ON THE BAND-SPECTRAL ESTIMATION OF BUSINESS CYCLE MODELS

Nikolay Iskrev

EEA-ESEM 2023
August 31, 2023

What?

What?

investigate the properties and performance of band-spectral estimators of BC/DSGE models

Why?

Why?

- what is band-spectral estimation?

Why?

- what is band-spectral estimation?
  - estimation in the frequency domain based on a subset (band) of frequencies

Why?

- what is band-spectral estimation?
  - estimation in the frequency domain based on a subset (band) of frequencies
- why is this useful?

Why?

- what is band-spectral estimation?
  - estimation in the frequency domain based on a subset (band) of frequencies
- why is this useful?
  - estimate relationships which vary across frequencies

Why?

- what is band-spectral estimation?
  - estimation in the frequency domain based on a subset (band) of frequencies
- why is this useful?
  - estimate relationships which vary across frequencies
  - estimate models which are a priori known to be unable to represent some frequencies

From “***Quantifying confidence***” by Angeletos, Collard, and Dellas (2018)

From “***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.*

From “***Quantifying confidence***” by Angeletos, Collard, and Dellas (2018)

*In a nutshell, there is a risk of **contamination** of the estimates of a model by frequencies that the model was **not designed** to capture.*

From “***Quantifying confidence***” by Angeletos, Collard, and Dellas (2018)

*In a nutshell, there is a risk of **contamination** of the estimates of a model by frequencies that the model was **not designed** to capture.*

- ACD estimate their model with band-spectral estimator using business cycle frequencies only

How?

How?

- Monte Carlo study, compare 3 maximum likelihood estimators (MLE)

How?

- Monte Carlo study, compare 3 maximum likelihood estimators (MLE)
  - Time domain, Kalman filter likelihood (**TD**)
  - Frequency domain, Whittle likelihood, using all frequencies (**FD**)
  - Frequency domain, Whittle likelihood, using business cycle frequencies only (**BC**)

How?

- Monte Carlo study, compare 3 maximum likelihood estimators (MLE)
  - Time domain, Kalman filter likelihood (**TD**)
  - Frequency domain, Whittle likelihood, using all frequencies (**FD**)
  - Frequency domain, Whittle likelihood, using business cycle frequencies only (**BC**)
- compare bias and estimation efficiency

How?

- Monte Carlo study, compare 3 maximum likelihood estimators (MLE)
  - Time domain, Kalman filter likelihood (**TD**)
  - Frequency domain, Whittle likelihood, using all frequencies (**FD**)
  - Frequency domain, Whittle likelihood, using business cycle frequencies only (**BC**)
- compare bias and estimation efficiency
- Main question: BC uses less information. How much less? Which parameters are more/less affected

Results

- no (good) reason to use full info FD (all freqs) instead of full info TD estimator

Results

- no (good) reason to use full info FD (all freqs) instead of full info TD estimator
- significantly less info in BC frequencies for all parameters

Results

- no (good) reason to use full info FD (all freqs) instead of full info TD estimator
- significantly less info in BC frequencies for all parameters
- we can reliably predict 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

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

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

Gaussian Likelihood Function

Replace $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$ with SYMMETRIC BLOCK CIRCULANT MATRIX $\boldsymbol{\Omega}_T(\boldsymbol{\theta})$

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

- $\boldsymbol{F}_T$ is Fourier transform matrix
- $\boldsymbol{S}_T(\boldsymbol{\theta})$ is block-diagonal matrix of spectral density of $\boldsymbol{y}_t$, evaluated at different frequencies.

$$\boldsymbol{s}(\boldsymbol{\theta}, \omega_i) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \mathrm{cov}_{\boldsymbol{y}}(\tau; \boldsymbol{\theta}) \exp(-\mathrm{i}\omega_i \tau), \quad \omega_i = \frac{2\pi(i-1)}{T}$$

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
- 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

