ON THE BAND-SPECTRAL ESTIMATION OF BUSINESS CYCLE MODELS

Nikolay Iskrev

IAAE 2024
June 26, 2024

What?

What?

- investigate the properties and performance of band-spectral estimators (**BSE**) applied to BC models

What?

- investigate the properties and performance of band-spectral estimators (**BSE**) applied to BC models
- Monte Carlo study:

What?

- investigate the properties and performance of band-spectral estimators (**BSE**) applied to BC models
- Monte Carlo study:
  - band-spectral estimator

What?

- investigate the properties and performance of band-spectral estimators (**BSE**) applied to BC models
- Monte Carlo study:
    - band-spectral estimator
        - frequency domain, Whittle likelihood

What?

- investigate the properties and performance of band-spectral estimators (**BSE**) applied to BC models
- Monte Carlo study:
  - band-spectral estimator
    - frequency domain, Whittle likelihood (BC frequencies)
  - full-spectrum (Whittle) estimator
  - exact likelihood (time domain) estimator

What?

- investigate the properties and performance of band-spectral estimators (**BSE**) applied to BC models
- Monte Carlo study:
    - band-spectral estimator
        - frequency domain, Whittle likelihood (BC frequencies)
    - full-spectrum (Whittle) estimator
    - exact likelihood (time domain) estimator
- assess approximation distortions and information loss

Why band-spectral estimation?

Why band-spectral estimation?

- estimate models which are a priori known to be unable to represent some frequencies

Why band-spectral estimation?

"***Quantifying confidence***" by Angeletos, Collard, and Dellas (2018)

The model described above – like other business-cycle models – cater to ***business-cycle phenomena*** *and therefore omit shocks and mechanisms that may account for medium- to long-run phenomena, such as trends in demographics and labor-market participation, structural transformation, regime changes in productivity growth or inflation, and so on.*

Why band-spectral estimation?

"***Quantifying confidence***" by Angeletos, Collard, and Dellas (2018)

The model described above – like other business-cycle models – cater to ***business-cycle phenomena*** *and therefore omit shocks and mechanisms that may account for medium- to long-run phenomena, such as trends in demographics and labor-market participation, structural transformation, regime changes in productivity growth or inflation, and so on.*

- ACD estimate their model with BSE using BC frequencies only

3 Questions

1 Time-domain vs frequency-domain **full info** estimators. Which is better?

3 Questions

1. Time-domain vs frequency-domain **full info** estimators. Which is better?

2. The BSE (BC) estimator uses less information than full info estimators. How much less? Which parameters are more/less affected?

3 Questions

1. Time-domain vs frequency-domain **full info** estimators. Which is better?

2. The BSE (BC) estimator uses less information than full info estimators. How much less? Which parameters are more/less affected?

3. Can we predict the answer to 2 without MC simulations?

Results

1. Time-domain vs frequency-domain **full info** estimators. Which is better?

Results

1. Time-domain vs frequency-domain **full info** estimators. Which is better?
   - no (good) reason to use full info FD (all freqs) instead of full info TD estimator

Results

1. Time-domain vs frequency-domain **full info** estimators. Which is better?
   - no (good) reason to use full info FD (all freqs) instead of full info TD estimator

   Note: full info, i.e. when misspecification is not a concern.

Results

1. Time-domain vs frequency-domain **full info** estimators. Which is better?
   - no (good) reason to use full info FD (all freqs) instead of full info TD estimator
2. The BSE (BC) uses less information than full info estimators. How much less? Which parameters are more/less affected?

Results

1. Time-domain vs frequency-domain **full info** estimators. Which is better?
   - no (good) reason to use full info FD (all freqs) instead of full info TD estimator
2. The BSE (BC) uses less information than full info estimators. How much less? Which parameters are more/less affected?
   - significantly less info in BC frequencies for **all** parameters

Results

1. Time-domain vs frequency-domain **full info** estimators. Which is better?
   - no (good) reason to use full info FD (all freqs) instead of full info TD estimator
2. The BSE (BC) uses less information than full info estimators. How much less? Which parameters are more/less affected?
   - significantly less info in BC frequencies for **all** parameters
3. Can we predict the answer to 2 without MC simulations?

Results

1. Time-domain vs frequency-domain **full info** estimators. Which is better?
    - no (good) reason to use full info FD (all freqs) instead of full info TD estimator
2. The BSE (BC) uses less information than full info estimators. How much less? Which parameters are more/less affected?
    - significantly less info in BC frequencies for **all** parameters
3. Can we predict the answer to 2 without MC simulations?
    - we can reliably predict the loss of efficiency

- Likelihood
- Monte Carlo: setup, results
- Conclusion

Gaussian Likelihood Function

Gaussian Likelihood Function

Let $\boldsymbol{y}_t$ be a stationary Gaussian process with zero mean.

$$\boldsymbol{Y}_T = \left(\boldsymbol{y}_1', \boldsymbol{y}_2', \ldots, \boldsymbol{y}_T'\right)' \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}_T(\boldsymbol{\theta})\right)$$

