On the band spectral estimation of business cycle models

Nikolay Iskrev

Abstract

In this paper, I evaluate the properties and performance of band spectral estimators applied to business cycle models. Band spectral methods are widely used to study frequency-dependent relationships among time series. In business cycle research, the Whittle likelihood approximation enables researchers to estimate models using only the frequencies those models are best suited to represent, such as the business cycle frequencies. Using the medium-scale model of Angeletos et al. (2018) as a data-generating process, I conduct a Monte Carlo study to assess the finite-sample properties of the band spectral maximum likelihood estimator (MLE) and compare its performance with that of the full-spectrum and exact time-domain MLEs. The results show that the band spectral estimator exhibits considerable biases and efficiency losses for most estimated parameters. Moreover, both the full-information and band spectral Whittle estimators perform poorly in contrast to the time domain estimator, which successfully recovers all model parameters. I demonstrate how these findings can be understood through the theoretical properties of the underlying model, and describe simple tools and diagnostics that can be used to detect potential problems in band spectral estimation for a wide class of macroeconomic models.

Keywords: DSGE models, frequency domain, Whittle likelihood, information content, Monte Carlo

JEL classification: C32, C51, C52, E32

nikolay.iskrev@bportugal.pt, Bank of Portugal, Av. Almirante Reis 71, 1150-012, Lisboa, Portugal

1 Introduction

Following the pioneering work of Hannan (1963), band spectral methods have become an important tool for analyzing dynamic models on the basis of a restricted band of frequencies. The rationale for band spectral estimation of macroeconomic models is straightforward: a model should not be forced to fit data it is not designed to explain. In particular, if a theoretical model lacks the features and mechanisms to account for data movements in some parts of the frequency range, those frequencies should be excluded from estimation. Failure to do so would distort parameter estimates by forcing them to accommodate empirical features outside the model's intended scope. Therefore, for example, a business cycle model known a priori to be ill-suited for explaining high- and low-frequency phenomena in the data should be estimated using only data components with business cycle periodicities. In other words, one should use band spectral estimation methods.

Band-spectral estimation via likelihood methods relies on Whittle (1953) approximation of the Gaussian likelihood in the frequency domain.[1] Hansen and Sargent (1980) showed how to apply the Whittle likelihood to estimate dynamic rational expectations models, with early examples of this approach provided by Altug (1989) and Christiano and Vigfusson (2003). Hansen and Sargent (1993) were also the first to leverage the Whittle likelihood for band spectral estimation of macroeconomic models, with a focus on understanding the implications of using seasonally adjusted data. Diebold et al. (1998) proposed a more general estimation approach where frequencies might be assigned different weights. The aim is to mitigate model misspecification and the band spectral maximum likelihood estimator, which they call Band-MLE, arises as a special case of their framework. Cogley (2001) also employed a band spectral likelihood approach to exclude low frequencies in the estimation of models with uncertain trend specifications.

Despite the compelling reasons for using band spectral estimation, particularly given the literature's focus on explaining business cycle phenomena, macroeconomic models are predominantly estimated in the time domain, incorporating information from all frequencies. Some notable exceptions are Qu and Tkachenko (2012a,b), Sala (2015), and, most recently, Angeletos et al. (2018), who employ likelihood-based estimation in the frequency domain using a subset of frequencies.

The main purpose of this paper is to evaluate the properties and performance of the band spectral estimator when applied to modern business cycle models. To that end, I conduct a Monte Carlo simulation study using the medium-scale dynamic stochastic general equilibrium (DSGE) model of Angeletos et al. (2018) as the data-generating process (DGP). Following the existing literature, I construct the band spectral maximum likelihood estimator (MLE) by restricting the Whittle likelihood to frequencies from the business cycle range of the spectrum. The finite-sample properties of the estimator are compared to those of the full-spectrum Whittle MLE and the time domain Gaussian MLE. The aim of this comparison is three-fold: First, since the Whittle likelihood approximates the exact Gaussian likelihood, it may exhibit finite-sample distortions, especially for highly persistent processes (see Hansen and Sargent (1980)). Since

[1] The band spectral MLE can be interpreted as a full information analogue of the band spectral linear regression proposed by Hannan (1963), and popularized in economics by Engle (1974, 1978).

no exact likelihood equivalent of the band spectral estimator exists in the current literature, comparing the properties of the full-information time and frequency domain estimators helps assess the nature and magnitude of the distortions involved in using the Whittle approximation.

Second, comparing the band spectral estimator to the efficient full-information one provides a benchmark for gauging the efficiency consequences of restricting information only to the business cycle part of the spectrum. Such a comparison is subject to the obvious caveat that the reason for using the band spectral estimator in the first place are the discrepancies between the model and the data in some parts of the spectrum. In the simulations, the model of Angeletos et al. (2018) is treated as the true DGP across all frequencies. Under this setup, the interpretation of the discrepancies is that, in the real world, the unobserved model-consistent data are contaminated by low- and high-frequency components, such as trends, measurement noise, etc, which are not accounted for in the model. In other words, this is framed as a missing data problem, where we compare the efficiency of the feasible band spectral estimator to that of the infeasible oracle estimator – one that represents a situation without missing data.[2] Third, knowing how important, under the true DGP, frequencies outside the business cycle range are is useful because it reveals the potential consequences of estimating misspecified models. The more informative a part of the spectrum, the more severe the distortions from misspecification in those frequencies. And, conversely, the effect of misspecification are small when the contaminated frequencies contribute a negligible amount of information. From this perspective, the simulation analysis is highly relevant since, as pointed out earlier, time domain estimation uses information from all frequencies and remains the dominant approach in business cycle research.

The simulation results are presented and discussed in Section 4, following an outline of the Angeletos et al. (2018) model in Section 2 and a brief introduction to the Whittle likelihood in Section 3. I find that both Whittle likelihood-based estimators exhibit severe biases under the original parametrization of the ACD model, with many parameters showing biases several orders of magnitude larger than the time domain MLE. I show that the cause of the bias is the extreme near–unit root levels of persistence of the model variables. Under an alternative less persistent parametrization, the biases of the full-spectrum and band spectral estimators decrease substantially, though remaining higher than the time domain estimator's bias. Furthermore, I find that band spectral estimation entails a significant efficiency loss compared to the full-information case, often exceeding 100% of the true values. This suggests that frequencies from the lower and higher end of the spectrum contain relevant information for all parameters, and, for some parameters, the information is essential. It also raises the question of whether we can predict the relative importance of different parts of the spectrum using the theoretical model alone without resorting to computationally expensive simulations. I explore this issue in Sections 5 and 6, where I examine the utility of the expected Fisher information matrix (FIM) and the implied Cramér-Rao lower bounds (CRLBs) as predictors of estimation uncertainty. In Section 5, I compare FIM-based predictions with the Monte Carlo simulation results for the full information estimator and the BC-frequency-only band spectral estimator. I find that the analytical approach

[2]This is analogous to treating seasonal frequencies as missing when estimating a model that abstracts from seasonality with data that exhibits seasonal patterns. Then, the oracle estimator is the one using seasonally-adjusted data. Insofar as such data is available, the oracle estimator is in fact feasible.

is generally highly accurate, both qualitatively (identifying which parameters may suffer greater or lesser information loss) and quantitatively (predicting the magnitude of the efficiency losses). In Section 6, I investigate whether the information missing from the BC frequency band is located mainly in the low or the high frequencies. I also show how the FIM formalism can be used to explain why certain frequencies are more informative than others. The analytical predictions are again validated against MC simulations and the results confirm the usefulness of the approach as a framework for understanding the distribution of parameter information across frequencies.

Consistent with common practice in structural macroeconomics, Whittle likelihood-based estimation is typically implemented within a Bayesian framework. This is the case in Qu and Tkachenko (2012a) and Sala (2015), who considered full- and band spectral estimation of DSGE models, and in Angeletos et al. (2018), who performed only band spectral estimation. Bayesian methods are also employed in Plagborg-Møller (2019), who developed a method for estimating structural impulse response functions using the full-spectrum Whittle likelihood. In contrast, the analysis in this paper is pursued entirely in a frequentist setting. This is appropriate since both issues I investigate – the quality of the Whittle approximation and the efficiency costs of the band spectral approach – relate solely to the likelihood function and may be obscured by introducing a prior distribution on the estimated parameters. Both Qu and Tkachenko (2012a) and Sala (2015) found that the estimation results are very different depending on whether estimation uses all or a subset of frequencies, in particular those with business cycle periodicities. They interpreted these differences as evidence of model misspecification, arguing that both approaches should yield similar results absent misspecification. In particular, Sala (2015, p.220) writes "In general, if no misspecification is present, parameter estimates will not depend on the frequencies used. Estimation on frequency bands would just be less efficient than estimation over the entire frequency domain." My simulation results challenge this interpretation, showing that finite-sample distortions in the Whittle likelihood alone can produce substantially different parameter estimates, even without misspecification. Furthermore, I show that, by introducing prior information about the estimated parameters, one may greatly underestimate the loss of sample information in the band spectral case. In frequentist settings this occurs when theoretical parameter restrictions – common in structural estimation – become binding due to the limited amount of information in the sample. While this also affects the full information case, the impact is generally much stronger when estimation is based on a subset of frequencies. Similar issues arise in fully Bayesian settings, where the relative importance of prior and sample information can be very different in the band spectral compared to the full information case. This is important to keep in mind as prior distributions are rarely tailored to the applications at hand and are typically specified to be in line with the rest of the literature, i.e. for models estimated in the time domain using all frequencies.[3]

This paper contributes to a small number of studies using simulations to investigate the properties of the MLE when applied to the estimation of DSGE models. Consistent with the findings of Iskrev (2010a), Schmitt-Grohé and Uribe (2012), Iskrev and Ritto (2016), and Adolfson

[3]Qu and Tkachenko (2012a) estimate the Smets and Wouters (2007) model using the same prior distribution. Sala (2015) describes the prior specification as "standard in the literature", referencing Justiniano et al. (2010). Similarly, Angeletos et al. (2018) describe their choice of priors as "broadly in line with the literature".

et al. (2019), I show that the time-domain MLE successfully recovers model parameters when the model is correctly specified for all frequencies. McDonald and Shalizi (2022) reach the opposite conclusion, stating that "even with centuries' worth of data, the model remains poorly estimated". None of these works considers estimation in the frequency domain. To my knowledge, Sala (2015) is the only previous study to investigate the finite-sample properties of the Whittle likelihood approach when applied to the estimation of DSGE models. As in this paper, he conducts Monte Carlo experiments comparing the time domain and frequency domain MLE. In addition to the business cycle frequency-only estimator, Sala (2015) also considers estimation using all except the low frequencies and all except the high frequencies.[4] His findings are remarkably different from mine. In particular, Sala (2015, Table 1, Appendix B) shows that the bias of the full-spectrum Whittle MLE is comparable to, and even slightly lower than, that of the time domain estimator. Furthermore, although he reports larger, on average, estimation uncertainty for the band spectral estimators, there are several instances where some of these estimators are more efficient than the full information ones. Consequently, he concludes (Sala (2015, footnote 13)) that "the frequency domain approximation, both on the whole spectrum and on subsets of frequencies, is remarkably good". Since no further details or discussion of these results are provided, it remains unclear whether the discrepancy with the present study's findings stems from differences in the properties of the data generating process used in Sala (2015), or variations in the design of the experiment, such as the number of replications, choice of optimizers, and so on.[5]

Due to high computation costs, Monte Carlo-style analysis is rarely performed when estimating structural macro models. Yet another contribution of this paper is demonstrating the usefulness of the CRLBs as predictors of the information content of a sample. In the context of time domain MLE, similar results were presented in Iskrev (2010a) where I run Monte Carlo simulations with the Smets and Wouters (2007) model to compare predicted and estimated parameter uncertainty. Likewise, Iskrev and Ritto (2016) found that using FIM-based measures to compare the relative informativeness of different subsets of observables yields an identical ranking to the one obtained from Monte Carlo simulations. This paper extends the analysis to the frequency domain with the aim of quantifying the relative informativeness of different subsets of frequencies. In this respect, the analysis relates to the question of whether different subsets of frequencies contain sufficient information to identify model parameters. Frequency domain conditions for local and global identification are provided in Qu and Tkachenko (2012a) and Qu and Tkachenko (2017). Their condition for local identification, in particular, is equivalent to the asymptotic FIM being of full rank.[6] Finding that the FIM-based approach provides a reliable alternative to Monte Carlo simulations is significant because its negligible computational costs enable comparative analyses across various scenarios, such as different parameterizations, selections of observed variables, frequency bands, sample sizes, and more. This exploration can be conducted before to taking models to data, thus helping researchers design their empirical investigations.

[4] Sala (2015) defines the business cycle range to include frequencies with period between 4 and 32 quarters, whereas, following Angeletos et al. (2018), I include only frequencies with period between 6 and 32 quarters.

[5] One potentially important factor is the different number of Monte Carlo replications – 100 in Sala (2015) whereas I use 1000.

[6] The condition in Qu and Tkachenko (2012b) is more general and allows for singular models. For non-singular models, such as the one of Angeletos et al. (2018), the equivalence holds.

2 The Model

The model is taken from Angeletos et al. (2018) (henceforth ACD). There are two reasons for choosing this model. First, it is similar in size and shares many features with other estimated medium-scale DSGE models in the contemporary literature. These include: a neoclassical growth core augmented with sticky prices, habit formation in consumption, adjustment costs in investment, monetary policy following a Taylor rule, and a number of exogenous shocks driving business cycle fluctuations. What sets ACD apart is the departure from the standard assumption of rational expectations and common information about the state of the economy. In their model, agents' beliefs about the expectations of other agents (higher-order beliefs) are subject to autonomous variation, called "confidence shock", which creates divergence between the two forms of beliefs. This generates exogenous variations in agents' expectations of the economic outcomes in the short-run, without altering their medium or long-run expectations, or their expectations of exogenous fundamentals at any horizon. ACD show that embedding this mechanism into an otherwise standard New Keynesian business cycle model improves the match with observed macroeconomic patterns. Estimating the model with U.S. data, they find that the confidence shock accounts for more than half of the volatility in the main macro aggregates at business cycle frequencies.

The second reason for selecting the ACD model is that the authors explicitly state that their model describes business cycle phenomena only, acknowledging its lack of features and mechanisms needed to account for low and high frequency properties of empirical time series. For this reason, the model is estimated in the frequency domain using only the business cycle frequencies. While focusing on business cycle fluctuations is common in the literature, models are typically estimated in the time domain using all frequencies. The few exceptions that use frequency domain estimation, such as Sala (2015) and Qu and Tkachenko (2012b), work with models not specifically designed to fit only the business cycle part of the spectrum.

The main methodological contribution of ACD is demonstrating how to introduce higher-order belief dynamics into macroeconomic models in a tractable way. Like most DSGE models, estimation uses a linear state space representation derived from log-linearizing the equilibrium conditions around steady state. For reference, the linearized equilibrium conditions of the ACD model are presented below. For more details on the model and solution method, readers should consult the original publication.

2.1 Linearized equilibrium conditions

The economy consists of a continuum of islands and a mainland. Each island contains a representative household and a continuum of monopolistically competitive firms producing differentiated commodities using household-supplied labor and capital. These commodities combine through a CES aggregator into an island-specific composite good, which in turn enters mainland final good production through another CES aggregator. The final good is used for both consumption and investment. The log-linearized equilibrium conditions, with variables expressed as log-deviations from steady-state values, are summarized as follows:

Optimal consumption allocation

$$\mathrm{E}_{it}\left[\zeta_t^c + \nu n_{it}\right] = \zeta_t^c - \frac{c_{it} - bC_{t-1}}{1-b} + \mathrm{E}_{it}\left[s_{it} + \varrho Y_t + (1-\varrho)y_{it} - n_{it}\right], \tag{2.1}$$

where c_{it} and C_t are consumption on island i and aggregate consumption, y_{it} and Y_t are the quantity of the final good produced in island i and aggregate output, n_{it} is hours worked, s_{it} denotes the realized markup in island i, and ζ_t^c is a preference shock. The parameter ν determines the inverse labor supply elasticity, while b and ϱ denote, respectively, the degree of habit persistence and the degree of substitutability across the islands' composite goods in the production of the final good.

Optimal investment decision

$$\begin{aligned}\mathrm{E}_{it}\left[\lambda_{it} + q_{it}\right] = {} & \mathrm{E}_{it}\left[\lambda_{it+1} + \beta(1-\delta)q_{it+1} + (1-\beta(1-\delta))(s_{it+1} + \varrho Y_{t+1}\right. \\ & \left. + (1-\varrho)y_{it+1} - u_{it+1} - k_{it+1})\right]\end{aligned} \tag{2.2}$$

where q_{it} is the price of capital, u_{it} is the rate of capital utilization, and λ_{it} is the marginal utility of consumption, given by

$$\lambda_{it} = \zeta_t^c - \frac{c_{it} - bC_{t-1}}{1-b} \tag{2.3}$$

The parameter β denotes the intertemporal discount rate in the utility function of the households, and δ denotes the depreciation rate.

Optimal bond holdings decision

$$R_t = \zeta_t^c - (1+\nu)n_{it} - s_{it} - \varrho Y_t - (1-\varrho)y_{it} - \mathrm{E}'_{it}[\lambda_{it+1} - \pi_{it+1}] \tag{2.4}$$

where R_t denotes the nominal interest rate and π_{it} denotes the inflation rate in island i.

Equilibrium price of capital

$$q_{it} = (1+\beta)\varphi\iota_{it} + \varphi\iota_{t-1} - \beta\varphi\,\mathrm{E}'_{it}\,\iota_{it+1} + \zeta_t^{IP} - \zeta_t^{IT} \tag{2.5}$$

where i_{it} denotes the level of investment, ζ_t^{IP} denotes the investment-specific technology shock, ζ_t^{IT} denotes the investment demand shock, and φ denotes a parameter governing investment adjustment costs.

Production function

$$y_{it} = \zeta_t^A + \alpha(u_{it} + k_{it}) + (1-\alpha)n_{it} \tag{2.6}$$

where k_{it} denotes the local capital stock, ζ_t^A denotes the level of aggregate TFP, and α represents

the share of capital in the production function. The capital accumulation equation is

$$k_{it+1} = (1-\delta)k_{it} + \delta(\zeta_t^{IT} + \iota_{it}), \tag{2.7}$$

and level of TFP is the sum of a permanent (a_t^p) and a transitory (a_t^τ) component:

$$\zeta_t^A = a_t^p + a_t^\tau, \tag{2.8}$$

Resource constraint

$$\varrho y_t + (1-\varrho)y_{it} = x_{it} + \alpha u_{it}, \tag{2.9}$$

where x_{it} denotes GDP on island i, given by

$$x_{it} = s_c c_{it} + (1-s_c-s_g)(\zeta_t^{IP} + \iota_{it}) + s_g G_t, \tag{2.10}$$

and G_t, s_c and s_g denote the level of government spending and the steady-state ratios of consumption and government spending to output. To ensure the existence of a balanced growth path, government spending is defined as

$$G_t = \zeta_t^g + \frac{1}{1-\alpha}a_t^p - \frac{\alpha}{1-\alpha}\zeta_t^{IP} \tag{2.11}$$

where ζ_t^g denotes the government spending shock.

Equilibrium utilization

$$\zeta_t^{IP} + \frac{1}{1-\psi}u_{it} = s_{it} + \varrho y_t + (1-\varrho)y_{it} - k_{it}, \tag{2.12}$$

where ψ denotes the capital utilization elasticity parameter.

Inflation rate

$$\pi_{it} = \frac{(1-\chi)(1-\beta\chi)}{\chi\,(1+\chi(1-\beta))}s_{it} + \frac{\beta\chi(1-\chi)\pi_t + \beta\chi\,\mathrm{E}'\,\pi_{it+1}}{\chi\,(1+\chi(1-\beta))}, \tag{2.13}$$

where Π_{it} denotes the aggregate inflation rate, and $(1-\chi)$ denotes the probability that a firm resets its price in a given period.

Monetary policy rule

$$R_t = \kappa_R R_{t-1} + (1-\kappa_R)(\kappa_\pi \pi_{it} + \kappa_y(x_{it} - x_{it}^F)) + \zeta_t^m \tag{2.14}$$

where x_{it}^F denotes the GDP that would be attained in a flexible-price allocation, ζ_t^m is a monetary policy shock, κ_π and κ_y are parameters determining the policy rate response to inflation and the

output gap, and κ_{Ri} controls the degree of interest-rate smoothing. The flexible-price allocations are obtained from equations (2.1) – (2.12) by setting the realized markup to zero ($s_{it} = 0$) and replacing R_t in (2.4) with the real interest rate.

It is worth pointing out that there are two different subjective expectation operators E_{it} and E'_{it} in the above conditions. In the model, each time period t is divided into two stages. In stage 1, inhabitants of each island receive an unbiased signal about that period's TFP level and form beliefs that firms and households on other islands receive a signal biased by the observed confidence shock ξ_t. In stage 2, the true state of nature and the realized value of economic activity are publicly revealed. ACD discuss two timing protocols for firms' and households' decisions: either supply is determined first with prices adjusting to meet demand, or demand is determined first with supply adjusting to meet it. The model presented above is estimated under the second assumption, as seen by the use of stage 1 expectations in the optimality conditions for consumption and saving in equations (2.1), (2.2), and stage 2 expectations in equations (2.4), (2.5), (2.13).

There are nine shocks in the model: a permanent (a_t^p) and a transitory (a_t^τ) TFP shock; a permanent (ζ_t^{IP}) and a transitory (ζ_t^{IT}) investment-specific shock; a news shock regarding future productivity (a_t^n); a discount-rate shock (ζ_t^c); a government-spending shock (ζ_t^g); a monetary policy shock (ζ_t^m); and a confidence shock (ξ_t). The latter shock, observed in stage 1 of each period, representing the perceived bias in other islands' signals about that period's TFP level. The permanent TFP shock is given by

$$a_t^p = a_{t-1}^p + a_{t-1}^n + \varepsilon_t^p, \tag{2.15}$$

and the permanent investment-specific shock follows a random walk

$$\zeta_t^{IP} = \zeta_{t-1}^{IP} + \varepsilon_t^{IP}, \tag{2.16}$$

where ε_t^p and ε_t^{IP} are i.i.d. innovations. All remaining shocks are stationary AR(1) processes.

The model is estimated using quarterly U.S. data for six variables: GDP, consumption, investment, hours worked, the inflation rate, and the federal funds rate. The sample period is 1960:Q1 - 2007:Q4. The model parameters are estimated with Bayesian methods using the frequency domain representation of the likelihood function. The estimated median of the posterior distribution is reported in Table 7.

3 The Whittle likelihood

3.1 General case

Let $\{\boldsymbol{y}_t\}_{t=1}^T$ denote a T-dimensional sample from a zero mean stationary Gaussian process with autocovariance function $\boldsymbol{\Gamma}(\tau;\boldsymbol{\theta}) = \mathrm{cov}\,(\boldsymbol{y}_{t+\tau}, \boldsymbol{y}_t)$. The log-likelihood function of $\boldsymbol{Y}_T =$

$(\boldsymbol{y}_1', \boldsymbol{y}_2', \ldots, \boldsymbol{y}_T')'$, up to an additive constant, is given by

$$\begin{aligned} \ell(\boldsymbol{\theta}; \boldsymbol{Y}_T) &= -\frac{1}{2}\log\det(\boldsymbol{\Sigma}_T(\boldsymbol{\theta})) - \frac{1}{2}\boldsymbol{Y}_T'\boldsymbol{\Sigma}_T^{-1}(\boldsymbol{\theta})\boldsymbol{Y}_T \quad (3.1) \\ &= -\frac{1}{2}\log\det(\boldsymbol{\Sigma}_T(\boldsymbol{\theta})) - \frac{1}{2}\operatorname{tr}\left(\hat{\boldsymbol{\Sigma}}_T\boldsymbol{\Sigma}_T^{-1}(\boldsymbol{\theta})\right) \quad (3.2) \end{aligned}$$

where $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$ is a block Toeplitz matrix, with blocks given by $\boldsymbol{\Gamma}(\tau;\boldsymbol{\theta})$ for $\tau \in \{0, 1, 2, \ldots, T-1\}$, and $\hat{\boldsymbol{\Sigma}}_T = \boldsymbol{Y}_T\boldsymbol{Y}_T'$ is the sample version of $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$.

Evaluating $\ell(\boldsymbol{\theta}; \boldsymbol{Y}_T)$ requires computing the determinant and inverse of $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$, which can be computationally prohibitive even for moderate sample sizes. To address this problem, Whittle (1953) introduced a spectral approximation of $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$ as a computationally efficient method for calculating the likelihood function of stationary Gaussian time series. The approximation exploits the fact that block Toeplitz matrices can be approximated by block circulant matrices,[7] whose eigenvalue decomposition can be computed very efficiently using the discrete Fourier transform (DFT). Specifically, it can be shown that for large T,[8]

$$\boldsymbol{\Sigma}_T(\boldsymbol{\theta}) \approx \boldsymbol{\Omega}_T(\boldsymbol{\theta}) = \boldsymbol{F}_T^*\boldsymbol{S}_T(\boldsymbol{\theta})\boldsymbol{F}_T \quad (3.3)$$

where $\boldsymbol{\Omega}_T$ is a symmetric block circulant matrix, $\boldsymbol{F}_T$ is an orthonormal matrix of Fourier transform coefficients, and $\boldsymbol{F}_T^*$ is the conjugate transpose of $\boldsymbol{F}_T$. The matrix $\boldsymbol{S}_T(\boldsymbol{\theta})$ is block diagonal with it's i-th block $\{\boldsymbol{S}_T(\boldsymbol{\theta})\}_{ii} = \boldsymbol{s}(\boldsymbol{\theta}, \omega_i)$ given by the spectral density matrix of $\boldsymbol{y}_t$ evaluated at the i-th Fourier frequency,

$$\boldsymbol{s}(\boldsymbol{\theta}, \omega_i) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty}\boldsymbol{\Gamma}(\tau;\boldsymbol{\theta})\exp(-\mathrm{i}\omega_i\tau), \quad \omega_i = \frac{2\pi(i-1)}{T} \quad (3.4)$$

The sample version of $\boldsymbol{s}(\boldsymbol{\theta}, \omega)$, called the periodogram of $\boldsymbol{Y}_T$, is defined as

$$\boldsymbol{I}_T(\omega) = \frac{1}{2\pi}\sum_{\tau=-(T-1)}^{T-1}\hat{\boldsymbol{\Gamma}}_{\boldsymbol{y}}(\tau)\exp(-\mathrm{i}\omega\tau) \quad (3.5)$$

where $\hat{\boldsymbol{\Gamma}}_{\boldsymbol{y}}(\tau) = \sum_{\tau=1}^{T-\tau}\boldsymbol{y}_{t+\tau}\boldsymbol{y}_t'$ is the sample autocovariance of $\boldsymbol{y}_t$ at lag τ and $\hat{\boldsymbol{\Gamma}}_{\boldsymbol{y}}(-\tau) = \hat{\boldsymbol{\Gamma}}_{\boldsymbol{y}}(\tau)'$. The periodogram can be efficiently calculated as

$$\boldsymbol{I}_T(\omega) = \frac{1}{2\pi T}\boldsymbol{J}_T(\omega)\boldsymbol{J}_T(\omega)^* \quad (3.6)$$

where $\boldsymbol{J}_T(\omega) = \sum_{t=1}^{T}\boldsymbol{y}_t\exp(-\mathrm{i}\omega t)$ is the DFT of $\boldsymbol{Y}_T$. Indeed, pre-multiplication of $\boldsymbol{Y}_T$ by the matrix $\boldsymbol{F}$ in (3.3) performs this transformation for $\omega \in \{0, 2\pi/T, \ldots, 2\pi(T-1)/T\}$.

[7] A block circulant matrix $\boldsymbol{A}$ has the following form

$$\boldsymbol{A} = \begin{bmatrix} A_0 & A_1 & A_2 & \cdots & A_{n-1} \\ A_{n-1} & A_0 & A_1 & \cdots & A_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ A_1 & A_2 & A_3 & \cdots & A_0 \end{bmatrix},$$

where the blocks have the same size. If $\boldsymbol{A}$ is symmetric, we have $A_{n-j} = A_j'$.

[8] See Gray (2006, Section 4.4) and the references therein.

The Whittle log-likelihood is obtained by replacing $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$ in (3.1)-(3.2) with $\boldsymbol{\Omega}_T(\boldsymbol{\theta})$, as follows:[9]

$$\ell_w(\boldsymbol{\theta}; \boldsymbol{I}_T) = -\frac{1}{2}\log\det(\boldsymbol{S}_T(\boldsymbol{\theta})) - \frac{1}{2}(\boldsymbol{F}_T\boldsymbol{Y}_T)^*\boldsymbol{S}^{-1}(\boldsymbol{\theta})(\boldsymbol{F}_T\boldsymbol{Y}_T) \tag{3.7}$$

$$= -\frac{1}{2}\sum_{j=1}^{T}\Big\{\log\det(\boldsymbol{s}(\boldsymbol{\theta},\omega_j)) + \boldsymbol{J}_T(\omega_j)^*\boldsymbol{s}^{-1}(\boldsymbol{\theta},\omega_j)\boldsymbol{J}_T(\omega_j)\Big\} \tag{3.8}$$

$$= -\frac{1}{2}\sum_{j=1}^{T}\Big\{\log\det(\boldsymbol{s}(\boldsymbol{\theta},\omega_j)) + \operatorname{tr}\Big(\boldsymbol{I}_T(\omega_j)\boldsymbol{s}^{-1}(\boldsymbol{\theta},\omega_j)\Big)\Big\} \tag{3.9}$$

This can be recognized as the log-likelihood function of a sample of T independent but not identically distributed $n_{\boldsymbol{y}}-$dimensional zero-mean complex Gaussian vectors, whose covariance matrices are given by the spectral density of $\boldsymbol{y}_t$ evaluated at the Fourier frequencies. Furthermore, since $\boldsymbol{I}_T(\omega_{T-j})$ and $\boldsymbol{s}(\boldsymbol{\theta},\omega_{T-j})$ are complex conjugates of the spectral density and periodogram evaluated at ω_j, only half of the terms in equations (3.8)-(3.9) need to be evaluated.

The Whittle log-likelihood is considerably simpler to evaluate than the expression in equations (3.1)-(3.2), as it avoids inverting potentially large covariance matrices and can leverage existing fast DFT algorithms. However, it only approximates the exact Gaussian log-likelihood because $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$ and $\boldsymbol{\Omega}_T(\boldsymbol{\theta})$ differ in finite samples. Moreover, other computationally efficient methods, such as the Kalman filter, can evaluate the exact likelihood for a large class of models without inverting large matrices. Thus, the computational advantage of the Whittle log-likelihood is less compelling now than when it was first proposed.[10]

Another appealing feature of the Whittle log-likelihood is its ability to facilitate estimation using only a subset of frequencies. This may be desirable in cases when data contains noise affecting only part of the periodogram, (e.g., the high frequencies) or when the theoretical model is intended to match data only within a specific frequency range. This feature is particularly relevant in macroeconomic research, where theoretical models often specifically target business cycle movements and are known to be misspecified at lower and higher frequencies. Consequently, one might prefer to fit the model to business cycle frequencies while ignoring those at the spectrum's lower and higher ends. In practice, this can be achieved by performing the summation in (3.8)–(3.9) over the frequencies of interest,

$$\ell_w(\boldsymbol{\theta}; \boldsymbol{I}_T^{\bar{\boldsymbol{\omega}}}) = -\frac{1}{2}\sum_{\omega\in\bar{\boldsymbol{\omega}}}\log\det(\boldsymbol{s}(\boldsymbol{\theta},\omega)) + \operatorname{tr}\Big(\boldsymbol{I}_T(\omega)\boldsymbol{s}^{-1}(\boldsymbol{\theta},\omega)\Big) \tag{3.10}$$

where $\bar{\boldsymbol{\omega}}$ denotes the set of included frequencies and is, in general, a set of disjoint intervals from $\{0, 2\pi/T, \ldots, 2\pi(T-1)/T\}$ such that if $\omega \in \bar{\boldsymbol{\omega}}$ then $2\pi - \omega \in \bar{\boldsymbol{\omega}}$.