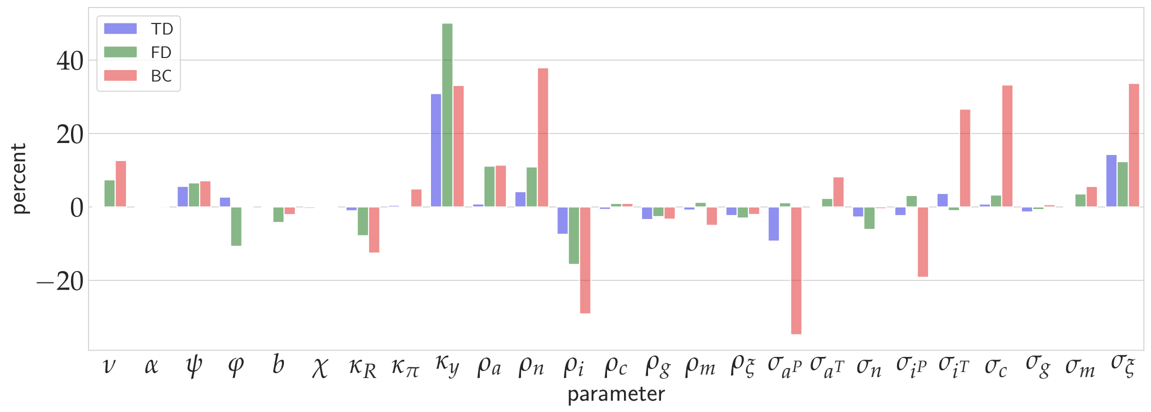
New Keynesian DSGE model from “Quantifying confidence” by Angeletos, Collard, and Dellas (2018)

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

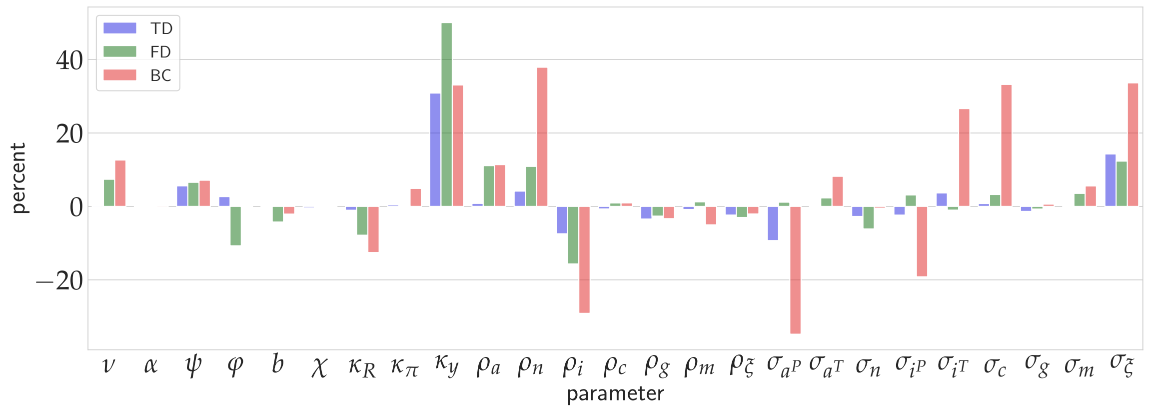
RESULTS

1000 replications

Bias (% of θ)

