On the band-spectral estimation of business cycle models

Nikolay Iskrev

Abstract

I evaluate the performance of the band-spectral estimation approach applied to business cycle models. Band-spectral methods are widely used to study frequency-dependent relationships among variables. In business cycle research, this approach permits the estimation of structural models on the basis of the frequencies they are best suited to represent, such as the business cycle frequencies. In particular, the frequency domain approximation of the likelihood function (the Whittle likelihood) can be used to estimate the parameters of fully-specified models on the basis of a targeted band of frequencies. Using the medium-scale model of Angeletos, Collard, and Dellas (Econometrica, 2018) as a data-generating process, I perform a Monte Carlo study to evaluate the finite-sample properties of the band-spectral maximum likelihood estimator (MLE) and to compare them to those of the full spectrum and the exact time-domain MLE. The results show that using the band-spectral estimator leads to considerable biases and efficiency losses for most estimated parameters. In fact, the performance of both Whittle likelihood-based estimators is found to be seriously deficient in terms of bias and accuracy, in contrast to that of the time domain estimator, which successfully recovers all model parameters. I show how these findings can be explained with the theoretical properties of the underlying model, and describe simple-to-use 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
The views expressed in this paper are those of its author and do not necessarily reflect the views of Banco de Portugal or the Eurosystem.

1 Introduction

Following the pioneering work of Hannan (1963), band-spectral methods have become an important tool for estimating dynamic models on the basis of a restricted band of frequencies. The rationale for a band-spectral estimation of structural macroeconomic models is simple: a model should not be forced to fit data it is not intended to explain. In particular, if a theoretical model lacks the features and mechanisms required for it to account for data movements in some parts of the frequency range, then those frequencies should be excluded from the estimation. Failure to do so would lead to distorted estimates of the model parameters, as they would be forced to accommodate empirical features outside of what the model is designed for. Therefore, for instance, a business cycle model that is a priori known to be incapable of 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.[1]

Hansen and Sargent (1993) were the first to use the Whittle likelihood for band-spectral estimation of macro models, with a focus on understanding the effect of using seasonally adjusted data.[2] Diebold et al. (1998) develop an estimation approach where different frequencies receive different weights in the estimation loss function, as a way to account for differential contamination of different frequency components due to model misspecification. The band-spectral likelihood estimation, which they call Band-MLE, emerges as a special case in that framework. Cogley (2001) also uses band-spectral likelihood estimation to exclude low frequencies in the context of estimation of models with uncertain trend specification. In spite of the compelling reasons for using band spectral estimation, given the common focus in the empirical literature on explaining business cycle phenomena, macroeconomic models are usually estimated in the time domain, which is tantamount to using all frequencies. Notable exceptions are Qu and Tkachenko (2012), Sala (2015), and, most recently, Angeletos et al. (2018), all of whom consider likelihood-based estimation in the frequency domain using a subset of frequencies.

[1] The same argument applies (among many others) to: (i) real business cycle models, which are not expected to explain nominal variables; (ii) closed economy models, which are not expected to account for international trade statistics; (iii) models without financial frictions or banks, which are not expected to fit financial variables; etc.

[2] Band-spectral MLE can be interpreted as a full information analogue of the band-spectral linear regression proposed by Hannan (1963) whose application in economics was popularized by Engle (1974).

The purpose of this paper is to evaluate the performance of the band-spectral likelihood estimation when applied to modern business cycle models. In particular, I use the medium-scale DSGE model of Angeletos et al. (2018) as a data-generating process in a Monte Carlo study that compares the finite-sample properties of the band-spectral maximum likelihood estimator (MLE) to those of the full-spectrum and the exact time-domain MLE. The results show that using the band-spectral estimator leads to considerable biases and efficiency losses for most estimated parameters. In fact, the performance of both Whittle likelihood-based estimators is found to be seriously deficient in terms of bias and accuracy, in contrast to that of the time domain estimator, which successfully recovers all model parameters. I show how these findings can be explained with the theoretical properties of the underlying model, and describe simple-to-use tools and diagnostics that can be used to detect potential problems in band-spectral estimation for a wide class of macroeconomic models.

The plan of the paper is as follows: Section 2 introduces the model of Angeletos et al. (2018). Section 3 gives an overview of the Whittle approximation of the Gaussian likelihood function. Section 4 provides the Monte Carlo simulation results. A summary of the main conclusions is given in Section 5.

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 of the features found in 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 usual assumption of rational expectations and common information about the state of the economy. In particular, in their model, agents' beliefs regarding the expectations of other agents (higher-order beliefs) are subject to autonomous variation, called "confidence shock", which causes divergence between the two forms of beliefs. This leads to exogenous variations in agents' expectations of the economic outcomes in the short-run, without altering their medium or long-run expectations of those outcomes, or the expectations of the exogenous fundamentals at any horizon. ACD show that embedding their mechanism in an otherwise standard New Keynesian business

cycle model can help better match observed patterns in macroeconomic data. 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 state explicitly that the model describes business cycle phenomena only, and lacks the features and mechanisms that would be required for it to account for the low and high frequency properties of empirical time series. For that reason, the model is estimated, in the frequency domain, using only the business-cycle frequencies. While a focus on the business cycle is common for the majority of studies in the literature, estimation is typically done in the time domain using all frequencies. The few exceptions, such as Sala (2015) and Qu and Tkachenko (2012), estimate in the frequency domain models that were originally developed without being specifically tailored to fit only the business cycle part of the spectrum.

The main methodological contribution of ACD is to show how to introduce higher-order belief dynamics into macroeconomic models in a tractable way. As in most of the DSGE literature, the model estimation is based on a linear state space representation obtained from the solution of the log-linear approximation around steady state of the model's equilibrium conditions. For reference, the linearized equilibrium conditions of the ACD model are presented below. For more details on the model and solution method the reader is referred to the original publication.

2.1 Linearized equilibrium conditions

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

Optimal consumption allocation

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

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, and the parameters b and ϱ denote the degree of habit persistence, and the degree of substitutability across the islands' composite goods in the production of the production of the final good, respectively.

Optimal investment decision

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

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

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

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

Optimal bond holdings decision

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

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

Equilibrium price of capital

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

where i_{it} denotes the level of investment, ζ_t^{IP} is investment-specific technology shock, ζ_t^{IT} is a shock shifting the demand for investment, and φ is a parameter governing the size of 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} is the local capital stock, ζ_t^A is the level of aggregate TFP, and α is 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 a government spending shock.

Equilibrium utilization

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

where ψ is a capital utilization elasticity parameter.

Inflation rate

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

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

Monetary policy rule

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

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 reaction 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, the inhabitants of each island receive an unbiased signal about the level of TFP in that period, and form beliefs that firms and households on other islands receive a signal that is biased by the confidence shock ξ_t, which is also observed. In stage 2, the true state of nature and the realized value of economic activity is publicly revealed. ACD discuss two protocols for the timing of decisions of firms and households, depending on whether supply is determined first and prices adjust to make demand meet supply, or whether demand is determined first and supply adjusts to meet demand. 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 later shock is an exogenous random variable observed in stage 1 of each period, representing the perceived bias in the other islands' signals about the level of TFP in that period. 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 US data for six variables: GDP, consumption, investment, hours worked, the inflation rate, and the federal fund rate. The sample period is 1960Q1 - 2007Q4. The model parameters are estimated with Bayesian methods using a frequency domain representation of the likelihood function. The estimated median of the posterior distribution is reported in Table 1.

Table 1: 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

3 The Whittle likelihood

3.1 General case

Let $\boldsymbol{Y}_T = (\boldsymbol{y}_1', \boldsymbol{y}_2', \ldots, \boldsymbol{y}_T')'$ be a T-dimensional sample from a zero mean stationary Gaussian process $\{\boldsymbol{y}_t\}_{t=-\infty}^{\infty}$ with autocovariance function $\boldsymbol{\Gamma}(\tau;\boldsymbol{\theta}) = \operatorname{cov}(\boldsymbol{y}_{t+\tau}, \boldsymbol{y}_t)$. Apart from an additive constant, the log-likelihood function of $\boldsymbol{Y}_T$ is given by

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

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 $\boldsymbol{\Gamma}(\tau;\boldsymbol{\theta})$.