[9]Note that this is a discretized version of the log-likelihood using the Riemann sum as an approximation of an integral in the original expression of Whittle (1953).

[10]There are, nevertheless, models for which the Whittle approximation remains the preferred approach for computational efficiency reasons.

3.2 Linearized DSGE models

Evaluating the Whittle log-likelihood function requires only computing the model-implied spectral density matrix of the observed variables at any frequency ω as a function of the model parameters $\boldsymbol{\theta}$. In general, a linearized DSGE model can be expressed as a recursive equilibrium law of motion through the following system of equations:

$$\begin{aligned} \boldsymbol{y}_t &= \boldsymbol{C}(\boldsymbol{\theta})\boldsymbol{v}_{t-1} + \boldsymbol{D}(\boldsymbol{\theta})\boldsymbol{u}_t && (3.11) \\ \boldsymbol{v}_t &= \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{v}_{t-1} + \boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{u}_t && (3.12) \\ \boldsymbol{u}_t &= \boldsymbol{G}(\boldsymbol{\theta})\boldsymbol{u}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}(\boldsymbol{\theta})\right) && (3.13) \end{aligned}$$

where $\boldsymbol{y}_t$ is a $n_{\boldsymbol{y}}$ vector of observed variables, $\boldsymbol{v}_t$ is a $n_{\boldsymbol{v}}$ vector of endogenous state variables, $\boldsymbol{u}_t$ is a $n_{\boldsymbol{u}}$ vector of exogenous state variables, and $\boldsymbol{\varepsilon}_t$ is a $n_{\boldsymbol{u}}$ vector of exogenous shocks. The matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, and $\boldsymbol{G}$ are functions of the structural parameters of the model, collected in the $n_{\boldsymbol{\theta}}$ vector $\boldsymbol{\theta}$.

Then, the spectral density matrix of the observed variables $\boldsymbol{y}_t$ is given by (see Uhlig (1999)):

$$\boldsymbol{s}_{\boldsymbol{yy}}(\boldsymbol{\theta}, \omega) = \frac{1}{2\pi}\boldsymbol{W}(\omega, \boldsymbol{\theta})\boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}(\boldsymbol{\theta})\boldsymbol{W}(\omega, \boldsymbol{\theta})^* \tag{3.14}$$

where

$$\boldsymbol{W}(\omega, \boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{C}(\boldsymbol{\theta})e^{-i\omega} & \boldsymbol{D}(\boldsymbol{\theta}) \\ \boldsymbol{I}_{n_{\boldsymbol{v}}} & \mathbf{O}_{n_{\boldsymbol{v}}, n_{\boldsymbol{u}}} \end{bmatrix} \begin{bmatrix} \left(\boldsymbol{I}_{n_{\boldsymbol{v}}} - \boldsymbol{A}(\boldsymbol{\theta})e^{-i\omega}\right)^{-1}\boldsymbol{B}(\boldsymbol{\theta})\left(\boldsymbol{I}_{n_{\boldsymbol{u}}} - \boldsymbol{G}(\boldsymbol{\theta})e^{-i\omega}\right)^{-1} \\ \left(\boldsymbol{I}_{n_{\boldsymbol{u}}} - \boldsymbol{G}(\boldsymbol{\theta})e^{-i\omega}\right)^{-1} \end{bmatrix}$$

Using the above expression for $\boldsymbol{s}_{\boldsymbol{yy}}(\boldsymbol{\theta}, \omega)$ in (3.9) or (3.10) yields the full spectrum or band spectral Whittle log-likelihood function of the set of observed variables $\boldsymbol{y}$.

4 Simulation Study

In this section, I use Monte Carlo simulations to investigate the finite sample performance of three estimators: (1) the time domain MLE using the exact likelihood, (2) the frequency domain MLE using all frequencies, and (3) the frequency domain MLE using only business cycle frequencies. The latter two estimators employ the Whittle approximation of the likelihood introduced in Section 3. The time domain MLE uses the exact Gaussian likelihood function evaluated via the Kalman filter. For brevity, I will refer to these estimators hereafter as TD, FD, and BC.

4.1 Setup

The Monte Carlo simulation proceeds as follows:

1. Solve the model from Section 2 using the ACD algorithm at the parameter values shown in Table 7.
2. Using the Gaussian linear state space representation of the model solution, generate sample trajectories of size T for the six observed variables: GDP (y), consumption (c), investment (i), hours worked (h), the inflation rate (π), and the federal funds rate (r).

3. Estimate the 25 free model parameters with the three estimators by maximizing their respective log-likelihood functions.
4. Repeat steps 2 and 3 N times.

The sample size is set to $T = 192$ observations, matching ACD's sample size. The simulation is initialized from the stationary distribution of the variables and the first 500 observations are discarded to eliminate the dependence on initial conditions. The number of replications is $N = 1000$. The numerical optimization in step (2) uses the true parameter values as a starting point and combines global and local optimization algorithms. This optimization process is identical for all three estimators.

4.2 Results: baseline parametrization

Table 1: Monte Carlo results

		Mean			Median			IQR		
parameter	true	TD	FD	BC	TD	FD	BC	TD	FD	BC
ν	0.28	0.28	0.33	0.42	0.28	0.33	0.40	[0.22, 0.33]	[0.26, 0.40]	[0.29, 0.52]
α	0.26	0.25	0.25	0.25	0.25	0.25	0.25	[0.25, 0.26]	[0.24, 0.26]	[0.24, 0.27]
ψ	0.50	0.51	0.45	0.43	0.50	0.44	0.38	[0.43, 0.58]	[0.36, 0.53]	[0.26, 0.54]
φ	3.31	3.25	2.31	2.50	3.17	2.09	2.11	[2.67, 3.71]	[1.45, 2.95]	[1.49, 3.09]
b	0.76	0.75	0.60	0.64	0.75	0.55	0.64	[0.73, 0.77]	[0.51, 0.69]	[0.55, 0.72]
χ	0.73	0.72	0.70	0.74	0.72	0.71	0.74	[0.71, 0.74]	[0.68, 0.73]	[0.70, 0.78]
κ_R	0.20	0.20	0.14	0.12	0.19	0.14	0.03	[0.14, 0.25]	[0.08, 0.20]	[0.00, 0.20]
κ_π	2.27	2.23	2.02	2.18	2.18	1.95	1.94	[1.96, 2.46]	[1.67, 2.33]	[1.42, 2.58]
κ_y	0.12	0.16	0.24	0.16	0.14	0.20	0.13	[0.11, 0.18]	[0.13, 0.31]	[0.08, 0.20]
ρ_a	0.41	0.40	0.58	0.58	0.41	0.62	0.63	[0.24, 0.56]	[0.38, 0.80]	[0.23, 0.97]
ρ_n	0.22	0.22	0.39	0.18	0.20	0.33	0.03	[0.07, 0.32]	[0.15, 0.60]	[0.00, 0.32]
ρ_i	0.37	0.35	0.38	0.27	0.35	0.37	0.26	[0.30, 0.40]	[0.31, 0.45]	[0.04, 0.46]
ρ_c	0.89	0.88	0.94	0.93	0.88	0.95	0.95	[0.86, 0.90]	[0.90, 0.99]	[0.90, 0.98]
ρ_g	0.79	0.76	0.76	0.73	0.77	0.77	0.74	[0.72, 0.80]	[0.71, 0.81]	[0.63, 0.86]
ρ_m	0.65	0.63	0.64	0.62	0.63	0.64	0.64	[0.59, 0.66]	[0.59, 0.68]	[0.55, 0.72]
ρ_ξ	0.83	0.81	0.77	0.85	0.82	0.79	0.88	[0.79, 0.84]	[0.72, 0.84]	[0.79, 0.93]
σ_{aP}	0.41	0.38	0.82	0.51	0.39	0.70	0.49	[0.33, 0.45]	[0.49, 1.08]	[0.20, 0.71]
σ_{aT}	0.35	0.34	0.36	0.41	0.34	0.34	0.40	[0.30, 0.39]	[0.27, 0.43]	[0.27, 0.55]
σ_n	0.38	0.38	0.26	0.60	0.38	0.29	0.56	[0.32, 0.44]	[0.14, 0.38]	[0.33, 0.83]
σ_{iP}	0.61	0.55	0.66	0.60	0.57	0.70	0.17	[0.31, 0.78]	[0.27, 1.00]	[0.03, 1.13]
σ_{iT}	5.80	5.77	4.20	5.57	5.56	3.79	4.54	[4.73, 6.63]	[2.57, 5.38]	[2.71, 7.04]
σ_c	0.36	0.36	1.03	0.60	0.35	0.53	0.48	[0.25, 0.46]	[0.35, 1.16]	[0.26, 0.77]
σ_g	1.71	1.68	1.69	1.71	1.69	1.70	1.69	[1.63, 1.74]	[1.62, 1.77]	[1.56, 1.85]
σ_m	0.31	0.30	0.32	0.33	0.30	0.31	0.31	[0.28, 0.32]	[0.28, 0.34]	[0.27, 0.37]
σ_ξ	0.61	0.66	0.45	0.36	0.60	0.41	0.17	[0.44, 0.81]	[0.26, 0.58]	[0.03, 0.48]

Note: Monte Carlo performance of the TD, FD, and BC estimators at the baseline parametrization of the ACD model. The estimates of the mean, median and interquartile range (IQR) are based on 1000 MC replications with a sample size of $T = 192$.

Table 1 presents the first set of results, reporting the mean, median, and interquartile range (IQR) for each estimated parameter. Comparing these statistics with the true parameter values (first column) reveals the relative performance of the three estimators. As expected, the TD estimator performs best in both point estimate accuracy and variability around true values. The two Whittle likelihood-based estimators perform significantly worse for most parameters, with the BC estimator generally showing lower accuracy and higher volatility. For instance, with

few exceptions, the BC estimator's IQR is the widest, while the TD estimator's is consistently the narrowest, often by a substantial margin. Despite their width, the FD and BC estimators' IQRs sometimes exclude the true parameter values. Specifically, both estimators' 25th percentile exceeds the true value of ρ_c, while their 75th percentile falls below the true values of σ_ξ, φ, and b. The FD estimator shows additional misalignments: its IQR lies above the true values of κ_y and σ_{a^P}, and below those of κ_R and σ_{i^T}.[11] In contrast, the TD estimates' IQR consistently includes true parameter values. Figure 1 illustrates these findings using boxplots that display, for each parameter, the distribution of estimates expressed as percentage deviations from their true values.

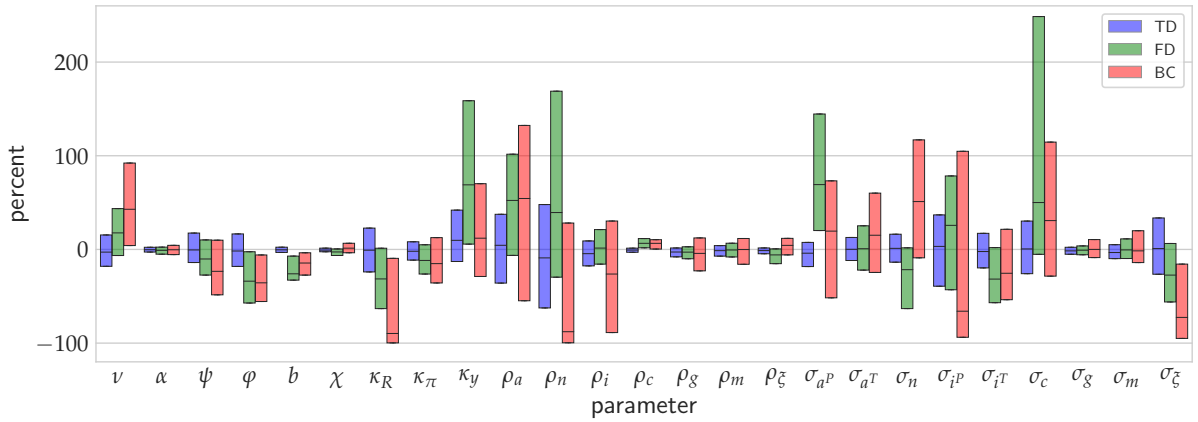


Figure 1: Boxplots of parameter estimate deviations (as percentages of the true values). Each box represents the interquartile range of the estimates obtained using the respective estimator. The vertical bars within each box indicate the median estimate. The results are based on 1000 MC replications with a sample size of $T = 192$.

These observations suggest that the Whittle likelihood-based estimators may exhibit a significant estimation bias. This is further corroborated by Table 2, which reports the bias, standard deviation (SD) and root mean squared error (RMSE) for each parameter across the 1000 replications. To enable comparisons across parameters, all statistics are expressed as percentages of their true values. On average, the bias of the TD estimates is about 4% in absolute value, compared to 30% and 21% for FD and BC, respectively. The TD estimator's largest bias is approximately 28% for κ_y. In contrast, FD shows the highest bias of about 187% for σ_c, with two other parameters (σ_{a^p} and κ_y) approaching 100% bias. The BC estimator's largest bias is also for σ_c, at slightly below 100%. The two Whittle likelihood-based estimators exhibit moderate agreement in identifying relatively more or less biased parameters, with a Spearman rank correlation of .47. In contrast, TD and FD show a negative correlation ($-.3$), while TD and BC exhibit a very weak positive correlation (.1). The SDs reported in the middle panel measures the variability of the estimates. On average, the TD estimates show roughly half the variability of the FD estimates and one-third that of the BC estimates. For individual parameters, the FD estimates are generally less variable than the BC estimates, while the TD estimates consistently exhibit the lowest standard deviation. Certain common patterns emerge across

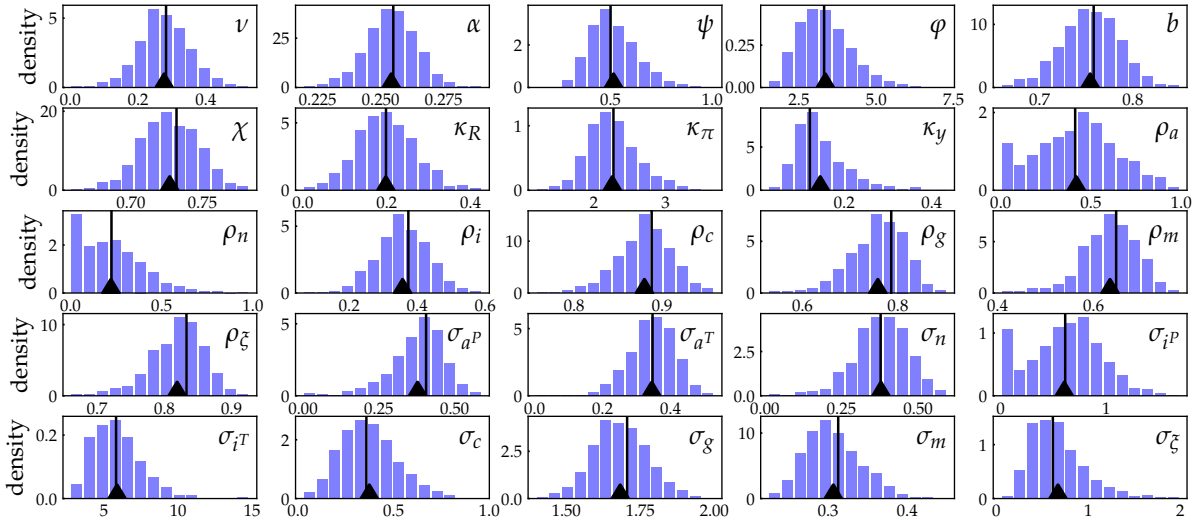
[11] Note that the figures in the table are rounded to the second digit. The true value of κ_R is 0.198 while the 75th percentile of FD is 0.197689.

Table 2: Monte Carlo results (cont.)

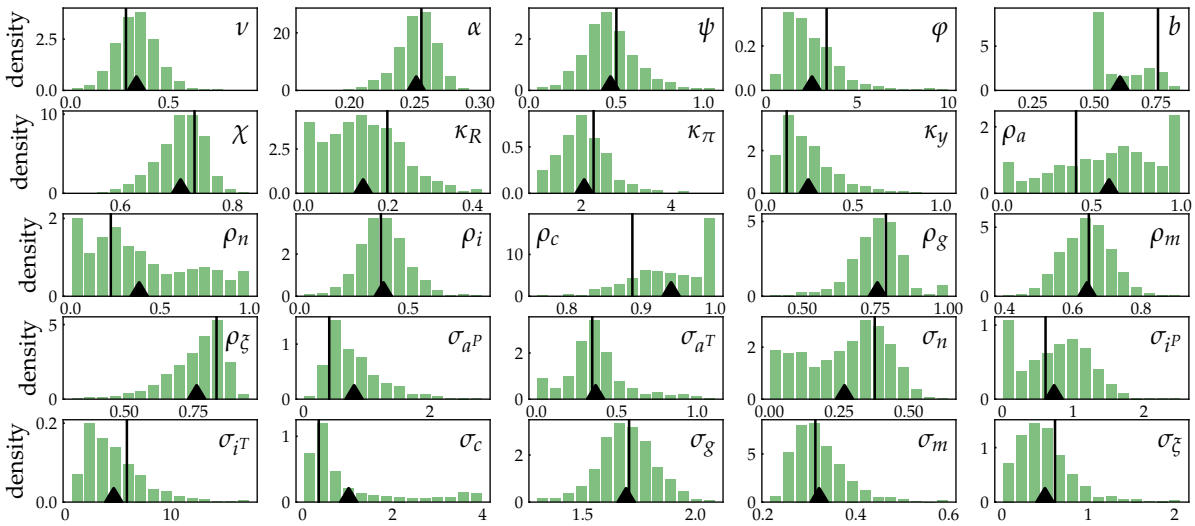
	Bias (%)			SD (%)			RMSE (%)		
parameter	TD	FD	BC	TD	FD	BC	TD	FD	BC
ν	-2.3	17.9	50.4	26.4	36.0	81.3	26.5	40.2	95.7
α	-0.3	-1.8	-0.4	4.2	5.9	7.0	4.2	6.2	7.1
ψ	2.5	-10.7	-13.7	25.1	27.5	49.4	25.2	29.5	51.3
φ	-2.0	-30.4	-24.6	24.4	33.2	44.4	24.5	45.0	50.8
b	-0.8	-21.0	-15.0	4.1	13.1	13.1	4.2	24.8	19.9
χ	-1.2	-3.8	1.0	2.9	5.4	8.4	3.1	6.6	8.5
κ_R	-0.4	-29.8	-41.7	37.5	42.7	80.7	37.5	52.1	90.8
κ_π	-1.9	-11.2	-3.8	16.8	21.8	46.3	16.9	24.5	46.5
κ_y	28.3	96.0	34.3	57.5	122.2	102.6	64.1	155.4	108.2
ρ_a	-2.5	41.7	40.1	55.3	70.3	91.2	55.4	81.7	99.7
ρ_n	-3.5	72.8	-20.1	77.4	128.0	112.7	77.5	147.3	114.4
ρ_i	-6.0	2.0	-26.5	20.5	28.1	62.0	21.3	28.2	67.4
ρ_c	-1.3	6.0	4.9	3.7	5.7	6.9	3.9	8.2	8.4
ρ_g	-3.4	-3.6	-6.5	7.5	11.1	23.8	8.2	11.7	24.6
ρ_m	-3.1	-1.7	-3.4	9.0	10.2	23.0	9.6	10.4	23.3
ρ_ξ	-2.4	-7.9	1.7	5.1	11.8	12.9	5.6	14.2	13.1
σ_{a^P}	-5.8	101.4	25.1	23.2	105.3	95.2	23.9	146.2	98.5
σ_{a^T}	-1.2	2.9	17.9	19.5	52.1	66.5	19.6	52.2	68.8
σ_n	-0.6	-29.9	58.8	24.2	37.8	96.7	24.2	48.2	113.2
σ_{i^P}	-10.0	7.5	-2.2	55.2	73.1	115.8	56.1	73.5	115.8
σ_{i^T}	-0.7	-27.6	-4.1	26.5	36.7	70.0	26.5	45.9	70.1
σ_c	1.3	189.7	67.0	42.4	304.7	134.0	42.4	358.9	149.8
σ_g	-1.2	-0.8	0.4	5.3	7.4	14.4	5.4	7.4	14.4
σ_m	-2.9	1.0	6.3	11.5	15.8	32.9	11.9	15.9	33.5
σ_ξ	7.8	-27.4	-41.8	49.9	42.7	77.3	50.5	50.7	87.9

Note: Average percentage bias, standard deviation and root mean square error relative to the modulus of the true parameter value. The results are based on 1000 MC replications with a sample size of $T = 192$.

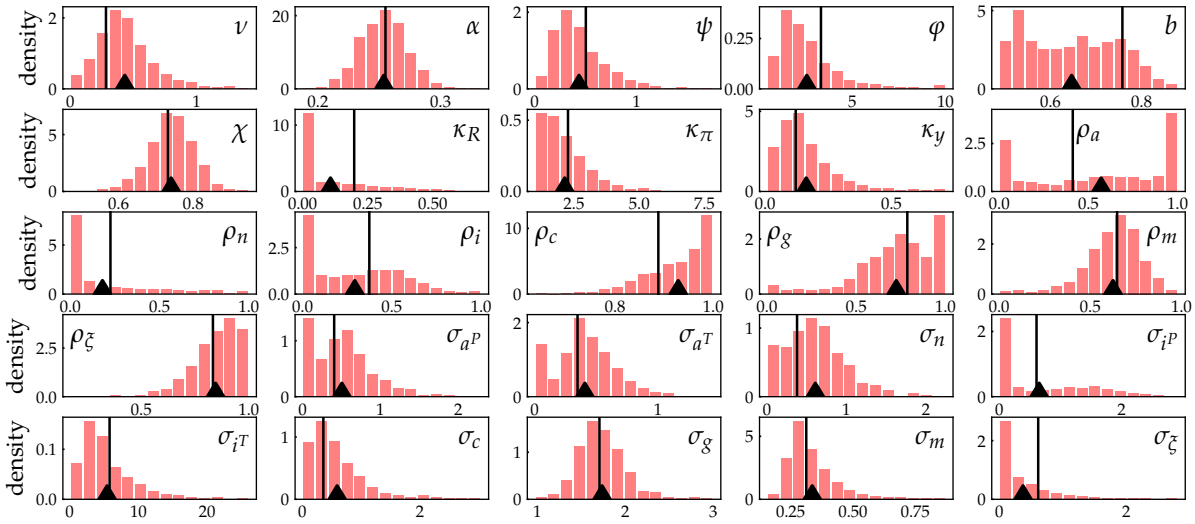
estimators regarding the relative variability of the parameters: the estimates of α, χ, ρ_c, and σ_g show the least variation across all estimators, while the estimates of κ_y, ρ_n, σ_c, and σ_{i^P} are among the most variable. The rank correlation between SDs of the TD and either one of the Whittle likelihood-based estimators is approximately .9, while that between the FD and the BC reaches .93. The last three columns of Table 2 show the normalized RMSE, which measures overall estimation accuracy by accounting for both bias and variability. The RMSE patterns largely mirror those of the standard deviations, but the TD estimator's superiority over the Whittle likelihood-based estimators is even more pronounced, reflecting the larger biases affecting the latter. Rank correlations across estimator pairs remain close to .9, slightly weaker than those for SDs. The parameters χ, ρ_ξ, σ_g, α, and ρ_c have the lowest RMSE across all estimators, while κ_y, ρ_n, σ_c, and σ_{i^P} have the highest. Notably, even TD's least accurately estimated parameters achieve lower RMSE than many FD or BC estimates. For instance, 10 of BC's 25 parameter estimates exceed TD's highest RMSE value, observed for ρ_n." Figure 2 displays histograms of the sampling distributions of the TD, FD, and BC estimates, further demonstrating the TD estimator's superiority over the Whittle likelihood-based estimators. The distributions of the TD estimates (panel (a)) are generally unimodal, approximately symmetric, and reasonably well centered at the true parameter values, with three exceptions: ρ_a, ρ_n, and σ_{i^P}. For these parameters, large concentrations of estimates occur at the lower boundary of the parameter



(a) Time domain MLE.



(b) Frequency domain MLE using all frequencies.



(c) Frequency domain MLE using only business cycle frequencies.

Figure 2: Sampling distributions of the estimated parameters. The ▲ symbol shows the estimate of the mean and the black vertical line indicates the true value. The results are based on 1000 MC replications with a sample size of $T = 192$.

space, resulting in negative mean and median biases (see Tables 1 and 2), with σ_{iP} being the most severely affected. The FD (panel (b)) and BC (panel (c)) estimators perform notably worse, with many parameters showing multimodality, skewness, and concentration of estimates at the boundaries, far from the true values. These departures from normality are more clearly seen in the Q-Q plots in Figure 11 of the Appendix. While exact normality is not expected with 192 observations, the TD estimates generally align better with the 45-degree line and the theoretical normal quantiles compared to the frequency domain estimators.

Discussion

The Monte Carlo results strongly favor the time domain estimator over both frequency domain estimators. The superiority of the exact likelihood-based estimator is not surprising, given the assumption of correct model specification. However, the notably poor performance of the Whittle likelihood-based estimators warrants further investigation, which is the focus of this section.

The discussion in Section 3 (see equations 3.3 – 3.9) outlines two equivalent interpretations of the Whittle likelihood approximation. First, it approximates the likelihood function of the time-domain data vector $\boldsymbol{Y}_T$ by replacing its covariance matrix $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$ with a block-circulant matrix $\boldsymbol{\Omega}_T(\boldsymbol{\theta})$. Alternatively, it approximates the likelihood function of the Fourier-transformed data vector $\boldsymbol{F}_T\boldsymbol{Y}_T$, replacing its covariance matrix $\boldsymbol{F}_T\boldsymbol{\Sigma}_T(\boldsymbol{\theta})\boldsymbol{F}_T^*$ with a block diagonal matrix $\boldsymbol{S}_T(\boldsymbol{\theta})$. The vector $\boldsymbol{Y}(\boldsymbol{\omega}) = \boldsymbol{F}_T\boldsymbol{Y}_T$, whose i-th block element is

$$\boldsymbol{y}(\omega_i) = [y(\omega_i), c(\omega_i), i(\omega_i), h(\omega_i), \pi(\omega_i), r(\omega_i)]', \tag{4.1}$$

is the frequency domain representation of the observed data, organized by frequency $\omega \in \{0, 2\pi/T, \ldots, 2\pi(T-1)/T\}$, rather than time $t \in \{1, 2, \ldots, T\}$. To obtain the band spectral likelihood function, one selects the frequencies of interest, i.e. the business cycle frequencies. This amounts to replacing the full sample Fourier matrix $\boldsymbol{F}_T$ with a Fourier matrix $\boldsymbol{F}_{bc}$ that extracts these frequencies from the original data vector. The exact band spectral likelihood function's covariance matrix is then $\boldsymbol{F}_{bc}\boldsymbol{\Sigma}_T(\boldsymbol{\theta})\boldsymbol{F}_{bc}^*$, while its Whittle approximation uses the corresponding submatrix of $\boldsymbol{S}_T(\boldsymbol{\theta})$.[12]

A block diagonal covariance matrix implies that observations at different frequencies are uncorrelated. For example, any two frequency components of output, $y(\omega_j)$ and $y(\omega_k)$, should be uncorrelated for all $\omega_j \neq \omega_k$, and this holds for both autocorrelations of individual variables and cross-correlations among them. A simple way to assess the appropriateness of the Whittle approximation is to check whether this is true for the correlation matrix in the exact frequency domain likelihood function. Figure 3 compares the respective correlation matrices of the full spectrum and band spectral likelihood functions. It is clear that the off-diagonal blocks are not all zero. The largest cross-frequency correlations (in absolute value) are observed between the frequency components of y, c, and, to a lesser extent, i. This is easier to see in panel (d) of the figure, which shows a large number of nonzero off-diagonal 3×3 blocks, corresponding to the correlation matrices of these variables at different frequencies from the BC part of the

[12] Let $\boldsymbol{K}$ be a selection matrix of zeros and ones such that $\boldsymbol{K}\boldsymbol{S}_T(\boldsymbol{\theta})\boldsymbol{K}'$ selects the BC frequency submatrix of $\boldsymbol{S}_T(\boldsymbol{\theta})$. Then $\boldsymbol{F}_{bc} = \boldsymbol{K}\boldsymbol{F}_T$.

spectrum. Note that all correlations smaller than .1 in absolute value have been zeroed-out for clarity. The strongest correlations are between different frequency components of c (up to .93), between c and y (up to .8), and between components of y (up to .73). The autocorrelations and cross-correlations involving i are weaker, around 0.4, while the variables h, π, and r exhibit no pairwise correlations exceeding 0.1. As explained in Section 3, the Whittle likelihood is valid as

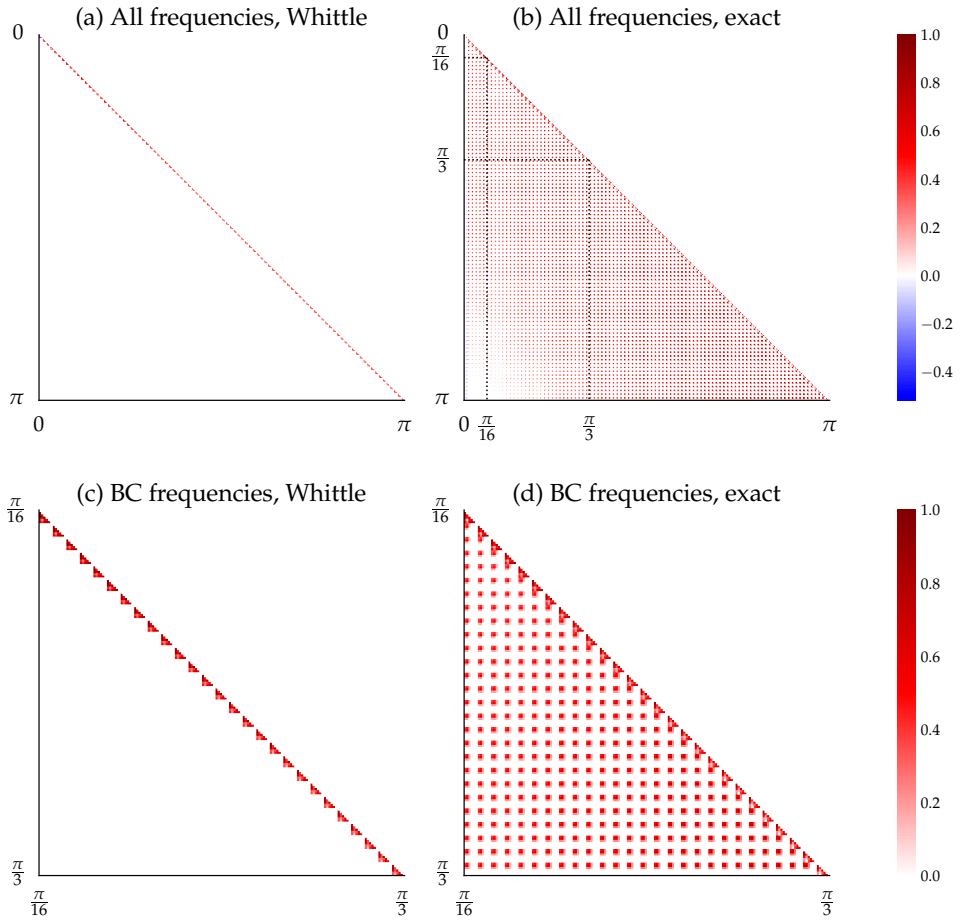


Figure 3: Correlation matrices in the Whittle approximation and the exact frequency domain likelihood function. Because of symmetry, only the lower triangles of the matrices and frequencies from the $[0, \pi]$ part of the spectrum are displayed. Frequencies between $\pi/16$ and $\pi/3$ are business cycle frequencies. Correlation coefficients smaller than .1 in absolute value have been zeroed out.

an asymptotic approximation of the true Gaussian likelihood function only when the process is stationary. For non-stationary processes, different frequency components of the periodogram lose their asymptotic independence. The strong correlation patterns observed in Figure 3 may therefore stem from the high persistence of the variables in the model. Indeed, while all eigenvalues of matrix $\boldsymbol{G}$ in (3.13) lie inside the unit circle – ensuring the stationarity of $\boldsymbol{y}_t$ – two of them equal 0.999999. This near-unit root characteristic arises from how the model in Section 2 is implemented in the code. While a_t^p and ζ_t^{IP} are modeled as permanent, they are represented in the code as AR(1) processes with autoregressive coefficients of 0.999999. As a result, these shocks are technically stationary but exhibit extreme persistence. Given the crucial role of a_t^p in driving the dynamics of y_t, c_t, and i_t,[13] these variables inherit its persistence. For example, their

[13]For instance, in the variance decompositions a_t^p contributes most of the volatility of these variables.

autocorrelations remain near .999 even at lag 1000, whereas the other observed variables – h_t, π_t, and r_t – show autocorrelations below .5 by the 10th lag.

