{z}_1, \ldots, \tilde{z}_{n/k_n}\}$ and
- Define $z_t = \tilde{z}_s$ if $t \in \mathbb{B}_s$.

- $\{z_t\}$ is now a sequence of Gaussian multipliers

Go back

- LSTAR Model (Terasvirta, 1994; Andrews and Cheng, 2013; Hill, 2017):

$$\varepsilon_t(\theta) = y_t - \beta y_{t-1} \times g(y_{t-d}, \pi) - \zeta y_{t-1},$$
$$g(z, \pi) = \frac{1}{1 + \exp\{-\pi_1(z - \pi_2)\}}$$

- Estimate with Least Squares:

$$Q_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t(\theta)^2/2$$

- Notation: $\theta = (\psi, \pi)$, $\psi = (\beta, \zeta)$, $\psi_n = (\beta_n, \zeta)$, $\psi_{0,n} = (0, \zeta)$

Go to STAR Model Detailed Assumptions | Go to Estimator Limiting Distributions | Return to example models

Recap - ARMA Model

$$y_t = (\beta + \pi)y_{t-1} + \varepsilon_t - \pi\varepsilon_{t-1}$$
$$\varepsilon_t(\theta) = y_t - \beta \sum_{j=0}^{\infty} \pi^j y_{t-j-1}$$

- Estimate with QML:

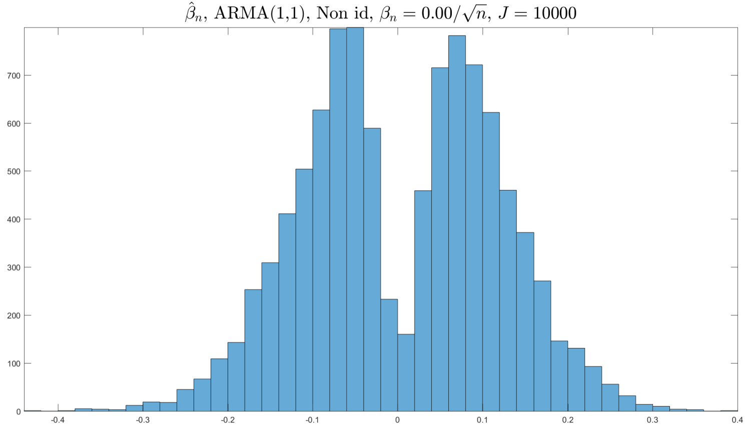
$$Q_n(\theta) = \frac{1}{2} log\zeta + \frac{1}{2\zeta}\frac{1}{n}\sum_{t=1}^{n}\left(y_t - \beta \sum_{j=0}^{t-1} \pi^j y_{t-j-1}\right)^2$$

- Notation: $\theta = (\psi, \pi)$, $\psi = (\beta, \zeta)$, $\psi_n = (\beta_n, \zeta)$, $\psi_{0,n} = (0, \zeta)$

Return to example models

Identification Robustness

Identification failure leads to non-standard distributions:

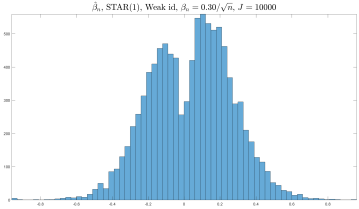


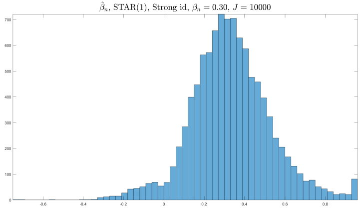
$\Rightarrow$ Non-standard Inference!

back to ARMA(1,1) $\hat{\beta}_n$

Identification Robustness

$y_t = \beta_n g(y_{t-1}, \pi) + \zeta_n y_{t-1} + \varepsilon_t$:





- $\Rightarrow$ Non-standard Inference!

