A bootstrap strategy for optimal bandwidth selection in Local Whittle estimation

Josu Arteche & Jesus Orbe
University of the Basque Country UPV/EHU
eman ta zabal zazu
UPV EHU

9th International Conference on Computational and Financial Econometrics (CFE 2015)
London, December 2015

Motivation

The Local Whittle estimator (LW) of the memory parameter has a very well developed asymptotic theory (Robinson 1995, Velasco 1999, Phillips and Shimotsu 2004, Shao and Wu 2007) .

Motivation

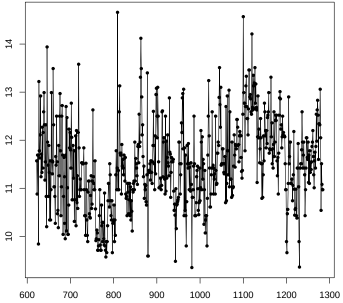
The Local Whittle estimator (LW) of the memory parameter has a very well developed asymptotic theory (Robinson 1995, Velasco 1999, Phillips and Shimotsu 2004, Shao and Wu 2007) .

Some empirical issues unsolved: different bandwidths may lead to completely different conclusions about the characteristics of the series.

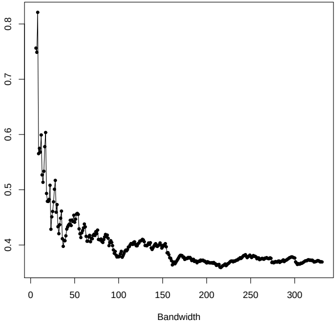
Motivation: Examples

Figure : Nile river annual minimum

(a) Nile minimum levels 622-1284



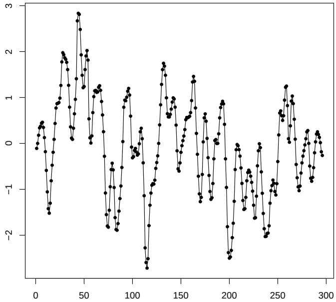
(b) LW estimates of d



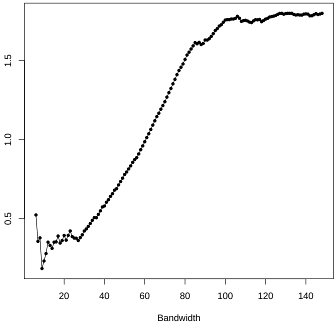
Motivation: Examples

Figure : Input gas rate

(a) Input gas rate



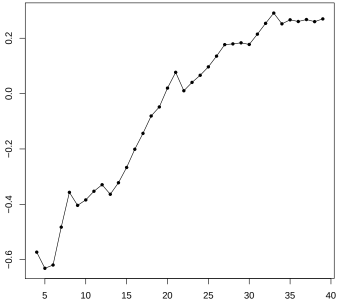
(b) LW estimates of d



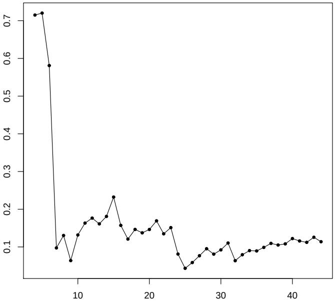
Motivation: Examples

Figure : First differenced GNP p.c. and Bond Yield (Nelson and Plosser)

(a) LW estimates GNP pc



(b) LW estimates Bond Yield



Motivation

Main question: Which bandwidth should we use?

Motivation

Main question: Which bandwidth should we use?

Answer: The bandwidth that minimizes the MSE of the LW estimator (optimal)

Local Whittle (LW) estimation

Applied to long memory processes with (pseudo)spectral density

$$f(\lambda) = |\lambda|^{-2d_0} g(\lambda) \quad \lambda \in [-\pi, \pi]$$

- d_0 is the memory parameter (parameter of interest).
- $g(\lambda)$ is positive and bounded and,

  $$g(\lambda) = g(0) + \Delta(\lambda)\,, \quad |\Delta(\lambda)| \leq C_1 |\lambda|^{\alpha}$$

  for $C_1 > 0$ and spectral smoothness parameter $\alpha > 0$ ($\alpha = 2$ in ARFIMA).
- α determines the rate of increase of the optimal bandwidth.