The log-likelihood is

$$\ell(\boldsymbol{\theta}; \boldsymbol{Y}_T) = -\frac{1}{2}\log\det(\boldsymbol{\Sigma}_T(\boldsymbol{\theta})) - \frac{1}{2}\boldsymbol{Y}_T'\boldsymbol{\Sigma}_T^{-1}(\boldsymbol{\theta})\boldsymbol{Y}_T \quad (1)$$

Gaussian Likelihood Function

Whittle approximation: replace $\boldsymbol{\Sigma}_T(\boldsymbol{\theta}) \approx \boldsymbol{\Omega}_T(\boldsymbol{\theta}) = \boldsymbol{F}_T^* \boldsymbol{S}_T(\boldsymbol{\theta}) \boldsymbol{F}_T$

$$\ell(\boldsymbol{\theta}; \boldsymbol{Y}_T) = -\frac{1}{2} \log\det(\boldsymbol{\Sigma}_T(\boldsymbol{\theta})) - \frac{1}{2} \boldsymbol{Y}_T' \boldsymbol{\Sigma}_T^{-1}(\boldsymbol{\theta}) \boldsymbol{Y}_T \quad (1)$$

$$\approx -\frac{1}{2} \log\det(\boldsymbol{S}_T(\boldsymbol{\theta})) - \frac{1}{2} (\boldsymbol{F}_T \boldsymbol{Y}_T)^* \boldsymbol{S}^{-1}(\boldsymbol{\theta}) (\boldsymbol{F}_T \boldsymbol{Y}_T) \quad (2)$$

$$\approx -\frac{1}{2} \sum_{\omega=\omega_1}^{\omega_T} \left\{ \log\det(\boldsymbol{s}(\boldsymbol{\theta}, \omega)) + \tilde{\boldsymbol{y}}(\omega)^* \boldsymbol{s}^{-1}(\boldsymbol{\theta}, \omega) \tilde{\boldsymbol{y}}(\omega) \right\} \quad (3)$$

- $\boldsymbol{F}_T$ is Fourier transform matrix
- $\boldsymbol{S}_T(\boldsymbol{\theta})$ is block-diagonal

Three estimators

- **TD** maximizes (full info, KF)

$$\ell(\boldsymbol{\theta}; \boldsymbol{Y}_T) = -\frac{1}{2}\log\det(\boldsymbol{\Sigma}_T(\boldsymbol{\theta})) - \frac{1}{2}\boldsymbol{Y}_T'\boldsymbol{\Sigma}_T^{-1}(\boldsymbol{\theta})\boldsymbol{Y}_T$$

- **FD** maximizes (full info, Whittle, all freqs)

$$\ell_w(\boldsymbol{\theta}; \boldsymbol{I}_T) = -\frac{1}{2}\sum_{all\ \omega}\left\{\log\det(\boldsymbol{s}(\boldsymbol{\theta},\omega)) + \tilde{\boldsymbol{y}}(\omega)^*\boldsymbol{s}^{-1}(\boldsymbol{\theta},\omega)\tilde{\boldsymbol{y}}(\omega)\right\}$$

- **BC** maximizes (limited info, Whittle, BC freqs - periodicity between 6 and 32 quarters)

$$\ell_w(\boldsymbol{\theta}; \boldsymbol{I}_T^{BC}) = -\frac{1}{2}\sum_{\omega\in\bar{\boldsymbol{\omega}}^{BC}}\left\{\log\det(\boldsymbol{s}(\boldsymbol{\theta},\omega)) + \tilde{\boldsymbol{y}}(\omega)^*\boldsymbol{s}^{-1}(\boldsymbol{\theta},\omega)\tilde{\boldsymbol{y}}(\omega)\right\}$$

MONTE CARLO

DGP

New Keynesian DSGE model from “Quantifying confidence” by Angeletos, Collard, and Dellas (2018)

DGP

New Keynesian DSGE model from "Quantifying confidence" by Angeletos, Collard, and Dellas (2018)

- sticky prices, habit formation in consumption, adjustment costs in investment, monetary policy following a Taylor rule

DGP

New Keynesian DSGE model from “Quantifying confidence” by Angeletos, Collard, and Dellas (2018)

- sticky prices, habit formation in consumption, adjustment costs in investment, monetary policy following a Taylor rule

- 9 shocks: permanent and transitory TFP, permanent and transitory ISP, intertemporal preference, government-spending, monetary policy, news about future TFP, **confidence**

DGP

New Keynesian DSGE model from “Quantifying confidence” by Angeletos, Collard, and Dellas (2018)

- sticky prices, habit formation in consumption, adjustment costs in investment, monetary policy following a Taylor rule
- 9 shocks: permanent and transitory TFP, permanent and transitory ISP, intertemporal preference, government-spending, monetary policy, news about future TFP, **confidence**
- the confidence shock represents perceived bias in the other agents’ expectations about the level of TFP in each period (higher-order beliefs)

DGP

New Keynesian DSGE model from "Quantifying confidence" by Angeletos, Collard, and Dellas (2018)

- sticky prices, habit formation in consumption, adjustment costs in investment, monetary policy following a Taylor rule
- 9 shocks: permanent and transitory TFP, permanent and transitory ISP, intertemporal preference, government-spending, monetary policy, news about future TFP, **confidence**
- the confidence shock represents perceived bias in the other agents' expectations about the level of TFP in each period (higher-order beliefs)
  - leads to waves of optimism (believing that others are optimistic) and pessimism (believing that others are pessimistic) that generate business cycle fluctuations unrelated to fundamentals