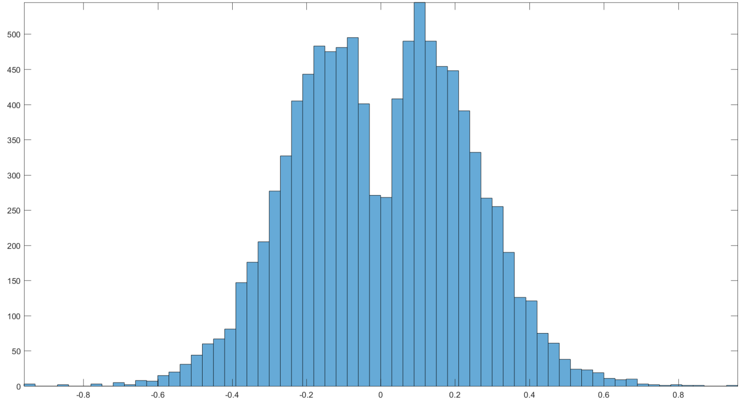
ARMA(1,1) $\hat{\beta}_n$ STAR(1) $\hat{\beta}_n$, non-id

- Current inference relies on calculation of distributions (Andrews and Ploberger, 1996; Andrews and Cheng, 2012; Cheng, 2015)
- Bootstrap has not been explored as a means of better approximating the finite sample distribution

Identification Robustness

Identification failure leads to non-standard finite sample distributions ($y_t = \beta_n g(y_{t-1}, \pi) + \zeta_n + \varepsilon_t$):

$\hat{\beta}_n$, STAR(1), Non id, $\beta_n = 0.00/\sqrt{n}$, $J = 10000$



$\Rightarrow$ Non-standard Inference!

back to STAR(1) $\hat{\beta}_n$

Identification Categories

$$n^{\alpha}\beta_n \to b \in \mathbb{R}^{k_\beta}$$

α	Identification Category of π
$\alpha \in [0, 1/2)$	(semi-) Strong
$\alpha \in [1/2, \infty)$	Weak

Return to Introduction

LSTAR Model Assumptions

a. True Parameter Space
   i. $\Theta^* = \{(\beta, \zeta, \pi) : \beta \in \mathcal{B}^*,\ \zeta \in \mathcal{Z}^*(\beta), \pi \in \Pi^*\}$ is compact.
   ii. $0 \in int(\mathcal{B}^*)$, $\Pi^* = \{(\pi_1, \pi_2) :\ \pi_1 \geq c\}$ for some $c > 0$. For some set $\mathcal{Z}_0^*$ and $\delta > 0$, $\mathcal{Z}^*(\beta) = Z_0^*\ \forall \beta$ s.t. $||\beta|| < \delta$.

b. Optimization Parameter Space
   i. $\Theta = \{(\beta, \zeta, \pi) : \beta \in \mathcal{B},\ \zeta \in \mathcal{Z}(\beta), \pi \in \Pi\}$ is compact, and $\Theta^* \subset int(\Theta)$.
   ii. For some set $\mathcal{Z}_0$ and $\delta > 0$, $\mathcal{Z}(\beta) = Z_0\ \forall \beta$ s.t. $||\beta|| < \delta$, and $\mathcal{Z}_0^* \subset int(\mathcal{Z}_0)$.

Assumptions continued...