Local Whittle (LW) estimation

The LW estimator of d_0 is $\hat{d} = argminR(d)$ for

$$R(d) = \log\left(\frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d} I_j\right) - \frac{2d}{m}\sum_{j=1}^{m}\log\lambda_j$$

where I_j is the periodogram of x_t, $t = 1, 2, ..., n$, at Fourier frequency $\lambda_j = 2\pi j/n$

$$I_j = I(\lambda_j) = \frac{1}{2\pi n}\left|\sum_{t=1}^{n} x_t \exp(-i\lambda_j t)\right|^2$$

and m is the bandwidth (number of frequencies used in the estimation).

LW estimation: optimal bandwidth

Writing $g(\lambda) = g(0)(1 + \tau\lambda^{\alpha} + o(\lambda^{\alpha}))$ the optimal (in a asymptotic MSE sense) bandwidth for $d < 0.75$ is

$$\boxed{m_{opt} = K(\alpha, \tau) n^{2\alpha/(1+2\alpha)}}$$

for

$$\boxed{K(\alpha, \tau) = \left\{ \frac{(\alpha+1)^4}{2\alpha^3\tau^2(2\pi)^{2\alpha}} \right\}^{1/(1+2\alpha)}}$$

Feasible strategies for optimal bandwidth selection

Valid in stationary and invertible contexts:

- **Plug-in bandwidth selection (Henry and Robinson, 1996)**: Imposes $\alpha = 2$ and estimate τ (and thus $K(\alpha, \tau)$).

Feasible strategies for optimal bandwidth selection

Valid in stationary and invertible contexts:

- **Plug-in bandwidth selection (Henry and Robinson, 1996)**: Imposes $\alpha = 2$ and estimate τ (and thus $K(\alpha, \tau)$).
- **Adaptive LW estimation (Andrews and Sun, 2004)**: adapts to un unknown spectral smoothness α but does not estimate the constant of proportionality $K(\alpha, \tau)$.

Feasible strategies for optimal bandwidth selection

Valid in stationary and invertible contexts:

- **Plug-in bandwidth selection (Henry and Robinson, 1996)**: Imposes $\alpha = 2$ and estimate τ (and thus $K(\alpha, \tau)$).
- **Adaptive LW estimation (Andrews and Sun, 2004)**: adapts to un unknown spectral smoothness α but does not estimate the constant of proportionality $K(\alpha, \tau)$.

Our proposal for optimal bandwidth selection adjusts the selection of bandwidth to unknown α and τ and is valid for stationary, non-stationary and non-invertible series.

Bootstrap for optimal bandwidth selection

Basic idea:

Bootstrap for optimal bandwidth selection

Basic idea:

Optimal bandwidth selected by minimising some approximation of the exact MSE of the LW estimator.

Bootstrap for optimal bandwidth selection

Basic idea:

Optimal bandwidth selected by minimising some approximation of the exact MSE of the LW estimator.

Main instrument:

The bootstrap approximation of the distribution of the LW estimator proposed by Arteche and Orbe (2015).

Bootstrap for optimal bandwidth selection

Basic idea:

> Optimal bandwidth selected by minimising some approximation of the exact MSE of the LW estimator.

Main instrument:

> The bootstrap approximation of the distribution of the LW estimator proposed by Arteche and Orbe (2015).

Rule:

> The optimal bandwidth is then selected as the value that minimises the bootstrap MSE, which is a fair approximation to the exact MSE.

Bootstrap for optimal bandwidth selection

Based on resampling a standardised periodogram. Two options:

Bootstrap for optimal bandwidth selection

Based on resampling a standardised periodogram. Two options:

- *Global bootstrap on Studentised periodogram:* $\boxed{\hat{v}_j^{(0)} = I_j/\hat{f}_j}$.

  $\hat{f}_j$ is the estimator in Hidalgo and Yajima (2002), which is consistent at every Fourier frequency (Arteche, 2015):

$$\hat{f}_j = \hat{f}(\lambda_j) = \frac{|\lambda_j|^{-2\hat{d}}}{2m^* + \mathbf{1}_{j>m^*}} \sum_{k=-m^*, \neq -j}^{m^*} |\lambda_j + \lambda_k|^{2\hat{d}} I(\lambda_j + \lambda_k)$$

  $\hat{v}_j^{(0)}$ is resampled over the whole band of Fourier frequencies in a global Efron's bootstrap strategy.

Bootstrap for optimal bandwidth selection

- *Local bootstrap on locally standardised periodogram:* $\boxed{\hat{v}_j^{(1)} = I_j \lambda_j^{2\hat{d}}}$. Avoids estimation of the sdf. Potential weak dependent components remain visible in $\hat{v}_j^{(1)}$ and a global resampling scheme is not adequate. The local bootstrap proposed by Paparoditis and Politis (1999) is used instead. The resampling width k_n needs to be selected.