Evaluating $\ell(\boldsymbol{\theta};\boldsymbol{Y}_T)$ requires computing the determinant and the inverse of $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$, which can be computationally prohibitive for even moderate sample sizes. To circumvent this problem, Whittle (1953) introduced a spectral approximation of $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$ as a computationally cheaper method for calculating the likelihood function of stationary Gaussian time series. It uses the fact that block Toeplitz matrices can be approximated by block circulant matrices,[3] whose eigenvalue decomposition can be computed very efficiently using the discrete Fourier transform (DFT). Specifically, it can be shown that for large T,

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

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 i-th block $\{\boldsymbol{S}_T(\boldsymbol{\theta})\}_{ii} = \boldsymbol{s}(\boldsymbol{\theta}, \omega_i)$ given by the spectral density matrix

[3] 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 $A_{n-j} = A_j'$.

of $\boldsymbol{y}_t$ evaluated at the $i-$th Fourier frequency,

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

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

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

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 calculated very efficiently as

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

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

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:[4]

$$\begin{aligned}
\ell_w(\boldsymbol{\theta}; \boldsymbol{I}_T) &= -\frac{1}{2} \log\det(\boldsymbol{S}_T(\boldsymbol{\theta})) - \frac{1}{2} (\boldsymbol{F}_T \boldsymbol{Y}_T)^* \boldsymbol{S}^{-1}(\boldsymbol{\theta}) (\boldsymbol{F}_T \boldsymbol{Y}_T) && (3.7) \\
&= -\frac{1}{2} \sum_{j=0}^{T-1} \log\det(\boldsymbol{s}(\boldsymbol{\theta}, \omega_j)) + \boldsymbol{J}_T(\omega)^* \boldsymbol{s}^{-1}(\boldsymbol{\theta}, \omega_j) \boldsymbol{J}_T(\omega) && (3.8) \\
&= -\frac{1}{2} \sum_{j=0}^{T-1} \log\det(\boldsymbol{s}(\boldsymbol{\theta}, \omega_j)) + \mathrm{tr}\left(\boldsymbol{I}_T(\omega_j) \boldsymbol{s}^{-1}(\boldsymbol{\theta}, \omega_j)\right) && (3.9)
\end{aligned}$$

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 (3.8)-(3.9) need to be evaluated.

The Whittle log-likelihood is clearly much easier to evaluate than the expression in (3.1)-(3.2) as it avoids inverting a potentially very large covariance matrix and can be

[4]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).

done by leveraging existing fast algorithms for performing DFT. On the other hand, it is only an approximation of the exact Gaussian log-likelihood because $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$ and $\boldsymbol{\Omega}_T(\boldsymbol{\theta})$ are not equal in finite samples. Moreover, other computationally efficient methods exist, such as the Kalman filter, which evaluate the exact likelihood for a large class of models also without having to invert large matrices. Therefore, the computational efficiency argument for using the Whittle log-likelihood is not as relevant now as it was when it was first proposed.[5]

Another appealing feature of the Whittle log-likelihood is that it allows for estimation to be based on only a subset of frequencies. This may be desirable in cases where the data is known to contain noise affecting only part of the periodogram, e.g. the high frequencies, or when the theoretical model is only intended to match the data at a narrow range of frequencies. This is particularly relevant in macroeconomic research where theoretical models are often specifically designed to explain business cycle movements in the data, and are known to be misspecified in terms of their implication for the lower and higher frequencies of the data. This implies that one would want to fit the model to business cycle frequencies and ignore those at the lower and higher end of the spectrum. In practice, this can be achieved by performing the summation in (3.8)-(3.9) over the frequencies of interest,

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

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 $\omega_j \in \bar{\boldsymbol{\omega}}$ then $2\pi - \omega_j \in \bar{\boldsymbol{\omega}}$.

3.2 Linearized DSGE models

In general, a linearized DSGE model can be expressed as a recursive equilibrium law of motion given by the following system of equations:

$$\boldsymbol{y}_t = \boldsymbol{C}(\boldsymbol{\theta})\boldsymbol{v}_{t-1} + \boldsymbol{D}(\boldsymbol{\theta})\boldsymbol{u}_t \tag{3.11}$$

$$\boldsymbol{v}_t = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{v}_{t-1} + \boldsymbol{B}(\boldsymbol{\theta})\boldsymbol{u}_t \tag{3.12}$$

$$\boldsymbol{u}_t = \boldsymbol{G}(\boldsymbol{\theta})\boldsymbol{u}_{t-1} + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim \mathcal{N}\left(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}(\boldsymbol{\theta})\right) \tag{3.13}$$

[5]There are nevertheless models where the Whittle approximation is often the preferred approach for efficiency-related reasons.

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 $\boldsymbol{\theta}$.

Evaluating the Whittle log-likelihood function requires the spectral density matrix of the observed variables $\boldsymbol{y}_t$, which 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

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

4 Simulation Study

This section presents 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 the business cycle frequencies. The latter two estimators are based on the Whittle approximation of the likelihood introduced in Section 3. The time domain MLE is based on the exact Gaussian likelihood function evaluated using the Kalman filter. For brevity, the three estimators will be referred to in the sequel as TD, FD, and BC.

4.1 Setup

The Monte Carlo simulation proceeds as follows:

1. Solve the model from Section 2 using the algorithm developed by Angeletos et al. (2018) at the parameter values shown in Table 1.

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 fund rate (r).

3. Estimate the 25 free model parameters with the three estimators by maximizing the respective log-likelihood function.

4. Repeat N times steps 2 and 3.

The sample size is set to $T = 192$ observations, which is the size of the sample used by ACD. The simulation is initialized from the stationary distribution of the variables and the first 500 observations are discarded to remove the dependence on initial conditions. The number of replications is $N = 1000$. The numerical optimization in step (2) takes the true values of the parameters as a starting point, and is performed by a combination of global and local optimization algorithms. The same process is applied in the optimization of the objective function of each estimator.

4.2 Results: baseline parametrization

The first set of results is presented in Table 2, which reports the mean, median, and interquartile range (IQR) for each estimated parameter. Comparing the values of these statistics to the true parameter values, reported in the first column of the table, is helpful to get a sense of the relative performance of the three estimators. As expected, the TD estimator performs best, both in terms of the accuracy of the point estimates, and variability around the true values. The two Whittle likelihood-based estimators are significantly worse for most parameters, with the BC estimator being generally both the less accurate and the more volatile of the two. For instance, with the exceptions of very few parameters, the BC estimator's IQR is the widest, while the TD one is the most narrow of the three, often by orders of magnitude. In spite of being very wide, the FD and BC estimators' IQR do not always include the true parameter values. In particular, for both estimators the 25 percentile is above the true value of ρ_c, while the 75 percentile is below the true values of σ_ξ, φ, and b. In the case of FD, there are additional two parameters (κ_y, and σ_{a^P}) for which the IQR is to the right of the true value, and additional two parameters (κ_R and σ_{i^T}) for which the IQR is to the left of the true value.[6] In contrast, the true parameter values are always within the IQR of the TD estimates. A graphical representation of these findings is presented in Figure 1, which displays boxplots of the deviations of the parameter estimates for each parameter as a percent of its true value.