These findings suggest an explanation for the poor performance of Whittle likelihood-based estimators: the underlying data generating process prevents accurate approximation by the Whittle likelihood, resulting in larger bias and RMSE compared to estimators using the true likelihood function. It remains unclear why some parameters are more affected by misspecification than others. This is likely tied to the distribution of parameter-specific information across different regions of the likelihood function and the extent to which those regions are influenced by misspecification. This is an interesting question that will be explored in future research. Another interesting question left for future work is whether the exact band spectral likelihood, alluded to above, can be a viable alternative to the Whittle likelihood approximation. Since the band spectral case is simply a linear transformation of the full information time domain Gaussian density, the exact band spectral likelihood function is, in principle, straightforward to compute. However, in practice, it is computationally very expensive, both in terms of time and memory consumption. This makes it impracticable for most cases of interest. It remains to be determined whether a more efficient implementation can be found.

The question I will focus on here is how the quality of the Whittle approximation is affected by less persistent processes for a_t^p and ζ_t^{IP}. To investigate this, I set their autoregressive coefficients to 0.9, matching the persistence of the next most persistent shock in the model – the a discount-rate/preference shock ζ_t^c (see Table 7). Under this parameterization, the persistence of y_t, c_t, and i_t is similar to, though slightly higher than, that of h_t, π_t, and r_t. Figure 4 shows the correlation matrices of the frequency domain likelihoods for this case. As before, the off-diagonal blocks are non-zero, but the correlations are noticeably weaker, with a maximum value of 0.48. The strongest correlations are again between different frequency components of c, followed by those of y and i. This reduced deviation from a block-diagonal correlation matrix suggests that the Whittle approximation may be more appropriate for the data generating process under this alternative parameterization. In the next section, I evaluate how this change affects the relative performance of the Whittle likelihood-based estimators.

4.3 Results: alternative parametrization

The results for the second set of simulations are organized similarly to the previous section. Table 3 reports the mean, median, and IQR for all parameters. While the TD estimator remains the most accurate and least volatile, the Whittle likelihood-based estimators show notable improvement along these dimensions. Notably, their IQRs now include the true values for all parameters. The TD estimator consistently produces the shortest IQRs, while the BC estimator yields the longest. However, the average difference between the TD and FD estimators' IQRs is only about 22%, in line with theory, given their asymptotic equivalence. In contrast, the BC estimator's IQRs are on average about 80% wider than those of FD and 130% wider than those of TD. Figure 5 provides a visual representation, displaying boxplots of the percent deviations of parameter estimates from their true values. A key observation is that all bars now cross the horizontal line at 0, and the TD and FD bars are much more comparable in length, in contrast to

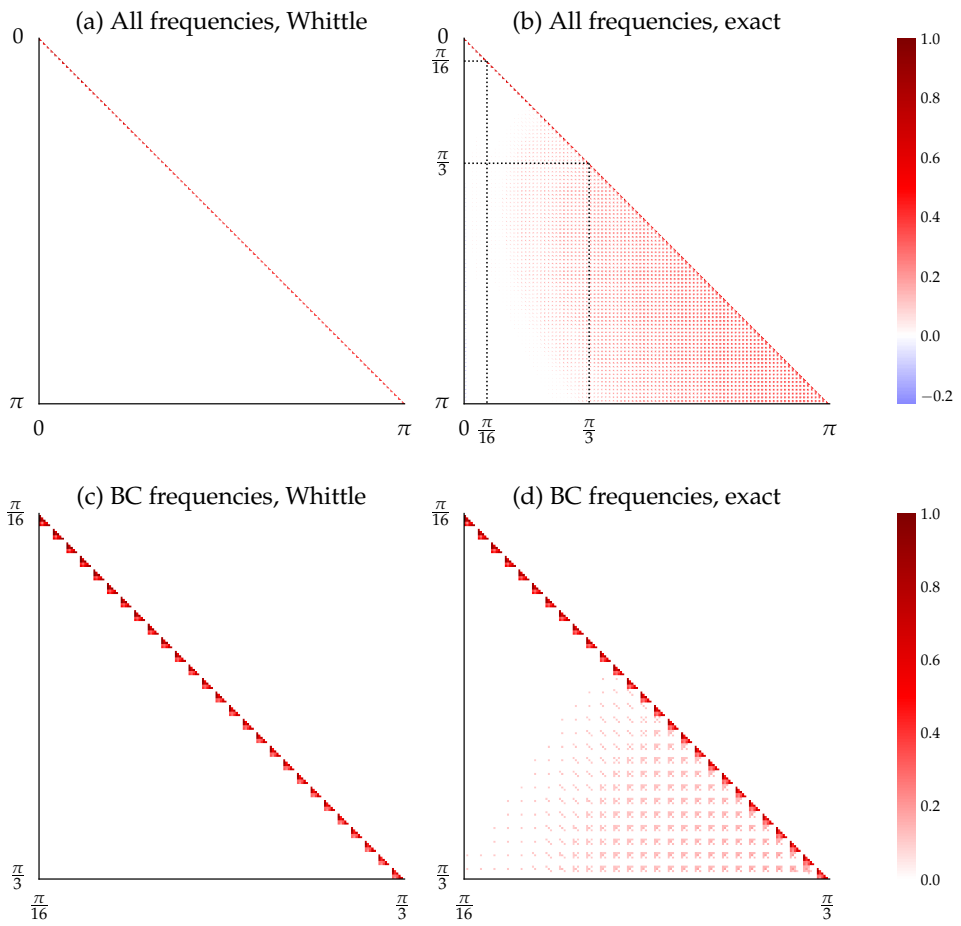


Figure 4: Correlation matrices in the Whittle approximation and the exact frequency domain likelihood function under the alternative parametrization for the a_t^p and ζ_t^{IP} processes.

Figure 1. Table 4 reports the bias, SD, and RMSE for all parameters, expressed as percentages

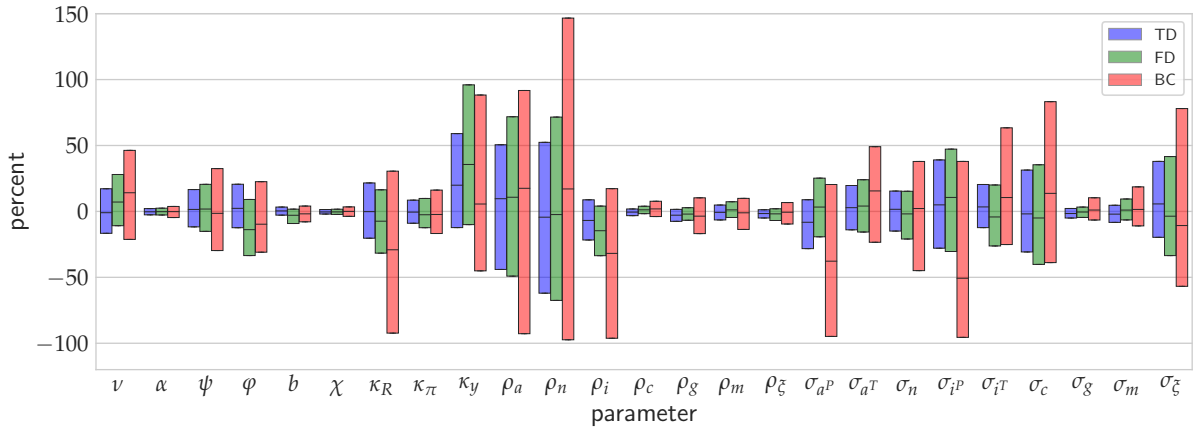


Figure 5: Alternative parametrization of the ACD model. See the Notes to Figure 1.

relative to their true values. The average absolute bias is approximately 4%, 7%, and 13% for the TD, FD, and BC estimates, respectively. While the TD estimator maintains the same average bias as in the baseline parameterization, the Whittle likelihood-based estimators show substantially reduced bias compared to their previous performance. Furthermore, there is now

Table 3: Monte Carlo results

		Mean			Median			IQR		
parameter	true	TD	FD	BC	TD	FD	BC	TD	FD	BC
ν	0.28	0.28	0.31	0.32	0.28	0.30	0.32	[0.24, 0.33]	[0.25, 0.36]	[0.22, 0.41]
α	0.26	0.25	0.25	0.25	0.25	0.25	0.25	[0.25, 0.26]	[0.25, 0.26]	[0.24, 0.26]
ψ	0.50	0.52	0.52	0.54	0.51	0.51	0.49	[0.44, 0.58]	[0.42, 0.60]	[0.35, 0.66]
φ	3.31	3.50	2.98	3.36	3.39	2.86	2.99	[2.90, 4.00]	[2.20, 3.61]	[2.29, 4.06]
b	0.76	0.76	0.73	0.74	0.76	0.74	0.74	[0.74, 0.78]	[0.69, 0.77]	[0.70, 0.79]
χ	0.73	0.73	0.73	0.73	0.73	0.73	0.73	[0.72, 0.74]	[0.72, 0.74]	[0.70, 0.76]
κ_R	0.20	0.20	0.18	0.16	0.20	0.18	0.14	[0.16, 0.24]	[0.14, 0.23]	[0.02, 0.26]
κ_π	2.27	2.27	2.25	2.35	2.26	2.21	2.22	[2.07, 2.46]	[1.99, 2.49]	[1.89, 2.64]
κ_y	0.12	0.16	0.19	0.17	0.15	0.16	0.13	[0.11, 0.19]	[0.11, 0.24]	[0.07, 0.23]
ρ_a	0.41	0.43	0.46	0.46	0.45	0.46	0.48	[0.23, 0.62]	[0.21, 0.71]	[0.03, 0.79]
ρ_n	0.22	0.23	0.26	0.32	0.21	0.22	0.26	[0.09, 0.34]	[0.07, 0.38]	[0.01, 0.55]
ρ_i	0.37	0.35	0.31	0.26	0.35	0.32	0.25	[0.29, 0.41]	[0.25, 0.39]	[0.01, 0.44]
ρ_c	0.89	0.88	0.90	0.90	0.88	0.90	0.90	[0.86, 0.90]	[0.87, 0.92]	[0.85, 0.96]
ρ_g	0.79	0.76	0.77	0.74	0.76	0.77	0.76	[0.73, 0.80]	[0.73, 0.81]	[0.65, 0.87]
ρ_m	0.65	0.64	0.65	0.62	0.64	0.65	0.64	[0.60, 0.68]	[0.62, 0.69]	[0.56, 0.71]
ρ_ξ	0.83	0.82	0.81	0.81	0.82	0.82	0.83	[0.79, 0.84]	[0.78, 0.85]	[0.75, 0.89]
σ_{aP}	0.41	0.35	0.41	0.27	0.37	0.42	0.25	[0.29, 0.44]	[0.33, 0.51]	[0.02, 0.49]
σ_{aT}	0.35	0.35	0.36	0.38	0.36	0.36	0.40	[0.30, 0.41]	[0.29, 0.43]	[0.27, 0.52]
σ_n	0.38	0.38	0.36	0.37	0.38	0.37	0.39	[0.32, 0.44]	[0.30, 0.44]	[0.21, 0.52]
σ_{iP}	0.61	0.61	0.65	0.48	0.64	0.68	0.30	[0.44, 0.85]	[0.43, 0.90]	[0.03, 0.84]
σ_{iT}	5.80	6.16	5.86	7.41	6.00	5.57	6.42	[5.09, 6.99]	[4.28, 6.97]	[4.35, 9.49]
σ_c	0.36	0.36	0.37	0.47	0.35	0.34	0.41	[0.25, 0.47]	[0.21, 0.48]	[0.22, 0.65]
σ_g	1.71	1.68	1.70	1.75	1.68	1.70	1.72	[1.62, 1.74]	[1.63, 1.76]	[1.59, 1.88]
σ_m	0.31	0.31	0.32	0.33	0.31	0.32	0.32	[0.29, 0.33]	[0.29, 0.34]	[0.28, 0.37]
σ_ξ	0.61	0.70	0.69	0.80	0.65	0.59	0.55	[0.49, 0.85]	[0.41, 0.87]	[0.26, 1.09]

Note: Alternative parametrization of the ACD model. See the Notes to Table 1.

stronger agreement among the three estimators regarding which parameters exhibit relatively higher or lower bias, as evidenced by the rank correlation coefficients: 0.56 between TD and FD estimators, and 0.89 between TD and BC estimators. The estimates of α and χ consistently show low bias across all estimators, while κ_y, ρ_i, ρ_n, and σ_ξ are consistently among the most biased parameters. With the substantially reduced biases, the SD and RMSE results now align more closely. The TD estimator's average RMSE is approximately 80% of the FD estimator's and 45% of the BC estimator's. The rank correlations between RMSEs are high: 0.98 between TD and FD estimators, and 0.95 between either of these and the BC estimator. The pattern of parameters with the lowest and highest RMSEs remains consistent with the previous findings (see Table 2). Specifically, χ, ρ_ξ, σ_g, α, ρ_c, and ρ_g exhibit the lowest RMSEs, while κ_y, ρ_a, ρ_n, σ_c, and σ_{iP}, and σ_ξ show the highest RMSE across all estimators.

Figure 6 displays histograms of the sampling distributions of the TD, FD, and BC estimates. As noted earlier, the relatively short sample means that we should not expect high accuracy in the normal approximation (see the Q-Q plots in Figure 12 of the Appendix). The differences between the sampling distributions of the time and frequency domain estimators reflect distortions caused by the Whittle approximation of the exact likelihood function. In this second parameterization, the improved quality of the Whittle approximation is evident in the substantially better sampling distributions for the FD and BC estimators across several parameters. The most striking improvements are seen for b (both the FD and BC estimators), and for κ_R, ρ_c and σ_c (FD estimator only). Less pronounced but still notable improvements appear in the sampling

Table 4: Monte Carlo results (cont.)

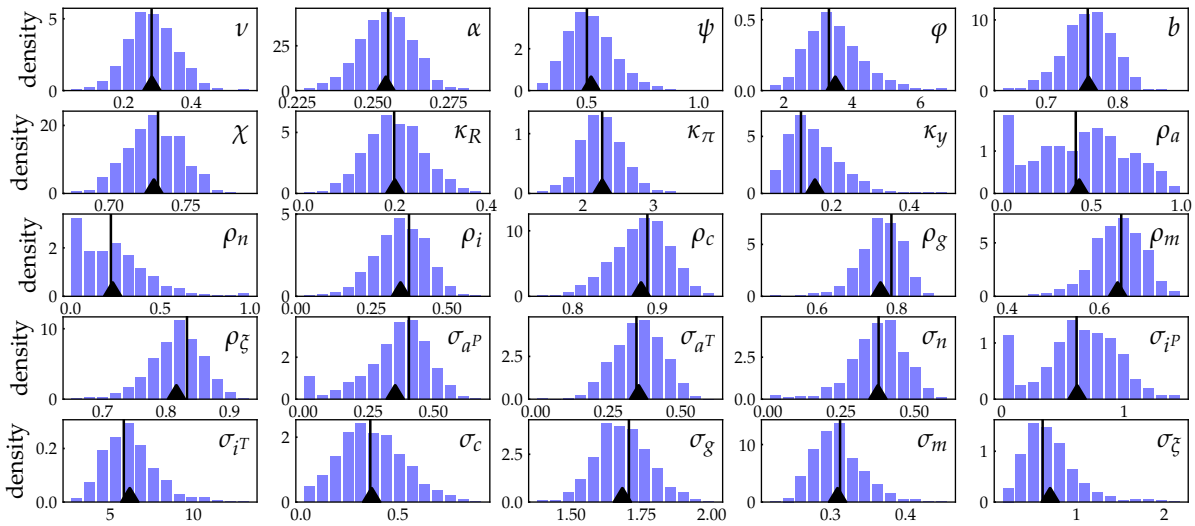
	Bias (%)			SD (%)			RMSE (%)		
parameter	TD	FD	BC	TD	FD	BC	TD	FD	BC
ν	0.2	8.3	14.1	25.8	31.1	53.8	25.8	32.1	55.6
α	-0.3	-0.2	-0.3	3.5	4.0	6.0	3.6	4.0	6.0
ψ	3.7	4.7	7.5	22.0	26.7	52.5	22.3	27.1	53.0
φ	5.6	-10.1	1.3	25.7	33.0	47.5	26.3	34.5	47.6
b	0.1	-4.3	-2.4	4.7	8.6	9.4	4.7	9.6	9.7
χ	-0.4	-0.3	-0.1	2.5	3.1	5.8	2.5	3.1	5.8
κ_R	0.7	-7.2	-19.9	32.4	36.2	72.5	32.4	36.9	75.2
κ_π	0.0	-0.7	3.5	14.0	17.9	31.2	14.0	17.9	31.4
κ_y	29.7	53.0	37.9	58.6	85.5	115.3	65.6	100.6	121.3
ρ_a	4.7	11.1	10.7	63.5	74.2	88.5	63.6	75.1	89.2
ρ_n	4.2	15.4	44.0	85.3	103.7	139.9	85.4	104.9	146.7
ρ_i	-7.5	-17.0	-29.6	23.3	29.2	61.2	24.5	33.8	67.9
ρ_c	-0.8	0.9	1.1	3.8	4.5	7.8	3.9	4.5	7.9
ρ_g	-3.5	-2.6	-5.3	7.2	7.6	22.2	8.0	8.1	22.8
ρ_m	-1.2	1.1	-3.9	8.5	9.3	22.3	8.6	9.3	22.6
ρ_ξ	-2.0	-3.2	-2.4	4.9	7.6	12.5	5.3	8.2	12.8
σ_{a^P}	-13.0	0.2	-32.8	32.6	36.8	59.0	35.1	36.8	67.4
σ_{a^T}	2.2	4.1	9.7	24.1	29.1	58.7	24.2	29.4	59.5
σ_n	-0.6	-5.9	-3.3	25.4	31.0	54.5	25.4	31.5	54.6
σ_{i^P}	0.7	7.3	-20.8	52.9	60.2	85.7	52.9	60.7	88.2
σ_{i^T}	6.2	0.9	27.6	27.0	38.0	77.4	27.7	38.0	82.1
σ_c	2.0	3.4	31.1	46.1	59.1	92.6	46.2	59.2	97.6
σ_g	-1.4	-0.4	2.4	5.6	5.9	13.9	5.8	6.0	14.1
σ_m	-1.1	2.6	6.8	10.5	13.2	26.6	10.5	13.5	27.5
σ_ξ	13.9	12.5	30.4	49.6	66.4	126.5	51.5	67.6	130.1

Note: Alternative parametrization of the ACD model. See the Notes to Table 2

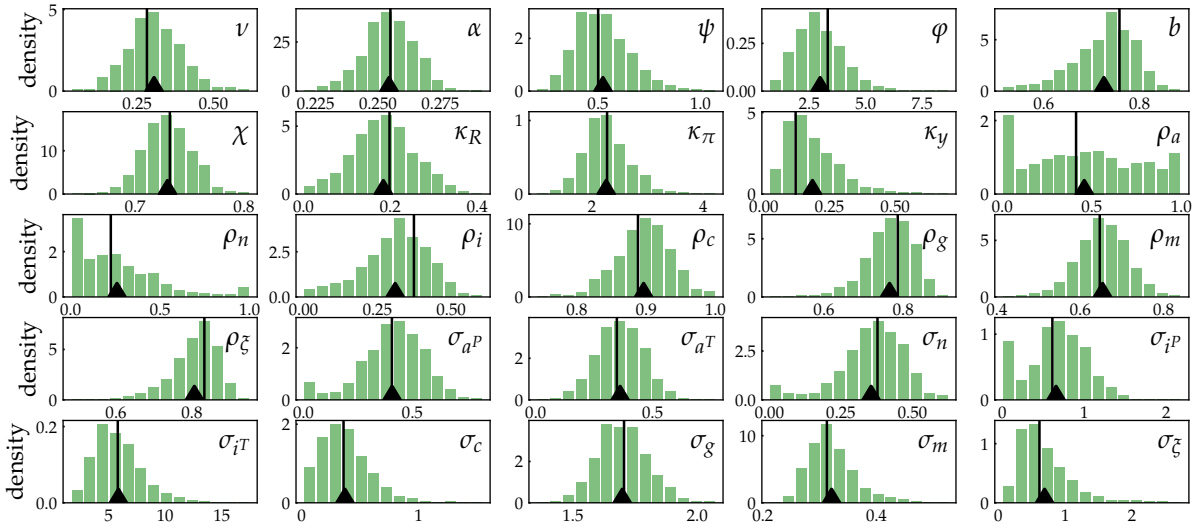
distributions of κ_π and ρ_ξ (both FD and BC estimators), as well as σ_{a^P}, σ_{a^T} and σ_n (FD estimator only).

The fact that, for the same set of parameters, the performance of the FD estimator improves notably while the BC estimator does not, underscores that Whittle likelihood-based band spectral estimators are susceptible to two distinct sources of distortion. The first, as discussed earlier, stems from the significant correlations among different frequency components of the observed variables. This distortion also affects the FD estimator, whose performance would otherwise align with that of the TD estimator. Therefore, comparing the sampling distributions of these two estimators reveals the extent to which different parameters are influenced by distortions from the Whittle approximation. A visual comparison of plots (a) and (b) in Figure 6 highlights substantial variation in this effect across parameters. For instance, the TD and FD estimates exhibit very similar sampling distributions for σ_g, ρ_n, and α, but noticeably different distributions for φ and b.

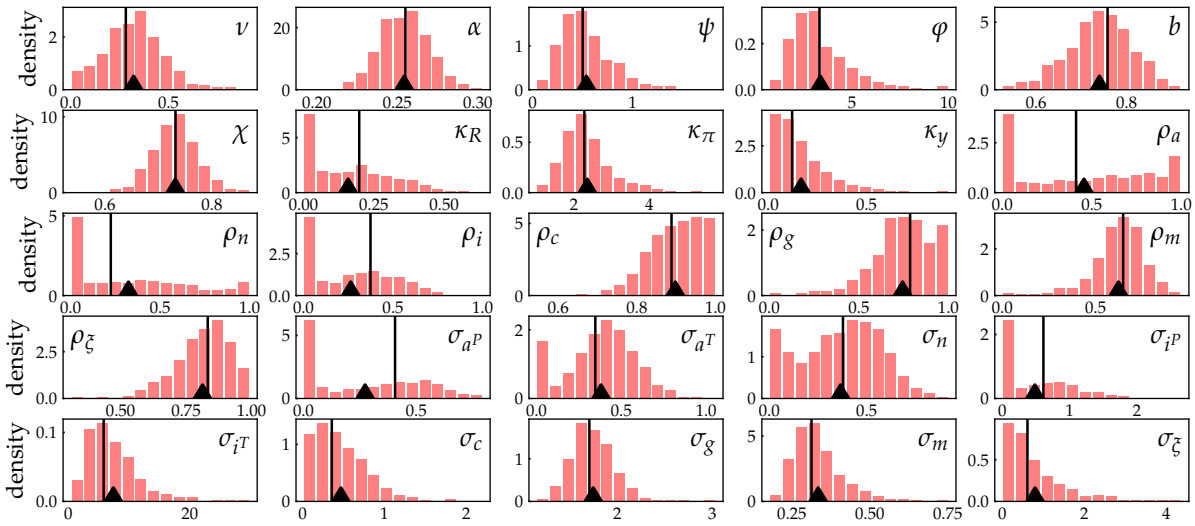
The second type of distortion arises from the reduced information content in the band spectral version of the likelihood, where $\bar{\boldsymbol{\omega}}$ in (3.10) represents only a subset of all frequencies. The BC estimator, in particular, suffers from this limitation, using less than 30% of all frequency domain observations – just 58 observations when $T = 192$. Absent such information loss, the BC estimator would deliver the same results as the FD estimator. This, of course, is only a theoretical possibility – in reality, all frequencies contain some information, and removing any of them would result in loss of information. Visual inspection of Figures 5 – 6 along with the



(a) Time domain MLE.



(b) Frequency domain MLE using all frequencies.



(c) Frequency domain MLE using only business cycle frequencies.

Figure 6: Sampling distributions of the estimated parameters in the alternative parametrization of the ACD model. See the Notes to Figure 2.

uncertainty statistics in Tables 3 and 4, reveals that the impact of this information loss varies across parameters. For example, for φ, b, and κ_π, the FD and BC estimators show relatively similar performance, suggesting minimal information loss from excluding low and high frequencies. In contrast, for parameters like ρ_g ρ_m, and ρ_i, the BC estimator shows considerably larger sampling uncertainty compared to the FD estimator, indicating greater information loss. Interestingly, some parameters exhibit markedly different sampling distributions between the FD and BC estimators despite having similar uncertainty statistics. This is evident in the distributions of ρ_i, ρ_n, κ_R, σ_a^P, and σ_n in (b) and (c) of Figure 6. Given that the TD and FD sampling distributions are fairly similar for these parameters, the observed differences between the FD and BC estimators are likely driven by significant information loss in the BC estimator, rather than distortions from the Whittle likelihood approximation.

The last observation notwithstanding, finite sample approximation errors inevitably affect the performance of Whittle likelihood-based estimators. In particular, MC estimates of uncertainty, such as those presented earlier in this section, may be unreliable and misleading. This complicates our ability to assess the true information loss from excluding certain frequencies – a crucial consideration when using band spectral estimation methods. The next section presents an alternative approach to quantifying information content and loss that is robust to distortions from Whittle likelihood approximations.

5 Can the loss of information be predicted?

This section considers an alternative method for assessing the information loss associated with band spectral estimation. The motivation for exploring an alternative to the Monte Carlo simulation-based approach is twofold. First, as discussed at the end of the previous section, the finite-sample Whittle likelihood's approximate nature can distort assessments of the true information loss in band spectral estimation. Second, the Monte Carlo method is computationally intensive, making it challenging – and potentially impractical – to conduct comparative analyses across different scenarios, such as varying parameterizations, sets of observed variables, frequency bands, or sample sizes, etc.

As noted earlier, band spectral estimators rely on a subset of the frequency domain observations. It is important to emphasize that this is not equivalent to having fewer observations in the time domain. Specifically, excluding some low and high frequency observations is not equivalent to trimming observations from the beginning and end of the time domain sample. This distinction arises because the information content of different frequencies is not uniform across parameters. For instance, omitting a small number of low frequencies might result in a disproportionately large loss of information for some parameters while having only a negligible impact on others. The extent of information loss depends on how parameter-specific information is distributed across the spectrum. This, in turn, is determined by the role each parameter plays in the theoretical model and its influence on the distribution of the observed variables. In this sense, the selection of frequency bands is akin to the choice of observed variables in estimation: both choices affect parameters differentially, as neither frequencies nor variables provide the same level of information about all model parameters.

Similarly to the effect of excluding observable variables, the information analysis in band spectral estimation can be framed as a missing data problem (see, e.g., Dempster et al. (1977), Palm and Nijman (1984), and Iskrev (2019)). In this framework, the loss of information is quantified by comparing the expected information content of complete and incomplete samples. In the current context, a complete sample refers to observing all frequencies, while an incomplete sample is the one restricted to a specific frequency band. The expected amount of information about the model parameters contained in a sample's distribution can be measures using the Fisher information matrix, which is discussed next.

5.1 Fisher information matrix and Cramér-Rao lower bounds

The Fisher information matrix (FIM) is defined as

$$\mathcal{I}_T(\boldsymbol{\theta}) = \mathrm{E}\left[\frac{\partial}{\partial \boldsymbol{\theta}}\ell(\boldsymbol{\theta}; \boldsymbol{Y}_T)\frac{\partial}{\partial \boldsymbol{\theta}'}\ell(\boldsymbol{\theta}; \boldsymbol{Y}_T)\right] \tag{5.1}$$

where the expectation is with respect to the distribution of $\boldsymbol{Y}_T$ given the value of $\boldsymbol{\theta}$. The partial derivative of the log-likelihood function with respect to $\boldsymbol{\theta}$, called the score function, has an expected value of zero and therefore $\mathcal{I}_T(\boldsymbol{\theta})$ is the covariance matrix of the score at $\boldsymbol{\theta}$. When the log-likelihood is twice continuously differentiable equation (5.1) can be rewritten as[14]

$$\mathcal{I}_T(\boldsymbol{\theta}) = -\mathrm{E}\left[\frac{\partial^2}{\partial \boldsymbol{\theta}\partial \boldsymbol{\theta}'}\ell(\boldsymbol{\theta}; \boldsymbol{Y}_T)\right] \tag{5.2}$$

Thus, the FIM characterizes the expected curvature of the log-likelihood function, showing how quickly it declines around its peak value for each parameter. Intuitively, this indicates how informative a sample $\boldsymbol{Y}_T$ is expected to be about any given parameter θ_i from $\boldsymbol{\theta}$. A sharply peaked log-likelihood with respect to θ_i implies that the set of plausible values for the parameter is narrow, allowing for precise estimation. Conversely a flatter log-likelihood suggests that many values of θ_i are nearly equally plausible, making precise estimation more challenging.

More formally, the use of FIM as a measure of the quantity of information about $\boldsymbol{\theta}$ is based on the fact that it sets a limit for the maximum attainable precision of unbiased estimators of $\boldsymbol{\theta}$. By the information inequality (Rao (1945), Cramér (1946)), the inverse of the FIM provides a lower bound – the Cramér-Rao lower bound (CRLB) – on the variance of any unbiased estimator of $\boldsymbol{\theta}$. Therefore, if $\hat{\boldsymbol{\theta}}$ is an unbiased estimator with covariance matrix $\mathrm{cov}\,(\hat{\boldsymbol{\theta}})$, then

$$\mathrm{cov}\,(\hat{\boldsymbol{\theta}}) - \mathcal{I}_T^{-1}(\boldsymbol{\theta}) \succeq 0 \tag{5.3}$$

where $\succeq 0$ means that the quantity is a positive semidefinite matrix. This implies that the square roots of the diagonal elements of the inverse FIM are CRLBs for the standard deviations of the elements of $\hat{\boldsymbol{\theta}}$, i.e.

$$\mathrm{std}\,(\hat{\theta}_i) \geq \mathrm{crlb}(\theta_i) = \sqrt{\left\{\mathcal{I}_T^{-1}(\boldsymbol{\theta})\right\}_{ii}} \tag{5.4}$$

[14] See Hansen (2022, Ch10.)