DGP

New Keynesian DSGE model from “Quantifying confidence” by Angeletos, Collard, and Dellas (2018)

- 25 estimated parameters
- six observed variables: GDP, consumption, investment, hours worked, inflation, and the federal funds rate
- T=192

DGP

New Keynesian DSGE model from “Quantifying confidence” by Angeletos, Collard, and Dellas (2018)

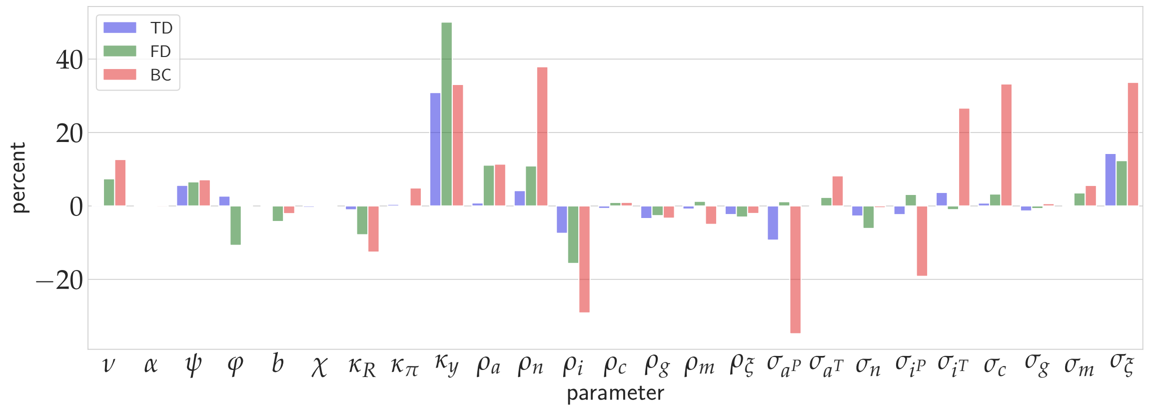
- 25 estimated parameters
- six observed variables: GDP, consumption, investment, hours worked, inflation, and the federal funds rate
- T=192

true DGP for **all** frequencies

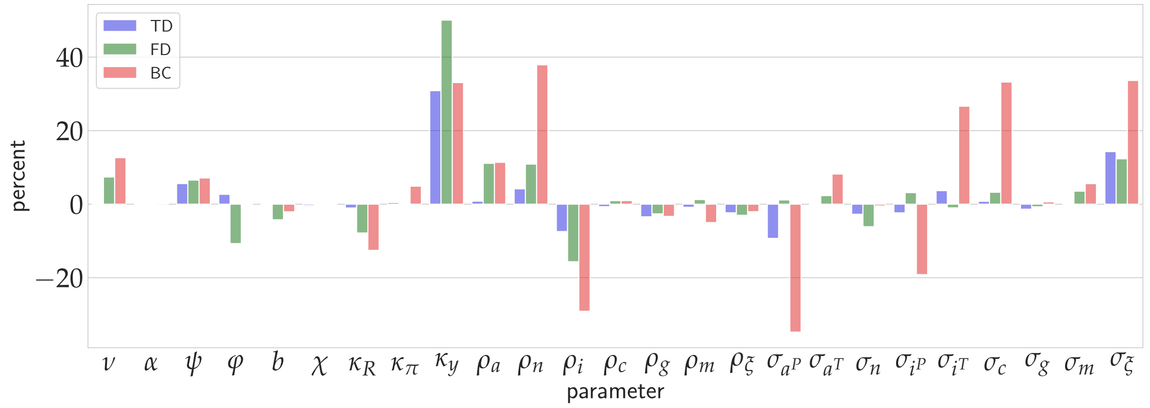
RESULTS

1000 replications

Bias (% of θ)

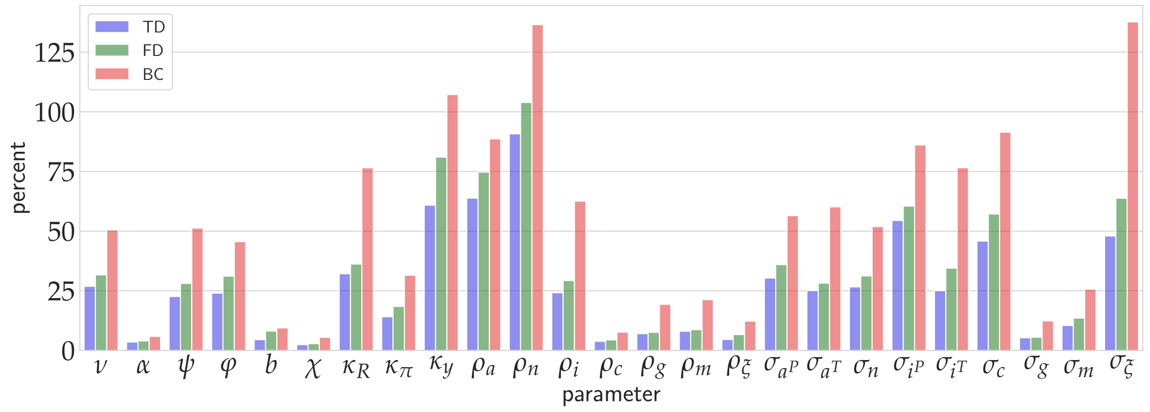


Bias (% of θ)