[6]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

		Mean			Median			IQR		
parameter	true	TD	FD	BC	TD	FD	BC	TD	FD	BC
ν	0.28	0.28	0.34	0.42	0.28	0.32	0.38	[0.23, 0.33]	[0.27, 0.40]	[0.28, 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	0.49	0.50	0.44	0.40	[0.43, 0.59]	[0.34, 0.54]	[0.26, 0.62]
φ	3.31	3.20	2.33	2.60	3.18	2.09	2.16	[2.63, 3.67]	[1.44, 2.95]	[1.56, 3.10]
b	0.76	0.75	0.60	0.64	0.75	0.54	0.64	[0.73, 0.77]	[0.51, 0.69]	[0.55, 0.72]
χ	0.73	0.72	0.70	0.75	0.72	0.71	0.75	[0.71, 0.74]	[0.68, 0.73]	[0.71, 0.79]
κ_R	0.20	0.20	0.14	0.13	0.19	0.13	0.03	[0.14, 0.25]	[0.07, 0.20]	[0.00, 0.22]
κ_π	2.27	2.24	2.04	2.60	2.18	1.95	2.10	[1.96, 2.49]	[1.67, 2.34]	[1.51, 3.02]
κ_y	0.12	0.15	0.24	0.18	0.14	0.20	0.13	[0.11, 0.18]	[0.13, 0.31]	[0.08, 0.21]
ρ_a	0.41	0.39	0.59	0.56	0.40	0.62	0.59	[0.23, 0.56]	[0.37, 0.82]	[0.18, 0.97]
ρ_n	0.22	0.21	0.39	0.19	0.19	0.33	0.04	[0.08, 0.31]	[0.16, 0.59]	[0.00, 0.34]
ρ_i	0.37	0.35	0.38	0.28	0.35	0.38	0.25	[0.30, 0.41]	[0.31, 0.45]	[0.03, 0.48]
ρ_c	0.89	0.88	0.94	0.94	0.88	0.95	0.95	[0.86, 0.90]	[0.90, 0.99]	[0.91, 0.98]
ρ_g	0.79	0.76	0.76	0.74	0.77	0.77	0.76	[0.73, 0.80]	[0.71, 0.81]	[0.65, 0.88]
ρ_m	0.65	0.63	0.64	0.64	0.63	0.64	0.66	[0.59, 0.66]	[0.60, 0.68]	[0.57, 0.74]
ρ_ξ	0.83	0.81	0.77	0.84	0.82	0.79	0.88	[0.79, 0.84]	[0.72, 0.84]	[0.78, 0.94]
σ_{a^P}	0.41	0.38	0.81	0.51	0.39	0.70	0.48	[0.32, 0.44]	[0.48, 1.07]	[0.20, 0.73]
σ_{a^T}	0.35	0.34	0.37	0.40	0.34	0.35	0.39	[0.30, 0.39]	[0.28, 0.43]	[0.26, 0.56]
σ_n	0.38	0.38	0.27	0.58	0.39	0.29	0.54	[0.33, 0.44]	[0.14, 0.38]	[0.31, 0.81]
σ_{i^P}	0.61	0.53	0.64	0.58	0.55	0.65	0.16	[0.30, 0.76]	[0.24, 0.98]	[0.03, 1.08]
σ_{i^T}	5.80	5.69	4.22	5.75	5.56	3.81	4.65	[4.69, 6.58]	[2.52, 5.44]	[2.78, 7.33]
σ_c	0.36	0.36	1.02	0.71	0.35	0.54	0.51	[0.25, 0.46]	[0.34, 1.11]	[0.28, 0.84]
σ_g	1.71	1.69	1.69	1.71	1.69	1.70	1.70	[1.63, 1.75]	[1.62, 1.77]	[1.56, 1.85]
σ_m	0.31	0.30	0.32	0.38	0.30	0.31	0.32	[0.28, 0.33]	[0.28, 0.34]	[0.27, 0.39]
σ_ξ	0.61	0.66	0.45	0.49	0.61	0.41	0.18	[0.44, 0.80]	[0.25, 0.58]	[0.04, 0.60]

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 and sample size of $T = 192$.

These observations suggest that the Whittle likelihood-based estimators may suffer from a significant estimation bias. This is confirmed in Table 3, which reports the bias, standard deviation (SD) and root mean squared error (RMSE) for each parameter over the 1000 replications. To avoid scale problems and facilitate comparison across parameters, the figures are displayed relative to the true parameter values and are expressed as a percentage. On average, the bias of the TD estimates is about 4% in absolute value, while for FD and BC it is around 30% and 21%, respectively. The most biased estimates obtained with the TD estimator are of κ_y, with bias of about 28%. In the case of FD, the largest in absolute value bias is about 187% and is with respect to σ_c, while two other parameters – σ_{a^p} and κ_y, have biases close to 100%. The estimates

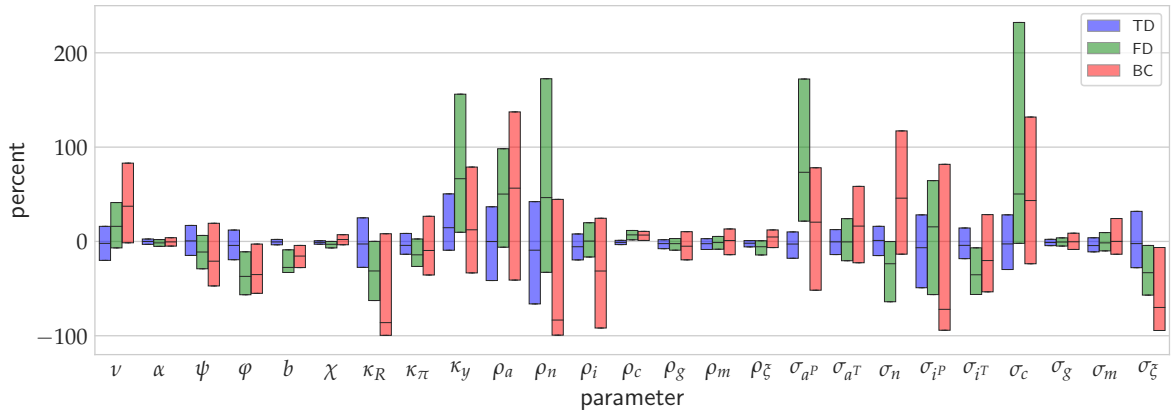

Figure 1: Boxplots of deviations of parameter estimates as a percentage of the true values. Each box shows the interquartile range of the estimates obtained with the indicated estimator. The vertical bars inside the boxes shows the median estimate. The results are based on 1000 MC replications and sample size of $T = 192$.

of σ_c are also the most biased ones for the BC estimator, with bias of slightly less than 100%. There is some agreement in terms of which parameters tend to be relatively more or less biased, but only for the two Whittle likelihood-based estimators. Using the rank (Spearman) correlation coefficient to quantify the overlap in terms of relative bias, the value for FD and BC is around .47, while the rank correlation between TD and FD is negative ($-.3$), and between TD and BC is positive but very weak (.1). The SD shown in the middle panel measures the variability of the estimates. On average, the TD estimates are about half as variable as the FD ones, and around one-third as variable as the BC estimates. Looking at individual parameters, the FD estimates are, with a few exceptions, less variable than the BC one, while the TD estimates always have the lowest standard deviation. There are also some common patterns with respect to the estimates of which parameters are relatively more or less variable: the estimates of α, χ, ρ_c, and σ_g are among the least variable for all estimators, while the estimates of κ_y, ρ_n, σ_c, and σ_{i^P} are among the most variable ones. The rank correlation between SDs of the TD and either one of the Whittle likelihood-based estimators is around .9, while that between FD and BC is .93. The last three columns of Table 3 show the normalized RMSE, which is a measure of the overall accuracy of the estimates, accounting for both bias and variability. The patterns are very similar to the ones observed in the estimated standard deviations, except for the even more pronounced relative superiority of the TD estimator over the Whittle likelihood-based ones, due to the larger bias contaminating

Table 3: Monte Carlo results (cont.)

	Bias (%)			SD (%)			RMSE (%)		
parameter	TD	FD	BC	TD	FD	BC	TD	FD	BC
ν	-1.8	19.5	49.2	26.1	38.7	90.5	26.1	43.3	103.0
α	-0.4	-2.0	-0.4	4.2	6.0	7.2	4.2	6.3	7.2
ψ	2.6	-10.3	-1.6	25.0	29.6	67.5	25.1	31.4	67.5
φ	-3.3	-29.8	-21.4	23.5	35.7	49.6	23.7	46.5	54.1
b	-0.9	-21.0	-15.0	4.0	13.3	13.2	4.2	24.8	20.0
χ	-1.2	-3.8	1.9	2.9	5.6	9.2	3.1	6.7	9.4
κ_R	-0.3	-29.2	-33.3	38.6	45.5	91.9	38.6	54.0	97.7
κ_π	-1.3	-10.2	14.6	17.1	25.0	74.5	17.2	27.0	76.0
κ_y	27.8	94.9	47.5	58.7	123.8	148.3	65.0	156.0	155.7
ρ_a	-4.6	42.4	35.6	56.0	72.4	91.5	56.2	83.9	98.2
ρ_n	-5.5	73.7	-15.8	74.7	127.9	114.9	74.9	147.6	116.0
ρ_i	-5.5	2.6	-25.6	20.8	28.5	64.5	21.6	28.7	69.4
ρ_c	-1.1	6.1	5.3	3.5	5.7	7.4	3.6	8.3	9.1
ρ_g	-3.1	-3.4	-5.3	7.1	11.1	23.4	7.8	11.6	24.0
ρ_m	-3.0	-1.7	-0.5	8.8	10.2	25.7	9.3	10.3	25.7
ρ_ξ	-2.5	-8.0	1.4	5.1	12.0	14.4	5.7	14.4	14.5
σ_{a^P}	-6.8	99.4	26.4	23.1	104.7	94.8	24.0	144.3	98.4
σ_{a^T}	-1.7	7.2	15.2	19.7	54.0	68.1	19.8	54.5	69.8
σ_n	0.3	-29.4	54.2	23.7	37.9	97.6	23.8	47.9	111.6
σ_{i^P}	-12.9	4.8	-5.0	54.5	73.7	114.0	56.0	73.8	114.1
σ_{i^T}	-2.0	-27.3	-1.0	24.9	37.8	75.7	25.0	46.7	75.7
σ_c	1.1	186.7	99.8	42.7	302.2	192.9	42.8	355.2	217.2
σ_g	-1.1	-1.1	0.4	5.3	7.3	14.4	5.4	7.4	14.4
σ_m	-2.7	1.2	22.0	11.0	16.9	74.2	11.4	17.0	77.4
σ_ξ	7.2	-26.6	-20.0	49.4	50.7	134.0	49.9	57.2	135.5