By comparing the CRLBs for complete and incomplete samples, we can quantify the information loss associated with the latter. This approach is particularly appealing for Gaussian models, where the required FIMs are straightforward to evaluate. Specifically, for a vector $\boldsymbol{Y}$ distributed as a (complex) Gaussian with mean zero and a covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, the (k,l)-th element of the FIM is given by[15]

$$\{\mathcal{I}(\boldsymbol{\theta})\}_{k,l} \propto \text{trace}\left(\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\frac{\partial}{\partial\theta_k}\boldsymbol{\Sigma}(\boldsymbol{\theta})\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\frac{\partial}{\partial\theta_l}\boldsymbol{\Sigma}(\boldsymbol{\theta})\right) \tag{5.5}$$

From this expression, it is clear that $\mathcal{I}(\boldsymbol{\theta})$ does not depend on a particular sample but rather on the general properties of $\boldsymbol{Y}$ – the set of observed variables and either the number of observations or the included frequency band, all encoded in the covariance matrix $\boldsymbol{\Sigma}$. In the full information time domain case, $\boldsymbol{Y}$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ correspond to $\boldsymbol{Y}_T$ and $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$, as described in Section 3.1. In the full information frequency domain case, $\boldsymbol{Y}$ is $\boldsymbol{Y}(\boldsymbol{\omega}) = \boldsymbol{F}_T\boldsymbol{Y}_T$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is $\boldsymbol{F}_T\boldsymbol{\Sigma}_T(\boldsymbol{\theta})\boldsymbol{F}_T^*$ (see Section 4.2). Because the Fourier transform matrix $\boldsymbol{F}_T$ is orthonormal, the expression in (5.5) is identical for both domains. This is unsurprising given that the two representations are equivalent and no information is lost or gained when transforming from one domain to the other.

Information loss generally does occur when using only a subset of frequencies, which corresponds to deleting rows of the $\boldsymbol{F}_T$ matrix. For instance, using only the BC frequencies modifies $\boldsymbol{Y}$ in (5.5) to $\boldsymbol{Y}(\boldsymbol{\omega}_{bc}) = \boldsymbol{F}_{bc}\boldsymbol{Y}_T$ and $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ to $\boldsymbol{F}_{bc}\boldsymbol{\Sigma}_T(\boldsymbol{\theta})\boldsymbol{F}_{bc}^*$ (see Section 4.2). This yields the FIM corresponding to the exact band spectral likelihood function. For the FIM associated with the Whittle approximation, the block diagonal matrix $\boldsymbol{S}_T(\boldsymbol{\theta})$ from Section 3.1 is used in place of $\boldsymbol{\Sigma}(\boldsymbol{\theta})$. This provides an asymptotic approximation of the exact FIM and offers a convenient way to compute the asymptotic covariance matrix of the MLE through its inverse. This relationship holds due to the asymptotic efficiency of the MLE, where (5.4) holds with equality as $T \to \infty$.

The full information and band spectral FIMs, along with their respective CRLBs, enable predictions along two dimensions: (1) the estimation uncertainty associated with each approach, and (2) the relative loss of information in the band spectral case. The next section presents these predictions for the ACD model parameters and compares them with the MC simulations results.

5.2 Monte Carlo results vs CRLB predictions

This section evaluates the accuracy of the CRLB as a predictor of estimation uncertainty. This evaluation is both necessary and important because, as its name suggests, the CRLB represents only a theoretical bound on uncertainty, leaving open the question of how closely it aligns with actual estimation uncertainty, particularly for relatively small sample sizes. Furthermore, the theoretical result applies only to unbiased estimators; however, as observed in the previous section, some parameter estimates exhibit non-negligible biases. I use the MC estimates of the sampling standard deviations of the exact and Whittle MLEs as proxies for true estimation uncertainty. Due to the issues with Whittle likelihood-based estimators discussed in Section 4.2, this section presents results only for the alternative parametrization described in Section 4.3. Results for the baseline parametrization are provided in Appendix C.

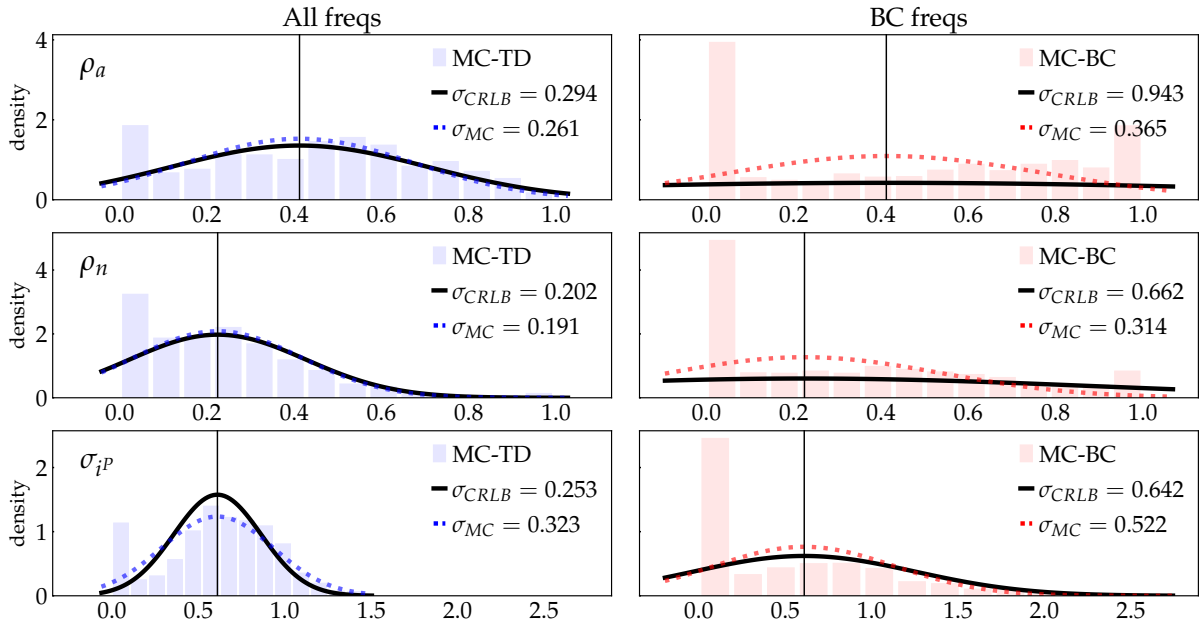
[15] See Kay (1993, Chapters 3, 15). The constant of proportionality depends on whether the distribution is the real or the complex Gaussian.

Table 5 presents two sets of results: Panel (a) reports the estimation uncertainty when using either the full spectrum or only BC frequencies. As before, estimation uncertainty is presented as a percentage of the absolute value of the true parameter values. Panel (b) shows the efficiency loss when using only BC frequencies, defined as the ratio of band spectral to full information standard deviations or CRLBs. Starting with estimation uncertainty, the CRLBs are generally very close to the TD estimator's standard deviations. For more than half of the parameters, the differences are under 10%, with the largest discrepancy being less than 25%. This occurs for σ_{a^T}, where the TD estimator's standard deviation is .132 compared to a CRLB of .1.

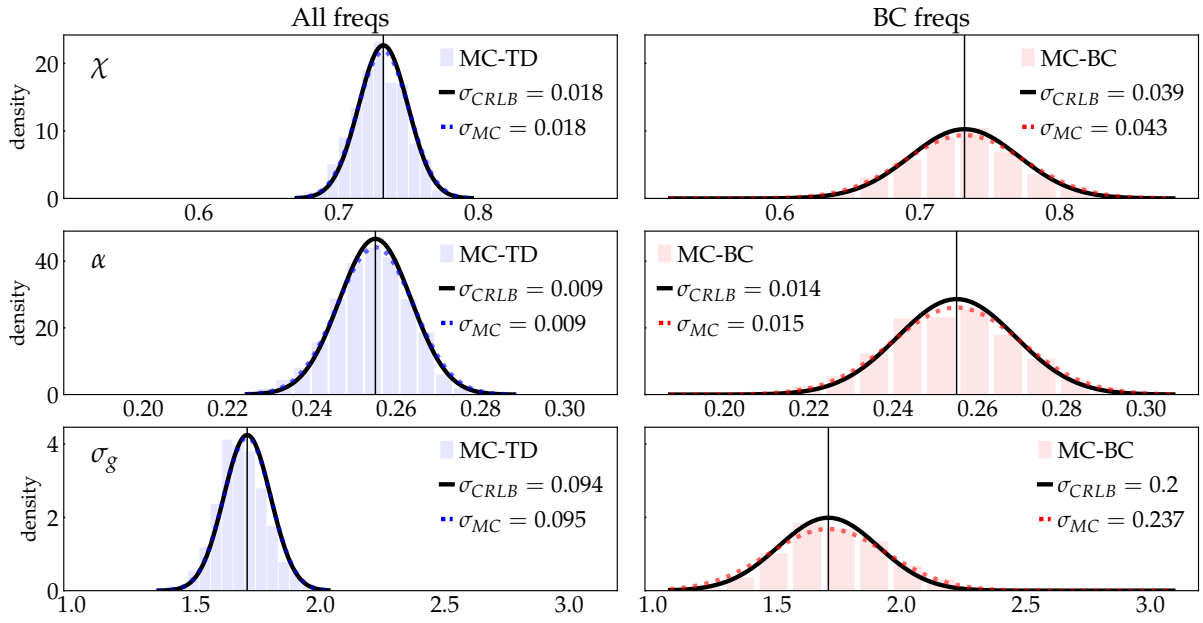
Except for two parameters (ρ_a and ρ_n), the CRLBs are smaller than the corresponding MC-TD standard deviations, consistent with the information inequality (5.4). The CRLB-based ranking of parameters by normalized estimation uncertainty closely aligns with the MC-TD results, with a rank correlation of .98. For the MC-FD results, the rank correlation with the CRLBs is slightly lower at .96. The numerical discrepancies are larger for the FD estimator, exceeding 20% for most parameters and reaching 50% for b. All CRLBs remain below their respective MC-FD standard deviations. In contrast, in the BC case, nine parameters violate the information inequality. For two of these parameters (ρ_a and ρ_n), the CRLBs are more than double the MC standard deviations. Significant discrepancies also arise for σ_n and κ_R, with differences of approximately 70% and 37%, respectively. Despite these exceptions, the CRLBs generally provide accurate predictions of estimation uncertainty in the band spectral case, with discrepancies under 20% for most parameters. The CRLB also accurately predicts the MC-implied parameter ranking by relative uncertainty, with a rank correlation of .96. As expected, both the MC and CRLB-based results indicate greater estimation uncertainty when information is restricted to the BC frequencies. This is reflected in relative efficiency values greater than one, as shown in panel (b) of Table 5. For completeness, the table reports relative efficiency values for MC-TD and MC-FD estimators, as well as those based on the finite sample and asymptotic CRLBs. The latter, denoted as $CRLB(\infty)$, as computed using the asymptotic band spectral and full information FIMs, as discussed earlier. While finite sample CRLBs provide the preferred measure, the asymptotic loss predictions are very similar, being only slightly larger across all parameters. In contrast, the MC-FD estimates of relative efficiency are lower than the MC-TD estimates due to consistently larger standard deviations in the full spectrum case. According to both MC and CRLB results, the efficiency loss is substantial, exceeding 100% for most parameters. While the CRLB predictions closely match the MC losses for most parameters, there are several notable exceptions. The MC results identify ρ_a, ρ_n, σ_{i^P}, and σ_{a^P} among the parameters with smallest efficiency losses ($40\% - 80\%$), yet CRLB-based predictions for these parameters are among the largest ($150\% - 230\%$). Similarly, CRLB predicts much larger efficiency losses for κ_R and σ_n (230% and 300%, respectively) compared to the MC results (120%).

What explains the few stark differences between predicted and estimated efficiency losses? As notes earlier, and evident from the last two columns in panel (a) of Table 5, the band spectral estimates of the six parameters listed above all violate the information inequality. In some cases, the MC standard deviations are significantly smaller than their respective CRLBs. For ρ_a and ρ_n this also occurs in the full information case, though to a much lesser extent.[16] Thus,

[16]The MC-TD standard deviations of ρ_a and ρ_n are about 13% and 6% larger than their CRLBs, respectively,



(a) Parameters with significant discrepancies between MC and CRLB-based efficiency loss values.



(b) Parameters with minor discrepancies between MC and CRLB-based efficiency loss values.

Figure 7: Each plot of the figure shows a histogram of the MC sampling distribution of the respective parameter, along with two gaussian densities. The standard deviations of these densities correspond to the CRLB (black) and either the MC-TD estimate (left, dotted-blue) or MC-BC estimate (right, dotted-red).

Table 5: Estimation uncertainty and relative efficiency: MC vs CRLBs

	(a) Estimation uncertainty (%)					(b) Relative efficiency			
	all freqs			BC freqs		BC vs all freqs			
parameter	MC-TD	MC-FD	CRLB	MC-BC	CRLB	MC-TD	MC-FD	CRLB	CRLB(∞)
ν	25.8	31.1	23.5	53.8	49.5	2.1	1.7	2.1	2.3
α	3.5	4.0	3.4	6.0	5.5	1.7	1.5	1.6	1.7
ψ	22.0	26.7	18.5	52.5	48.5	2.4	2.0	2.6	3.1
φ	25.7	33.0	22.6	47.5	37.9	1.8	1.4	1.7	1.8
b	4.7	8.6	4.4	9.4	8.2	2.0	1.1	1.9	1.9
χ	2.5	3.1	2.4	5.8	5.3	2.3	1.9	2.2	2.4
κ_R	32.4	36.2	30.3	72.5	99.3	2.2	2.0	3.3	3.4
κ_π	14.0	17.9	11.7	31.2	24.8	2.2	1.7	2.1	2.3
κ_y	58.6	85.5	44.6	115.3	106.7	2.0	1.3	2.4	3.0
ρ_a	63.5	74.2	71.4	88.5	229.0	1.4	1.2	3.2	3.4
ρ_n	85.3	103.7	90.0	139.9	295.1	1.6	1.3	3.3	3.4
ρ_i	23.3	29.2	23.0	61.2	75.6	2.6	2.1	3.3	3.4
ρ_c	3.8	4.5	3.5	7.8	8.6	2.1	1.7	2.4	3.0
ρ_g	7.2	7.6	5.9	22.2	15.2	3.1	2.9	2.6	2.9
ρ_m	8.5	9.3	7.6	22.3	16.7	2.6	2.4	2.2	2.3
ρ_ξ	4.9	7.6	4.3	12.5	10.9	2.6	1.7	2.6	3.1
σ_{a^P}	32.6	36.8	24.6	59.0	74.4	1.8	1.6	3.0	3.1
σ_{a^T}	24.1	29.1	23.5	58.7	70.4	2.4	2.0	3.0	3.1
σ_n	25.4	31.0	23.3	54.5	92.6	2.2	1.8	4.0	4.2
σ_{i^P}	52.9	60.2	41.5	85.7	105.3	1.6	1.4	2.5	2.9
σ_{i^T}	27.0	38.0	23.1	77.4	61.6	2.9	2.0	2.7	2.8
σ_c	46.1	59.1	43.1	92.6	85.1	2.0	1.6	2.0	2.3
σ_g	5.6	5.9	5.5	13.9	11.7	2.5	2.3	2.1	2.2
σ_m	10.5	13.2	9.9	26.6	21.4	2.5	2.0	2.2	2.2
σ_ξ	49.6	66.4	37.8	126.5	97.5	2.5	1.9	2.6	3.0

Note: Estimation uncertainty is reported in terms of MC standard deviation or CRLB as a percentage of the modulus of the true parameter value. Relative efficiency is defined as the ratio of the standard deviation or CRLB using BC frequencies only to that using all frequencies. All CRLBs are computed with the exact FIM for $T = 192$ except in the column labeled CRLB(∞) where the bounds are computed using the asymptotic FIM.

the differences arise primarily from disagreement between the two approaches in assessing of estimation uncertainty in the band spectral case.

To explore this further, panel (a) of Figure 7 reproduces the MC sampling distributions of ρ_a, ρ_n, and σ_{i^P}, adding two Gaussian density curves for each parameter. Both densities are centered on the true parameter values, with standard deviations corresponding to the CRLBs (in black) and the respective MC-TD and MC-BC estimates (in blue and red, respectively). The plots reveal two key reasons for the failure of the information inequality: (1) the low curvature of the likelihood with respect to these parameters, and (2) the influence of prior restrictions on their values. Consequently, the simulations underestimate sampling uncertainty for these parameters, particularly in the band spectral case where information is scarcer.

This is most evident for ρ_a and ρ_n, where the CRLBs predict an almost flat likelihood in the BC frequencies. While MC simulations confirm these predictions, they also reflect the parameters' prior bounds – $(0, 1)$ for ρ_a and ρ_n, and $[0, \infty)$ for σ_{i^P}. These restrictions cause bunching of MC estimates at the parameter space boundaries and result in a smaller increase in estimation uncertainty in the band spectral case compared to what would be expected based

whereas the MC-BC standard deviations exceed the corresponding band spectral CRLBs by over 150% and 110%.

on the reduction in likelihood curvature alone. The other parameters with large discrepancies between MC estimates and CRLB-predicted efficiency losses – κ_R, σ_{a^P}, and σ_n – exhibit similar patterns. This is illustrated in Figure 13 in Appendix C), where the same combination of reduced curvature and prior constraints explains the observed discrepancies.

Panel (b) of Figure 7 demonstrates that prior restrictions have much weaker effects on estimated uncertainty when the sample is sufficiently informative about a parameter. This is shown for three parameters with bounds identical to those in panel (a): $(0, 1)$ for χ and α, and $[0, \infty)$ for σ_g. Since the MC estimates concentrate away from the parameter space boundaries, the estimated standard deviations exceed their respective CRLBs, consistent with the information inequality. The differences between predicted and realized estimation uncertainty remain relatively small, particularly in the full information case, resulting in accurate – though slightly underestimated – predictions of the band spectral estimator's efficiency losses (see panel (b) of Table 5).

The results for all parameters are reported in Figure 13 of Appendix C. They provide additional evidence supporting the observation that the FIM-based predictions are highly accurate whenever parameter constraints are non-binding during estimation. Since these constraints represent a form of non-sample information, this suggests that the FIM-based approach might be more reliable than the MC-based method for assessing the loss of sample information associated with band spectral estimators. This possibility could be tested by increasing the amount of sample information, such as by using larger sample sizes in the simulations. Larger samples should reduce estimation uncertainty, thereby diminishing the relative influence of a priori parameter restrictions. If the FIM-based predictions are indeed accurate, we should observe convergence between the MC-based estimates and their CRLB-based counterparts as sample sizes frow.

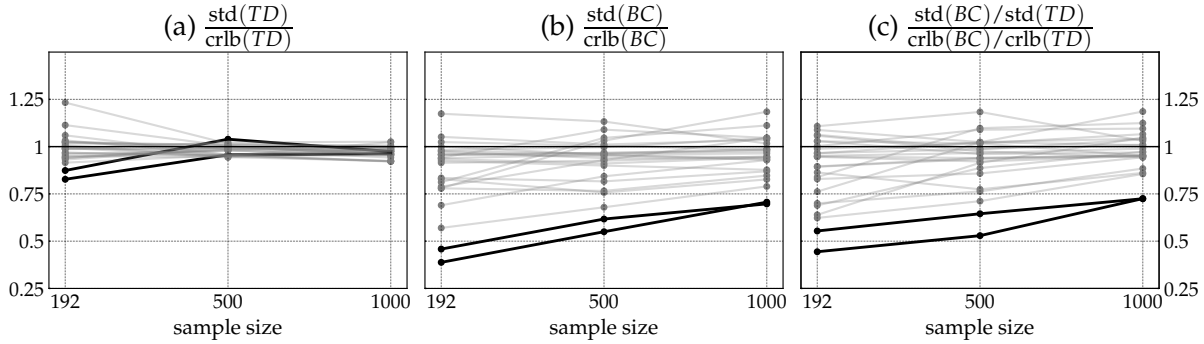


Figure 8: Each point represents the ratio of estimated and predicted uncertainty for a given parameter and sample size, in either the full information case (panel (a)) or the BC frequencies case (panel (b)), or the ratio of estimated to predicted relative efficiency (panel (c)). The SDs are estimated using 1000 MC replications, while the CRLBs are computed from the expected FIM. In all cases, the largest discrepancies between estimated and predicted values occur for the same two parameters, ρ_a, ρ_n, which are indicated in bold.

I repeat the MC simulations exercise from Section 4.1 with larger samples of $T = 500$ and $T = 1000$. Figure 8 shows how the ratios between estimated and predicted quantities evolve – parameter uncertainty in both full information and bans-spectral cases (panels (a) and (b)), and their respective relative efficiency values (panel (c)). In all cases, the values become more

concentrated around 1 as T increases, indicating convergence between predictions and estimates. Naturally, this convergence is more pronounced in the full information case compared to the band spectral case.

In the band spectral case, however, several parameters continue to exhibit significantly lower estimated uncertainty relative to the predicted values, even when $T = 1000$. This suggests that parameter restrictions remain materially binding for these parameters. The most striking examples are ρ_a and ρ_n, where the estimated uncertainty in the BC case is approximately 30% smaller than the predicted uncertainty. As shown in the sampling distribution plots in the Appendix (Figure 14), this discrepancy arises because the BC frequencies likelihood remains very flat for these parameters even at $T = 1000$. Nonetheless, panel (b) of Figure 8 shows a significant improvement compared to the original small-sample case, where the predicted uncertainty exceeded the estimated uncertainty by more than a factor of two.

The observed convergence between estimated and predicted uncertainty results in convergence between estimated and predicted relative efficiency values, as depicted in panel (c) of the figure. Apart from the two aforementioned parameters, ρ_a and ρ_n, the largest discrepancy between predicted and estimated values in either direction is at most 18%, indicating strong agreement. This is further illustrated in Figure 9, which compares estimated and predicted relative efficiency values for sample sizes of $T = 192$ and $T = 1000$.

When $T = 1000$, the largest remaining discrepancy, after ρ_a and ρ_n, is for σ_{a^P}, where the estimated relative efficiency is 3.7, compared to the predicted value of 3.1. On the opposite extreme, σ_n exhibits a predicted loss of 4.15, exceeding the estimated value of 3.7. Overall, the rank correlation between estimated and predicted relative efficiency improves significantly from .16 when $T = 192$ to .76 when $T = 1000$, indicating better concordance between predictions and estimates as the sample size increases. At the same time, the sample size has minimal impact on the magnitude of the predicted relative efficiency values. For instance, the rank correlation between CRLB-192 and CRLB-1000 is .95. This suggests that the FIM-based approach reliably captures the efficiency losses incurred by different parameters under band spectral estimation. I apply this insight in the next section to examine what information is preserved and what is lost when estimating the model using only business cycle frequencies.

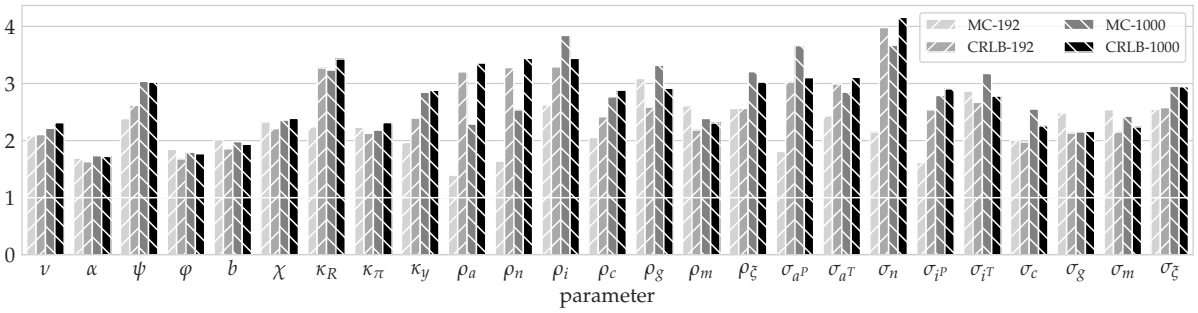


Figure 9: Predicted (CRLB-T) and estimated (MC-T) relative efficiency for $T = 192$ and $T = 1000$. Relative efficiency is defined as the ratio of the MC standard deviation or CRLB obtained using only BC frequencies to that obtained using all frequencies.

6 Information in (and not in) the business cycle frequencies

The BC frequencies account for less than 30% of all frequency domain observations. Thus, it is not surprising that the band spectral estimation results in significant efficiency losses compared to the full information case. At the same time, as discussed earlier, some parameters experience much greater losses than others. For instance, the efficiency losses for parameters such as σ_n, ρ_i, ρ_n, ρ_a, and κ_R are more than twice as large as those for α, φ, and b. The explanation for these differences must be that, for the parameters in the former group, either the low or high frequencies – or both – contribute a lot of information that is not captured within the BC frequencies alone.

The FIM formalism can help clarify which explanation is correct. Figure 10 compares the predicted relative efficiency of three band spectral estimators — based on BC frequencies only, Low-plus-BC frequencies, and High-plus-BC frequencies -– relative to the full information case.[17] The results show that the large information losses in the BC-frequency-only case primarily stems from the absence of high-frequency information. Specifically, for σ_n, ρ_i, ρ_n, ρ_a, and κ_R, the high-plus-BC frequencies estimator is nearly as efficient as the full information one. High frequencies also account for the bulk of the missing information for most of the remaining shock volatility parameters – σ_{a^P}, σ_{a^T}, σ_{i^T}, σ_g, and σ_m. In contrast, only a few parameters derive significant amount of unique information from the low frequencies, with ψ, κ_y, ρ_c, σ_{i^P} being the most notable exceptions.

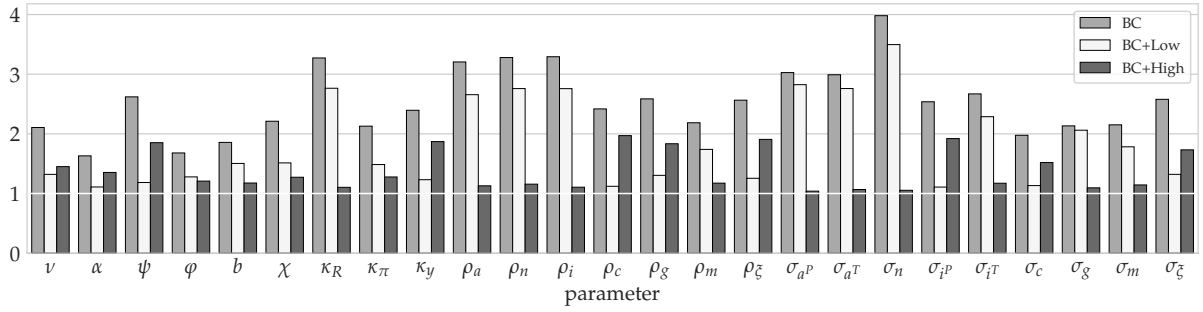


Figure 10: Predicted relative efficiency of the three band spectral estimators. The relative efficiency is defined as the ratio of the CRLB value using a specific frequency band to the CRLB value using all frequencies. The sample size is $T = 192$.

Why do some parameters benefit more from information in the higher frequencies whereas others are better informed by frequencies at the lower end of the spectrum? Macro modelers may be able to develop intuition about the answer to this question by considering how the functional roles of different parameters in a model translate to the statistical properties of the model variables. For example, parameters that determine a model's steady state are likely to derive their information primarily from lower frequencies. Conversely, parameters governing the volatility of exogenous disturbances tend to be better informed through high-frequency information.

Qualitative explanations based on intuition and experience, while useful, have their limitations

[17] Figure 15 compares the predicted relative efficiency in the Low, BC, and High frequency band cases.

and may be incorrect. In modern macroeconomic models, parameters often cannot be neatly categorized as predominantly influencing a specific band of frequencies. Even parameters directly linked to steady states or volatilities are typically related to – and therefore derive information from – frequencies beyond just the lower or higher ends of the spectrum. For instance, we observe that the efficiency gains from incorporating low-frequency information exceed those from high-frequency information for σ_{iP}, σ_c, and σ_ξ. This contrasts with the other volatility parameters, for which the results align with the previously mentioned intuition. A key challenge in informal analysis is properly accounting for parameter interdependence. Even when it is possible to establish that a parameter affects mainly, for instance, the long-run dynamics, it is usually much harder to determine how distinct the effect is from that of other parameters. Due to parameter interdependence, the dimensions of data most sensitive to a parameter are not necessarily the most informative ones. Ignoring these complexities can lead to misleading conclusions.

The FIM-based approach offers a formal framework to address these challenges. The marginal CRLB for any parameter θ_i can be decomposed into two factors: one capturing the likelihood sensitivity to θ_i, and another reflecting parameter interdependence. More formally, it can be shown that (see Iskrev (2010a,b))

$$\operatorname{crlb}(\theta_i) = \underbrace{\frac{1}{\sqrt{\{\mathcal{I}_T(\boldsymbol{\theta})\}_{ii}}}}_{\operatorname{crlb}(\theta_i|\boldsymbol{\theta}_{-i})} \times \underbrace{\frac{1}{\sqrt{1-\mathbf{R}_i^2}}}_{\operatorname{SIF}(\theta_i)} \tag{6.1}$$

where $\operatorname{crlb}(\theta_i|\boldsymbol{\theta}_{-i})$ is the conditional CRLB of θ_i given all other parameters $\boldsymbol{\theta}_{-i} = \{\theta_j \mid j \neq i\}$, which depends only on the inverse sensitivity of the likelihood function to θ_i. The second term, the standard deviation inflation factor (SIF), is an increasing function of the multiple correlation coefficient $\mathbf{R}_i$ between the scores with respect to θ_i and $\boldsymbol{\theta}_{-i}$.[18]

The interpretation of (6.1) is straightforward. Like the marginal CRLB, the conditional CRLB represents the predicted estimation uncertainty for θ_i when all other parameters are known. In this case, there is no parameter interdependence, and the conditional CRLB depends solely on the sensitivity of the likelihood function with respect to θ_i: higher sensitivity implies lower conditional uncertainty, and vice versa.The effect of parameter interdependence is captured by $\operatorname{SIF}(\theta_i)$. The more similar the effect of θ_i on the likelihood to that of other estimated parameters, the less precise its estimation will be. Since $\mathbf{R}_i^2 \in [0,1]$, $\operatorname{SIF}(\theta_i)$ is bounded from below by 1, where no parameter interdependence affects θ_i and the marginal CRLB equals the conditional. At $\mathbf{R}_i^2 = 1$, θ_i becomes unidentified as it cannot be distinguished from a combination of parameters in $\boldsymbol{\theta}_{-i}$.

The decomposition in (6.1) offers a straightforward explanation of why certain frequencies are predicted to be more informative than others: it is because of higher sensitivity, lower parameter interdependence, or a combination of both. Table 6 illustrates this by presenting

[18] In linear regression context, this is equivalent to the square root of the variance inflation factor (VIF) introduced by Marquaridt (1970), where $\mathbf{R}_i$ is the multiple correlation coefficient between the i-th regressor and the other regressors. The term SIF appears in Fox and Monette (1992) as *standard-error inflation factor* and is similarly defined as the square root of the VIF.