Return to LSTAR

LSTAR Model Assumptions

c. Under H_0, $E(\varepsilon_t|x_t) = 0$ a.s. and $E(\varepsilon_t^2|x_t) = \sigma^2 \in (0, \infty)$ a.s.
d. $E[\varepsilon_t(\psi_0, \pi)d_{\psi,t}(\pi)] = 0$ for a unique $\psi_0 = (0', \zeta_0')' \in int(\Psi^*)$, or $E[\varepsilon_t(\theta_0)d_{\theta,t}(\pi)] = 0$ for a unique $\theta_0 = (\beta_0', \zeta_0', \pi_0')' \in int(\Psi^* \times \Pi^*) = int(\Theta^*)$.
e. y_t is strictly stationary, L_p-bounded for some $p > 8$, and β-mixing with mixing coefficients $\beta_l = O(l^{-1p(1-p)-\iota})$ for some $q > p$ and $\iota > 0$.
f. $g(\cdot, \pi)$ is Borel measurable for each π and twice continuously differentiable in π. $g(z_t, \pi)$ is a non-degenerate random variable for each $\pi \in \Pi$. $E[\sup_{\pi \in \Pi} ||(\partial/\partial\pi)^j g(z_t, \pi)||^8] < \infty$ for $j = 0, 1, 2$.
g. Long Run Variances
    i. Under weak identification, $\liminf_{n\to\infty} \inf_{\alpha,r,\pi} E[(r' \sum_{i=1}^m \alpha_i G_n^\psi(\pi_i))^2] > 0$.
    ii. Under strong identification, $\liminf_{n\to\infty} \inf_r E[(r' G_n^\theta)^2] > 0$.
    iii. $\inf_{r,\pi} E[(r' d_{\psi,t}(\pi))^2] > 0$ and $\inf_{r,\beta,\pi} E[(r' B^{-1}(\beta) d_{\theta,t}(\pi))^2] > 0$.

Return to LSTAR

Estimator Limiting Distributions

Under **Weak Identification**:

$$\begin{pmatrix} n^{1/2}(\hat{\psi}_n - \psi_n) \\ \hat{\pi}_n \end{pmatrix} \xrightarrow{d} \begin{pmatrix} \tau(\pi^*(\gamma_0, b); \gamma_0, b) \\ \pi^*(\gamma_0, b) \end{pmatrix}$$

where

- $\tau(\pi; \gamma_0, b) = -H^{-1}(\pi; \gamma_0)(G(\pi; \gamma_0) + K(\pi; \gamma_0)b) - (b, 0)$
- $\pi^*(\gamma_0, b) = \underset{\pi \in \Pi}{\operatorname{argmin}} \ \xi(\pi; \gamma_0, b)$
- and $\xi(\pi; \gamma_0, b) = -\frac{1}{2}(G(\pi) + K(\pi, \pi_0)b)' H^{-1}(\pi)(G(\pi) + K(\pi, \pi_0)b)$.

Weak ID Objects

Under **Strong Identification**:

$$n^{1/2} B(\beta_n)(\hat{\theta}_n - \theta_n) \xrightarrow{d} J^{-1} G^{\theta}$$

- where $B(\beta) = \begin{pmatrix} I_{d_\psi} & 0_{d_\psi \times d_\pi} \\ 0_{d_\psi \times d_\pi} & ||\beta|| \cdot I_{d_\pi} \end{pmatrix}$ and $G^{\theta} \sim N(0, V)$

Return to LSTAR

LSTAR Model - Weak Identification Objects

Define $d_{\psi,t}(\pi) = \frac{\partial}{\partial\psi}\varepsilon_t(\psi,\pi) = (X_t' g(Z_t,\pi), X_t')'$.

$G(\cdot;\gamma_0)$ is a mean-zero Gaussian process with covariance kernel

$$\Omega(\pi,\tilde{\pi};\gamma_0) = E_{\gamma_0}[\varepsilon_t^2 d_{\psi,t}(\pi) d_{\psi,t}(\tilde{\pi})']$$

and

$$K(\pi;\gamma_0) = -E_{\gamma_0}[d_{\psi,t}(\pi) d_{\psi,t}(\pi_0)' \cdot S_\beta'], \qquad S_\beta = [I_{d_\beta} : 0] \in \mathbb{R}^{d_\beta \times d_\psi}$$

$$H(\pi;\gamma_0) = E_{\gamma_0}[d_{\psi,t}(\pi) d_{\psi,t}(\pi)']$$

$$D^\psi(h,\pi) = E_{\gamma_0}[d_{\psi,t}(\pi)\varepsilon_{t-h}(\psi_{0,n},\pi) + d_{\psi,t-h}(\pi)\varepsilon_t(\psi_{0,n},\pi)]$$