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

the latter. The rank correlations are slightly weaker compared to the ones observed for SDs, but remain close to .9 for each pair of estimators. The parameters with the lowest RMSE across all estimators are χ, ρ_ξ, σ_g, α, and ρ_c, while the least accurately estimated parameters in all cases are κ_y, ρ_n, σ_c, and σ_{i^P}. Note, however, that even the least accurately estimated parameters with the TD estimator have lower RMSE than many of FD or BC estimates. For instance, in the case of BC, 11 of the 25 parameters have larger RMSE than the TD estimates of ρ_n, which has the highest RMSE with that estimator.

Figures 2 - 4 display histograms of the sampling distributions of the TD, FD, and BC estimates. They provide further evidence for the superiority of the TD over the

Whittle likelihood-based estimators. In particular, the distributions of the TD estimates are, with a few exceptions, unimodal, approximately symmetric, and reasonably well centered at the true parameter values. The exceptions are ρ_a, ρ_n, and σ_{iP}, all of which have large number of parameter estimates at the lower end of the boundary of the parameter space. As seen in tables 2 and 3, this leads to negative mean and median bias for those parameters, with σ_{iP} being the most affected of the three parameters. The results for the FD and BC estimators are clearly much worse, with multimodality, skewness, and concentration of estimates at the boundaries and away from the true values affecting many more parameters. The departures from normality are easier to see from the Q-Q plots presented in Figure 5. The parameter estimates are normalized by subtracting the respective means and dividing them by the respective standard deviations of each estimator. While it cannot be expected that the normal approximation will be particularly accurate for a sample of 192 observations, many of the TD estimates fit relatively well with the 45 degree line and are generally much closer to the theoretical quantiles of the normal distribution, compared to either one of the frequency domain estimators.

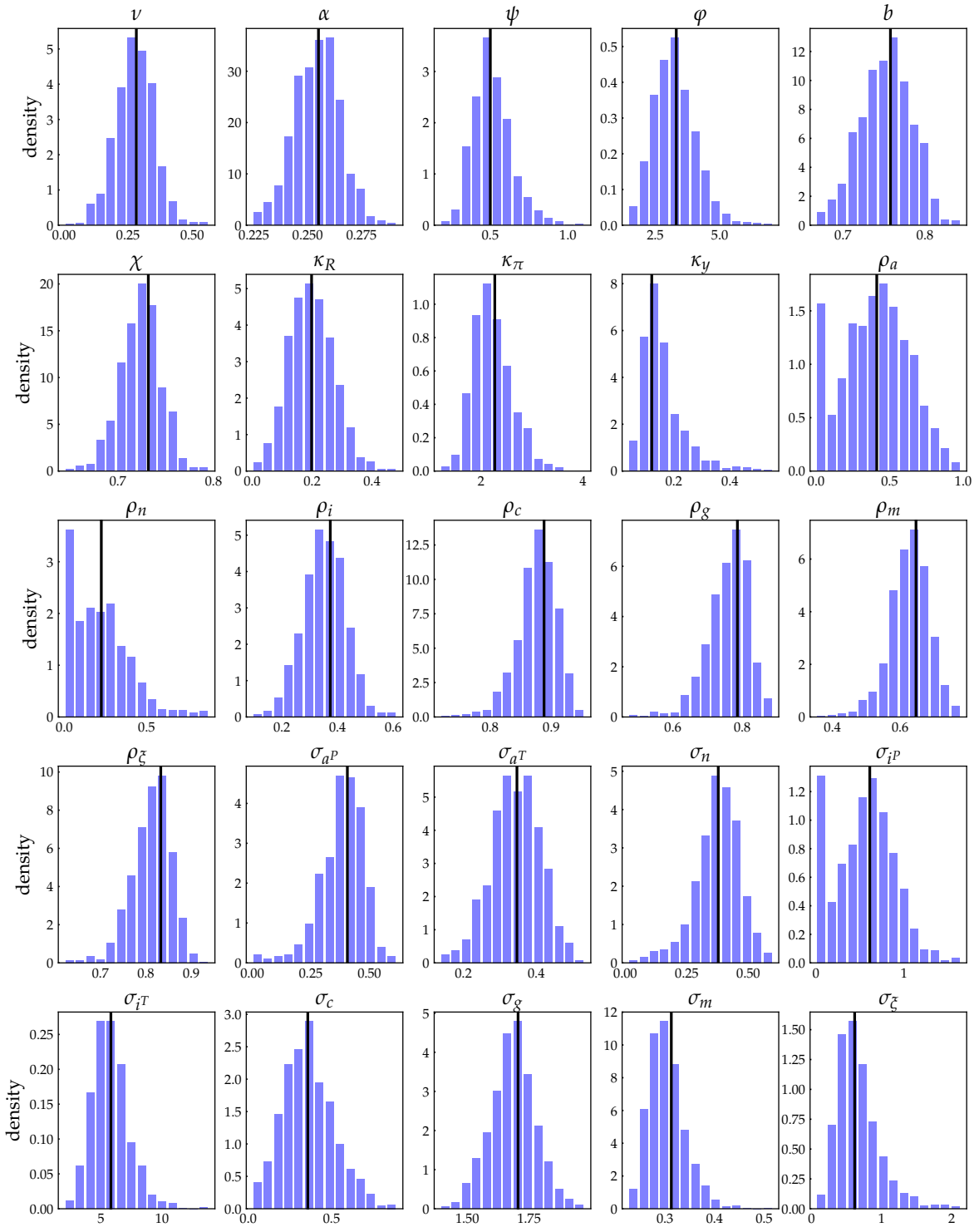

Figure 2: Sampling distributions of the estimated parameters, time domain MLE. The black vertical lines indicate true values. The results are based on 1000 MC replications and sample size of $T = 192$.