Table 6: Decomposing predicted relative efficiency

parameter	low	BC	high
ν	3.4 = 3.0 × 1.13	2.1 = 1.7 × 1.22	3.8 = 1.4 × 2.80
α	1.7 = 2.3 × 0.73	1.6 = 1.7 × 0.94	2.8 = 1.4 × 1.93
ψ	1.8 = 3.3 × 0.52	2.6 = 1.8 × 1.45	4.1 = 1.3 × 3.15
φ	3.9 = 5.6 × 0.69	1.7 = 1.8 × 0.91	2.5 = 1.2 × 2.03
b	6.8 = 4.9 × 1.40	1.9 = 1.9 × 0.98	2.2 = 1.2 × 1.78
χ	4.9 = 3.6 × 1.38	2.2 = 2.0 × 1.12	2.6 = 1.2 × 2.11
κ_R	12.5 = 5.5 × 2.28	3.3 = 2.4 × 1.34	1.8 = 1.1 × 1.64
κ_π	4.3 = 4.2 × 1.03	2.1 = 1.9 × 1.11	2.3 = 1.2 × 1.87
κ_y	4.2 = 4.4 × 0.95	2.4 = 1.9 × 1.27	3.6 = 1.2 × 2.93
ρ_a	42.0 = 24.7 × 1.70	3.2 = 3.2 × 1.01	2.2 = 1.1 × 2.06
ρ_n	16.9 = 2.2 × 7.71	3.3 = 1.5 × 2.15	2.6 = 1.6 × 1.56
ρ_i	11.4 = 3.4 × 3.35	3.3 = 1.2 × 2.65	2.8 = 2.0 × 1.44
ρ_c	3.0 = 3.1 × 0.98	2.4 = 1.9 × 1.29	3.8 = 1.3 × 2.97
ρ_g	8.2 = 1.2 × 6.74	2.6 = 2.3 × 1.14	6.9 = 2.7 × 2.55
ρ_m	13.6 = 12.0 × 1.14	2.2 = 2.3 × 0.95	2.2 = 1.1 × 1.95
ρ_ξ	3.9 = 2.4 × 1.59	2.6 = 1.9 × 1.37	4.4 = 1.4 × 3.25
σ_{a^P}	18.5 = 4.0 × 4.59	3.0 = 1.9 × 1.58	1.4 = 1.2 × 1.18
σ_{a^T}	88.2 = 52.2 × 1.69	3.0 = 3.1 × 0.95	1.7 = 1.1 × 1.65
σ_n	25.9 = 2.9 × 9.04	4.0 = 1.5 × 2.58	1.3 = 1.5 × 0.90
σ_{i^P}	1.7 = 1.3 × 1.34	2.5 = 2.1 × 1.22	3.9 = 3.0 × 1.29
σ_{i^T}	10.0 = 7.1 × 1.41	2.7 = 1.9 × 1.38	2.4 = 1.2 × 2.03
σ_c	1.9 = 4.0 × 0.47	2.0 = 1.8 × 1.07	3.8 = 1.2 × 3.04
σ_g	27.4 = 4.1 × 6.74	2.1 = 1.9 × 1.15	3.2 = 1.2 × 2.56
σ_m	7.1 = 4.5 × 1.58	2.2 = 1.9 × 1.14	1.7 = 1.2 × 1.35
σ_ξ	3.3 = 4.0 × 0.83	2.6 = 1.8 × 1.40	4.1 = 1.2 × 3.26

Note: The predicted relative efficiency is decomposed using (6.1) into a relative sensitivity factor and a relative interdependence factor.

the decomposition of predicted relative efficiency across the low, BC, and high-frequency bands relative to the full information case.[19] Several results from the table are worth highlighting.

First, the decompositions confirm the earlier observation that sensitivity and informativeness are distinct. For 16 of the 25 parameters, the frequency bands with the lowest marginal and conditional CRLBs differ. Notably, for ψ and σ_c, low frequencies are most informative yet least sensitive, while high frequencies exhibit the opposite pattern. Second, the first term in the decomposition is always greater than 1, while the second term can be less than 1. In other words, there is always some loss of information due to lower sensitivity in the band spectral case relative to the full information.[20] The effect of parameter interdependence can go in either direction, and, in particular, the SIF may be smaller in certain frequency bands than in the full spectrum. This occurs when a parameter's impact on the likelihood is markedly less correlated with that of other parameters in certain parts of the spectrum than in others. The most notable examples are the aforementioned ψ and σ_c whose low frequency SIF values are approximately half their full spectrum counterparts.[21] Third, the decompositions help develop a more nuanced understanding of the relative efficiency of the BC-based band-spectral estimator. For example, the substantial efficiency losses for σ_n, ρ_n, and ρ_i primarily result from stronger parameter interdependence

[19] Decompositions for the BC+Low and the BC+High frequency bands are provided in Table 9 in Appendix D.

[20] In principle, the relative sensitivity factor for a given frequency band could equal 1 if the parameter's effect on the likelihood is entirely confined to that band.

[21] In terms of correlation coefficients, ψ shows values of .933 (low frequencies) and .982 (full spectrum), while σ_c shows .965 and .992, respectively. Both parameters have multiple correlation coefficients near .999 in the high-frequency band.

in the BC frequencies, compared to the full spectrum. In contrast, ρ_a and σ_{a^T} show similar losses, but due to lower sensitivity in the BC band. Low sensitivity is also the dominant factor in the case of α, ψ, b, and ρ_m, all of which are less affected by interdependence issues in the BC frequency band compared to the full spectrum. For all other parameters, both factors contribute substantially to the band spectral estimator's reduced efficiency.

Lastly, the results in Table 6 clarify the effects of parameter restrictions in estimation. Specifically, because of parameter interdependence, binding restrictions on a parameter's permissible values reduce estimation uncertainty not only for that parameter but also for other related parameters. This explains why the information inequality may fail to hold, even in large samples, for parameters whose own constraints are not binding (see panel (b) of Figure 8). A notable example is the pair ρ_n and σ_n, whose scores have a pairwise correlation coefficient of .92 in the BC frequencies.[22] Since the constraints on ρ_n remain strongly binding in the band spectral case, the estimation uncertainty of σ_n is also reduced, even though its own restrictions are less relevant when $T = 1000$ (see Figure 14 in Appendix C). However, this does not occur to the same extent in the full spectrum, as the pairwise correlation coefficient there is only .43. Consequently, as shown in Figure 9 that σ_n is, in the large sample case, among the few parameters for which the CRLB-predicted efficiency loss is greater than the MC-estimated one.

The preceding discussion illustrates two complementary uses of the FIM-based analysis: (1) determining which frequency bands are most informative about each parameter, and (2) explaining why this is the case. Of course, the results must be interpreted only as predictions that are conditional on the model being the true DGP for all frequencies. As in the previous section, the validity and accuracy of these predictions can be tested. Specifically, the results in Figure 10 can be compared with MC estimates of the relative efficiency loss of the three band-spectral estimators shown in the figure. Similarly, simulations can be used to test the predictions in Table 6 regarding the relative sensitivity of different frequency bands with respect to each parameter. In this latter case, the simulations should be conducted by estimating parameters one at a time, i.e. fixing all other parameters at their true values and treating them as known during the estimation.

Appendix D provides a comprehensive overview of the results from such simulations. As in Section 5.2, the overall conclusion is that the CRLBs are generally very reliable predictors of the relative loss of efficiency. In the vast majority of cases, the MC results fully align with the predictions in Figure 10 regarding whether the low or high frequencies contribute more of the information missing in the BC-frequency-only case. The predictions are also generally accurate in terms of the magnitude of efficiency losses. However, as before, prior restrictions on parameter values sometimes lead to underestimation of uncertainty and, consequently, distort MC-based efficiency losses. As in Section 5.2, increasing the sample size to 1000 mitigates the effects of these restrictions and improves agreement between predicted and estimated efficiency losses.

The results in Appendix D also present concrete evidence for another source of distortions in the MC results stemming, from the misspecified nature of the Whittle likelihood function. Under correct specification, adding more frequencies to the band spectral estimator cannot increase estimation uncertainty. However, for several parameters in Figure 21 of the Appendix (see ρ_n

[22] See Section D.1 in the Appendix for details on score collinearity for all parameters and frequency bands.

and σ_n when $T = 192$, and ρ_i and σ_i^T when $T = 1000$), the BC-frequency-only estimator is more efficient than the BC+Low frequency estimator. Even more striking examples of this abnormality can be seen in the conditional estimation results for ν, α, ψ, ρ_i, and σ_ξ, where the BC+High frequency estimator is significantly less efficient than the BC-frequency-only estimator (see Figure 27 of the Appendix). This is in sharp contrast to the CRLB predictions for these parameters, according to which the BC+High frequency estimator is nearly as efficient as the full information estimator. While the FIM-based predictions can be inaccurate for various reasons, obtaining less precise estimates as more information is used can only be explained by optimizing the wrong objective function.[23]

Comparing the predicted and estimated conditional uncertainty of the aforementioned parameters, rather than the efficiency loss, shows that the CRLB predictions are very accurate for all estimators except the one based on the BC+High frequency band. For this estimator, the estimated uncertainty is much larger than predicted (see Figure 32 in the Appendix). Following the earlier discussion of the conditional CRLB, an obvious explanation for this difference is that the Whittle likelihood is misspecified in ways that make it considerably less sensitive to these parameters in the high-frequency region of the sample. To verify this conjecture, I employ the exact band spectral likelihood function discussed toward the end of Section 4.2. Note that although a naive evaluation of the function is computationally very demanding, estimating parameters individually remains feasible within a reasonable amount of time. The results, presented in Section D.2.3 of Appendix D, show nearly perfect alignment between CRLB predictions and MC estimates across all frequency bands, both in terms of estimation uncertainty and efficiency losses.

7 Concluding Comments

In this paper, I have focused on two key issues arising in the application of band spectral methods to estimate business cycle models: (1) the finite-sample distortions due to the approximate nature of the Whittle likelihood, currently the predominant method for estimating economic models in the frequency domain; and (2) the potentially significant information loss when estimation relies on only a subset of frequencies. The results from extensive Monte Carlo (MC) simulations, using the model of Angeletos et al. (2018) as a data-generating process, showed that the band spectral estimator exhibits considerable bias and efficiency loss compared to the full information case. Whereas the bias is caused by the poor performance of the Whittle approximation when data is persistent, the loss of precision is mainly due to the low information content of the business cycle frequencies. In addition to MC analysis, I assessed the reliability of analytical FIM-based predictions of estimation uncertainty and efficiency loss. The results suggest that this approach can serve as a useful alternative to conducting MC simulations, both for identifying the most informative parts of the spectrum and for understanding why certain frequency bands provide more information for some parameters than others.

The findings presented in the paper have several implications for the broader literature. First, when model misspecification is not a concern and the choice is merely between full information

[23] See Meng and Xie (2014).

estimation in the time domain versus the frequency domain, the time domain approach is clearly superior. Although the Whittle likelihood distortions may be smaller with less persistent data, approximation errors are unlikely to vanish entirely for the sample sizes commonly used in business cycle research. Second, obtaining different results when estimating a model using different frequency bands, or with time-domain methods, does not necessarily indicate model misspecification, contrary to the prevailing wisdom among empirical macroeconomists. An alternative explanation for such discrepancies is that both the size of the Whittle likelihood distortions and the impact of information loss typically vary across parameters and across parts of the spectrum. Third, any form of out-of-sample information on the estimated parameters – in the form of a prior distribution, theoretical restrictions on the parameter space, or assigned values of calibrated parameters, will likely have a much greater impact on band spectral estimators compared to full information ones. This is important to keep in mind when comparing results obtained using different estimation methods. Finally, frequency-domain FIM analysis, like the one developed in this paper, can also benefit researchers who estimate structural macro models in the time domain. In particular, quantifying the information content of different frequency bands will enhance their understanding of where in the data, according to a given model, information about parameters predominantly comes from. This, in turn, will make the estimation of such models more transparent.

There are several promising directions for future research. A key priority should be to develop more reliable methods for band spectral estimation of macroeconomic models. One approach could be to devise a computationally efficient implementation of the exact likelihood estimator outlined in this paper. Another is to improve upon the standard Whittle likelihood approximation, for instance along the lines of Sykulski et al. (2019). Beyond refining the conventional band spectral analysis, it may also be worthwhile to explore methods that target different frequency bands for different variables. Such estimators would allow for more efficient use of sample information, particularly in cases where the frequency-specific mismatch between model and data varies across observed variables. Similarly, the FIM-based analysis could be extended to incorporate variable-specific frequency bands, thereby providing a more nuanced view of the interplay between parameter sensitivity and interdependence and enhancing researchers' understanding of the origins of parameter information. Finally, it is important to investigate how model-dependent or parameter-specific the conclusions I have reached in this study are. As with any simulation-based analysis, it is unrealistic to expect that the results will be uniformly applicable under all circumstances. Conducting additional simulations with diverse models and parameter configurations would be valuable for assessing the robustness and generalizability of my findings.

References

Adolfson, Malin, Stefan Laséen, Jesper Lindé, and Marco Ratto (2019) "Identification versus misspecification in New Keynesian monetary policy models," *European Economic Review*, 113, 225–246.

Altug, Sumru (1989) "Time-to-Build and Aggregate Fluctuations: Some New Evidence," *International Economic Review*, 30 (4), 889–920, http://ideas.repec.org/a/ier/iecrev/v30y1989i4p889-920.html.

Angeletos, George-Marios, Fabrice Collard, and Harris Dellas (2018) "Quantifying confidence," *Econometrica*, 86 (5), 1689–1726.

Christiano, Lawrence J. and Robert J. Vigfusson (2003) "Maximum likelihood in the frequency domain: the importance of time-to-plan," *Journal of Monetary Economics*, 50 (4), 789–815, http://ideas.repec.org/a/eee/moneco/v50y2003i4p789-815.html.

Cogley, Timothy (2001) "Estimating and testing rational expectations models when the trend specification is uncertain," *Journal of Economic Dynamics and Control*, 25 (10), 1485–1525.

Cramér, Harald (1946) *Mathematical methods of statistics*, 26: Princeton university press.

Dempster, Arthur P, Nan M Laird, and Donald B Rubin (1977) "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the royal statistical society. Series B*, 1–38.

Diebold, Francis X, Lee E Ohanian, and Jeremy Berkowitz (1998) "Dynamic equilibrium economies: A framework for comparing models and data," *The Review of Economic Studies*, 65 (3), 433–451.

Engle, Robert F. (1974) "Band Spectrum Regression," *International Economic Review*, 15 (1), 1–11, http://www.jstor.org/stable/2526084.

Engle, Robert F (1978) "Testing price equations for stability across spectral frequency bands," *Econometrica: Journal of the Econometric Society*, 869–881.

Fox, John and Georges Monette (1992) "Generalized collinearity diagnostics," *Journal of the American Statistical Association*, 87 (417), 178–183.

Gray, Robert M (2006) "Toeplitz and circulant matrices: A review," *Foundations and Trends in Communications and Information Theory*, 2 (3), 155–239.

Hannan, Edward J (1963) "Regression for time series with errors of measurement," *Biometrika*, 50 (3/4), 293–302.

Hansen, Bruce (2022) *Probability and Statistics for Economists*: Princeton University Press.

Hansen, Lars Peter and Thomas J Sargent (1980) "Formulating and estimating dynamic linear rational expectations models," *Journal of Economic Dynamics and control*, 2, 7–46.

——— (1993) "Seasonality and approximation errors in rational expectations models," *Journal of Econometrics*, 55 (1-2), 21–55.

Iskrev, Nikolay (2010a) "Evaluating the strength of identification in DSGE models. An a priori approach," working paper series, Banco de Portugal.

——— (2010b) "Local identification in DSGE models," *Journal of Monetary Economics*, 57 (2), 189–202, http://ideas.repec.org/a/eee/moneco/v57y2010i2p189-202.html.

——— (2019) "On the sources of information about latent variables in DSGE models," *European Economic Review*, 119, 318–332.

Iskrev, Nikolay and Joao Ritto (2016) "Choosing the variables to estimate singular DSGE models: Comment,"Technical report.

Justiniano, Alejandro, Giorgio E Primiceri, and Andrea Tambalotti (2010) "Investment shocks and business cycles," *Journal of Monetary Economics*, 57 (2), 132–145.

Kay, Steven M (1993) "Fundamentals of statistical signal processing, volume i: Estimation theory (v. 1)," *PTR Prentice-Hall, Englewood Cliffs*.

Marquaridt, Donald W (1970) "Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation," *Technometrics*, 12 (3), 591–612.

McDonald, Daniel J and Cosma Rohilla Shalizi (2022) "Empirical macroeconomics and dsge modeling in statistical perspective," *arXiv preprint arXiv:2210.16224*.

Meng, Xiao-Li and Xianchao Xie (2014) "I got more data, my model is more refined, but my estimator is getting worse! Am I just dumb?" *Econometric Reviews*, 33 (1-4), 218–250.

Palm, Franz C and Th E Nijman (1984) "Missing observations in the dynamic regression model," *Econometrica: journal of the Econometric Society*, 1415–1435.

Plagborg-Møller, Mikkel (2019) "Bayesian inference on structural impulse response functions," *Quantitative Economics*, 10 (1), 145–184.

Qu, Zhongjun and Denis Tkachenko (2012a) "Frequency domain analysis of medium scale DSGE models with application to Smets and Wouters (2007)," in *DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments*, 319–385: Emerald Group Publishing Limited.

——— (2012b) "Identification and frequency domain quasi-maximum likelihood estimation of linearized dynamic stochastic general equilibrium models," *Quantitative Economics*, 3 (1), 95–132, http://ideas.repec.org/a/ecm/quante/v3y2012i1p95-132.html.

——— (2017) "Global identification in DSGE models allowing for indeterminacy," *The Review of Economic Studies*, 84 (3), 1306–1345.

Rao, Calyampudi Radhakrishna (1945) "Information and accuracy attainable in the estimation of statistical parameters," *Bulletin of the Calcutta Mathematical Society*, 37 (3), 81–91.

Sala, Luca (2015) "DSGE models in the frequency domains," *Journal of Applied Econometrics*, 30 (2), 219–240.

Schmitt-Grohé, Stephanie and Martín Uribe (2012) "What's News in Business Cycles," *Econometrica*, 80 (6), 2733–2764, 10.3982/ECTA8050.

Smets, Frank and Rafael Wouters (2007) "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *The American Economic Review*, 97 (3), 586–606.

Sykulski, Adam M, Sofia C Olhede, Arthur P Guillaumin, Jonathan M Lilly, and Jeffrey J Early (2019) "The debiased Whittle likelihood," *Biometrika*, 106 (2), 251–266.

Uhlig, Harald (1999) "A toolkit for analyzing nonlinear dynamic stochastic models easily," *Computational Methods for the Study of Dynamic Economies*, 30–61.

Whittle, P. (1953) "The Analysis of Multiple Stationary Time Series," *Journal of the Royal Statistical Society. Series B (Methodological)*, 15 (1), pp. 125–139, http://www.jstor.org/stable/2983728.

Appendix

A Appendix to Section 2

A.1 Angeletos et al. (2018) model

Table 7: Parameter values, ACD (2018) model

	parameter	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

B Appendix to Section 4

B.1 Q-Q plots

The Q-Q plots compare the quantiles of the sampling distributions of the estimated parameters to those of the standard normal distribution. Parameter estimates are normalized by subtracting their respective means and dividing by their respective standard deviations. The black lines represent the 45-degree line.

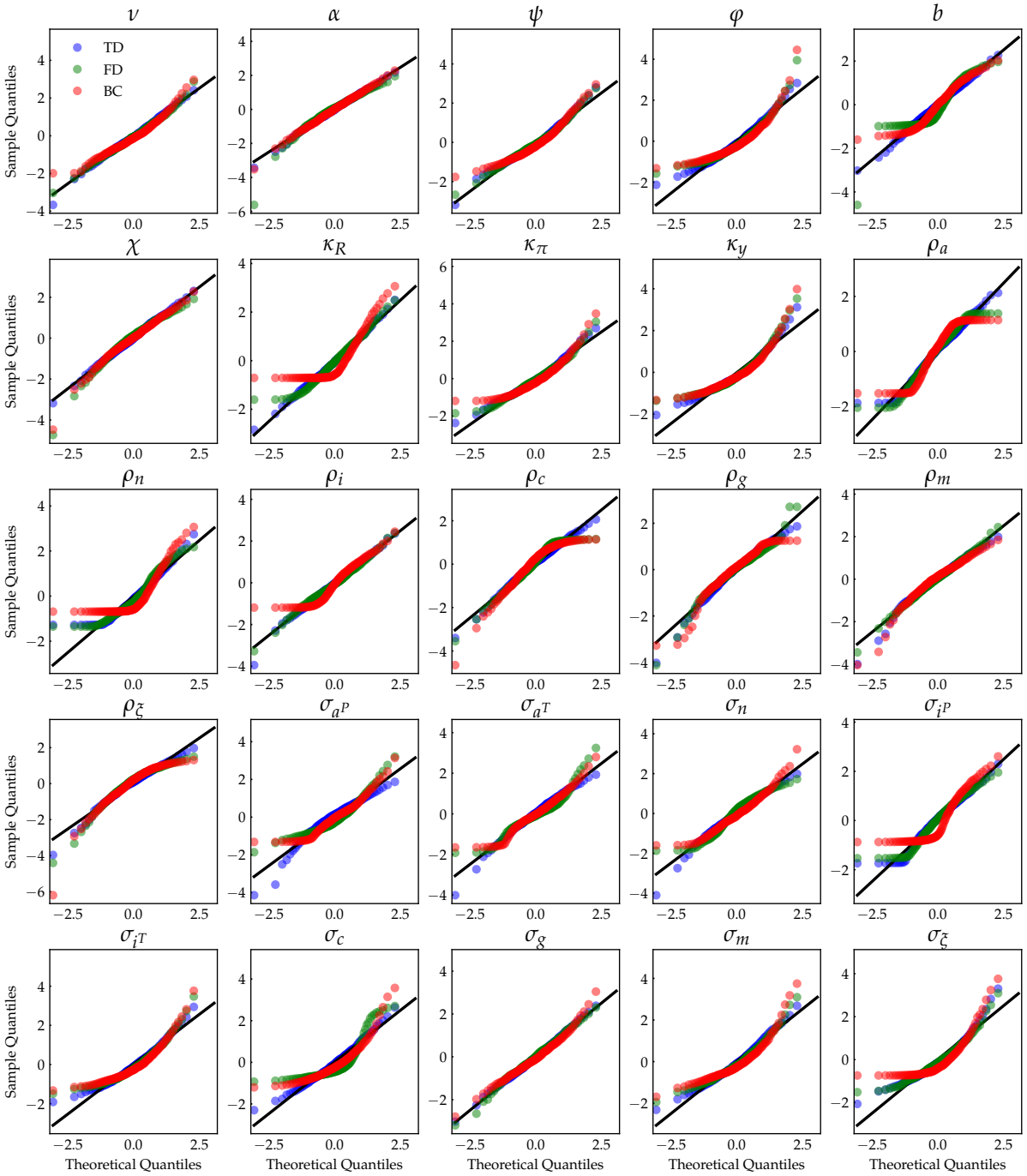


Figure 11: Baseline parametrization. The results are based on 1000 MC replications with a sample size of $T = 192$.

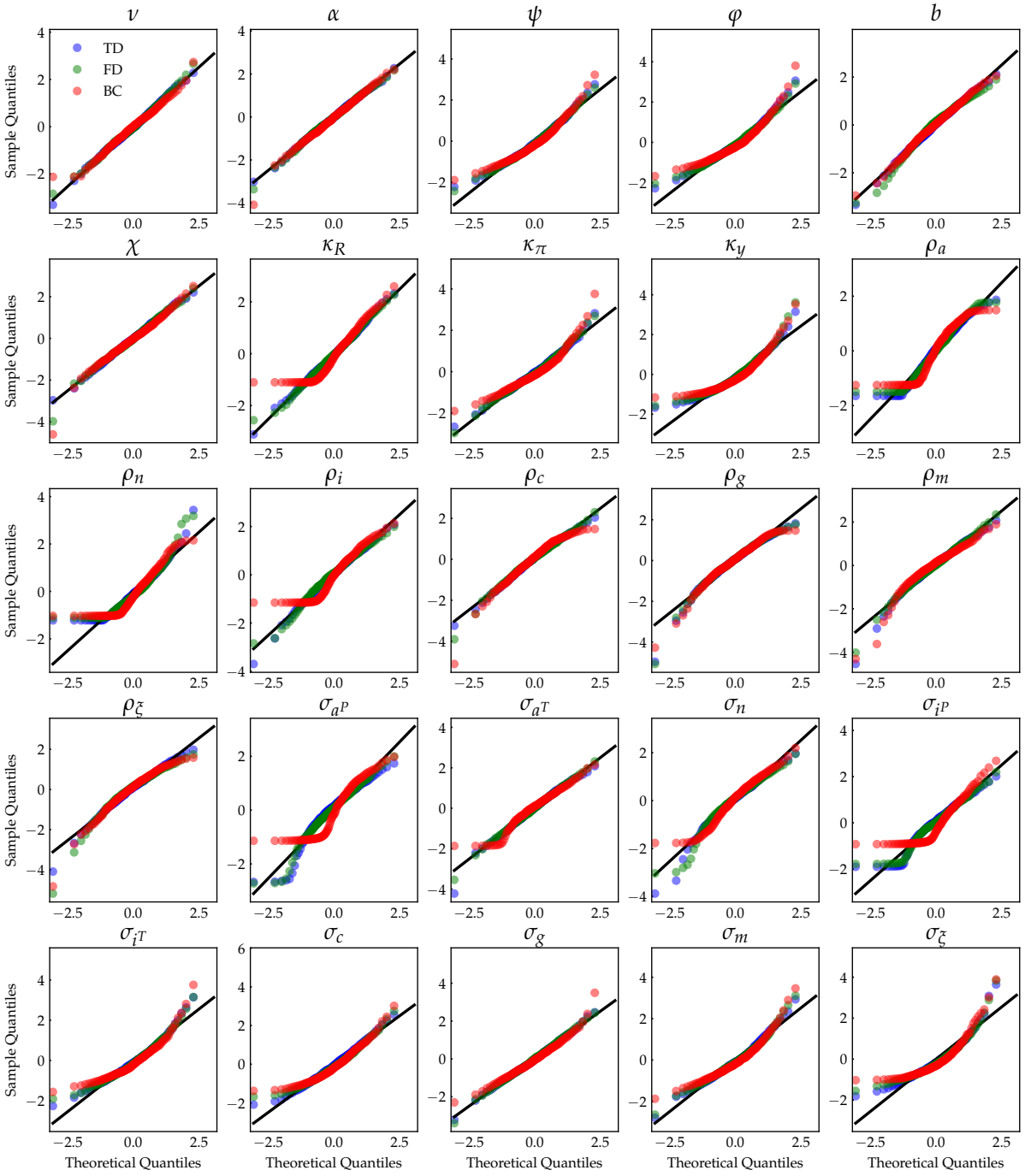


Figure 12: Alternative parametrization of the ACD model. The results are based on 1000 MC replications with a sample size of $T = 192$.

C Appendix to Section 5

This appendix contains additional results complementing Section 5 in the main text.

C.1 Results for the baseline parametrization

Table 8: Estimation uncertainty and efficiency loss: MC vs CRLBs

	estimation uncertainty					efficiency loss			
	all freqs			BC freqs		BC vs all freqs			
parameter	MC-TD	MC-FD	CRLB	MC	CRLB	MC-TD	MC-FD	CRLB	CRLB(∞)
ν	26.7	39.3	25.7	76.9	55.6	2.9	2.0	2.2	2.3
α	4.0	6.0	3.8	7.4	6.3	1.8	1.2	1.6	1.7
ψ	24.0	31.7	21.7	49.2	57.6	2.1	1.6	2.7	3.0
φ	26.2	44.7	24.8	48.8	42.6	1.9	1.1	1.7	1.8
b	4.2	13.7	4.0	13.1	7.6	3.1	1.0	1.9	2.0
χ	2.7	5.7	2.6	8.2	6.4	3.0	1.5	2.4	2.6
κ_R	35.0	44.3	33.8	76.2	118.6	2.2	1.7	3.5	3.7
κ_π	15.6	25.2	14.4	42.2	32.7	2.7	1.7	2.3	2.5
κ_y	48.2	131.8	44.1	96.7	106.2	2.0	0.7	2.4	3.0
ρ_a	54.4	71.0	57.1	91.4	185.6	1.7	1.3	3.3	3.4
ρ_n	77.4	126.2	85.7	116.3	279.1	1.5	0.9	3.3	3.4
ρ_i	19.8	30.0	19.7	65.7	63.7	3.3	2.2	3.2	3.4
ρ_c	3.5	5.7	3.5	7.1	8.9	2.0	1.2	2.5	3.2
ρ_g	7.3	11.4	6.2	28.1	15.4	3.9	2.5	2.5	2.8
ρ_m	8.8	11.1	8.0	23.9	18.1	2.7	2.1	2.3	2.4
ρ_ξ	4.8	12.4	4.4	13.1	11.4	2.7	1.1	2.6	3.0
σ_{a^P}	22.5	105.7	18.5	94.5	73.1	4.2	0.9	4.0	4.1
σ_{a^T}	19.2	55.3	18.4	71.0	66.2	3.7	1.3	3.6	3.7
σ_n	23.2	38.2	23.5	100.9	109.2	4.3	2.6	4.6	4.8
σ_{i^P}	56.7	78.1	44.1	122.5	208.5	2.2	1.6	4.7	6.1
σ_{i^T}	28.2	46.9	26.6	69.2	66.6	2.5	1.5	2.5	2.6
σ_c	42.2	301.9	44.1	136.6	86.0	3.2	0.5	2.0	2.2
σ_g	5.5	7.3	5.5	16.0	11.9	2.9	2.2	2.2	2.2
σ_m	11.3	17.6	11.2	34.3	25.0	3.0	1.9	2.2	2.4
σ_ξ	49.2	53.1	42.3	82.8	107.8	1.7	1.6	2.6	2.8

Note: Baseline parametrization. Estimation uncertainty is reported as the MC standard deviation or CRLB, expressed as percentage of the absolute value of the true parameter value. Efficiency loss is defined as the ratio of the standard deviations or CRLBs obtained using BC frequencies only vs. using all frequencies. CRLBs are computed with the exact FIM for $T = 192$ whereas CRLB(∞) is computed with the asymptotic FIM.