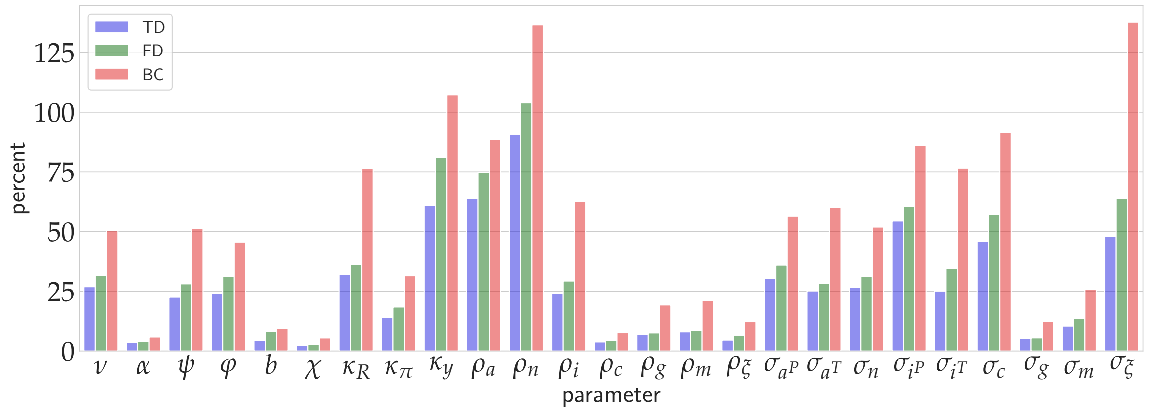


- 4%, 7%, 13%
- rank corr: 0.57 (TD, FD), 0.84 (TD, BC), 0.66 (FD, BC)

Efficiency (std as % of θ)



Efficiency (std as $\%$ of θ)



- 27%, 32%, 53%
- rank corr: 0.98 (TD, FD), 0.95 (TD, BC)

Can the efficiency loss be predicted?

Can the efficiency loss be predicted?

- Cramér-Rao lower bound (CRLB): if $\hat{\boldsymbol{\theta}}$ is unbiased, then

$$\mathrm{std}_{\hat{\theta}_i} \geq \sqrt{\{\mathcal{I}^{-1}(\boldsymbol{\theta})\}_{ii}} = \mathrm{crlb}_{\hat{\theta}_i} \tag{4}$$

  $\mathcal{I}(\boldsymbol{\theta})$ is the FIM

Can the efficiency loss be predicted?

- Cramér-Rao lower bound (CRLB): if $\hat{\boldsymbol{\theta}}$ is unbiased, then

$$\mathrm{std}_{\hat{\theta}_i} \geq \sqrt{\{\mathcal{I}^{-1}(\boldsymbol{\theta})\}_{ii}} = \mathrm{crlb}_{\hat{\theta}_i} \tag{4}$$

  $\mathcal{I}(\boldsymbol{\theta})$ is the FIM

- predicted efficiency loss

$$\frac{\mathrm{crlb}(BC)}{\mathrm{crlb}(TD)} \tag{5}$$

Can the efficiency loss be predicted?

- Cramér-Rao lower bound (CRLB): if $\hat{\boldsymbol{\theta}}$ is unbiased, then

$$\mathrm{std}_{\hat{\theta}_i} \geq \sqrt{\{\mathcal{I}^{-1}(\boldsymbol{\theta})\}_{ii}} = \mathrm{crlb}_{\hat{\theta}_i} \tag{4}$$

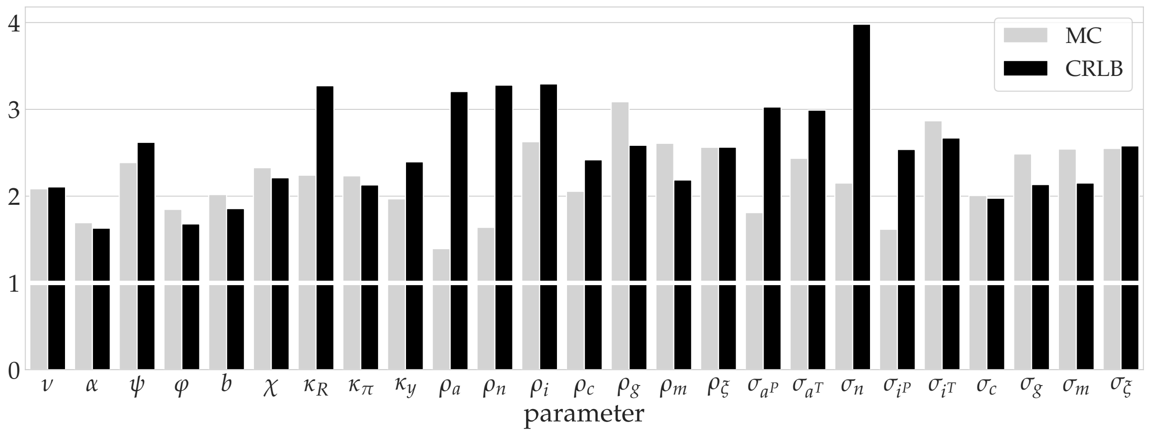
$\mathcal{I}(\boldsymbol{\theta})$ is the FIM