Return to Estimator Limiting Distributions

Return to Test Statistic Limiting Distributions

LSTAR Sample Weak Identification Objects

$$\hat{H}_n(\pi) = \frac{1}{n}\sum_{t=1}^{n} d_{\psi,t}(\pi) d_{\psi,t}(\pi)'$$

$$\hat{K}_n(\pi;\gamma_0) = -\frac{1}{n}\sum_{t=1}^{n} d_{\psi,t}(\pi) x_t g(z_t,\pi_0)$$

and

$$G_n(\pi) \equiv \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left\{ m_t^{\psi}(\pi) - E_{\gamma_n}[m_t^{\psi}(\pi)]\right\}$$

$$m_t^{\psi}(\pi) \equiv m_t(\psi_{0,n},\pi) = \varepsilon_t(\psi_{0,n},\pi) d_{\psi,t}(\pi)$$

Return to Estimator Limiting Distributions

Return to Test Statistic Limiting Distributions

Test Statistic Limits

$\{\mathcal{Z}^{\psi}(h,\pi) : h \in \mathbb{N}, \pi \in \Pi\}$ is a

- Gaussian process with mean $\lim_{n\to\infty} \sqrt{n} E_{\gamma_n}(r_t^{2,\psi,n}(h,\pi)) < \infty$ and
- covariance kernel $\lim_{n\to\infty} \frac{1}{n} \sum_{s,t=1}^{n} E_{\gamma_n}[r_s^{1,\psi,n}(h,\pi)\ r_t^{1,\psi,n}(\tilde{h},\tilde{\pi})]$.
- where $r_t^{1,\psi,n}(h,\pi) = \frac{\varepsilon_t \varepsilon_{t-h} - E[\varepsilon_t \varepsilon_{t-h}]}{E[\varepsilon_t^2]} - \frac{\mathcal{D}(h,\pi)' H^{-1}(\pi;\gamma_0)\left(m_t^{\psi}(\pi) - E_{\gamma_n}[m_t^{\psi}(\pi)]\right)}{E[\varepsilon_t^2]}$ and
- $r_t^{2,\psi,n}(h,\pi) =$ $\frac{\varepsilon_t(\psi_{0,n},\pi)\varepsilon_{t-h}(\psi_{0,n},\pi) - \varepsilon_t \varepsilon_{t-h}}{E[\varepsilon_t^2]} - \frac{\mathcal{D}(h,\pi)' H^{-1}(\pi;\gamma_0)\left(\beta_n \frac{\partial}{\partial \tilde{\beta}} E_{\tilde{\gamma}_n}[m_t^{\psi}(\pi)]\right)}{E[\varepsilon_t^2]}$.

$\{\mathcal{Z}^{\theta}(h) : h \in \mathbb{N}\}$ is a

- zero mean Gaussian process with
- covariance kernel $\lim_{n\to\infty} \frac{1}{n} \sum_{s,t=1}^{n} E[r_s^{\theta}(h) r_t^{\theta}(\tilde{h})]$.
- where $r_t^{\theta}(h) = \frac{\varepsilon_t \varepsilon_{t-h} - E[\varepsilon_t \varepsilon_{t-h}] - \mathcal{D}^{\theta}(h)' J^{-1}(\gamma_0) m_t^{\theta}}{E[\varepsilon_t^2]}$.