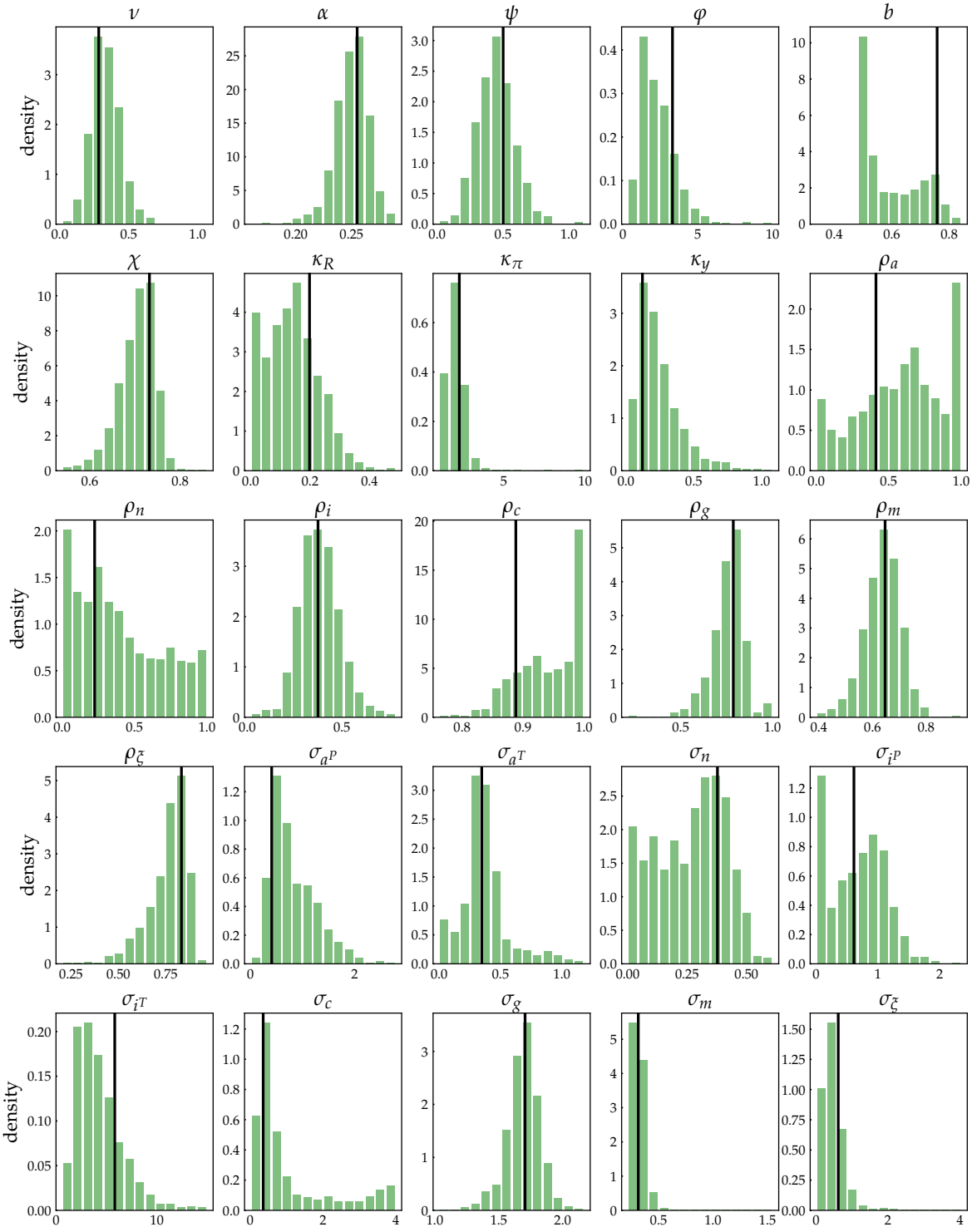

Figure 3: Sampling distributions of the estimated parameters, frequency domain MLE using all frequencies. The black vertical lines indicate true values. The results are based on 1000 MC replications and sample size of $T = 192$.

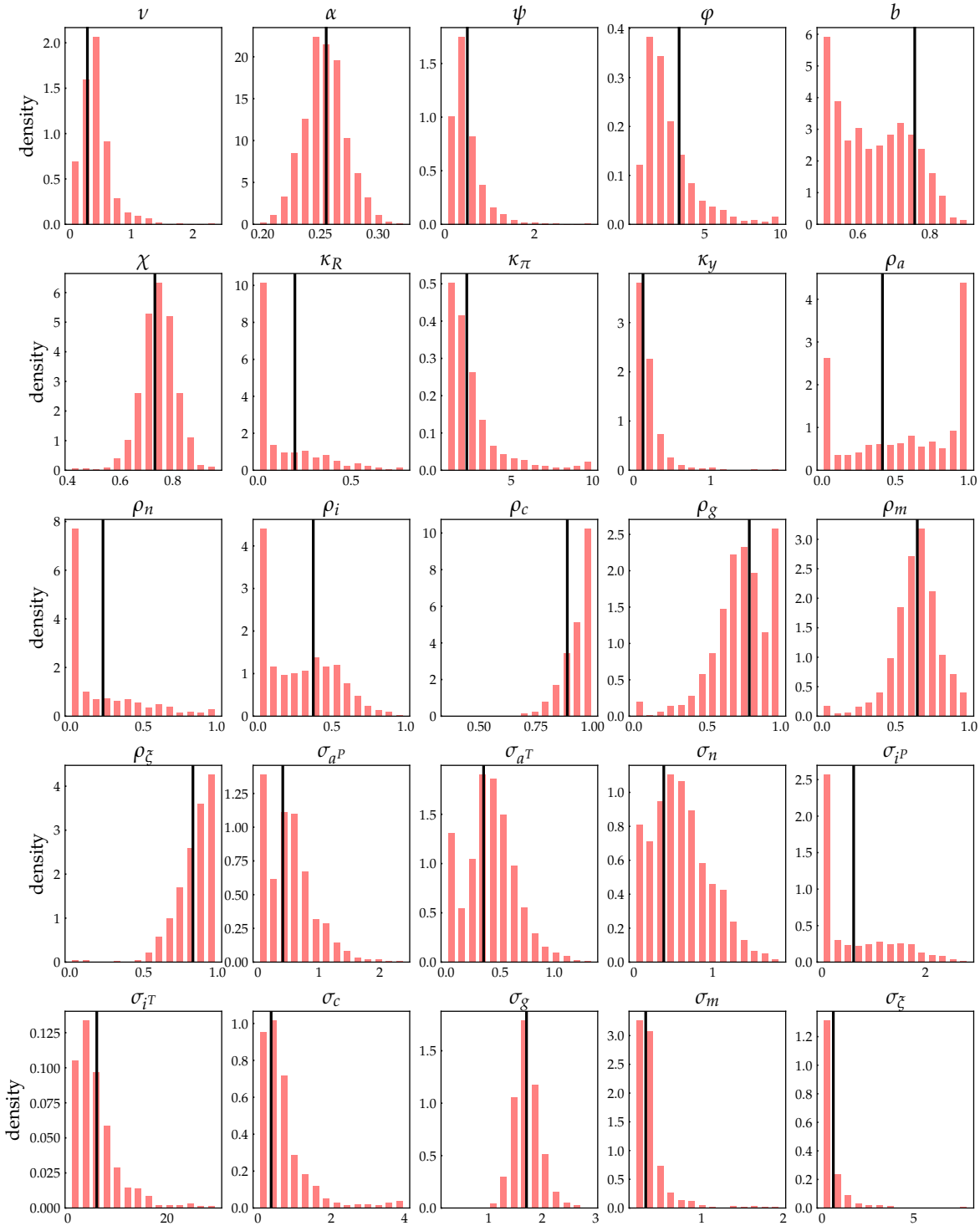

Figure 4: Sampling distributions of the estimated parameters, frequency domain MLE using only business cycle frequencies. The black vertical lines indicate true values. The results are based on 1000 MC replications and sample size of $T = 192$.