C.2 MC vs CRLBs: alternative parametrization

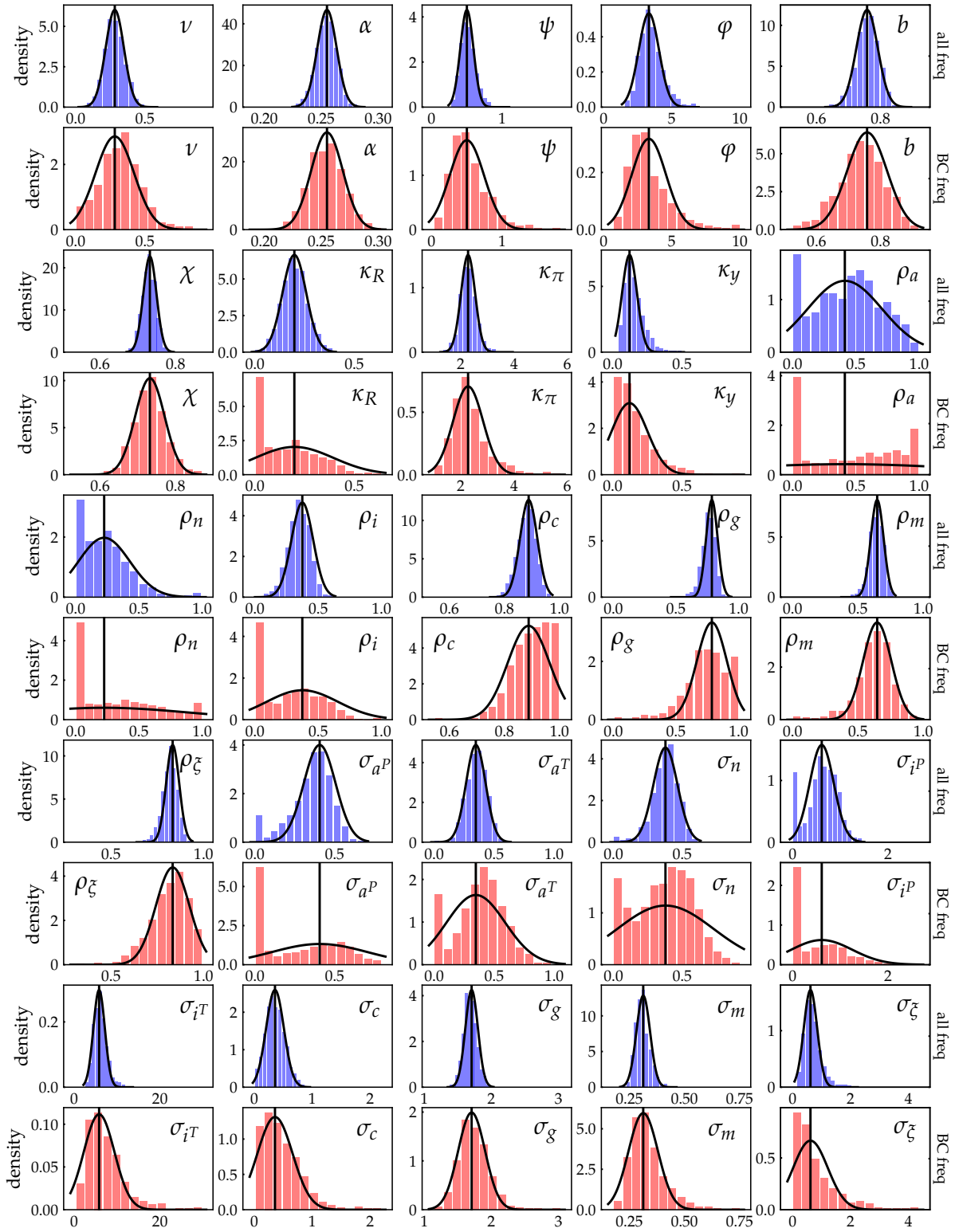


Figure 13: Each plot of the figure shows a histogram of the MC sampling distribution for the respective parameter, along with a gaussian density curve centered at the true parameter value and with a standard deviation equal to the CRLB. The sample size is $T = 192$.

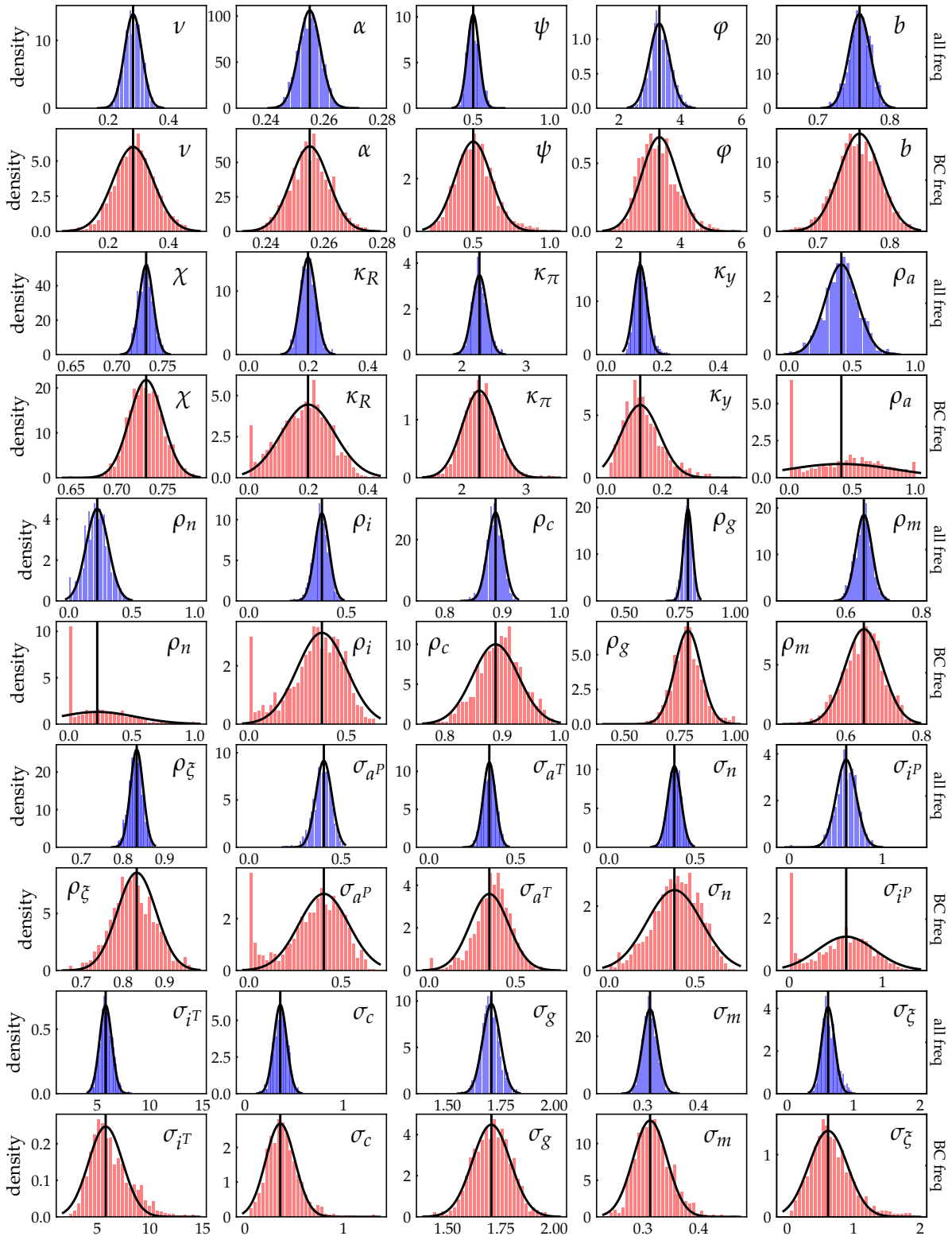


Figure 14: See the Notes to Figure 13. The sample size is $T = 1000$.

D Appendix to Section 6

This appendix contains additional results complementing Section 6 in the main text.

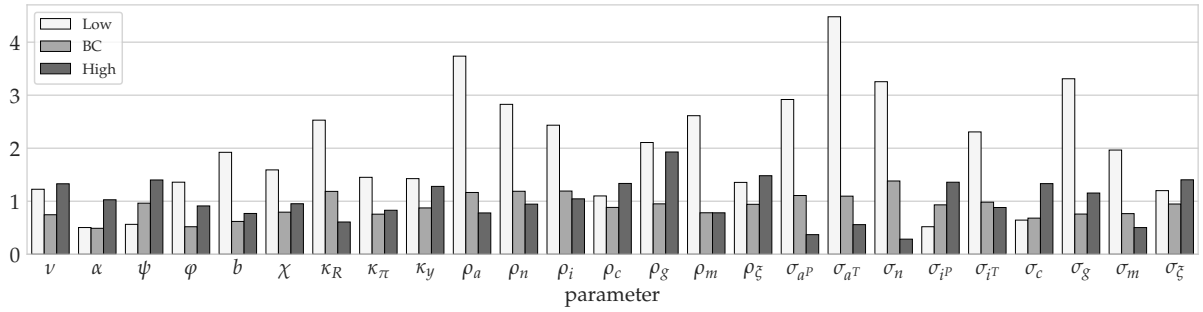


Figure 15: Log of the predicted relative efficiency of different band-spectral estimators. The relative efficiency is defined as the ratio of the CRLB value using a specific band of frequencies to the CRLB value using all frequencies. The sample size is $T = 192$.

Table 9: Decomposition of the predicted relative efficiency

parameter	Low	BC	High	BC+Low	BC+High
ν	$3.4 = 3.0 \times 1.13$	$2.1 = 1.7 \times 1.22$	$3.8 = 1.4 \times 2.80$	$1.3 = 1.5 \times 0.89$	$1.5 = 1.1 \times 1.37$
α	$1.7 = 2.3 \times 0.73$	$1.6 = 1.7 \times 0.94$	$2.8 = 1.4 \times 1.93$	$1.1 = 1.4 \times 0.81$	$1.4 = 1.1 \times 1.21$
ψ	$1.8 = 3.3 \times 0.52$	$2.6 = 1.8 \times 1.45$	$4.1 = 1.3 \times 3.15$	$1.2 = 1.6 \times 0.74$	$1.9 = 1.0 \times 1.77$
φ	$3.9 = 5.6 \times 0.69$	$1.7 = 1.8 \times 0.91$	$2.5 = 1.2 \times 2.03$	$1.3 = 1.7 \times 0.73$	$1.2 = 1.0 \times 1.19$
b	$6.8 = 4.9 \times 1.40$	$1.9 = 1.9 \times 0.98$	$2.2 = 1.2 \times 1.78$	$1.5 = 1.8 \times 0.85$	$1.2 = 1.0 \times 1.15$
χ	$4.9 = 3.6 \times 1.38$	$2.2 = 2.0 \times 1.12$	$2.6 = 1.2 \times 2.11$	$1.5 = 1.7 \times 0.88$	$1.3 = 1.0 \times 1.22$
κ_R	$12.5 = 5.5 \times 2.28$	$3.3 = 2.4 \times 1.34$	$1.8 = 1.1 \times 1.64$	$2.8 = 2.2 \times 1.24$	$1.1 = 1.0 \times 1.08$
κ_π	$4.3 = 4.2 \times 1.03$	$2.1 = 1.9 \times 1.11$	$2.3 = 1.2 \times 1.87$	$1.5 = 1.7 \times 0.86$	$1.3 = 1.0 \times 1.24$
κ_y	$4.2 = 4.4 \times 0.95$	$2.4 = 1.9 \times 1.27$	$3.6 = 1.2 \times 2.93$	$1.2 = 1.7 \times 0.71$	$1.9 = 1.0 \times 1.82$
ρ_a	$42.0 = 24.7 \times 1.70$	$3.2 = 3.2 \times 1.01$	$2.2 = 1.1 \times 2.06$	$2.7 = 3.1 \times 0.85$	$1.1 = 1.0 \times 1.13$
ρ_n	$16.9 = 2.2 \times 7.71$	$3.3 = 1.5 \times 2.15$	$2.6 = 1.6 \times 1.56$	$2.8 = 1.3 \times 2.20$	$1.2 = 1.1 \times 1.03$
ρ_i	$11.4 = 3.4 \times 3.35$	$3.3 = 1.2 \times 2.65$	$2.8 = 2.0 \times 1.44$	$2.8 = 1.2 \times 2.37$	$1.1 = 1.0 \times 1.05$
ρ_c	$3.0 = 3.1 \times 0.98$	$2.4 = 1.9 \times 1.29$	$3.8 = 1.3 \times 2.97$	$1.1 = 1.6 \times 0.70$	$2.0 = 1.1 \times 1.86$
ρ_g	$8.2 = 1.2 \times 6.74$	$2.6 = 2.3 \times 1.14$	$6.9 = 2.7 \times 2.55$	$1.3 = 1.1 \times 1.21$	$1.8 = 1.7 \times 1.05$
ρ_m	$13.6 = 12.0 \times 1.14$	$2.2 = 2.3 \times 0.95$	$2.2 = 1.1 \times 1.95$	$1.7 = 2.3 \times 0.77$	$1.2 = 1.0 \times 1.17$
ρ_ξ	$3.9 = 2.4 \times 1.59$	$2.6 = 1.9 \times 1.37$	$4.4 = 1.4 \times 3.25$	$1.3 = 1.5 \times 0.85$	$1.9 = 1.1 \times 1.74$
σ_{a^P}	$18.5 = 4.0 \times 4.59$	$3.0 = 1.9 \times 1.58$	$1.4 = 1.2 \times 1.18$	$2.8 = 1.7 \times 1.63$	$1.0 = 1.0 \times 1.01$
σ_{a^T}	$88.2 = 52.2 \times 1.69$	$3.0 = 3.1 \times 0.95$	$1.7 = 1.1 \times 1.65$	$2.8 = 3.1 \times 0.88$	$1.1 = 1.0 \times 1.07$
σ_n	$25.9 = 2.9 \times 9.04$	$4.0 = 1.5 \times 2.58$	$1.3 = 1.5 \times 0.90$	$3.5 = 1.4 \times 2.58$	$1.1 = 1.1 \times 0.99$
σ_{i^P}	$1.7 = 1.3 \times 1.34$	$2.5 = 2.1 \times 1.22$	$3.9 = 3.0 \times 1.29$	$1.1 = 1.1 \times 1.05$	$1.9 = 1.7 \times 1.11$
σ_{i^T}	$10.0 = 7.1 \times 1.41$	$2.7 = 1.9 \times 1.38$	$2.4 = 1.2 \times 2.03$	$2.3 = 1.9 \times 1.23$	$1.2 = 1.0 \times 1.16$
σ_c	$1.9 = 4.0 \times 0.47$	$2.0 = 1.8 \times 1.07$	$3.8 = 1.2 \times 3.04$	$1.1 = 1.7 \times 0.67$	$1.5 = 1.0 \times 1.47$
σ_g	$27.4 = 4.1 \times 6.74$	$2.1 = 1.9 \times 1.15$	$3.2 = 1.2 \times 2.56$	$2.1 = 1.7 \times 1.22$	$1.1 = 1.0 \times 1.06$
σ_m	$7.1 = 4.5 \times 1.58$	$2.2 = 1.9 \times 1.14$	$1.7 = 1.2 \times 1.35$	$1.8 = 1.7 \times 1.02$	$1.1 = 1.0 \times 1.11$
σ_ξ	$3.3 = 4.0 \times 0.83$	$2.6 = 1.8 \times 1.40$	$4.1 = 1.2 \times 3.26$	$1.3 = 1.7 \times 0.79$	$1.7 = 1.0 \times 1.68$

Note: The predicted relative efficiency is decomposed using (6.1) into a relative parameter sensitivity factor and a relative parameter interdependence factor.

D.1 Predicted parameter interdependence

Figures 16 – 20 display the highest pairwise and multiple correlation coefficients for the score of each parameter with the scores of 2 and 3 other parameters, as well as the multiple correlation coefficient with the scores of all other parameters. Results are presented for the low, BC, and high frequencies, as well as for the full spectrum.

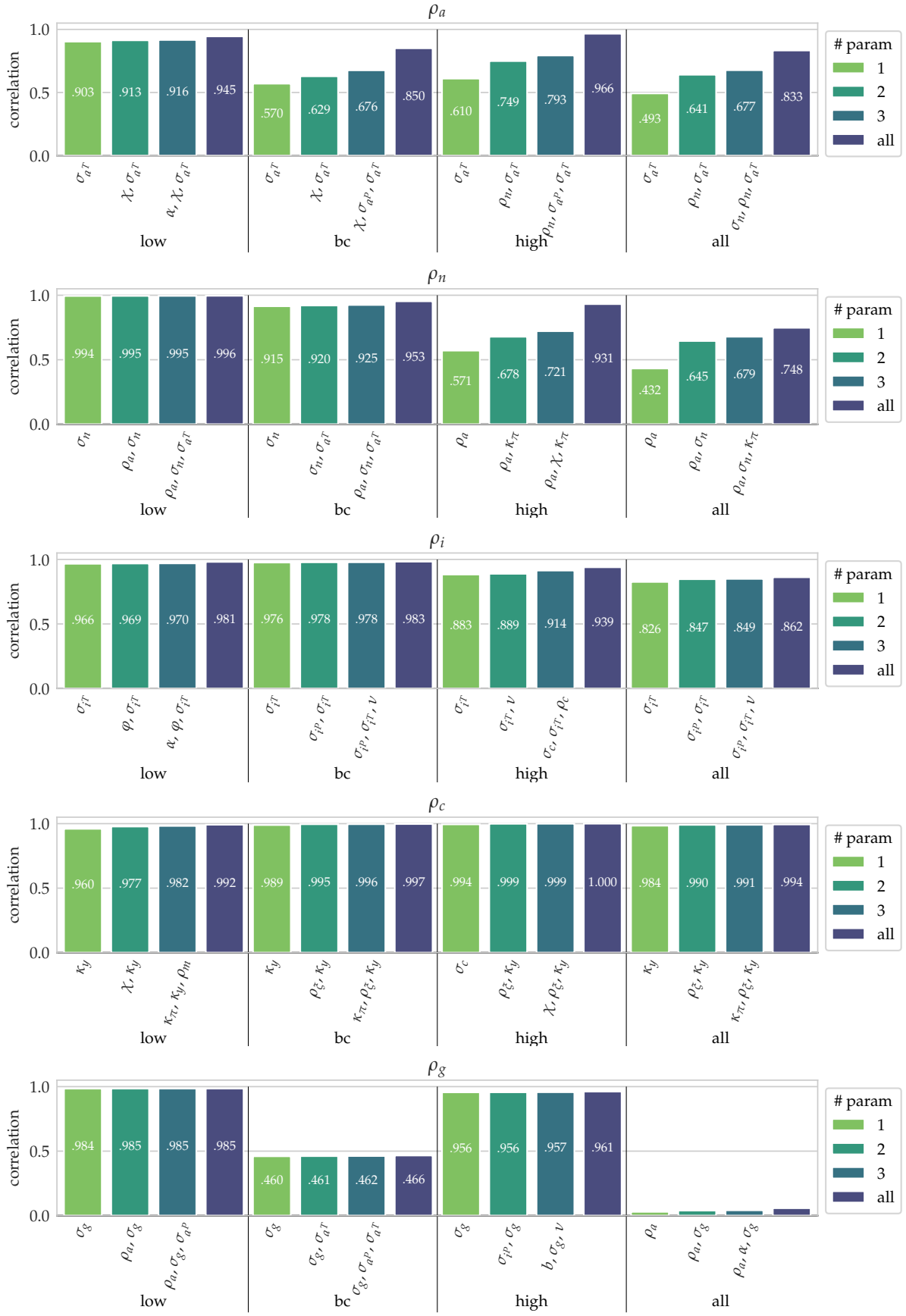


Figure 16: Highest score collinearity ($\max_j \{\mathrm{corr}\,(\frac{\partial \ell}{\partial \theta_i}, \frac{\partial \ell}{\partial \boldsymbol{\theta}_j})\}$) for different numbers of parameters ($\dim(\boldsymbol{\theta}_j) = 1$, 2, 3, and 24 (labeled "all")) and different frequency bands ("low", "bc", "high", and "all").

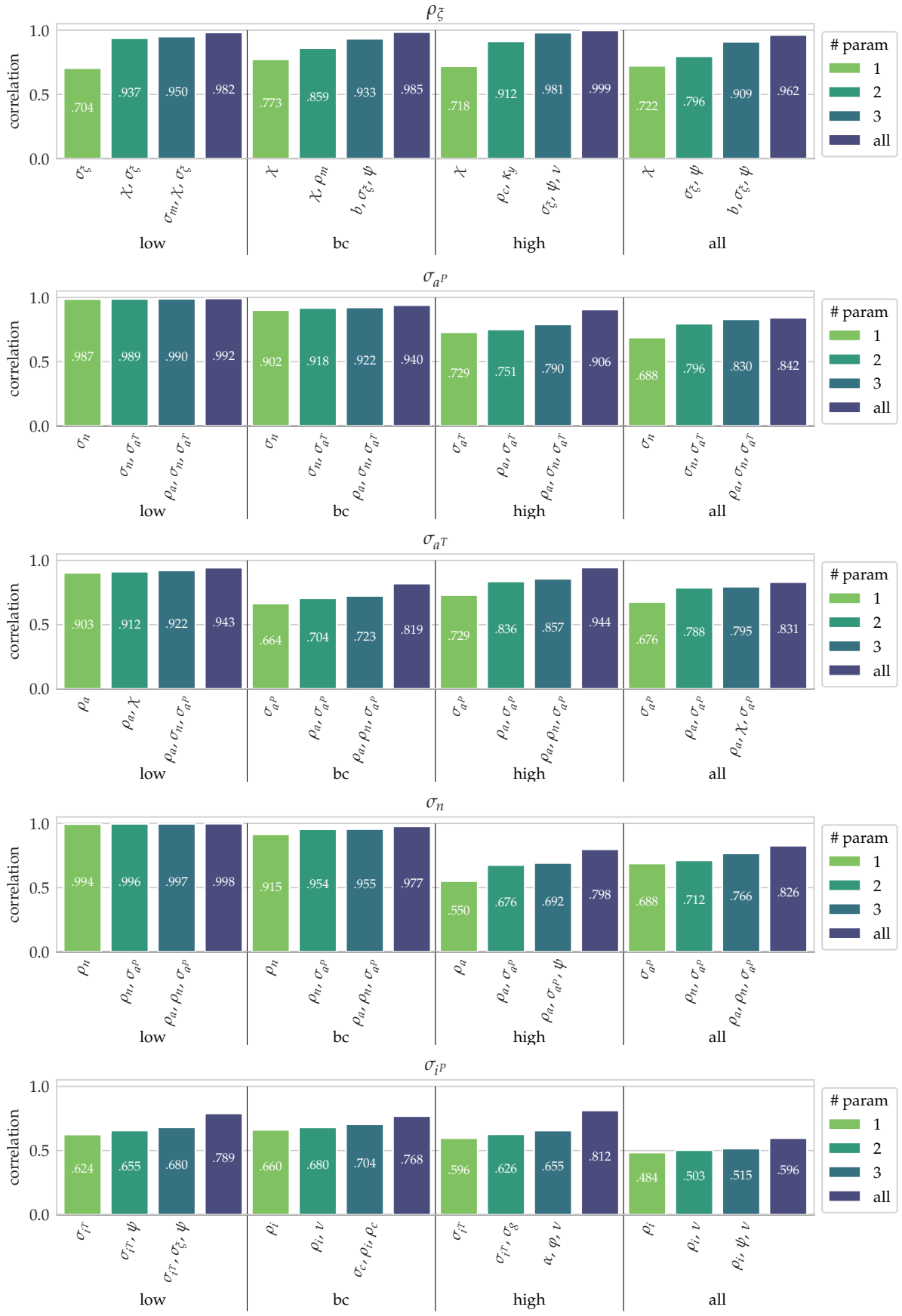


Figure 17: See Note to Figure 16.

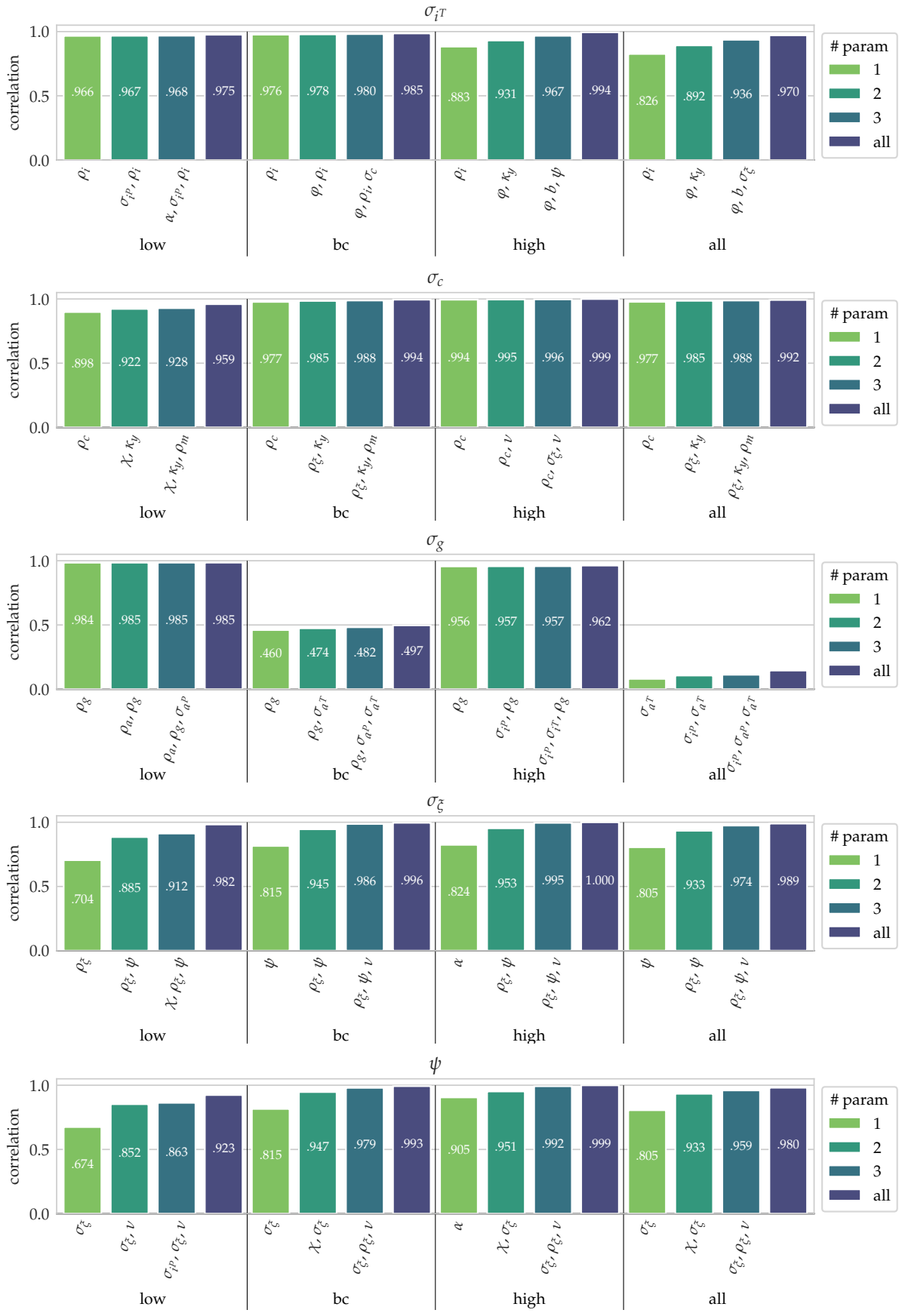


Figure 18: See Note to Figure 16.

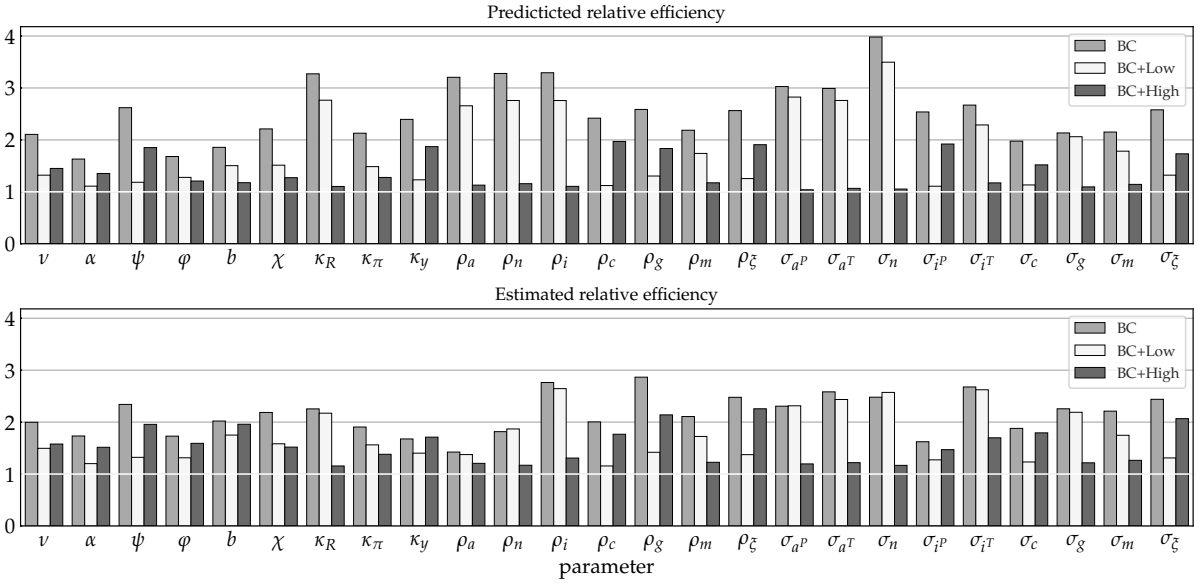
Figure 19: See Note to Figure 16.

Figure 20: See Note to Figure 16.

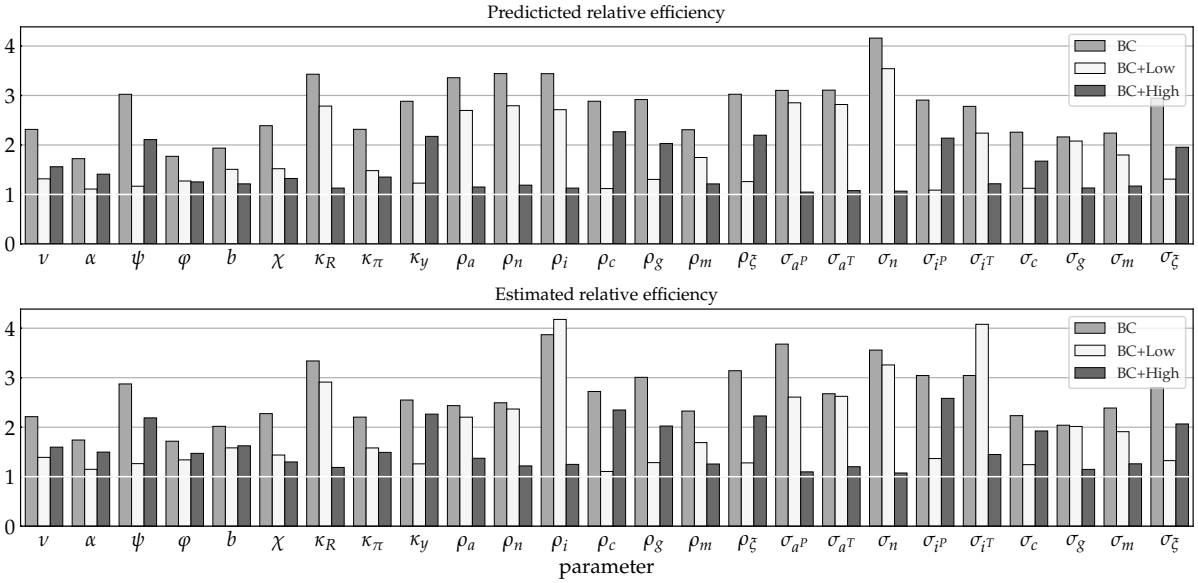
D.2 More Monte Carlo simulations

This section compares the CRLB-based predictions with the MC estimates of the efficiency loss for three band spectral estimators: using BC, BC+Low, and BC+High frequencies. Subsection D.2.1 presents the results for the case of joint estimation of all model parameters, while Subsection D.2.2 shows the results for the conditional case, where one parameter is estimated at a time.

D.2.1 Joint estimation: BC vs BC+Low vs BC+High



(a) T=192.



(b) T=1000.

Figure 21: Predicted (top panel of each subplot) and MC-estimated (bottom panel) relative efficiency of the three band spectral estimators. The relative efficiency is defined as the ratio of the MC standard deviation or CRLB using a band of frequencies to the value using all frequencies.