Go to LSTAR Weak ID Objets
Return to Test Statistics

Step 1. Compute

$$\hat{H}_n(\pi) = \frac{1}{n}\sum_{t=1}^{n} d_{\psi,t}(\pi) d_{\psi,t}(\pi)'$$

$$\hat{K}_n(\pi;\gamma_0) = \frac{1}{n}\sum_{t=1}^{n} d_{\psi,t}(\pi) x_t' g(z_t,\pi_0)$$

$$\hat{G}_n^{(bs)}(\pi) = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} z_t \Big\{ d_{\psi,t}(\pi)\varepsilon_t(\hat{\psi}_{0,n}(\pi),\pi) - \frac{1}{n}\sum_{t=1}^{n}\left[d_{\psi,t}(\pi)\varepsilon_t(\hat{\psi}_{0,n}(\pi),\pi)\right]\Big\}$$

Define

$$\xi_n^{(bs)}(\pi;\gamma_0,b) = \\ -\frac{1}{2}\Big(\hat{G}_n^{(bs)}(\pi) + \hat{K}_n(\pi;\gamma_0)b\Big)'(\hat{H}_n(\pi))^{-1}\Big(\hat{G}_n^{(bs)}(\pi) + \hat{K}_n(\pi;\gamma_0)b\Big),$$

and compute $\pi^*_{(bs)}(\gamma_0,b) = \operatorname*{argmin}_{\pi\in\Pi}\ \xi_n^{(bs)}(\pi;\gamma_0,b)$.

Bootstrap under Weak Identification - Algorithm

Step 2. Use $\pi^*_{(bs)}(\gamma_0, b)$ and $\hat{\psi}_{0,n} = (0', \hat{\zeta}_n')'$ to compute

$$\mathcal{G}_n(\pi^*_{(bs)}) = (\hat{H}_n(\pi^*_{(bs)}))^{-1}\left[m_t(\hat{\psi}_{0,n}, \pi^*_{(bs)}) - \frac{1}{n}\sum_{t=1}^{n} m_t(\hat{\psi}_{0,n}, \pi^*_{(bs)})\right]$$

and

$$\hat{\mathcal{D}}_n(h, \pi^*_{(bs)}) =$$
$$\frac{1}{n}\sum_{t=1+h}^{n}[d_{\psi,t}(\pi^*_{(bs)})\varepsilon_{t-h}(\hat{\psi}_{0,n}, \pi^*_{(bs)}) + d_{\psi,t-h}(\pi^*_{(bs)})\varepsilon_t(\hat{\psi}_{0,n}, \pi^*_{(bs)})].$$

Define

$$\hat{\mathcal{E}}_{t,h}(\psi, \pi) = \varepsilon_t(\psi, \pi)\varepsilon_{t-h}(\psi, \pi) - \mathcal{G}_n(\pi^*_{(bs)})'\hat{\mathcal{D}}_n(h, \pi^*_{(bs)})$$
$$- \frac{1}{n}\sum_{t=1+h}^{n}[\varepsilon_t(\psi, \pi)\varepsilon_{t-h}(\psi, \pi) - \varepsilon_t\varepsilon_{t-h}],$$

Step 3. Use the draws $\{z_t\}$ to define

$$\hat{\rho}_n^{(w)}(h;\gamma_n,b) = \frac{1}{n^{-1}\sum_{t=1}^{n}\varepsilon_t^2(\hat{\theta}_n)} \times \Bigg\{ \frac{1}{n}\sum_{t=1+h}^{n} z_t\Big(\hat{\mathcal{E}}_{t,h}(\hat{\psi}_{0,n},\pi^*_{(bs)}) - \frac{1}{n}\sum_{t=1+h}^{n}\hat{\mathcal{E}}_{t,h}(\hat{\psi}_{0,n},\pi^*_{(bs)})\Big) - ((\hat{H}_n(\pi^*_{(bs)}))^{-1}\hat{K}_n(\pi;\gamma_n)\frac{b}{\sqrt{n}})'\hat{\mathcal{D}}_n(h,\pi^*_{(bs)}) + \frac{1}{n}\sum_{t=1+h}^{n}\left[\varepsilon_t(\hat{\psi}_{0,n},\pi)\varepsilon_{t-h}(\hat{\psi}_{0,n},\pi) - \varepsilon_t\varepsilon_{t-h}\right]\Bigg\}.$$

Define the bootstrapped test statistic

$$\hat{\mathcal{T}}_n^{(w)}(\gamma_n,b) = \max_{1\le h\le \mathcal{L}_n} |\sqrt{n}\;\hat{\rho}_n^{(w)}(h;\gamma_n,b)|$$

Bootstrap under Weak Identification - Algorithm

Step 4. The final step is discussed under critical value computation.