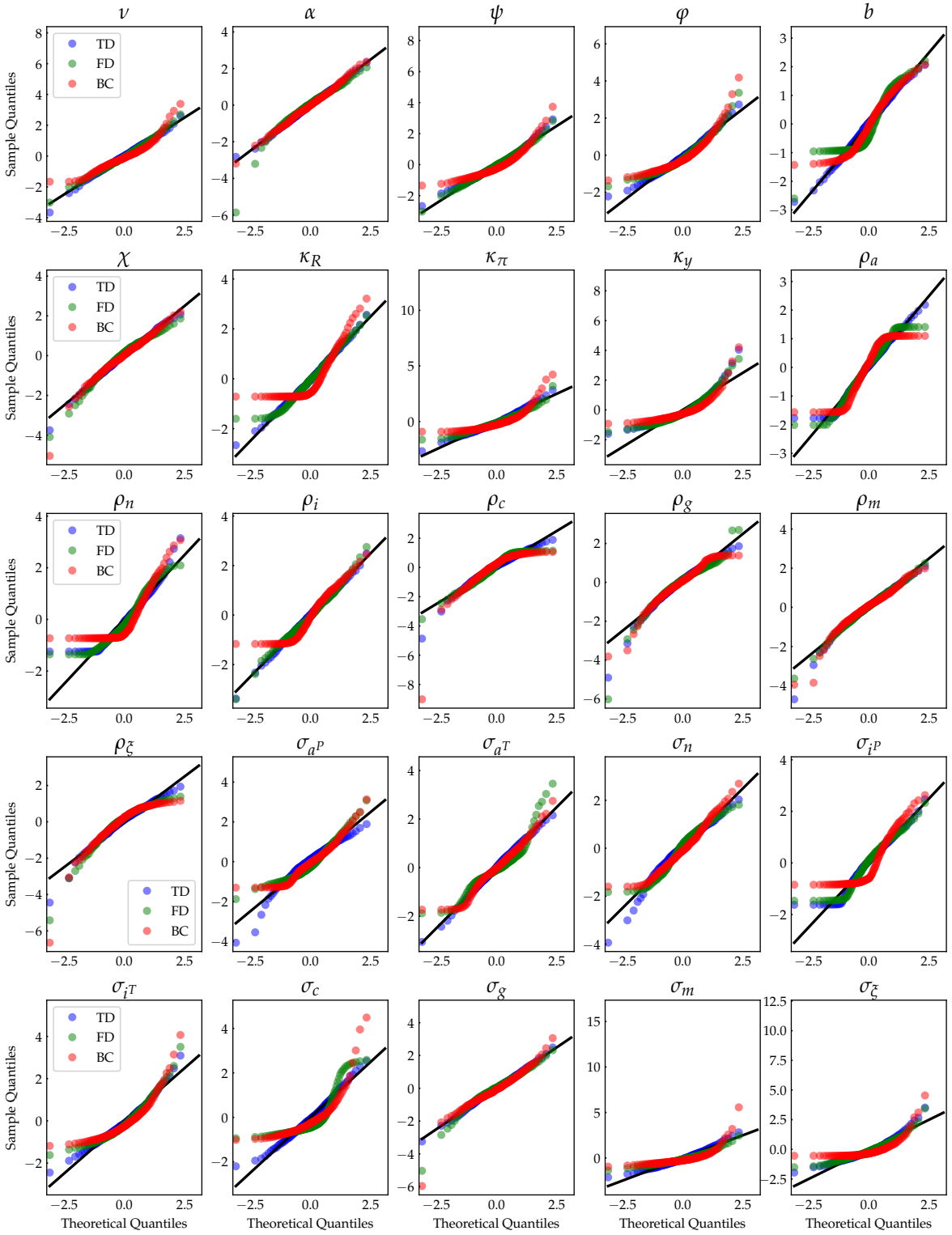

Figure 5: Q-Q plots showing the quantiles of the standardized sampling distributions of the estimated parameters vs the quantiles of the standard normal distribution. The black line shows the 45 degree line. The results are based on 1000 MC replications and sample size of $T = 192$.

Discussion

Taken together, the Monte Carlo results strongly favor the time domain estimator over the two frequency domain estimators. As the analysis was done under the assumption of correct model specification, the dominance of the exact likelihood function-based estimator is not surprising. Nevertheless, the very poor performance of the Whittle likelihood-based estimators merits further investigation, which is the goal of this section.

From the discussion in Section 3 (see in particular equations 3.3 – 3.9), we see that there are two equivalent ways to think about the approximation involved in constructing the Whittle likelihood function. First, we have an approximation of the likelihood function of the time-domain data vector $\boldsymbol{Y}_T$, whose covariance matrix $\boldsymbol{\Sigma}_T(\boldsymbol{\theta})$ is replaced by a block-circulant matrix $\boldsymbol{\Omega}_T(\boldsymbol{\theta})$. Alternatively, we can see it as an approximation of the likelihood function of the Fourier-transformed data vector $\boldsymbol{F}_T\boldsymbol{Y}_T$, whose covariance matrix $\boldsymbol{F}_T\boldsymbol{\Sigma}_T(\boldsymbol{\theta})\boldsymbol{F}_T^*$ is replaced by 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 $\boldsymbol{y}(\omega_i) = [y(\omega_i), c(\omega_i), i(\omega_i), h(\omega_i), \pi(\omega_i), r(\omega_i)]'$, is the frequency domain representation of the observed data, collecting observations indexed by frequency $\omega \in \{0, 2\pi/T, \ldots, 2\pi(T-1)/T\}$, instead of time $t \in \{1, 2, \ldots, T\}$. Then, to obtain the band-spectral likelihood function one has to select the frequencies of interest, i.e. the business cycle frequencies. This amounts to replacing the full sample Fourier matrix $\boldsymbol{F}_T$ with another Fourier matrix $\boldsymbol{F}_{bc}$, which extracts the business cycle frequencies from the original data vector. The covariance matrix of the exact band-spectral likelihood function is then given by $\boldsymbol{F}_{bc}\boldsymbol{\Sigma}_T(\boldsymbol{\theta})\boldsymbol{F}_{bc}^*$, while the Whittle approximation is obtained by using the respective submatrix of $\boldsymbol{S}_T(\boldsymbol{\theta})$.[7]

Having a block diagonal covariance matrix means that observations at different frequencies are uncorrelated. For example, the correlation between $y(\omega_j)$ and $y(\omega_k)$ is zero for all frequencies $\omega_j \neq \omega_k$; similarly for the autocorrelations of the other observed variables as well as the 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 6 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. A closer examination reveals that the largest, in absolute value, cross-frequency correlations are between different frequency components of y, c, and, to a lesser extent, i. This is easier to see in

[7]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 $\boldsymbol{F}_{bc} = \boldsymbol{K}\boldsymbol{F}_T$.

panel (d) of Figure 6, which shows that there is a large number of nonzero off-diagonal 3×3 blocks corresponding to correlation matrices of these variables at different frequencies from the BC part of the spectrum. Note that, to avoid visual clutter, all correlations smaller than .1 in absolute value have been zeroed-out. The largest correlations tend to be between different frequency components of c, followed by cross-frequency correlations between c and y, and correlations between frequency components of y. In particular, the largest correlation coefficients are .93 for c, 0.8 for c with y, and .73 for y. The auto or cross-correlations of i tend to be weaker, with values around 0.4. None of the other three variables – h, π, or r have pairwise correlation coefficients larger than 0.1.

As mentioned in Section 3, the Whittle likelihood is valid as an asymptotic approximation of the true Gaussian likelihood function only when the process is stationary. In particular, if a process is non-stationary, different frequency components of the periodogram will not be asymptotically independent. This suggests that the strong correlation patterns observed in Figure 6 may be the consequence of a high degree of persistence of the variables in the model. Indeed, while all eigenvalues of matrix $\boldsymbol{A}$ in (3.12) are inside the unit circle, thus ensuring the stationarity of $\boldsymbol{y}_t$, two of them are equal to 0.999999. The reason for this can be traced to a modelling choice made by Angeletos et al. (2018) in their implementation of the model from Section 2 in code: the authors parameterize the permanent shocks a_t^p and ζ_t^{IP} as AR(1) processes with autoregressive coefficients equal to 0.999999. Thus, the two shocks are stationary but extremely persistent. Because of the very important role a_t^p in particular plays for the dynamics of the y_t, c_t, and i_t,[8] these three variables inherit its persistence. For example, their autocorrelations remain around .999 even at lag 1000; as a comparison, the other three observed variable – h_t, π_t, and r_t, all have autocorrelations of less than .5 by the 10th lag.

These observations provide a tentative explanation for the poor performance of the Whittle likelihood-based estimators: the underlying data generating process is such that the true likelihood function cannot be approximated well by the Whittle likelihood. This leads to much worse results, both in terms of bias and RMSE, compared to the estimator using the correct likelihood function. It remains an open question why the impact of misspecification is stronger for some parameters than others. Naturally, this must be related to where in the likelihood the information about different parameters originates and how affected by misspecification those likelihood regions are. This is an interesting