- is

$$\frac{\mathrm{crlb}(BC)}{\mathrm{crlb}(TD)} \approx \frac{\mathrm{std}(BC)}{\mathrm{std}(TD)} \quad ?$$

Efficiency loss

How much less information in the BC frequencies?



What’s wrong with ρ_a (and ρ_n, κ_R, σ_n, ...)?

What's wrong with ρ_a (and ρ_n, κ_R, σ_n, ...)?

- estimated efficiency loss:

$$\frac{\widehat{\mathrm{std}}(BC)}{\widehat{\mathrm{std}}(TD)} = \frac{88.5}{63.5} = 1.4$$

- predicted efficiency loss:

$$\frac{\mathrm{crlb}(BC)}{\mathrm{crlb}(TD)} = \frac{229}{71.4} = 3.2$$

What's wrong with ρ_a (and ρ_n, κ_R, σ_n, ...)?

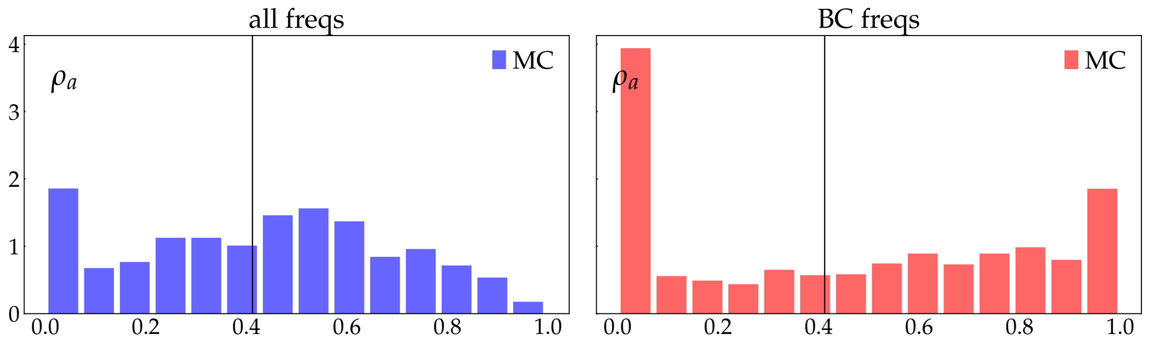
short answer:

- MC overestimates sample information, esp. in BC band.
- Thus, MC underestimates the efficiency loss (loss of sample info).

$$\frac{\mathrm{std}(BC) << \mathrm{crlb}(BC)}{\mathrm{std}(TD) < \mathrm{crlb}(TD)}$$

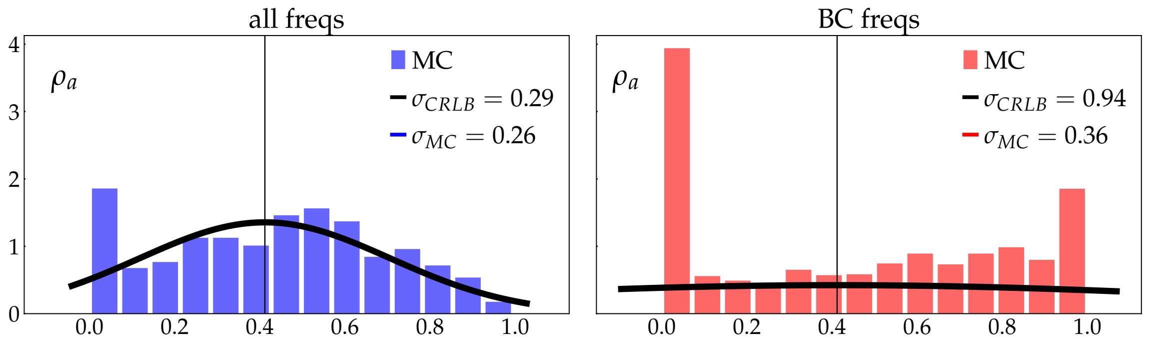
What's wrong with ρ_a (and ρ_n, κ_R, σ_n, ...)?