Go to Critical Value Computation

Return to Bootstrap Summary

Bootstrap under Strong Identification - Algorithm

- Simpler construction (no bias terms)
- Simpler procedure (no π^* and no nuisance parameters).

1. Compute $\hat{J}_n(\hat{\theta}_n) = \frac{1}{n}\sum_{t=1}^{n} B(\hat{\beta}_n)^{-1} d_{\theta,t}(\hat{\theta}_n) d_{\theta,t}(\hat{\theta}_n)' B(\hat{\beta}_n)^{-1\prime}$, $\hat{\mathcal{D}}_n^{\theta}(h, \hat{\theta}_n) = \frac{1}{n}\sum_{t=1+h}^{n} B(\hat{\beta}_n)^{-1}[d_{\theta,t}(\hat{\theta}_n)\varepsilon_{t-h}(\hat{\theta}_n) + d_{\theta,t-h}(\hat{\theta}_n)\varepsilon_t(\hat{\theta}_n)]$, and $m_t^{\theta}(\theta) = d_{\theta,t}(\hat{\theta}_n)\varepsilon_t(\hat{\theta}_n)$.
2. Define $\hat{\mathcal{E}}_{t,h}(\theta) = \varepsilon_t(\theta)\varepsilon_{t-h}(\theta) - (\hat{\mathcal{D}}_n^{\theta}(h,\theta))'(\hat{J}_n(\hat{\theta}_n))^{-1} B(\hat{\beta}_n)^{-1} m_t^{\theta}(\theta)$.
3. Use the draws $\{z_t\}$ to define

$$\hat{\rho}_n^{(s)}(h) = \frac{1}{n^{-1}\sum_{t=1}^{n}\varepsilon_t^2(\hat{\theta}_n)} \times \left\{ \frac{1}{n}\sum_{t=1+h}^{n} z_t \left(\hat{\mathcal{E}}_{t,h}(\hat{\theta}_n) - \frac{1}{n}\sum_{t=1+h}^{n} \hat{\mathcal{E}}_{t,h}(\hat{\theta}_n) \right) \right\}$$

4. Define the bootstrapped test statistic $\hat{\mathcal{T}}_n^{(s)} = \max_{1 \le h \le \mathcal{L}_n} |\sqrt{n}\, \hat{\rho}_n^{(s)}(h)|$.

Go to Critical Value Computation

Return to Bootstrap Summary

Critical Value Computation - Algorithm

1. For $k = w, s$, repeat the above procedures $i = 1, \ldots, M$ times: $\{\hat{\mathcal{T}}^{(s)}_{n,i}\}_{i=1}^{M}$ and $\{\hat{\mathcal{T}}^{(w)}_{n,i}(\gamma_n, b)\}_{i=1}^{M}$.
2. Define the order statistics $\{\hat{\mathcal{T}}^{(k)}_{n,(i)}\}_{i=1}^{M}$ such that $\hat{\mathcal{T}}^{(k)}_{n,(1)} \leq \hat{\mathcal{T}}^{(k)}_{n,(2)} \leq \cdots \leq \hat{\mathcal{T}}^{(k)}_{n,(M)}$.
3. The approximate α-level critical values are
   - strong id: $\hat{c}^{(s)}_{n,1-\alpha} = \hat{\mathcal{T}}^{(s)}_{n,[(1-\alpha)\cdot M]}$
   - weak id: $\hat{c}^{(w)}_{n,1-\alpha}(\gamma_n, b) = \hat{\mathcal{T}}^{(w)}_{n,[(1-\alpha)\cdot M]}(\gamma_n, b)$
     - Consistent estimators are not available for (π_n, b)
     - In practice, $c^{(w)}_{n,1-\alpha} = \sup_{\pi_n, b} c^{(w)}_{n,1-\alpha}(\gamma_n, b)$.