Bootstrap for optimal bandwidth selection

1. Obtain $\hat{v}_j^{(i)}$, $i = 0, 1$, with a bandwidth m_1 for the LW estimate, $\hat{d}_1$, and m^* for $\hat{f}_j$ (only needed for $\hat{v}_j^{(0)}$).
2. Select a resampling width $k_n \in \mathcal{N}$, $k_n \leq [n/2]$ (only for $\hat{v}_j^{(1)}$).
3. For $\overline{m} < [n/2]$, generate B bootstrap series $\hat{v}_{bj}^{*(i)}$ $b = 1, 2, ..., B$, $j = 1, ..., \overline{m}$ by resampling $\hat{v}_j^{(0)}$ in $(0, \pi]$ or $\hat{v}_j^{(1)}$ in $[\lambda_j \pm \lambda_{k_n}]$.
4. Generate B bootstrap samples for the periodogram $I_{bj}^{*(1)} = \lambda_j^{-2\hat{d}_1} \hat{v}_{bj}^{*(1)}$, $I_{bj}^{*(0)} = \hat{f}_j \hat{v}_{bj}^{*(0)}$ for $b = 1, 2, ..., B$ and $j = 1, ..., \overline{m}$.
5. Obtain B bootstrap LW estimates for each bandwidth $m \in [\underline{m}, \overline{m}]$ $\hat{d}_b^{*(i)}(m)$, $b = 1, ..., B$ by minimising $R(d)$ with I_j replaced by $I_{bj}^{*(i)}$.

Bootstrap for optimal bandwidth selection

The optimal bandwidth is selected in $[\underline{m}, \overline{m}]$ as follows.

6. Calculate the bootstrap MSE for every $m \in [\underline{m}, \overline{m}]$,

$$MSE^*(m) = \frac{1}{B}\sum_{b=1}^{B}(\hat{d}_b^{*(i)}(m) - \hat{d}_1)^2$$

7. Chose $\hat{m}_1$ such that $MSE^*(\hat{m}_1) \leq MSE^*(m)$ for all $m \in [\underline{m}, \overline{m}]$.

Bootstrap for optimal bandwidth selection

The optimal bandwidth is selected in $[\underline{m}, \overline{m}]$ as follows.

6. Calculate the bootstrap MSE for every $m \in [\underline{m}, \overline{m}]$,

$$MSE^*(m) = \frac{1}{B}\sum_{b=1}^{B}(\hat{d}_b^{*(i)}(m) - \hat{d}_1)^2$$

7. Chose $\hat{m}_1$ such that $MSE^*(\hat{m}_1) \leq MSE^*(m)$ for all $m \in [\underline{m}, \overline{m}]$.

8. Replace m_1 with $\hat{m}_1$ in step 1 and iterate until

$$\frac{MSE^*(\hat{m}_i) - MSE^*(\hat{m}_{i-1})}{MSE^*(\hat{m}_{i-1})} > \delta$$

for some small (in absolute value) $\delta < 0$ stopping criterion.

$\hat{m}_{i-1}$ is the estimated optimal bandwidth.

Bootstrap bandwidth selection: Remarks

Remark 1: selection of the starting bandwidth m_1 is not important because of the updating (advise: start with a low m_1).

Bootstrap bandwidth selection: Remarks

Remark 1: selection of the starting bandwidth m_1 is not important because of the updating (advise: start with a low m_1).

Remark 2: $[\underline{m}, \overline{m}]$ can be selected as large as desired but can be shortened based on the characteristics of the series (for example after visual inspection of $\hat{v}_j^{(1)}$).

Bootstrap bandwidth selection: Remarks

Remark 1: selection of the starting bandwidth m_1 is not important because of the updating (advise: start with a low m_1).

Remark 2: $[\underline{m}, \overline{m}]$ can be selected as large as desired but can be shortened based on the characteristics of the series (for example after visual inspection of $\hat{v}_j^{(1)}$).

Remark 3: The resampling width k_n for the local bootstrap or m^* for the estimation of the sdf can be chosen accordingly to the form of $\hat{v}_j^{(1)}$: the higher their structure (more different from the periodogram of a white noise) the lower k_n and m^* should be chosen.