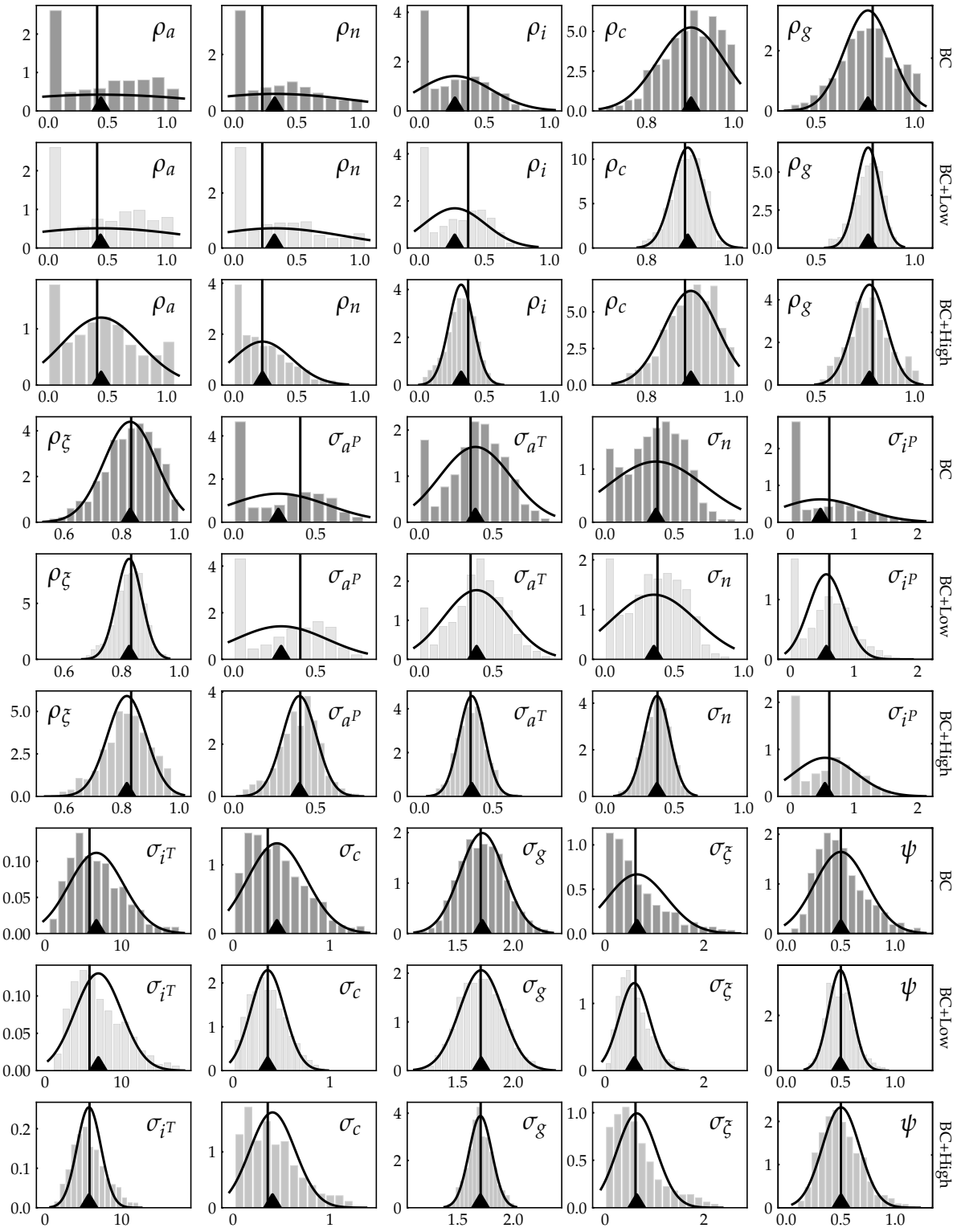


Figure 22: Joint estimation using the Whittle likelihood. The figure shows MC sampling distributions along with gaussian densities centered on the true values, with standard deviations equal to the marginal CRLB. The sample size is $T = 192$.

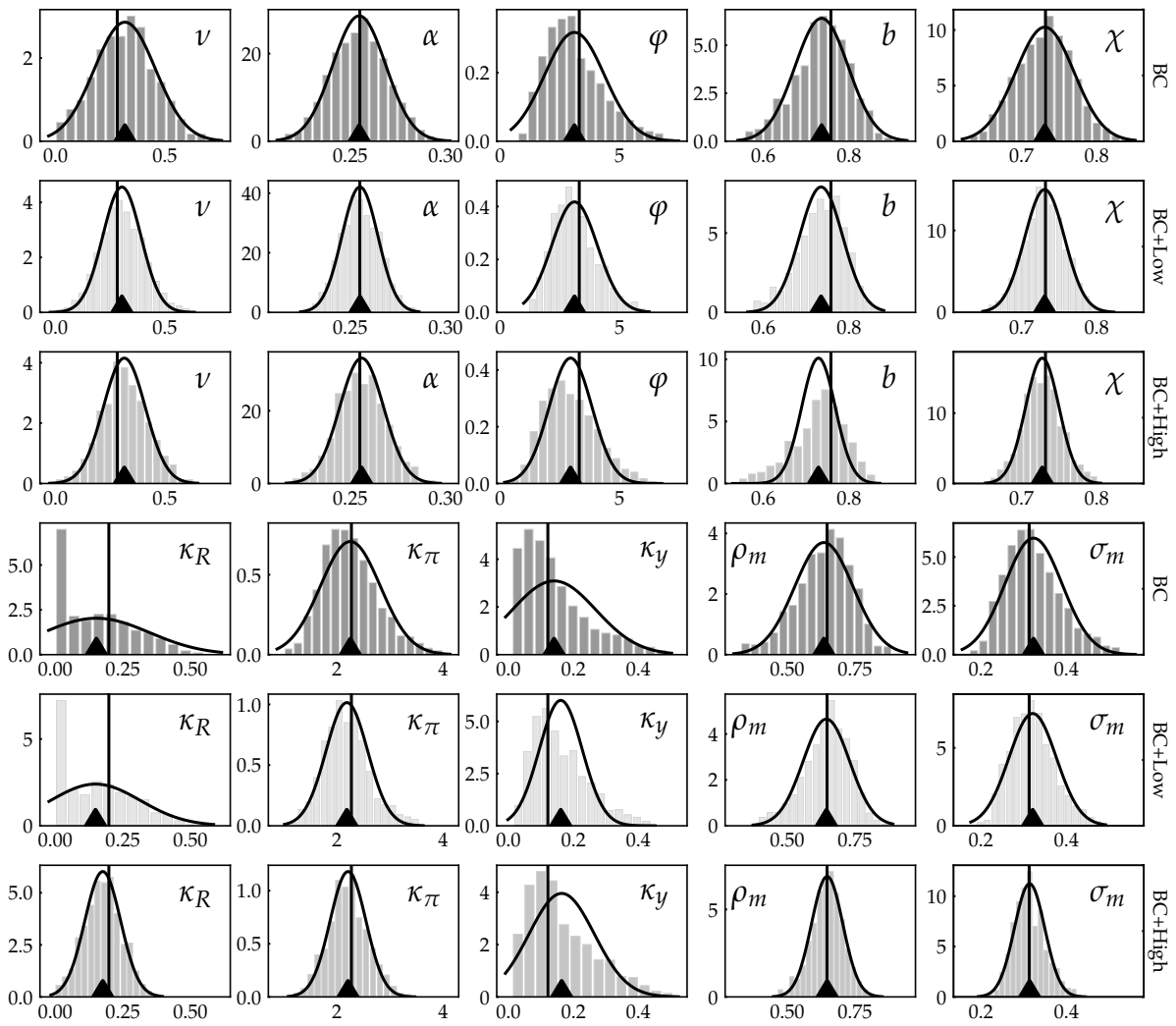


Figure 23: Continue Figure 22.

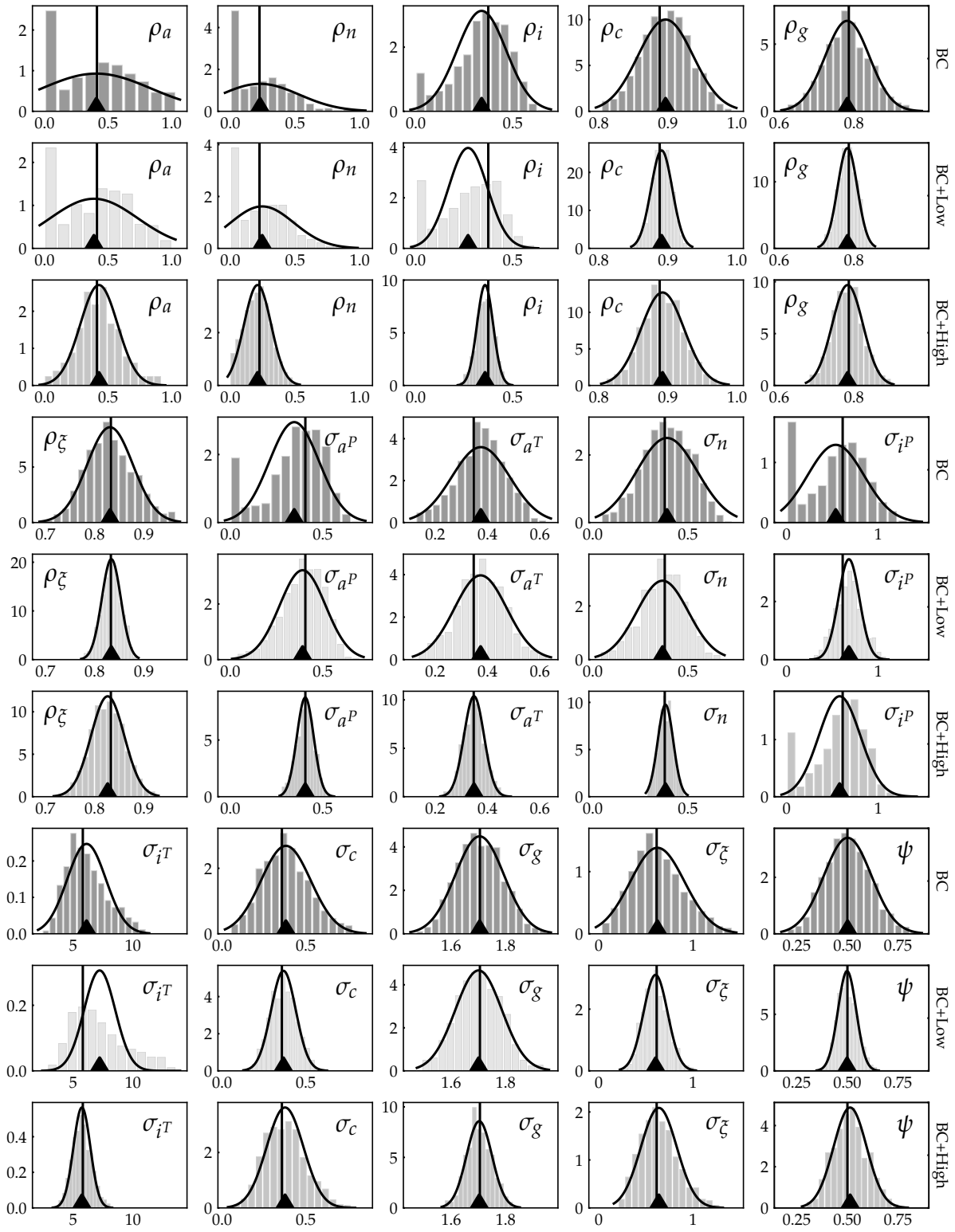


Figure 24: See the note to Figure 22. The sample size is $T = 1000$.

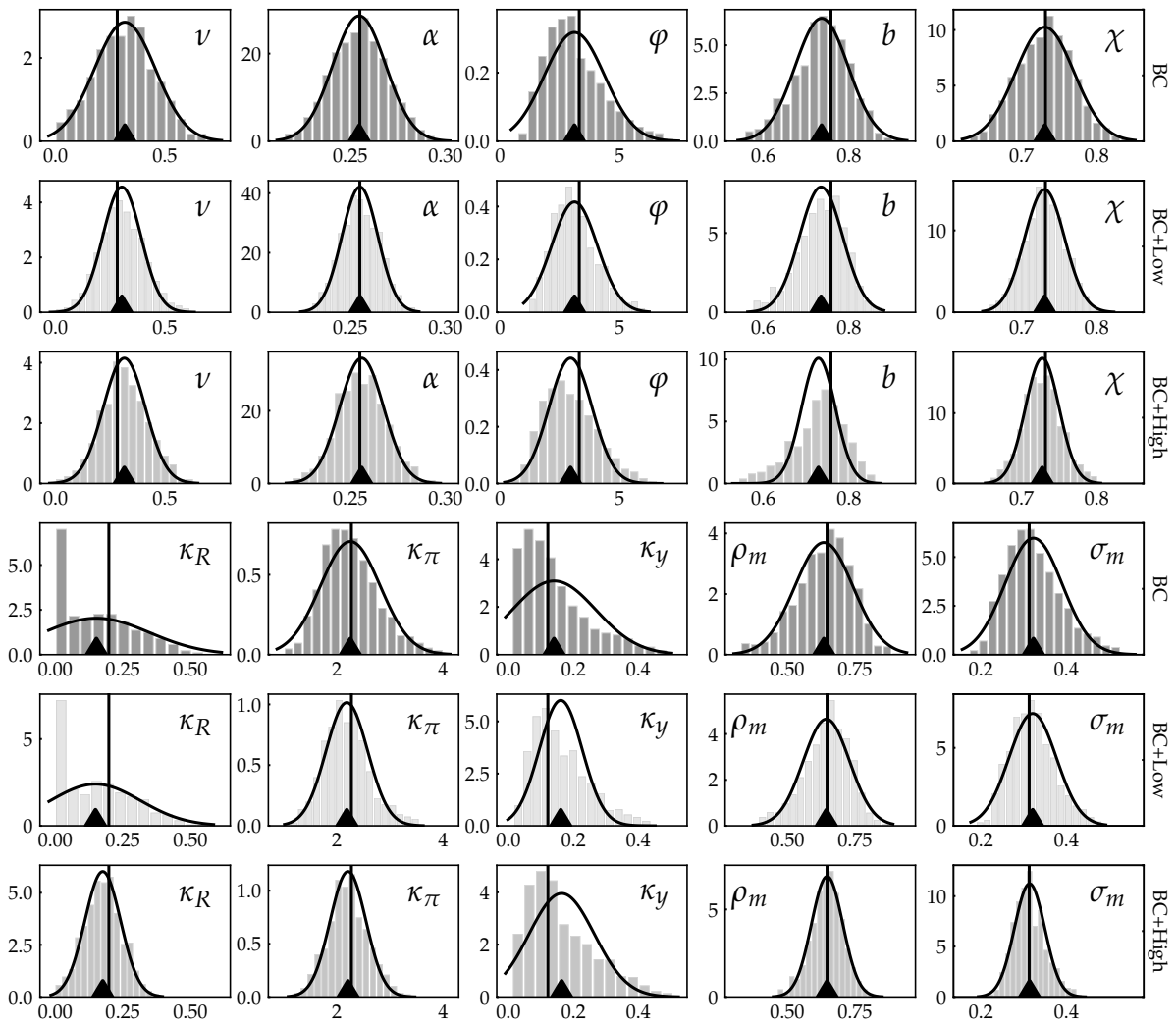
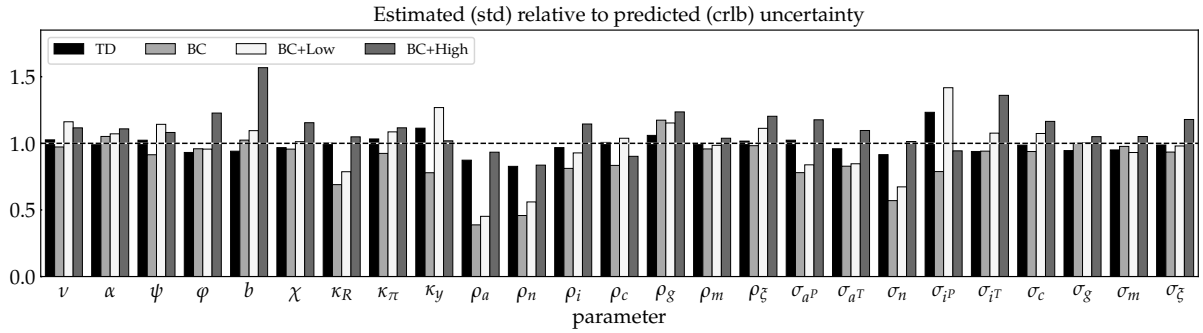
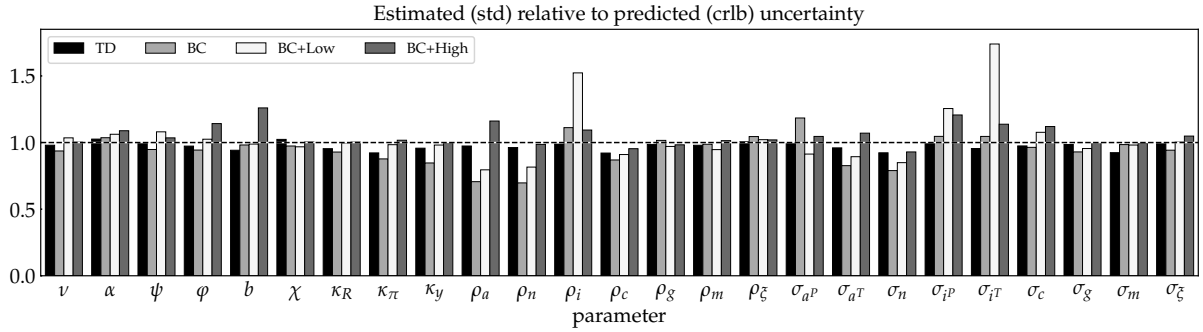


Figure 25: Continue Figure 24.



(a) T=192.

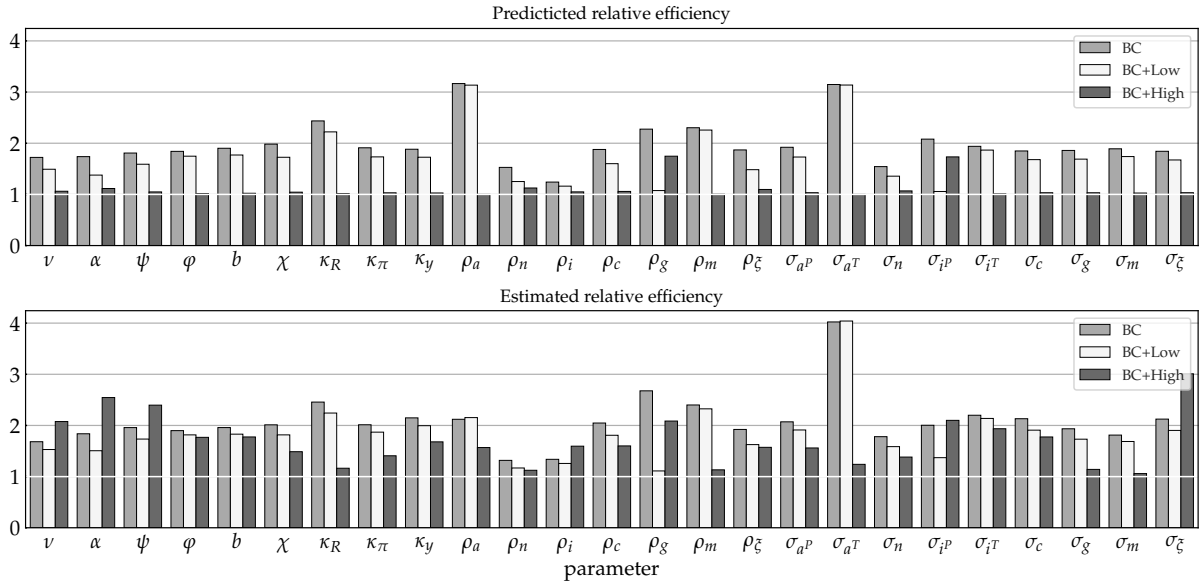


(b) T=1000.

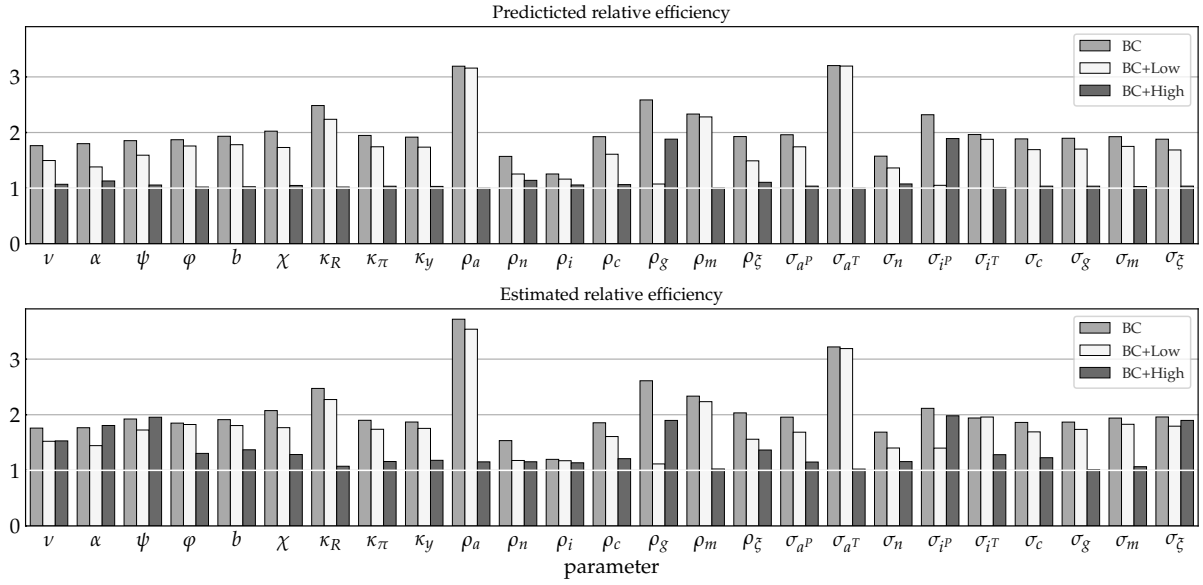
Figure 26: Ratios of Whittle-based unconditional MC standard deviations to unconditional CRLBs. TD represents the exact time domain MLE, while the other three are band spectral Whittle estimators using frequencies from the BC, BC+Low, and BC+High frequency bands.

D.2.2 Conditional estimation: BC vs BC+Low vs BC+High

The following results are for conditional estimation, where only one parameter is estimated at a time while the others are treated as known. The predicted efficiency loss is based on the conditional CRLBs in the band spectral and full information cases (see equation (6.1)).



(a) T=192.



(b) T=1000.

Figure 27: Predicted (top panel of each subplot) and MC-estimated (bottom panel) relative efficiency of the three band spectral estimators.

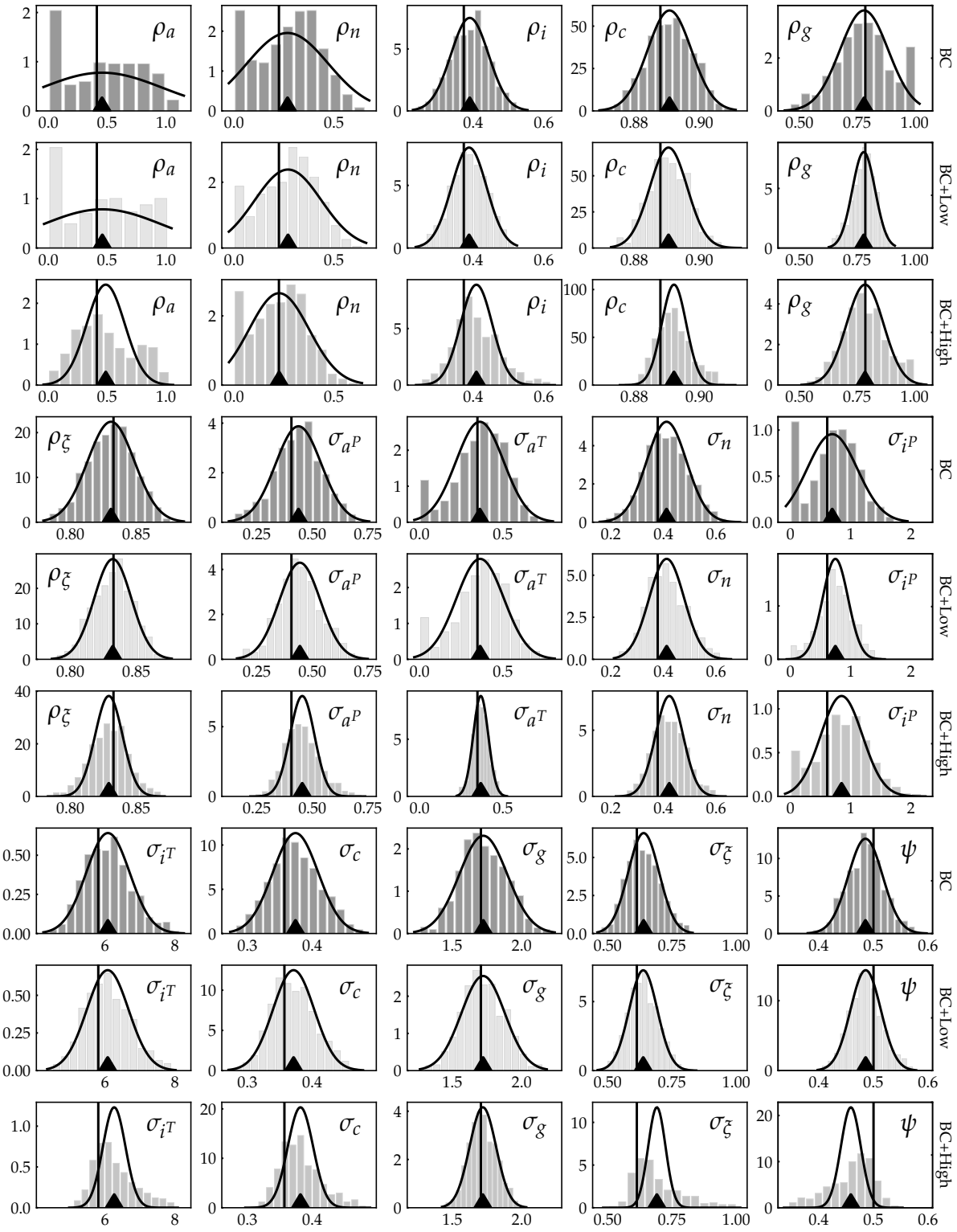


Figure 28: Conditional estimation using the Whittle likelihood. The figure shows MC sampling distributions and gaussian densities centered on the true values with standard deviation equal to the conditional CRLB. The sample size is $T = 192$.

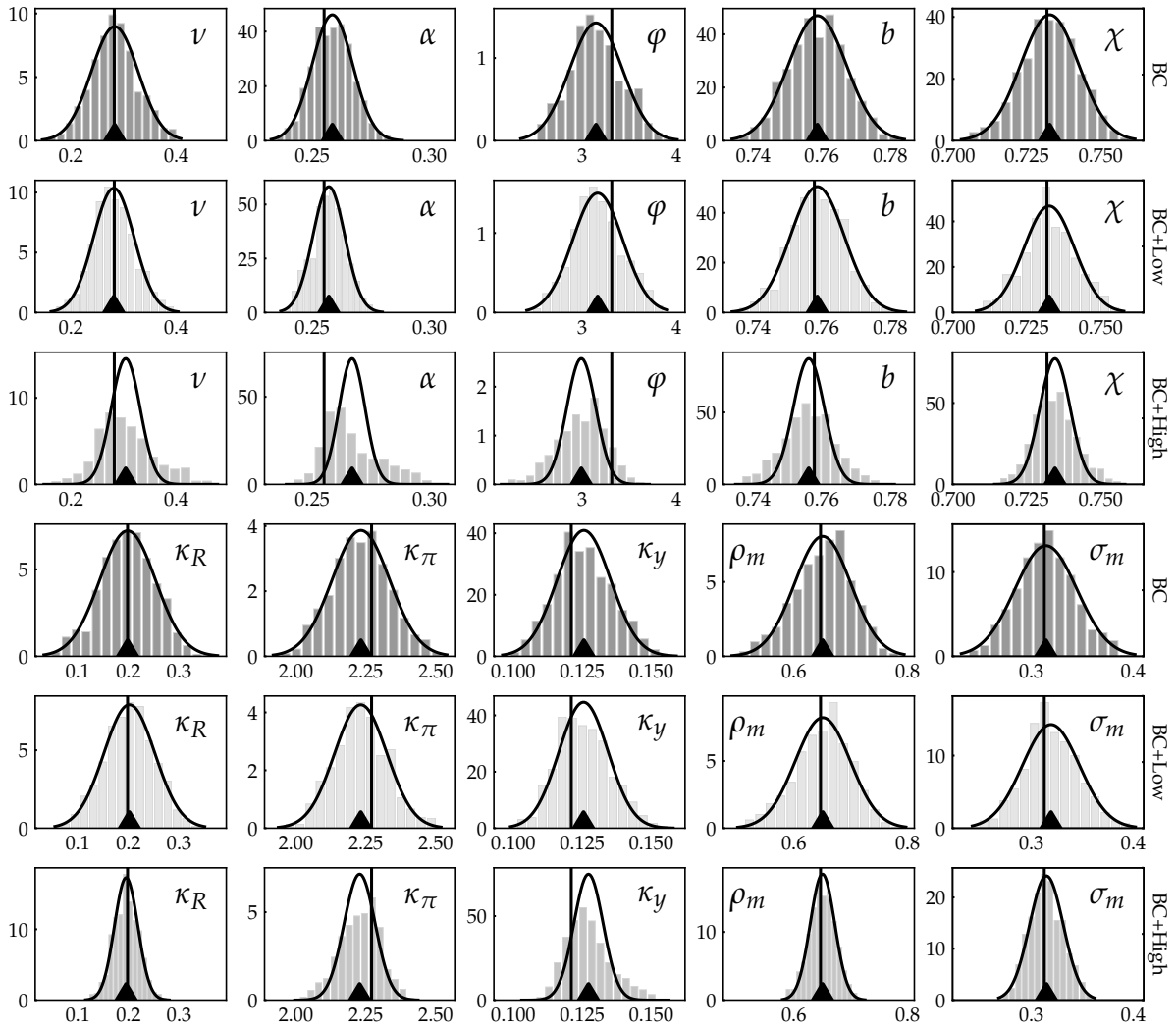


Figure 29: Continue Figure 28.