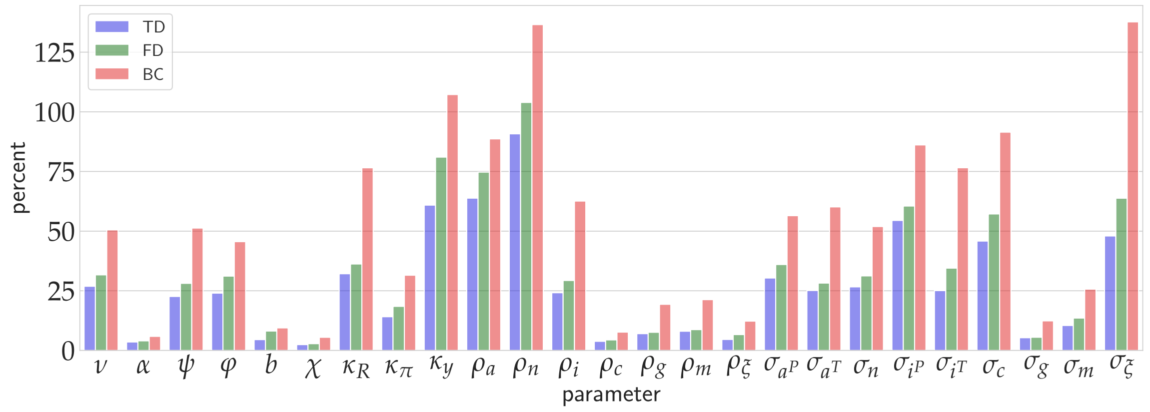


Bias (% of θ)

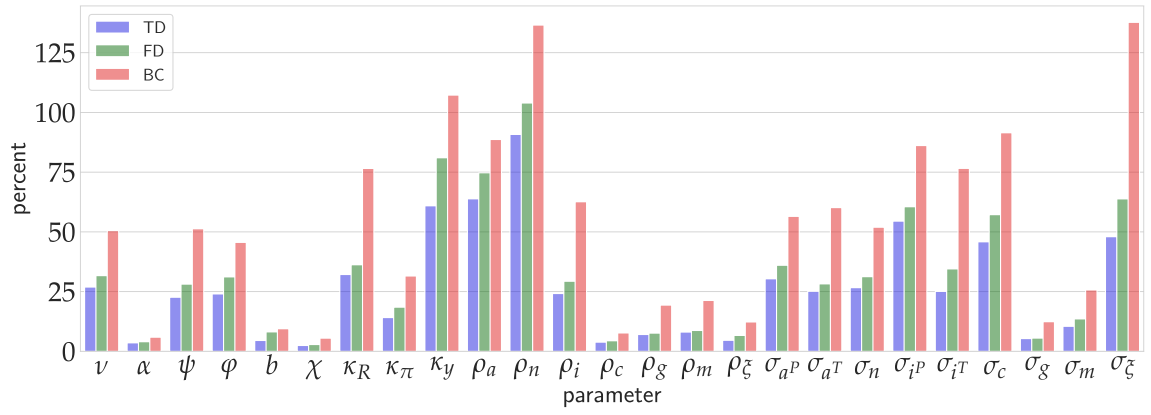


- 4%, 7%, 13%
- rank corr: 0.57 (TD, FD), 0.84 (TD, BC), 0.66 (FD, BC)

Efficiency (std as $\%$ of θ)



Efficiency (std as % of θ)

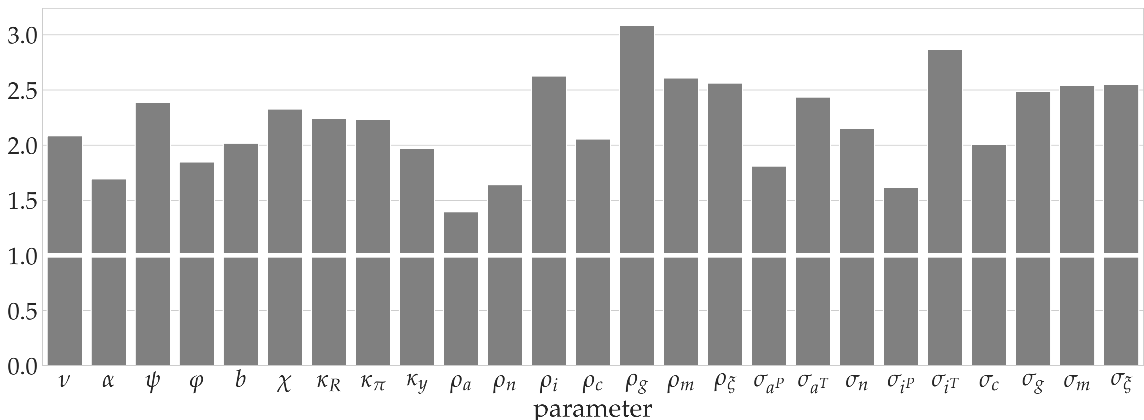


- 27%, 32%, 53%
- rank corr: 0.98 (TD, FD), 0.95 (TD, BC)

Efficiency loss

How much less information in the BC frequencies?