Monte Carlo: Models considered

1000 replications of series of length 512 in five different models:

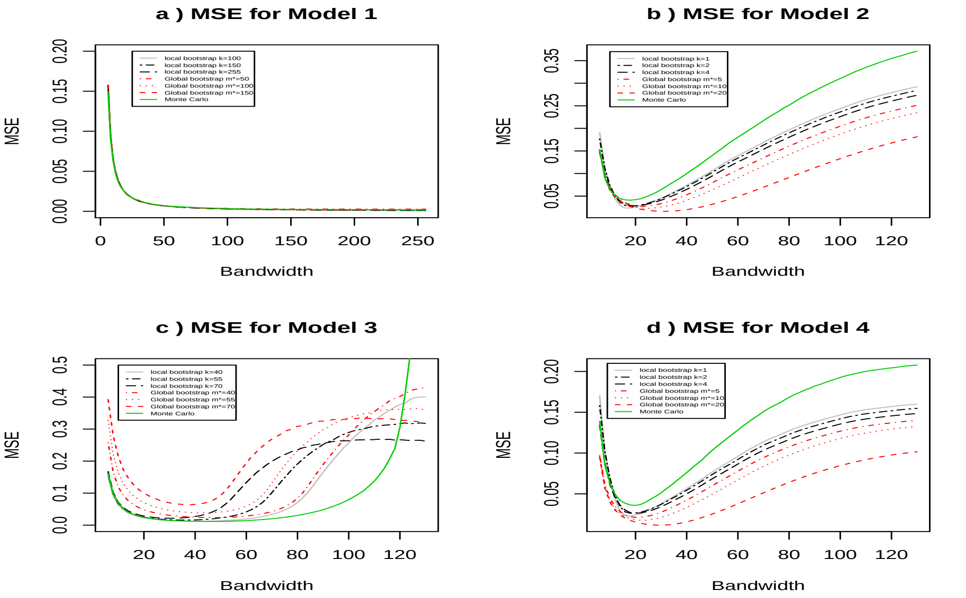
- Model 1: $(1 - 0.1L)(1 - L)^{0.4} x_t = u_t$
- Model 2: $(1 - 0.8L)(1 - L)^{0.4} x_t = u_t$
- Model 3: $(1 - 0.1L + 0.9L^2)(1 - L)^{0.4} x_t = u_t$
- Model 4: $(1 - 0.8L)(1 - L)^{0.8} x_t = u_t$
- Model 5: $(1 - 0.8L)(1 - L)^{-0.7} x_t = u_t$

for u_t standard normal.

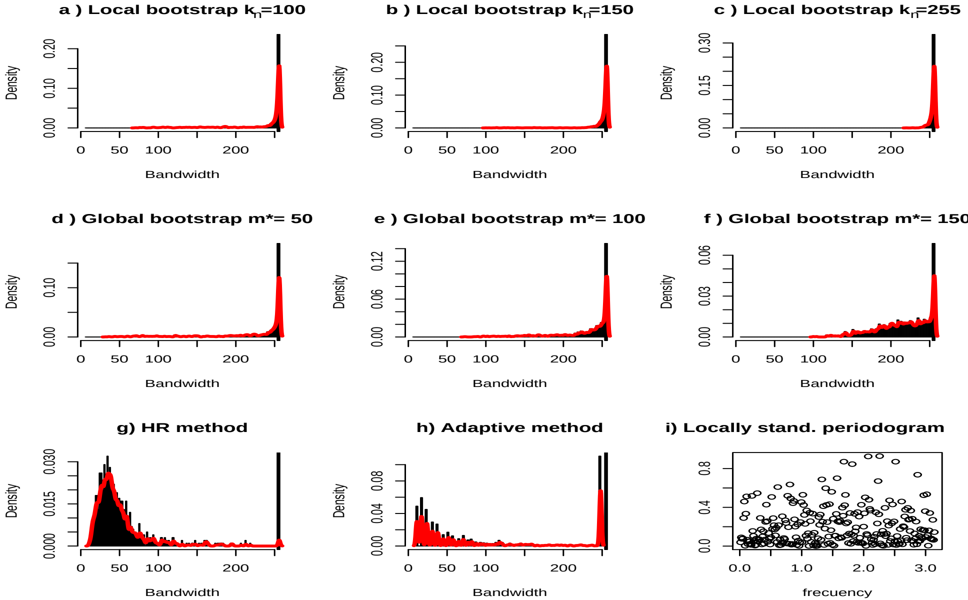
Monte Carlo: design of the bootstrap

- $m_1 = 15$.
- $\underline{m} = 6$ for every model and $\overline{m} = [n/2]$ for Model 1 and $\overline{m} = 130$ for the rest of models.
- $m^* = 50, 100, 150$ and $k_n = 100, 150, 255$ for Model 1,
  $m^* = 5, 10, 20$ and $k_n = 1, 2, 4$ for Models 2,4,5 and
  $m^* = 40, 55, 70$ and $k_n = 40, 55, 70$ for Model 3.
- $B = 200$.