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

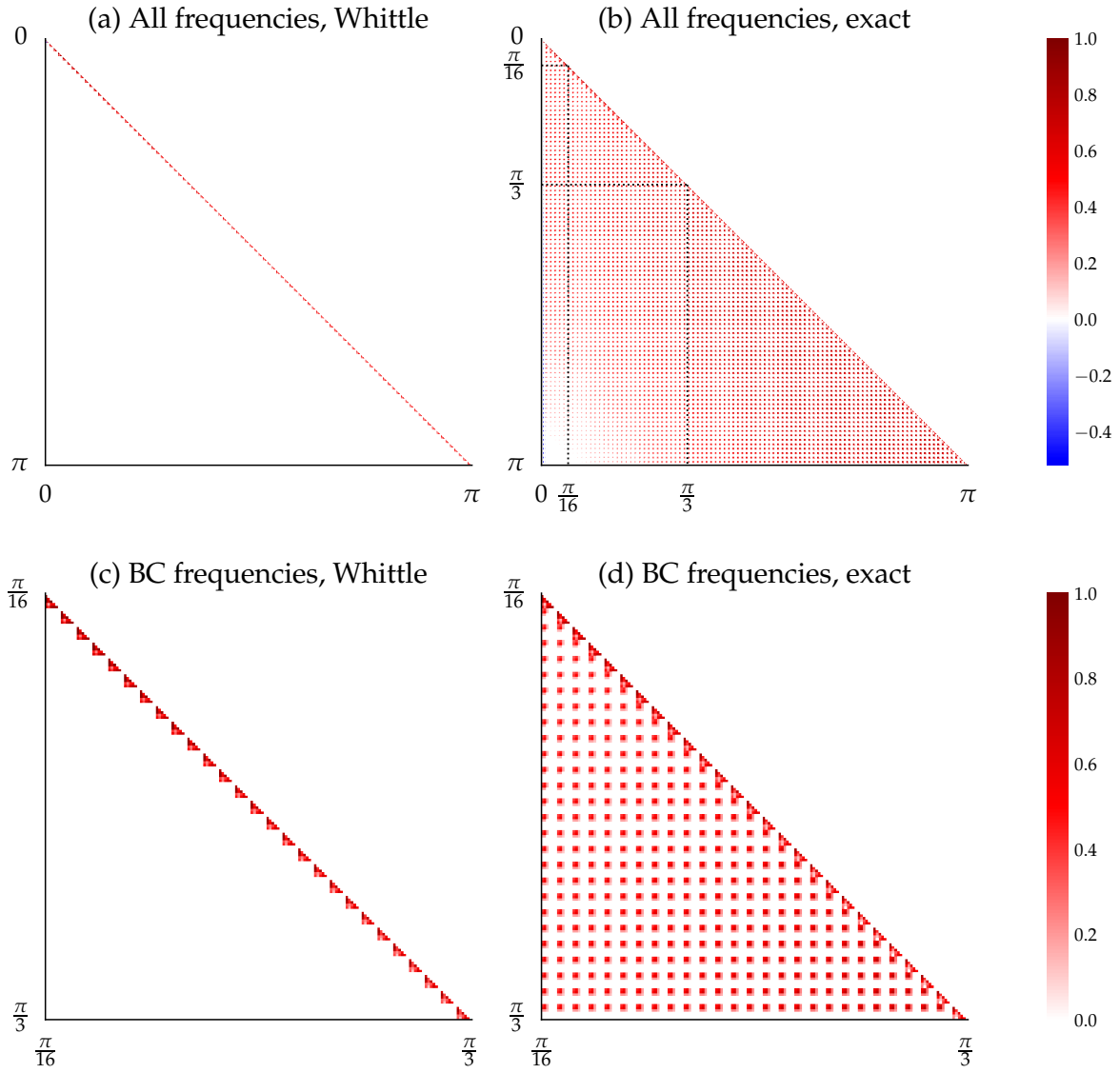

Figure 6: Correlation matrices in the Whittle approximation and the exact frequency domain likelihood function. Because of symmetry, only frequencies from the $[0, \pi]$ part of the spectrum are displayed. Frequencies between $\pi/16$ and $\pi/3$ are business cycle frequencies. Correlation coefficients smaller than .1 in absolute value have been zeroed-out.

question that will be pursued in future research.

An interesting question one may ask is whether the quality of the Whittle approximation would improve significantly with less persistent a_t^p and ζ_t^{IP} processes. To investigate this, I set the value of the autoregressive coefficients of their AR(1) processes to 0.9, which is on par with the value of the next most persistent shock in the model – the

preference shock (see Table 1). In that case the degree of persistence of y_t, c_t, and i_t is very similar and slightly greater than that of h_t, π_t, and r_t. The correlation matrices of the frequency domain likelihoods under that parametrization are shown in Figure 7. As before, it is clear that the off-diagonal blocks are not all zero. However, correlations are now significantly weaker, with a maximum value of 0.48. Again, the largest correlations tend to be between different frequency components of c. This suggests that the Whittle approximation should be more appropriate for the data generating process implied by this alternative parametrization. In the following section I evaluate the effect this has on the relative performance of the Whittle likelihood-based estimators.

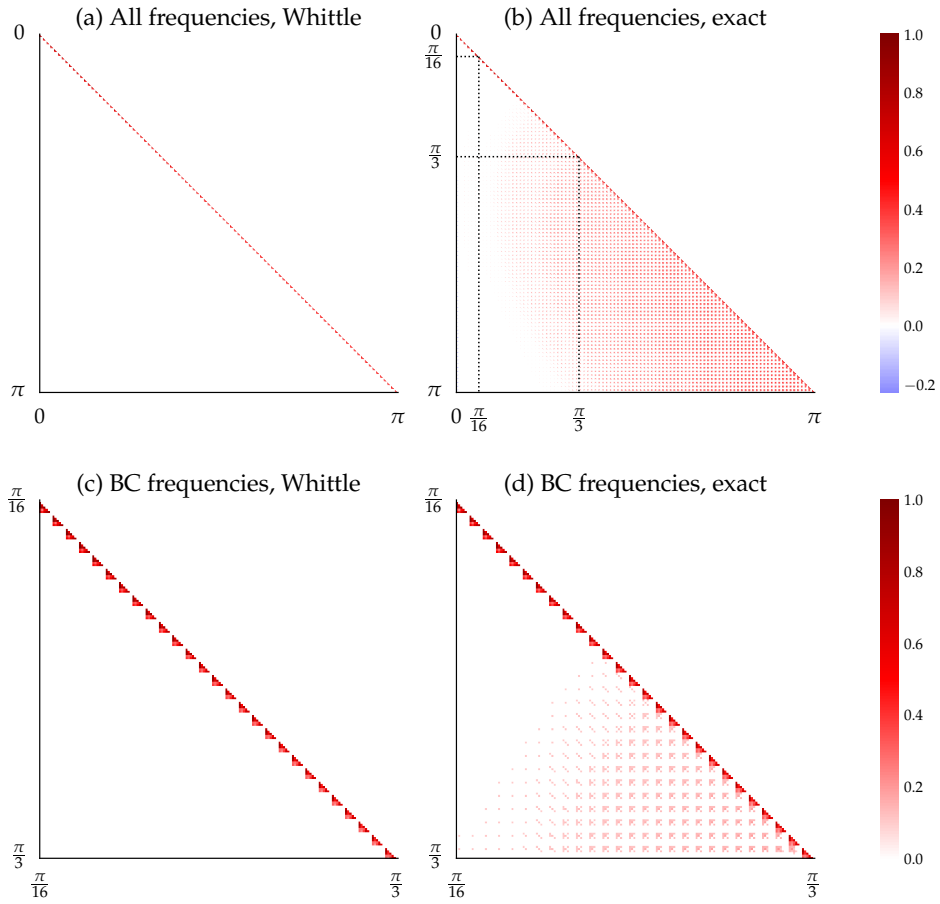

Figure 7: Correlation matrices in the Whittle approximation and the exact frequency domain likelihood function under the alternative parametrization for a_t^p and ζ_t^{IP}

Table 4: Monte Carlo results

	Bias (%)			SD (%)			RMSE (%)		
parameter	TD	FD	BC	TD	FD	BC	TD	FD	BC
ν	-0.4	6.9	10.9	27.0	32.1	51.6	27.0	32.8	52.8
α	-0.2	-0.1	-0.3	3.5	4.0	6.0	3.5	4.0	6.0
ψ	6.3	7.7	13.5	23.7	30.5	59.2	24.5	31.4	60.8
φ	3.7	-10.3	1.9	26.8	32.8	48.8	27.1	34.4	48.9
b	-0.0	-4.4	-2.0	4.6	8.5	9.8	4.6	9.6	10.0
χ	-0.3	-0.2	0.2	2.5	2.9	5.9	2.5	2.9	5.9
κ_R	-0.9	-7.3	-10.2	32.5	37.9	79.2	32.5	38.6	79.9
κ_π	0.6	1.1	12.2	14.4	24.3	50.8	14.4	24.3	52.2
κ_y	31.3	51.3	47.8	62.4	87.4	144.7	69.8	101.3	152.4
ρ_a	1.2	11.6	11.1	63.8	75.2	88.6	63.8	76.1	89.3
ρ_n	4.2	12.5	37.7	91.1	106.1	137.4	91.2	106.9	142.5
ρ_i	-7.3	-15.4	-29.2	24.2	29.7	63.0	25.3	33.4	69.4
ρ_c	-0.7	1.0	1.1	3.9	4.6	7.8	3.9	4.7	7.9
ρ_g	-3.4	-2.6	-4.0	7.0	7.6	20.3	7.8	8.0	20.7
ρ_m	-0.9	1.2	-4.3	8.1	8.8	22.7	8.1	8.8	23.1
ρ_ξ	-2.5	-3.2	-2.9	4.8	7.3	13.5	5.4	7.9	13.8
σ_{a^P}	-9.5	1.2	-34.1	30.6	36.4	56.7	32.0	36.4	66.1
σ_{a^T}	-0.0	2.6	7.5	25.1	28.5	59.9	25.1	28.6	60.3
σ_n	-2.7	-6.4	-1.3	26.7	31.6	52.9	26.9	32.3	52.9
σ_{i^P}	-2.6	3.4	-17.3	54.7	60.8	87.3	54.7	60.9	89.0
σ_{i^T}	4.6	-0.7	29.1	27.5	35.8	80.5	27.9	35.8	85.5
σ_c	1.1	9.7	39.0	46.3	86.0	113.1	46.3	86.5	119.6
σ_g	-1.4	-0.7	1.0	5.3	5.5	13.0	5.5	5.6	13.0
σ_m	-0.0	4.0	12.2	10.4	14.8	47.6	10.4	15.3	49.1
σ_ξ	18.8	17.1	59.6	70.5	83.7	196.8	73.0	85.4	205.6