Return to Critical Value Overview

Return to Bootstrap Overview

ICS Critical Values

The statistic used for category selection is

$$\mathcal{A}_n = (n\hat{\beta}_n'\hat{\Sigma}_{\beta\beta,n}^{-1}\hat{\beta}_n/d_\beta)^{1/2}$$

where $\hat{\Sigma}_{\beta\beta,n}$ is the upper left $d_\beta \times d_\beta$ block of $\hat{\Sigma}_n = \hat{J}_n^{-1}\hat{V}_n\hat{J}_n^{-1}$, the estimator of the strong identification covariance matrix $\Sigma(\gamma_0) = J^{-1}VJ^{-1}$. Let $\{\kappa_n : \ n \geq 1\}$ be a sequence of constants s.t.

$$\kappa_n \to \infty \qquad \text{and} \qquad \kappa_n/n^{1/2} \to 0$$

The ICS critical value is $c_{1-\alpha}^{(ICS)} = \begin{cases} c_{1-\alpha}^{(LF)} & \text{if } \mathcal{A}_n \leq \kappa_n \\ c_{1-\alpha}^{(s)} & \text{if } \mathcal{A}_n > \kappa_n \end{cases}$

Return to CV Overview

Why not just test a Taylor Expansion of the model?

- Some tests for linearity against STR alternatives use a Taylor expansion. (e.g. (Luukkonen et al., 1988)).
- This adds a remainder term to the error.
- Taylor Expansion of $g(z, \pi)$ around $\pi_1 = 0$ (not $\pi_{1,0}$):

$$y_t = \zeta y_{t-1} + \beta y_{t-1} g(z_t, \pi) + \varepsilon_t \quad \Rightarrow \quad \begin{aligned} y_t &= \alpha_0 y_{t-1} + \alpha_1 y_{t-1} z_t + \nu_t \\ \nu_t &= \varepsilon_t + \beta y_{t-1} R(z_t, \pi) \end{aligned}$$

- H_0 says nothing about β or $\pi \Rightarrow E[\nu_t \nu_{t-h}] \neq 0$.

Adendum

- Future Questions:
  - Longer Expansion: must make $R(z_t, \pi)$ and $R(z_t, \pi)^2$ disappear
  - What about other models (e.g. ARMA)?

Return to Introduction

Return to White Noise Test Recap

Adendum to Taylor Expansion method

- Under non-identification, $\beta_n = 0$
- Under weak identification $\beta_n \to 0$, $\sqrt{n}\beta_n \to b < \infty$

$$
\begin{aligned}
\text{Under } H_0, \quad & \sqrt{n}E_{\gamma_n}[\nu_t \nu_{t-h}] \\
&= \sqrt{n}E[\varepsilon_t \varepsilon_{t-h}] + \sqrt{n}E_{\gamma_n}[\beta_n y_{t-1}\varepsilon_{t-h}R(z_t,\pi)] \\
&\quad + \sqrt{n}E_{\gamma_n}[\beta_n y_{t-1-h}\varepsilon_t R(z_{t-h},\pi)] \\
&\quad + \sqrt{n}E_{\gamma_n}[\beta_n y_{t-1} y_{t-1-h} R(z_t,\pi)\beta_n R(z_{t-h},\pi)] \\
&\to bE[y_{t-1}\varepsilon_{t-h}R(z_t,\pi)] + bE[y_{t-1-h}\varepsilon_t R(z_{t-h},\pi)]
\end{aligned}
$$

- Potential for new robust test:
  - Use Taylor expansion under weak identification
  - Use original model under strong identification
  - eliminate the need to sup over nuisance parameters in practice
  - but still not clear how to use this method for other models...

Return to Taylor Expansion Page