Monte Carlo: exact and bootstrap MSE



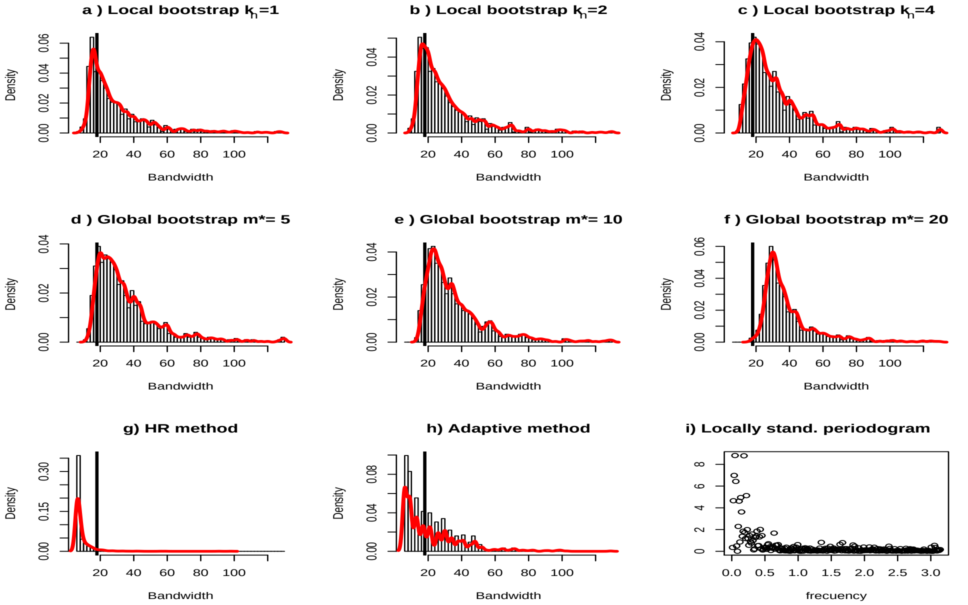
Monte Carlo: histograms and pdf of estimated optimal bandwidths in Model 1



Monte Carlo: results for Model 1

Selection strategy	MSE	m-opt (mean)	m-opt (sd)	m-opt (med)
Monte Carlo	0.00165	255		
Henry-Robinson	0.01683	49.3	35.1	40
Adaptive	0.02806	102.7	92.1	58
local bootstrap k_n=100	0.00262	233.8	42.9	254
local bootstrap k_n=150	0.00186	249.0	21.3	256
local bootstrap k_n=255	0.00168	254.1	3.6	256
global bootstrap m*=50	0.00362	221.0	56.3	250
global bootstrap m*=100	0.00274	227.8	39.6	246
global bootstrap m*=150	0.00245	216.2	33.6	222

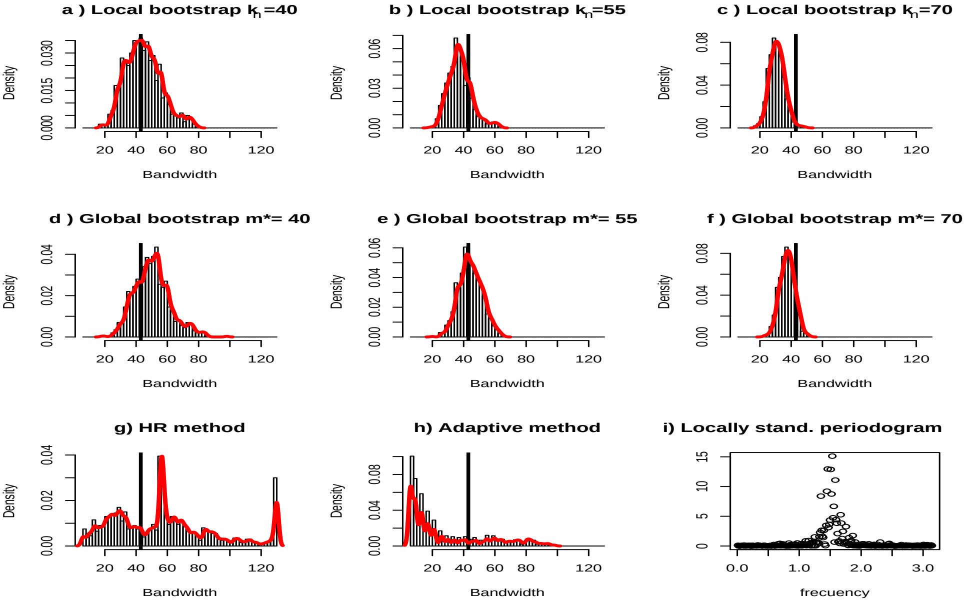
Monte Carlo: histograms and pdf of estimated optimal bandwidths in Model 2