MC-estimated marginal distribution of $\hat{\rho}_a$



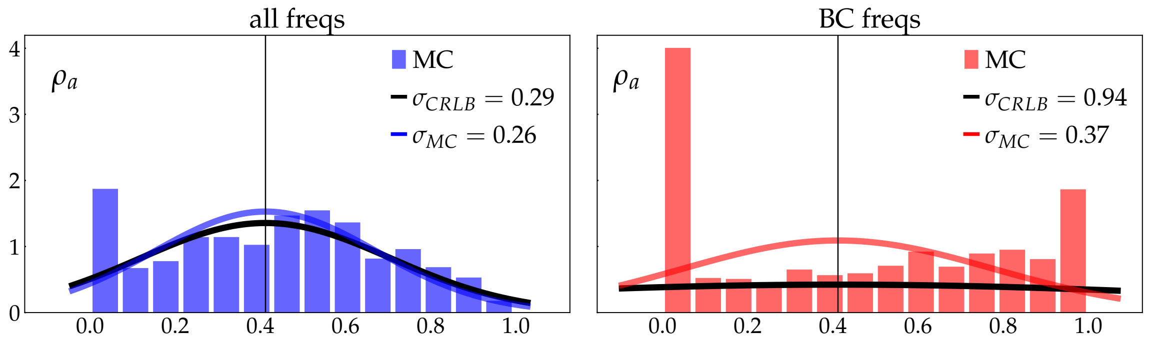
What's wrong with ρ_a (and ρ_n, κ_R, σ_n, ...)?

MC-estimated vs CRLB-predicted marginal distribution of $\hat{\rho}_a$



What's wrong with ρ_a (and ρ_n, κ_R, σ_n, ...)?

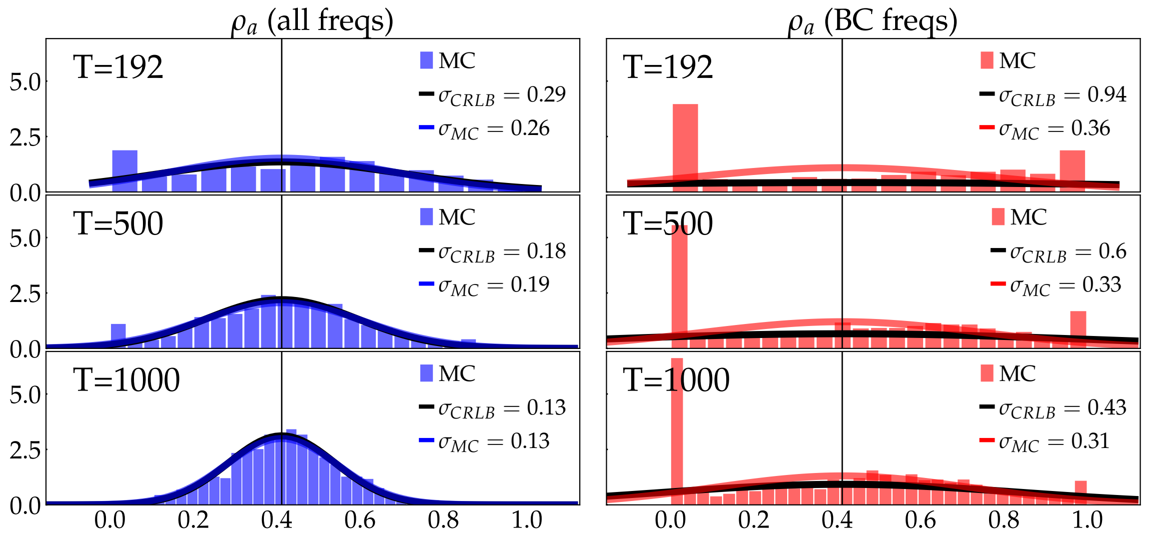
MC-estimated vs CRLB-predicted marginal distribution of $\hat{\rho}_a$



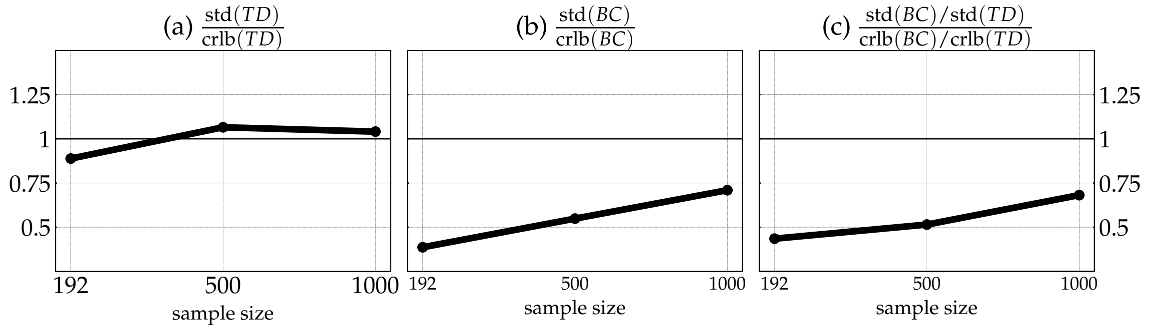
MC underestimates uncertainty (overestimates information **contained in the sample**) due to

- **flat likelihood** (in-sample information deficiency)
- **parameter constraints** (out-of-sample parameter information)