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



Can efficiency loss be predicted?

Fisher information approach

Can efficiency loss be predicted?

Fisher information approach

- FIM (expected information about $\boldsymbol{\theta}$ in the sample)

$$\mathcal{I}(\boldsymbol{\theta}) = \mathrm{E}\left[\frac{\partial}{\partial \boldsymbol{\theta}}\ell(\boldsymbol{\theta}; \boldsymbol{Y})\frac{\partial}{\partial \boldsymbol{\theta}'}\ell(\boldsymbol{\theta}; \boldsymbol{Y})\right] = -\mathrm{E}\left[\frac{\partial^2}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\ell(\boldsymbol{\theta}; \boldsymbol{Y})\right] \tag{1}$$

Can efficiency loss be predicted?
Fisher information approach

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

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

Can efficiency loss be predicted?
Fisher information approach

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

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

- predicted efficiency loss

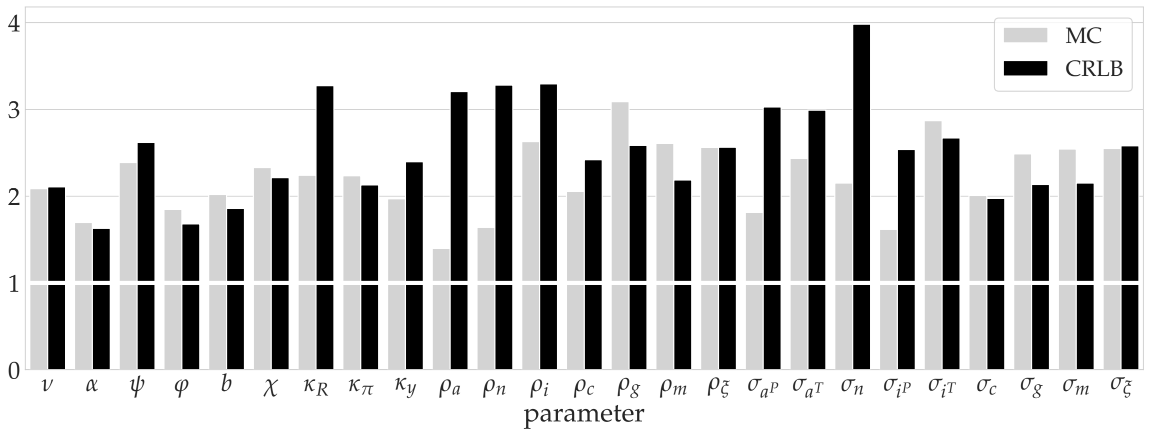
$$\frac{\mathrm{crlb}(BC)}{\mathrm{crlb}(TD)} \tag{2}$$

Can efficiency loss be predicted?
Fisher information approach

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, ...)?

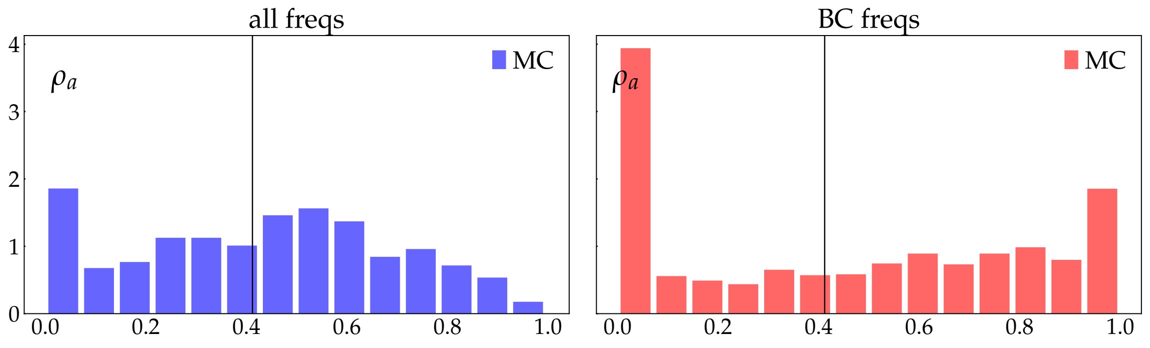
short answer:

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

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

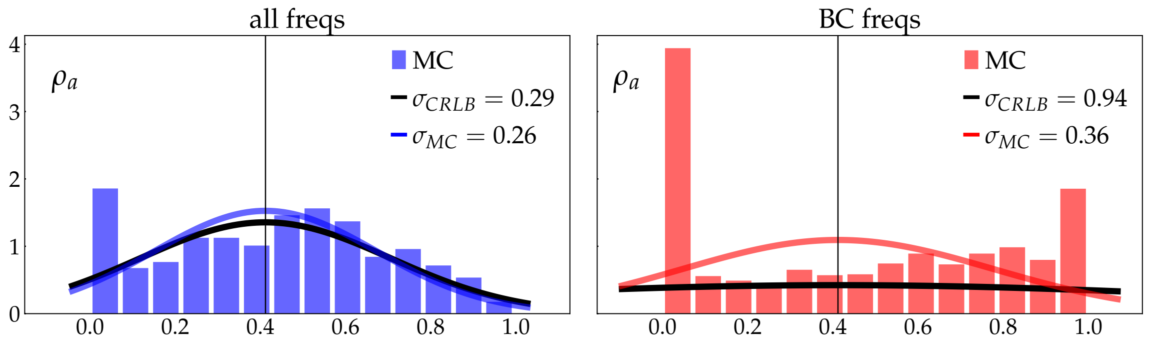
What's wrong with ρ_a (and ρ_n, κ_R, σ_n, ...)?

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



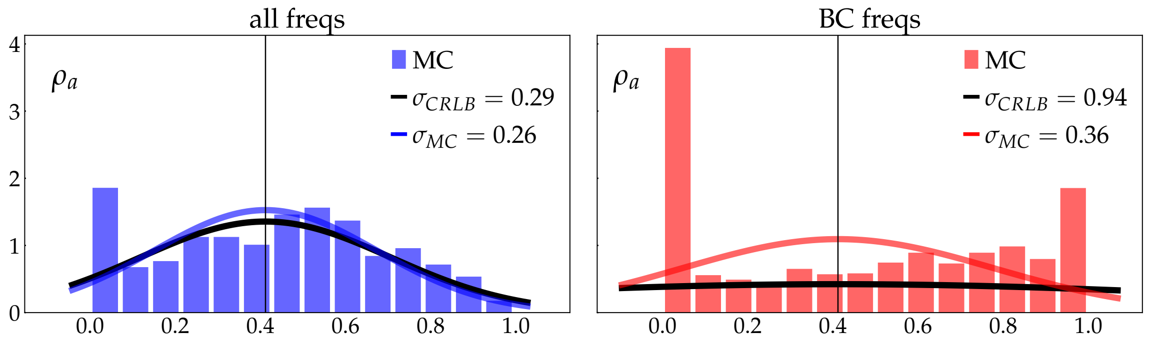
What's wrong with ρ_a (and ρ_n, κ_R, σ_n, ...)?