Monte Carlo: results for Model 2

Selection strategy	MSE	m-opt (mean)	m-opt (sd)	m-opt (med)
Monte Carlo	0.04101	18		
Henry-Robinson	0.09654	8.9	6.4	7
Adaptive	0.12989	22.3	18.3	17
local bootstrap k_n=1	0.05296	30.1	19.2	24
local bootstrap k_n=2	0.05542	31.5	19.6	26
local bootstrap k_n=4	0.06010	32.9	19.8	26
global bootstrap m*=5	0.06597	36.0	19.5	30
global bootstrap m*=10	0.06755	36.4	19.0	30
global bootstrap m*=20	0.07986	39.7	17.1	34

Monte Carlo: histograms and pdf of estimated optimal bandwidths in Model 3



Monte Carlo: results for Model 3

Selection strategy	MSE	m-opt (mean)	m-opt (sd)	m-opt (med)
Monte Carlo	0.01148	45		
Henry-Robinson	0.10954	60.4	33.2	57
Adaptive	0.10943	29.3	26.4	17
local bootstrap k_n=40	0.01346	45.5	11.2	44
local bootstrap k_n=55	0.01293	38.1	7.5	38
local bootstrap k_n=70	0.01421	31.2	4.8	30
global bootstrap m*=40	0.01446	51.4	11.6	50
global bootstrap m*=55	0.01220	45.0	7.3	44
global bootstrap m*=70	0.01248	38.2	4.8	38

Monte Carlo: results for Model 4

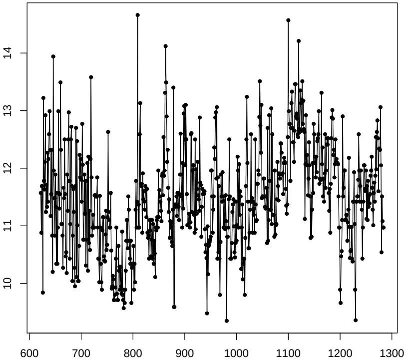
Selection strategy	MSE(LW)	m-opt (mean)	m-opt (sd)	m-opt (med)
Monte Carlo	0.03588	20		
Henry-Robinson	0.12031	8.1	6.7	6
Adaptive	0.11562	29.3	30.5	17
local bootstrap k_n=1	0.04714	37.9	30.9	26
local bootstrap k_n=2	0.05036	40.3	32.2	28
local bootstrap k_n=4	0.05202	42.5	33.2	30
global bootstrap m*=5	0.05441	41.2	29.9	32
global bootstrap m*=10	0.05560	42.4	31.2	32
global bootstrap m*=20	0.06130	46.3	31.5	34

Monte Carlo: results for Model 5

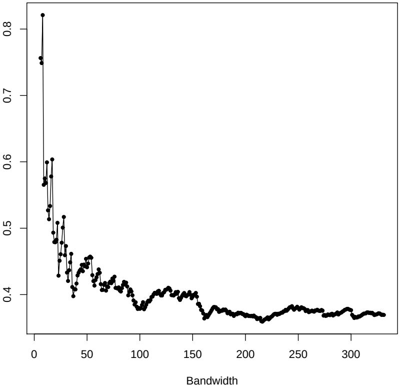
Selection strategy	MSE(LW)	m-opt (mean)	m-opt (sd)	m-opt (med)
Monte Carlo	0.05238	18		
Henry-Robinson	0.22071	67.1	25.7	60
Adaptive	0.14670	22.6	17.8	17
local bootstrap k_n=1	0.06680	32.2	19.6	28
local bootstrap k_n=2	0.07104	33.7	19.7	28
local bootstrap k_n=4	0.07487	34.6	19.7	30
global bootstrap m*=5	0.08550	37.9	19.1	34
global bootstrap m*=10	0.08724	38.3	18.6	34
global bootstrap m*=20	0.09696	40.3	16.2	36