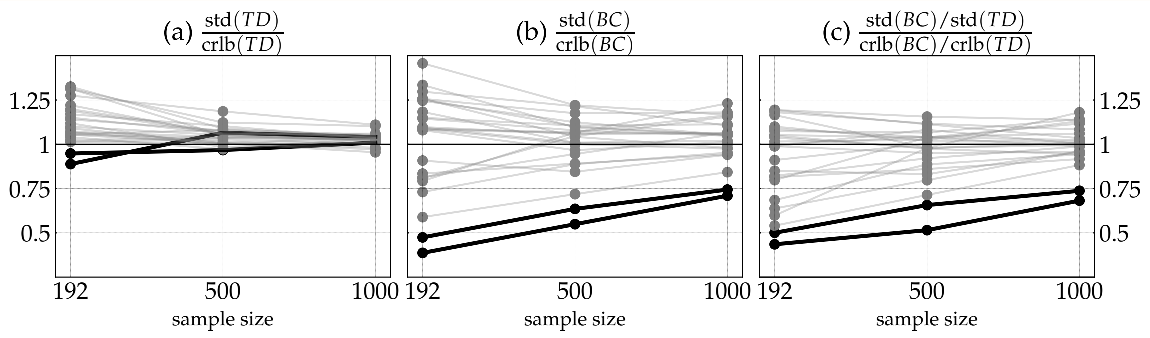
ρ_a: MC vs CRLB as T increases



MC vs CRLB as T increases



MC vs CRLB as T increases



Why compare band-limited to full-info estimation?

Why compare band-limited to full-info estimation?

- full info better, use always

Why compare band-limited to full-info estimation?

- ~~full info better, use always~~

Why compare band-limited to full-info estimation?

1 efficiency benchmark: efficient but **infeasible** estimator (model good, data wrong)

Why compare band-limited to full-info estimation?

1. efficiency benchmark: efficient but **infeasible** estimator (model good, data wrong)
2. measure of contaminated information when misspecification is ignored

Why compare band-limited to full-info estimation?

1. efficiency benchmark: efficient but **infeasible** estimator (model good, data wrong)

2. measure of contaminated information when misspecification is ignored

> *The model described above – **like other business-cycle models** – cater to business-cycle phenomena ... (ACD, 2018)*

Why compare band-limited to full-info estimation?

1. efficiency benchmark: efficient but **infeasible** estimator (model good, data wrong)

2. measure of contaminated information when misspecification is ignored

   *The model described above – **like other business-cycle models** – cater to business-cycle phenomena ... (ACD, 2018)*

   99.9% of BC models are estimated in the time-domain ...

Why compare band-limited to full-info estimation?

1. efficiency benchmark: efficient but **infeasible** estimator (model good, data wrong)
2. measure of contaminated information when misspecification is ignored
3. transparency: where does information about θ_i come from?

Why compare band-limited to full-info estimation?

1. efficiency benchmark: efficient but **infeasible** estimator (model good, data wrong)
2. measure of contaminated information when misspecification is ignored
3. transparency: where does information about θ_i come from?
   - model property

Why compare band-limited to full-info estimation?

1. efficiency benchmark: efficient but **infeasible** estimator (model good, data wrong)
2. measure of contaminated information when misspecification is ignored
3. transparency: where does information about θ_i come from?
   - model property
   - which frequencies are most informative and why

Conclusion

- Full information TD estimator (KF) is superior to the full information FD estimator
  - lower bias and estimation uncertainty

Conclusion

- Full information TD estimator (KF) is superior to the full information FD estimator
  - lower bias and estimation uncertainty
- band-spectral estimation significantly less efficient
  - a lot of info outside the BC freqs for all parameters

Conclusion

- Full information TD estimator (KF) is superior to the full information FD estimator
  - lower bias and estimation uncertainty
- band-spectral estimation significantly less efficient
  - a lot of info outside the BC freqs for all parameters
- FIM analysis is useful to assess the loss of information in band-spectral estimation
  - (relative) CRLBs accurately predict (relative) estimation uncertainty

Some Implications

Some Implications

- Evidence for misspecification
    - estimating a model over different frequency bands leads to different estimates (Qu and Tkachenko (2012), Sala (2015))

Some Implications

- Evidence for misspecification
  - estimating a model over different frequency bands leads to different estimates (Qu and Tkachenko (2012), Sala (2015))
  - might be true even if the model is **not** misspecified
    - bias
    - information deficiencies

Some Implications

- Calibration
  - weakly identified parameters are often calibrated

Some Implications

- Calibration
  - weakly identified parameters are often calibrated
  - may have to calibrate (many) other parameters for band-spectral estimation

Some Implications

- Bayesian estimation and importance of priors

Some Implications

- Bayesian estimation and importance of priors
    - the same prior is much more informative with band-spectral estimation

APPENDIX

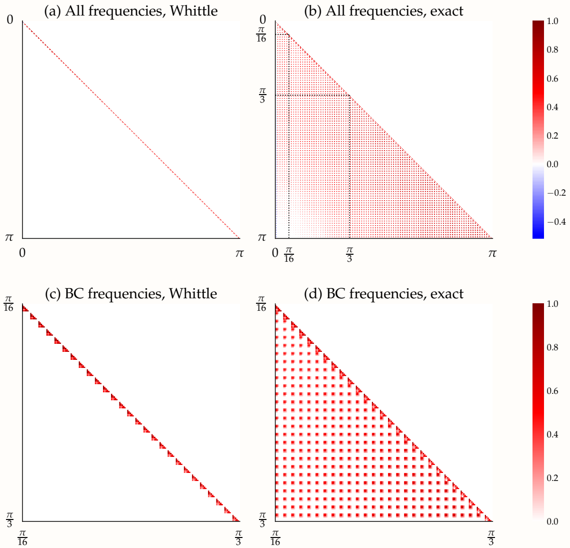
Sala (2015)