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



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

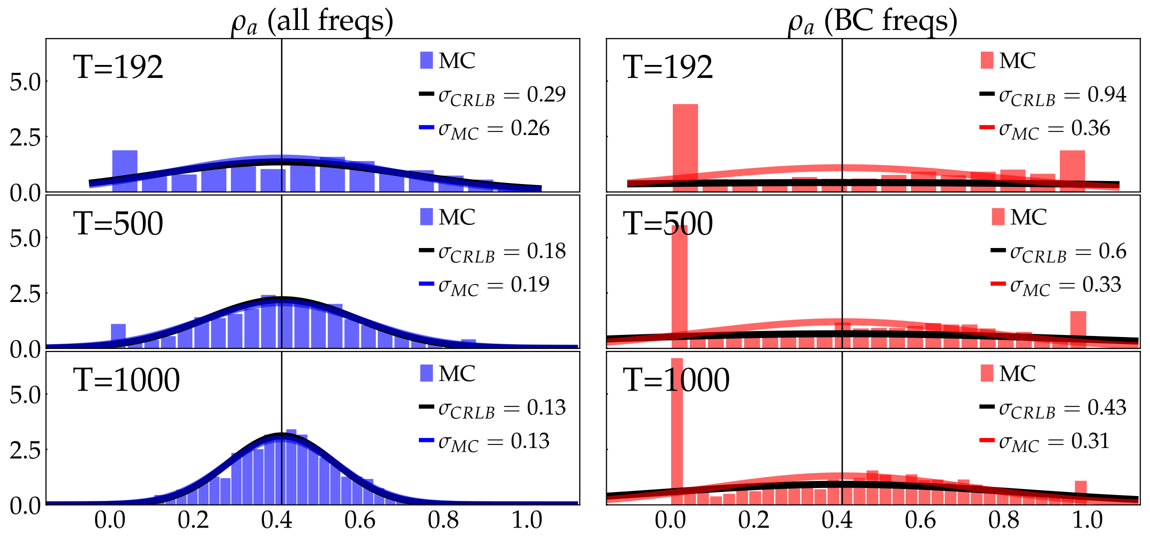
MC-estimated vs CRLB-predicted marginal distribution of $\hat{\rho}_a$



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

- flat likelihood
- parameter constraints

ρ_a: MC vs CRLB as T increases



ρ_a: MC vs CRLB for $T = 1000$ ($T = 192$)

- estimated/predicted **uncertainty**: full info

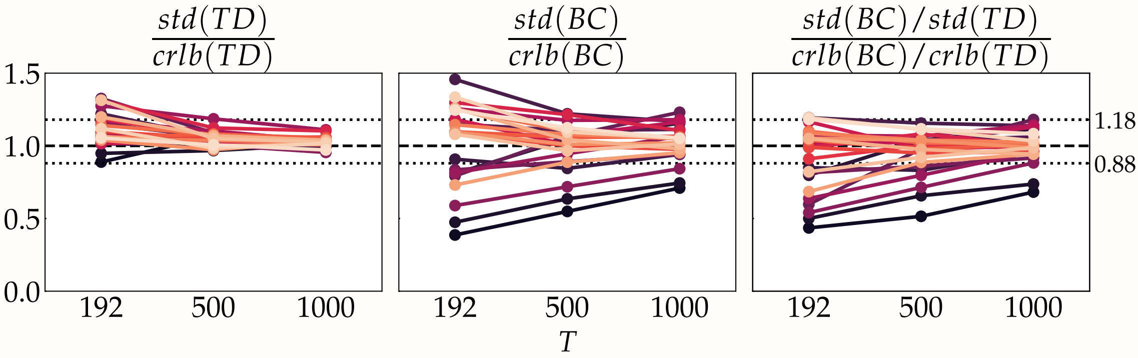
$$\mathrm{std}(TD) \big/ \mathrm{crlb}(TD) = 1.04 \ \ (0.89)$$

- estimated/predicted **uncertainty**: BC freqs

$$\mathrm{std}(BC) \big/ \mathrm{crlb}(BC) = 0.71 \ \ (0.39)$$

- estimated/predicted **relative efficiency**

$$\frac{\mathrm{std}(BC)}{\mathrm{std}(TD)} \Bigg/ \frac{\mathrm{crlb}(BC)}{\mathrm{crlb}(TD)} = 0.68 \ \ (0.44)$$



- CRLB is good predictor of **estimation uncertainty** if

$$\text{std} \,/\, \text{crlb} \approx 1$$

- CRLB is good predictor of **relative efficiency** if

$$\frac{\text{std}(BC)}{\text{std}(TD)} \Big/ \frac{\text{crlb}(BC)}{\text{crlb}(TD)} \approx 1$$

Conclusion

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

Conclusion

- Full information TD estimator (KF) is superior to FD estimator (Whittle likelihood)
  - 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 FD estimator (Whittle likelihood)
  - 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

APPENDIX

Some Implications

Some Implications

- time domain (full info) estimation of BC models

Some Implications

- time domain (full info) estimation of BC models
  - contaminated information?

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

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

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*”

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

◂ BACK

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