Nile river annual minimum

a) Nile minimum levels 622-1284

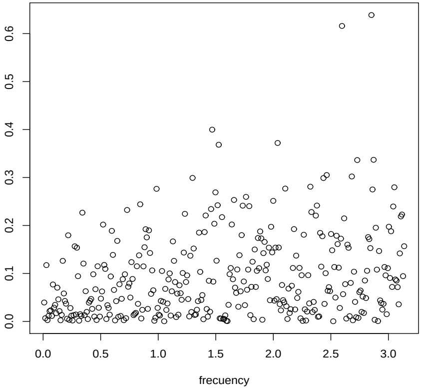


b) LW estimates

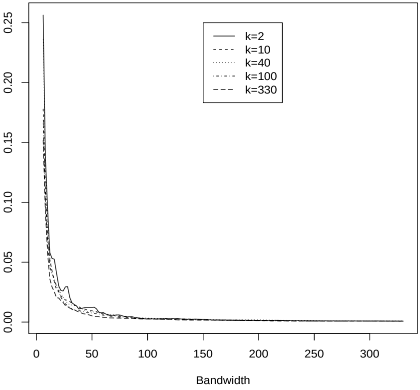


Nile river annual minimum

a) Locally standardised periodogram



b) Bootstrap MSEs

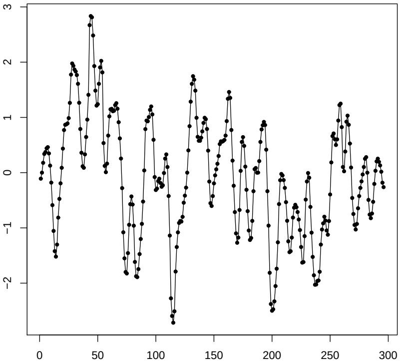


Nile river annual minimum

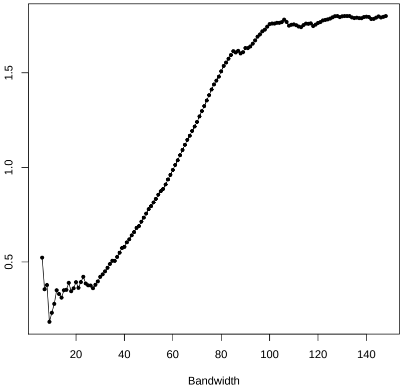
m-selection strategy	m-opt.	$\hat{d}$	$\hat{sd}$	$CI_{0.95}$	
Henry-Robinson	53	0.457	0.078	0.289	0.605
Adaptive	320	0.373	0.029	0.308	0.415
local bootstrap k_n=2	330	0.370	0.028	0.312	0.419
local bootstrap k_n=10	330	0.370	0.028	0.307	0.421
local bootstrap k_n=40	326	0.372	0.029	0.300	0.417
local bootstrap k_n=100	328	0.371	0.029	0.303	0.415
local bootstrap k_n=330	330	0.370	0.028	0.314	0.419

Input gas rate

a) Input gas rate

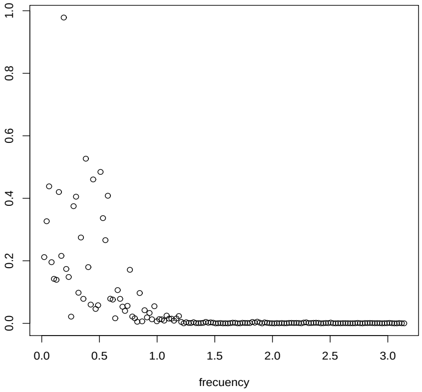


b) LW estimates

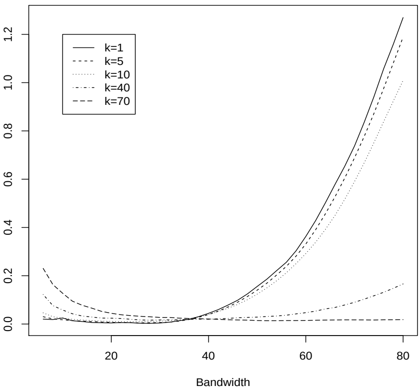


Input gas rate

a) Locally standardised periodogram



b) Bootstrap MSEs



Input gas rate

m-selection strategy	m-opt.	$\hat{d}$	$\hat{sd}$	$CI_{0.95}$	
Henry-Robinson	6	0.523	0.337	0.006	0.749
Adaptive	34	0.489	0.101	0.351	0.659
local bootstrap k_n=1	28	0.379	0.114	0.285	0.494
local bootstrap k_n=5	26	0.375	0.119	0.258	0.557
local bootstrap k_n=10	24	0.386	0.126	0.240	0.602
local bootstrap k_n=40	30	0.421	0.109	0.243	0.702
local bootstrap k_n=70	52	0.814	0.079	0.610	1.083