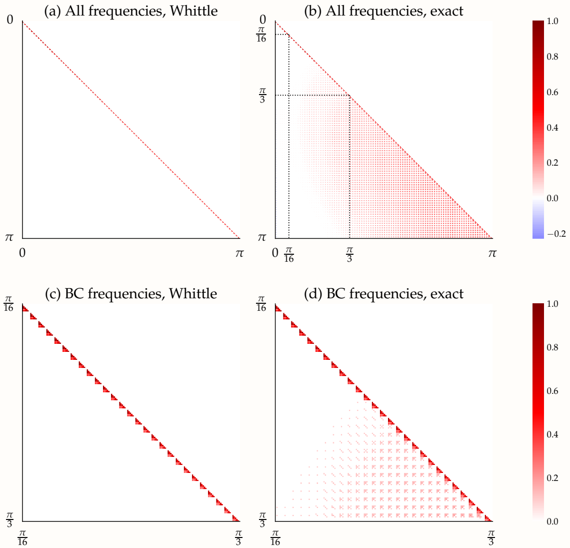
- Monte Carlo experiment with NK DSGE model
- 100 samples $T = 170$
- KF, All, Low-Pass, High-Pass, BC
- "*In sum, the evidence shows that, when using the DSGE model as data-generating process, maximum likelihood in the frequency domain is **equivalent** to maximum likelihood in the time domain, and that the precision of the estimates is still **very good** when estimation is performed on frequency bands*"

$$\boldsymbol{A} = \begin{bmatrix} A_0 & A_1 & \cdots & A_1' \\ A_1' & A_0 & \cdots & A_2' \\ \vdots & \vdots & \ddots & \vdots \\ A_1 & A_2 & \cdots & A_0 \end{bmatrix},$$

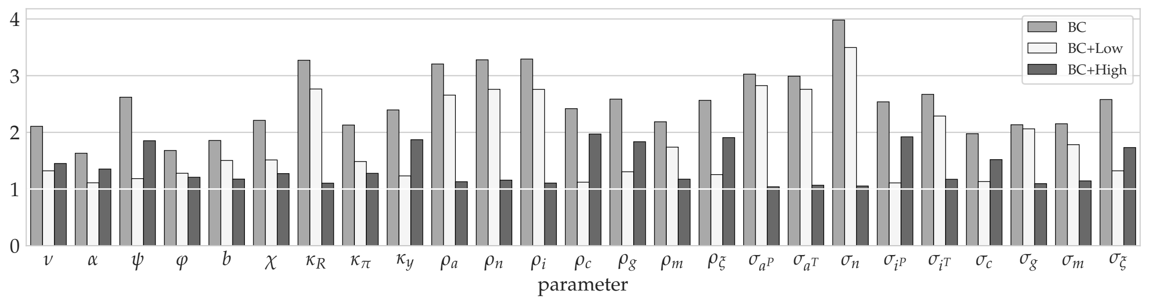
Table: posterior median

ψ	utilization elasticity	0.500
ν	inverse labor supply elasticity	0.282
α	capital share	0.255
φ	investment adjustment costs	3.312
b	habit persistence	0.758
χ	Calvo parameter,	0.732
κ_R	Taylor rule smoothing,	0.198
κ_π	Taylor rule inflation,	2.271
κ_y	Taylor rule output,	0.121
ρ_m	AR mon. policy	0.647
ρ_a	AR transitory TFP component	0.412
ρ_n	AR news	0.224
ρ_i	AR transitory investment-specific technology	0.374
ρ_c	AR preference	0.888
ρ_g	AR government spending	0.786
ρ_ξ	AR confidence	0.833
σ_a^P	std. permanent TFP component	0.406
σ_a^T	std. transitory TFP component	0.347
σ_n	std. news	0.378
σ_i^P	std. permanent investment-specific technology	0.610
σ_i^T	std. transitory investment-specific shocks	5.805
σ_c	std. preference	0.357
σ_g	std. government spending	1.705
σ_ξ	std. confidence	0.613
σ_m	std. mon. policy	0.313


(a) All frequencies, Whittle
(b) All frequencies, exact
(c) BC frequencies, Whittle
(d) BC frequencies, exact
0
π
$\frac{\pi}{16}$
$\frac{\pi}{3}$
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4



(a) All frequencies, Whittle
(b) All frequencies, exact
(c) BC frequencies, Whittle
(d) BC frequencies, exact
0
π
π/16
π/3
1.0
0.8
0.6
0.4
0.2
0.0
−0.2



4
3
2
1
0
BC
BC+Low
BC+High
ν α ψ φ b χ κ_R κ_π κ_y ρ_a ρ_n ρ_i ρ_c ρ_g ρ_m ρ_ξ σ_{a^P} σ_{a^T} σ_n σ_{i^P} σ_{i^T} σ_c σ_g σ_m σ_ξ
parameter