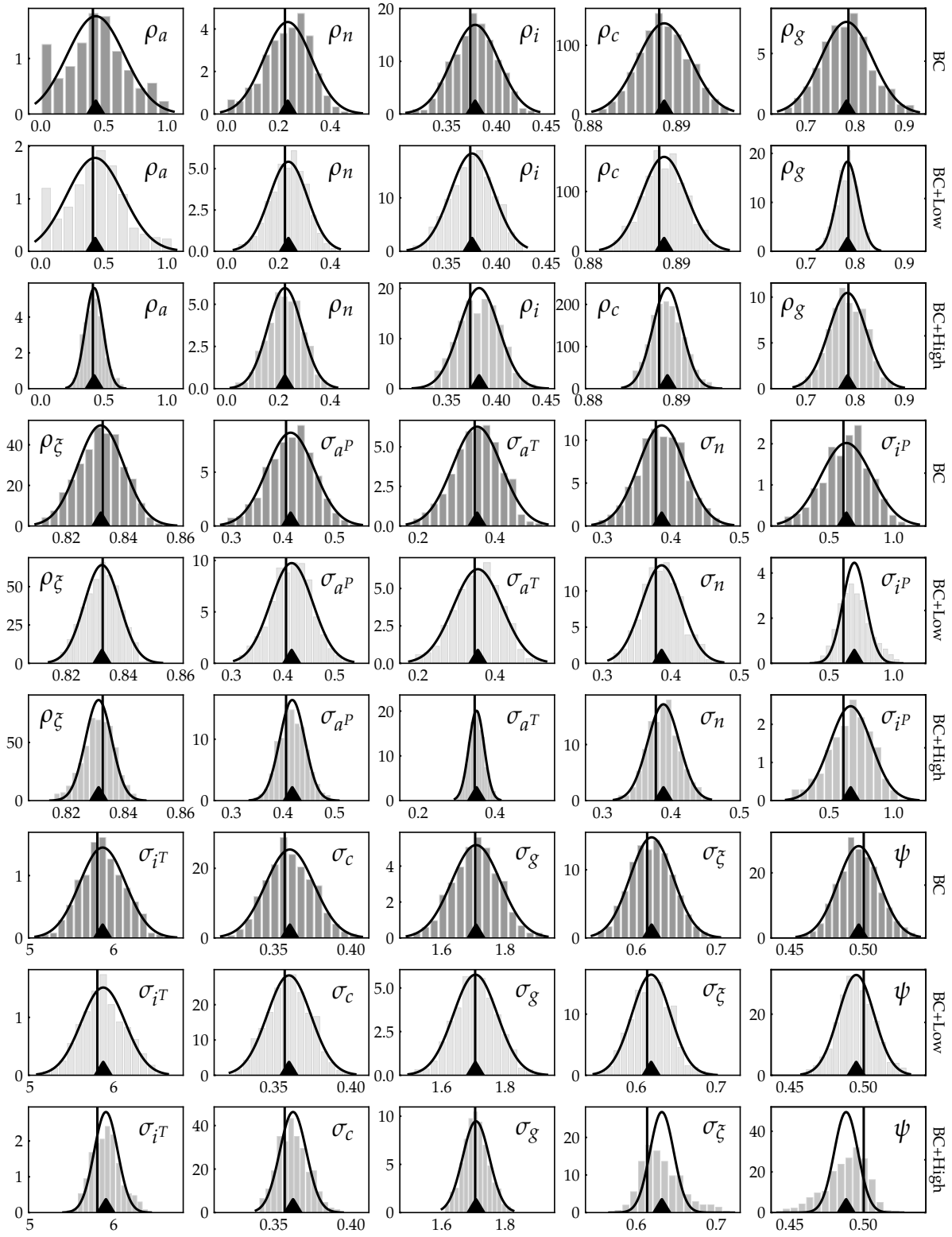


Figure 30: See the note to Figure 28. The sample size is $T = 1000$.

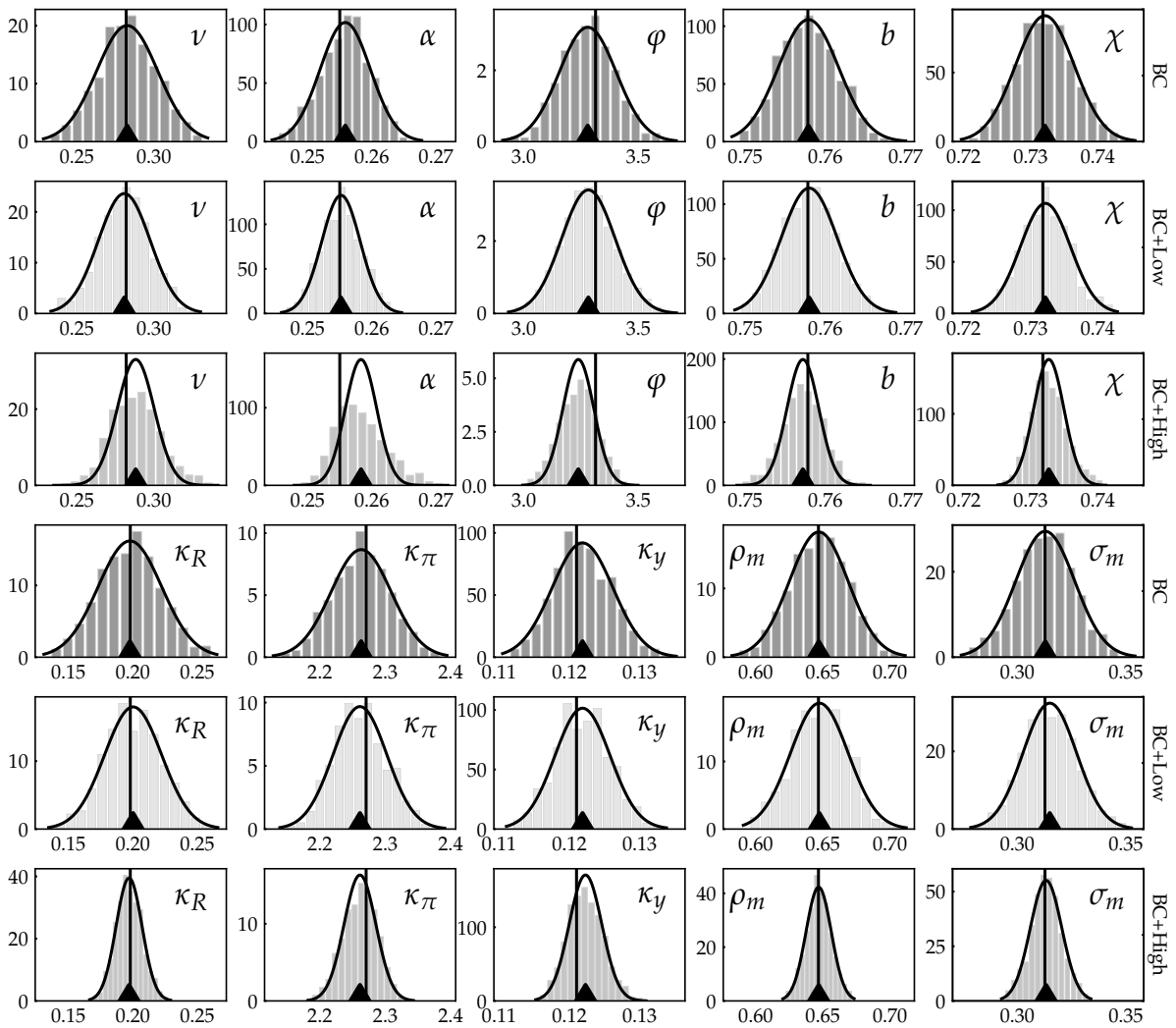
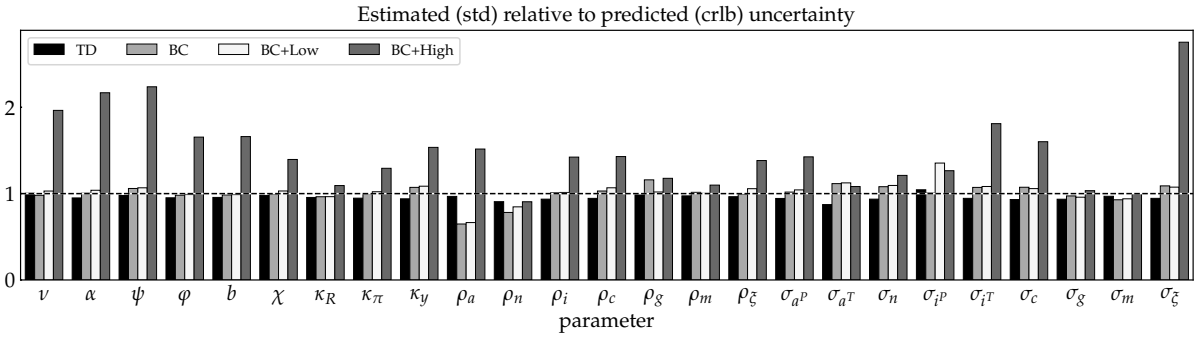
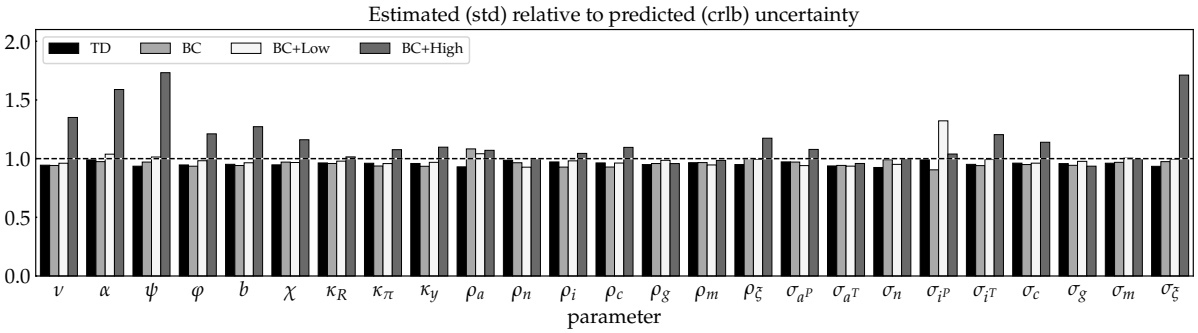


Figure 31: Continue Figure 30.



(a) T=192.

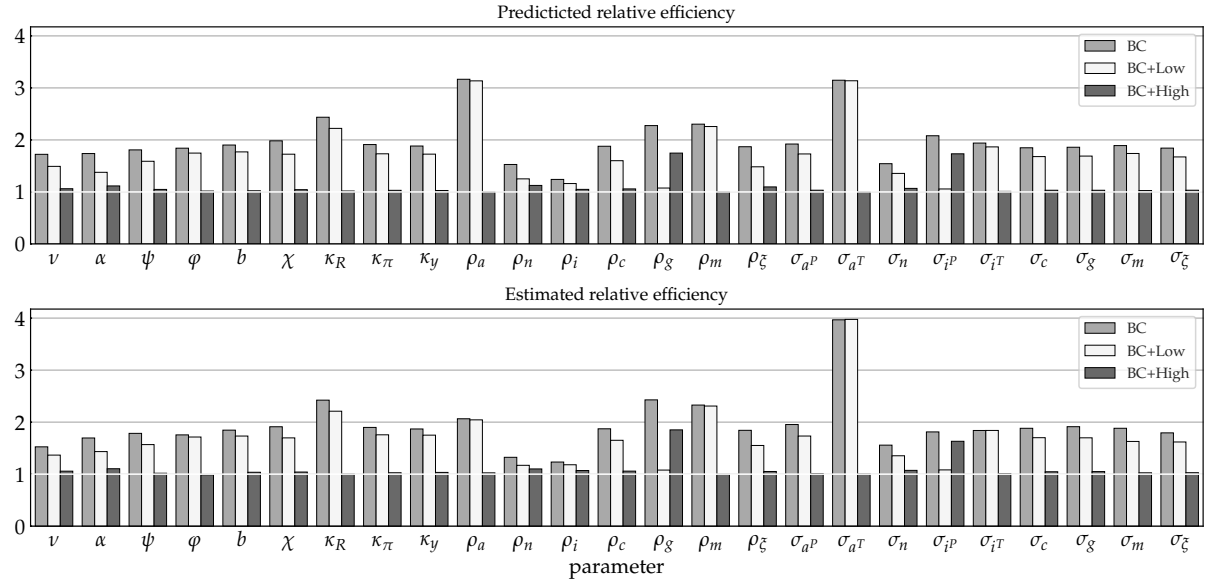


(b) T=1000.

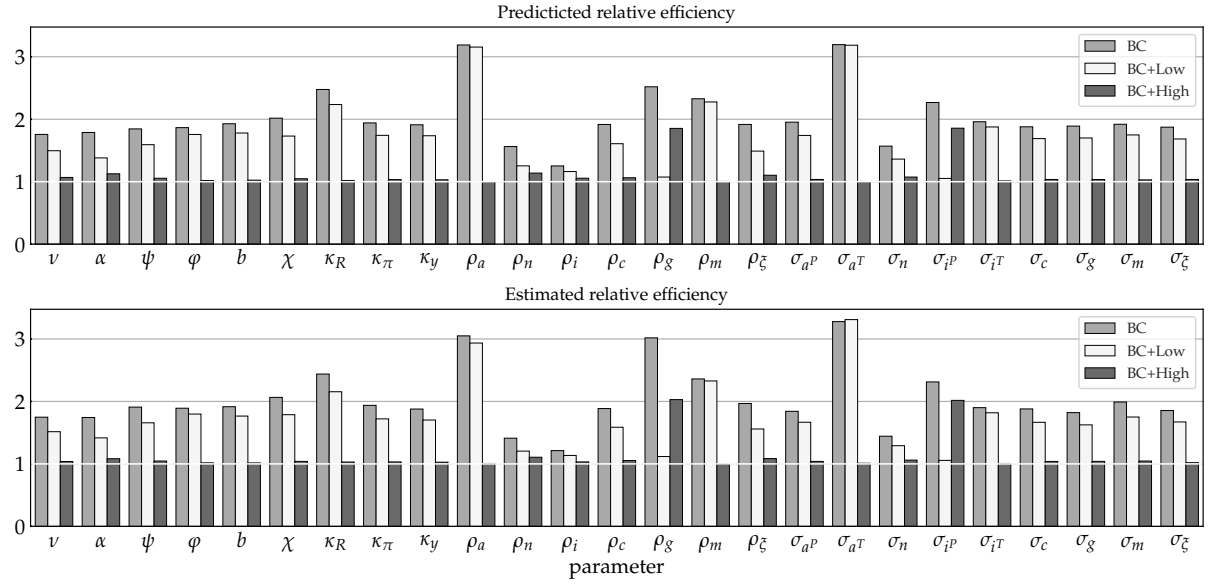
Figure 32: Ratios of Whittle-based conditional MC standard deviations to conditional CRLBs. TD represents the exact time domain MLE, while the other three are band spectral Whittle estimators using frequencies from the BC, BC+Low, and BC+High frequency bands.

D.2.3 Conditional estimation using the exact band spectral likelihood: BC vs BC+Low vs BC+High

The following results are for conditional estimation using the exact band spectral likelihood function instead of the Whittle approximation.



(a) T=192.



(b) T=500.

Figure 33: Predicted (top panel of each subplot) and MC-estimated (bottom panel) relative efficiency of the three band spectral estimators.

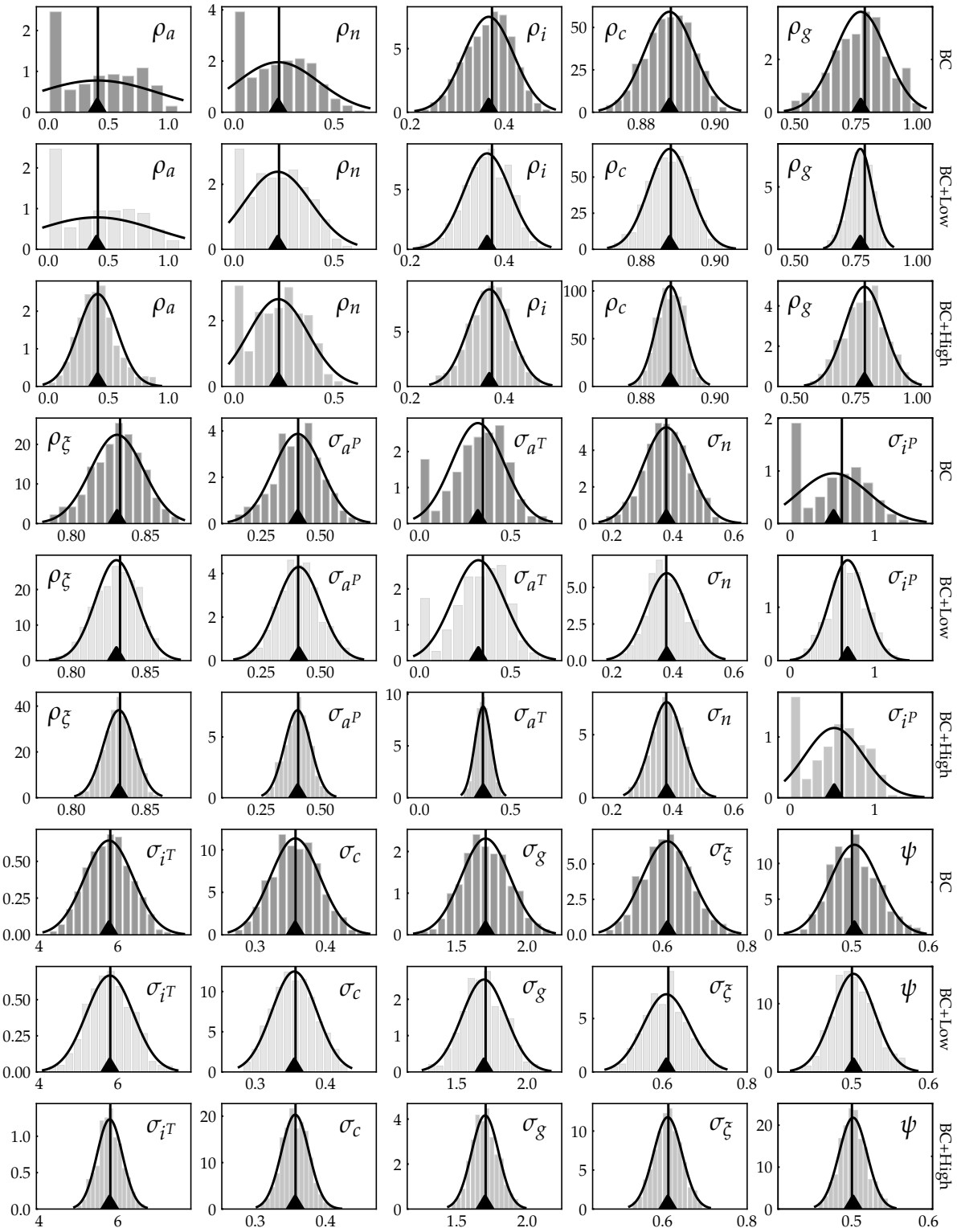


Figure 34: Conditional estimation using the exact likelihood. The figure shows MC sampling distributions and gaussian densities centered on the true values with standard deviation equal to the conditional CRLB. The sample size is $T = 192$.

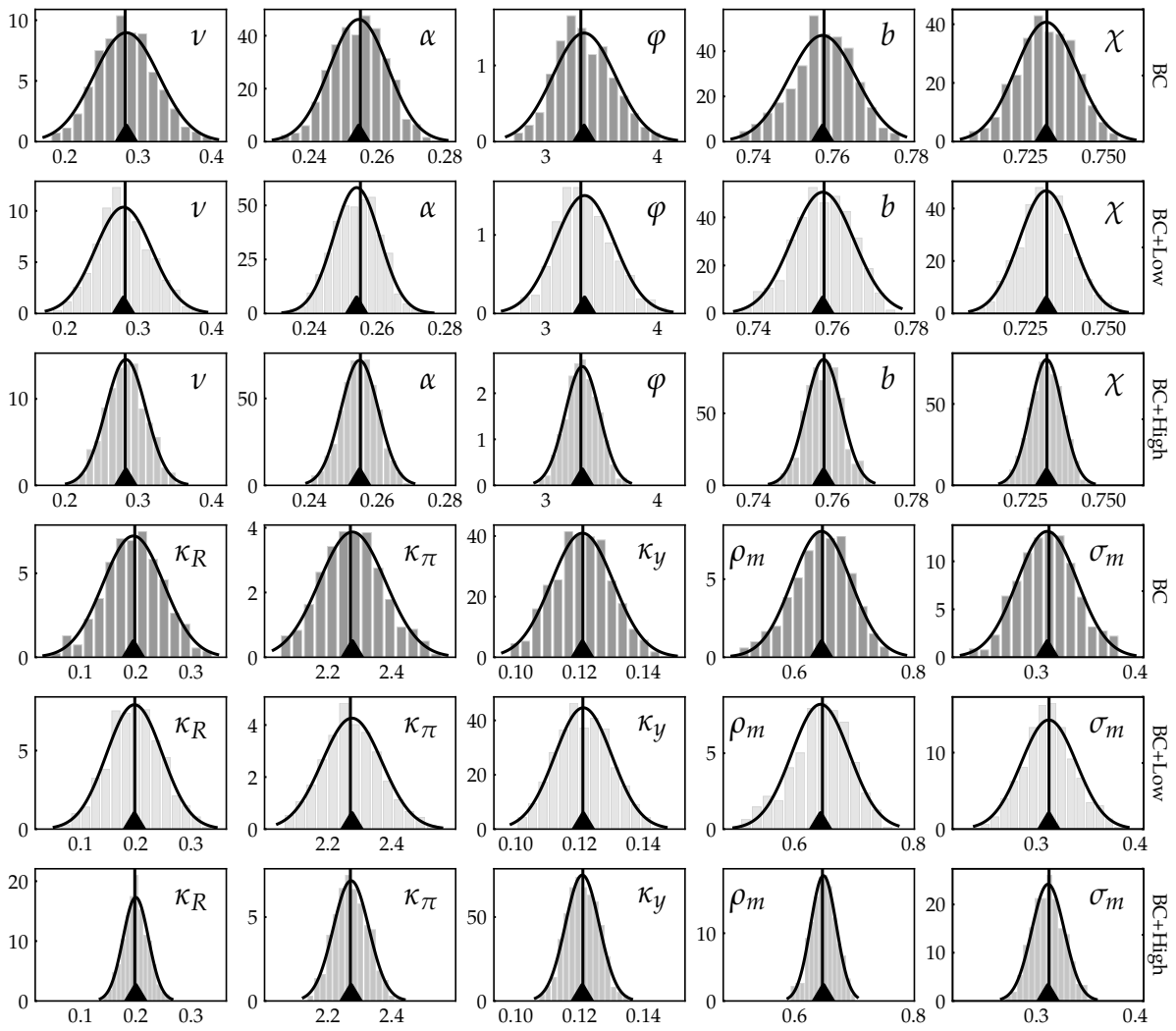


Figure 35: Continue Figure 34

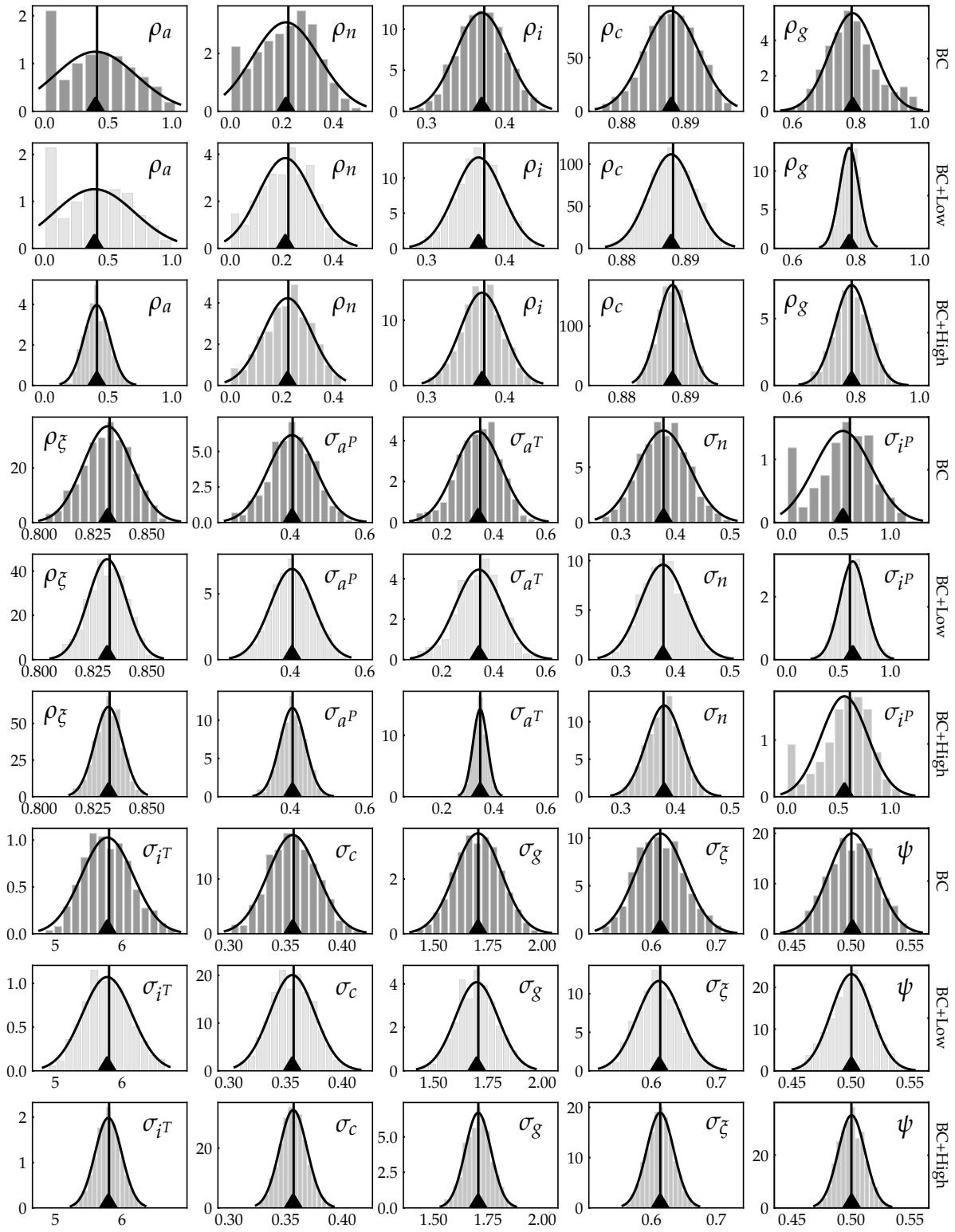


Figure 36: See the note to Figure 34. The sample size is $T = 500$.

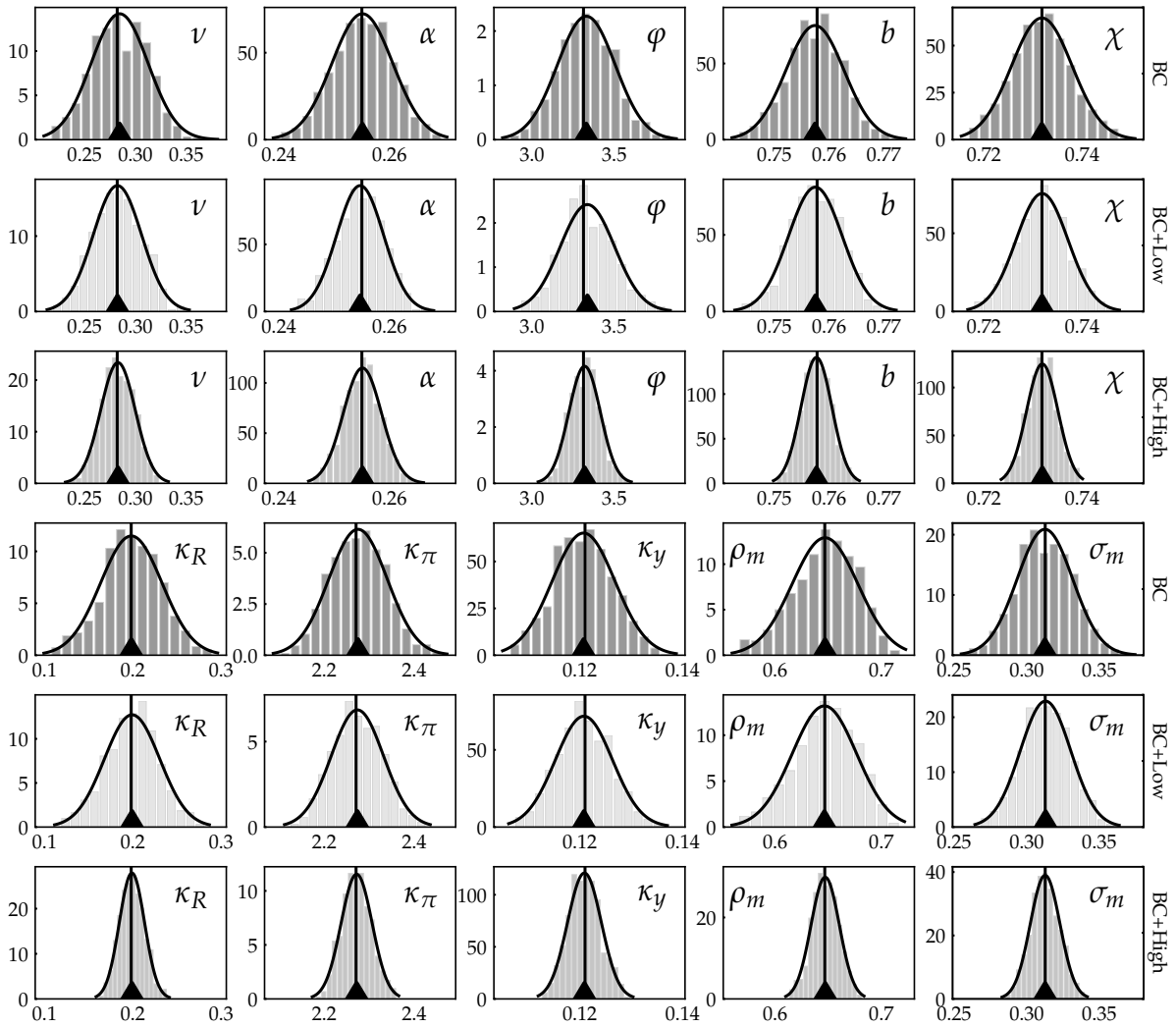
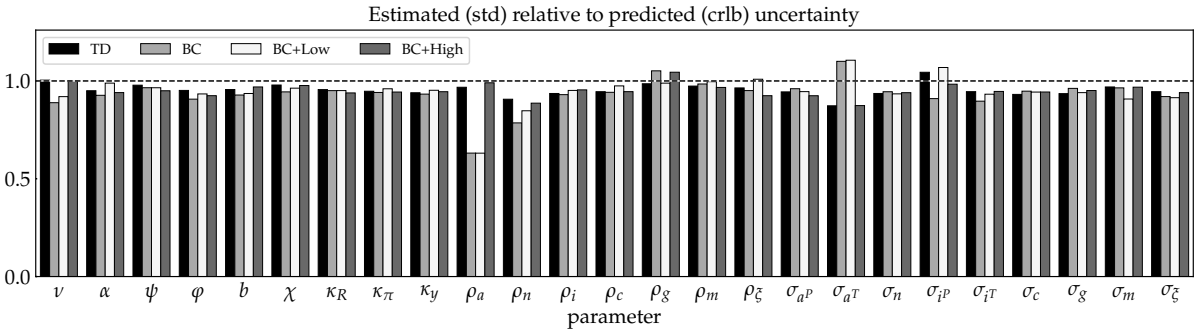
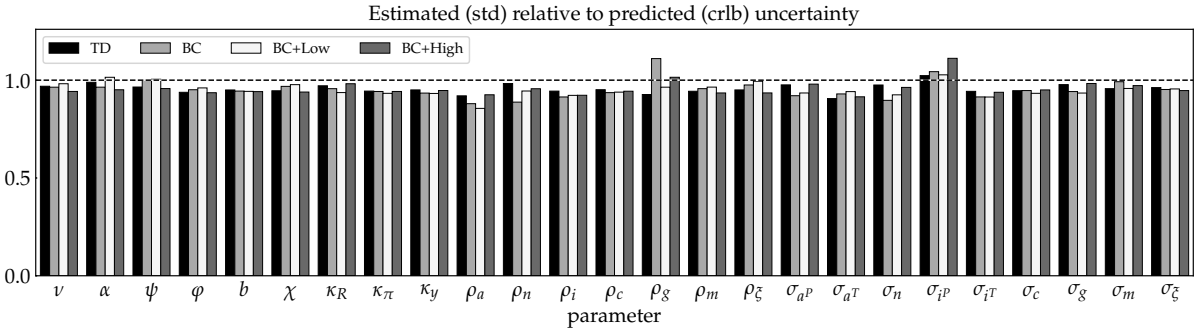


Figure 37: Continue Figure 36



(a) T=192.



(b) T=500.

Figure 38: Ratios of exact conditional MC standard deviations to conditional CRLBs. TD represents the exact time domain MLE, while the other three are exact band spectral likelihood estimators using frequencies from the BC, BC+Low, and BC+High frequency bands.