Macroeconomic series in Nelson and Plosser (1982)

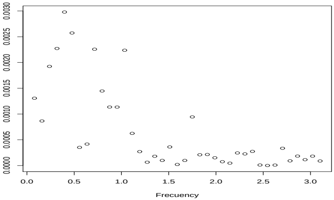
	k_n	m-opt	Local $\hat{d}$	Boots. $CI_{0.95}$	
CPI	12	46	0.435	0.290	0.610
Employment	9	11	-0.338	-0.593	-0.034
GNP deflator	39	49	0.417	0.257	0.562
GNP per capita	2	12	-0.329	-0.479	-0.162
Ind. production	12	17	-0.400	-0.623	-0.174
Bond Yield	35	44	0.114	-0.101	0.284
Money stock	2	21	0.321	-0.054	0.568
Nominal GNP	7	21	0.273	0.014	0.579
Real wages	35	43	0.172	0.047	0.305
Real GNP	2	12	-0.327	-0.551	-0.161
S&P500	11	30	-0.087	-0.268	0.112
Unemployment	2	32	0.604	0.395	0.774
Velocity	47	59	0.068	-0.061	0.188
Wages	8	36	0.393	0.131	0.641

Note: All the series are in first differences except unemployment.

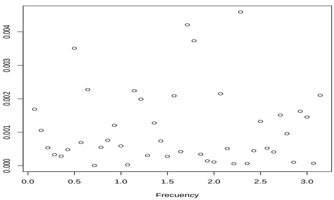
k_n selected among values $2, [n/10]$ and $[2n/5]$.

Macroeconomic series in Nelson and Plosser (1982)

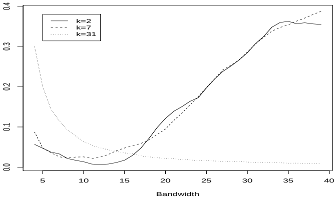
a) $\hat{v}_j^{(1)}$ in GNP percapita



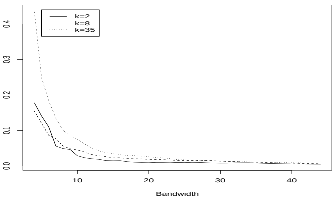
b) $\hat{v}_j^{(1)}$ in Bond Yield



c) Local bootstrap MSE



d) Local bootstrap MSE



Conclusion

- Stationarity, non-stationarity and even no mean reversion can be concluded for the same series and with the same statistical tools by just modifying the bandwidth selected.

Conclusion

- Stationarity, non-stationarity and even no mean reversion can be concluded for the same series and with the same statistical tools by just modifying the bandwidth selected.
- A simple procedure for optimal bandwidth selection is proposed based on minimising a bootstrap approximation of the MSE.

Conclusion

- Stationarity, non-stationarity and even no mean reversion can be concluded for the same series and with the same statistical tools by just modifying the bandwidth selected.
- A simple procedure for optimal bandwidth selection is proposed based on minimising a bootstrap approximation of the MSE.
- Valid for stationary, non-stationary and non-invertible series.

Conclusion

- Stationarity, non-stationarity and even no mean reversion can be concluded for the same series and with the same statistical tools by just modifying the bandwidth selected.
- A simple procedure for optimal bandwidth selection is proposed based on minimising a bootstrap approximation of the MSE.
- Valid for stationary, non-stationary and non-invertible series.
- Much better performance than the plug-in method of Henry and Robinson (1996) and the adaptive criteria of Andrews and Sun (2004).

Conclusion

- Stationarity, non-stationarity and even no mean reversion can be concluded for the same series and with the same statistical tools by just modifying the bandwidth selected.
- A simple procedure for optimal bandwidth selection is proposed based on minimising a bootstrap approximation of the MSE.
- Valid for stationary, non-stationary and non-invertible series.
- Much better performance than the plug-in method of Henry and Robinson (1996) and the adaptive criteria of Andrews and Sun (2004).
- All require prior selection of some quantities. But our proposal is quite robust to it as long as sensible values are used, which in any case can be selected based on the series at hand $\Rightarrow$ data-driven.


Thank You