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

4.3 Results: alternative parametrization

The results for the second set of simulations are organized similarly to the previous section. Table 4 reports the mean, median, and IQR for all parameters. While the TD estimator still performs best in terms of accuracy and volatility, the Whittle likelihood-based estimators are now much more competitive along these dimensions. In particular, unlike before, the IQRs now include the true values for all parameters. Without any exception, the lengths of the TD estimator's IQRs are the shortest and those of the BC estimators' IQRs – the longest. However, the difference, on average, between TD and FD estimators' IQRs is only about 22%, which is in line with what one may expect from theory, given that the two estimators are asymptotically equivalent. On the other hand,

the BC estimator's IQRs are on average 80% wider than those of FD, and about 130% wider than then TD estimator's IQRs. These findings is also easy to see from Figure 8, which displays boxplots of the percent deviations of the parameter estimates from the true values.

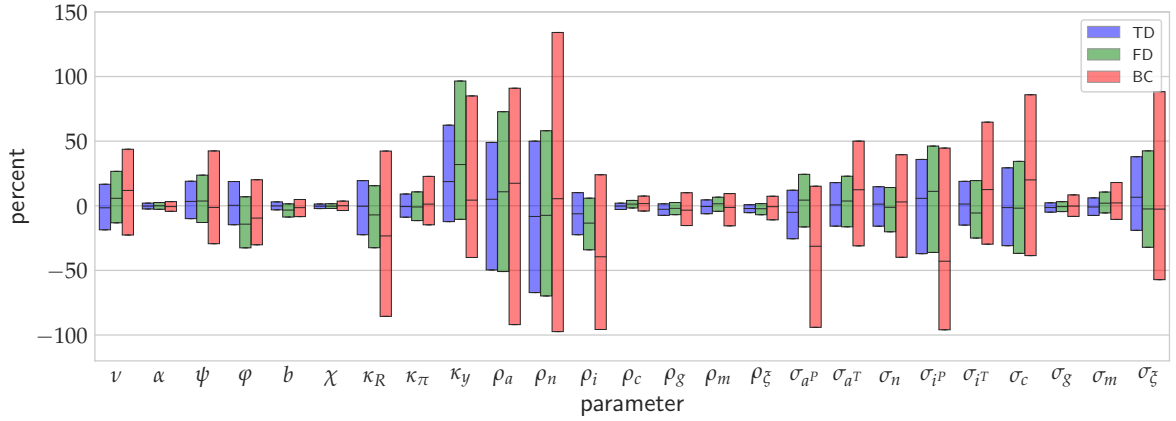

Figure 8: Alternative parametrization of the ACD model. See Notes to Figure 1.

Table 5 reports the bias, standard deviation and root mean squared error for all parameters shown, as before, as percentages relative to the true values. On average, the absolute value of bias of the TD, FD and BC estimates is about 4%, 7% and 16% respectively. Note that this is the same bias for the TD estimator as in the baseline model, while for the Whittle likelihood-based estimators it is significantly smaller than before. Moreover, we see that there is much stronger agreement among the three estimators in terms of which parameters tend to be relatively more or less biased: the values of the rank correlation coefficients between biases are 0.56 in the case of TD and FD estimators, and 0.89 in the case of the TD and BC estimators. We can also confirm that the estimates of α and χ are among the least biased across all estimators, while κ_y, ρ_i, ρ_n, and σ_ξ are common among the most biased parameters. Given the relatively much smaller biases, the results for SD and RMSE are now much more similar. On average, the RMSE of the TD estimator is about 80% of the FD and 45% of the BC estimator, respectively. The rank correlation between the RMSEs of the TD and FD estimator is 0.98, while that of either one of these estimators with the BC one is 0.95 The results in terms of which parameters have lowest or highest RMSE across all estimators are very similar to what was found before (see Table 3). In particular, χ, ρ_ξ, σ_g, α, ρ_c, and ρ_g are the parameters with the lowest RMSE, while κ_y, ρ_a, ρ_n, σ_c, and σ_{i^P}, and σ_ξ have the highest RMSE

across all estimators.

Table 5: Monte Carlo results (cont.)

	Bias (%)			SD (%)			RMSE (%)		
parameter	TD	FD	BC	TD	FD	BC	TD	FD	BC
ν	-0.4	6.9	10.9	27.0	32.1	51.6	27.0	32.8	52.8
α	-0.2	-0.1	-0.3	3.5	4.0	6.0	3.5	4.0	6.0
ψ	6.3	7.7	13.5	23.7	30.5	59.2	24.5	31.4	60.8
φ	3.7	-10.3	1.9	26.8	32.8	48.8	27.1	34.4	48.9
b	-0.0	-4.4	-2.0	4.6	8.5	9.8	4.6	9.6	10.0
χ	-0.3	-0.2	0.2	2.5	2.9	5.9	2.5	2.9	5.9
κ_R	-0.9	-7.3	-10.2	32.5	37.9	79.2	32.5	38.6	79.9
κ_π	0.6	1.1	12.2	14.4	24.3	50.8	14.4	24.3	52.2
κ_y	31.3	51.3	47.8	62.4	87.4	144.7	69.8	101.3	152.4
ρ_a	1.2	11.6	11.1	63.8	75.2	88.6	63.8	76.1	89.3
ρ_n	4.2	12.5	37.7	91.1	106.1	137.4	91.2	106.9	142.5
ρ_i	-7.3	-15.4	-29.2	24.2	29.7	63.0	25.3	33.4	69.4
ρ_c	-0.7	1.0	1.1	3.9	4.6	7.8	3.9	4.7	7.9
ρ_g	-3.4	-2.6	-4.0	7.0	7.6	20.3	7.8	8.0	20.7
ρ_m	-0.9	1.2	-4.3	8.1	8.8	22.7	8.1	8.8	23.1
ρ_ξ	-2.5	-3.2	-2.9	4.8	7.3	13.5	5.4	7.9	13.8
σ_{a^P}	-9.5	1.2	-34.1	30.6	36.4	56.7	32.0	36.4	66.1
σ_{a^T}	-0.0	2.6	7.5	25.1	28.5	59.9	25.1	28.6	60.3
σ_n	-2.7	-6.4	-1.3	26.7	31.6	52.9	26.9	32.3	52.9
σ_{i^P}	-2.6	3.4	-17.3	54.7	60.8	87.3	54.7	60.9	89.0
σ_{i^T}	4.6	-0.7	29.1	27.5	35.8	80.5	27.9	35.8	85.5
σ_c	1.1	9.7	39.0	46.3	86.0	113.1	46.3	86.5	119.6
σ_g	-1.4	-0.7	1.0	5.3	5.5	13.0	5.5	5.6	13.0
σ_m	-0.0	4.0	12.2	10.4	14.8	47.6	10.4	15.3	49.1
σ_ξ	18.8	17.1	59.6	70.5	83.7	196.8	73.0	85.4	205.6

Note: Alternative parametrization of the ACD model. See Notes to Table 3

Figures 9 - 12 show histograms of the sampling distributions of the TD, FD, and BC estimates, and Q-Q plots for their departures from normality. As explained before, it cannot be expected that the normal approximation will be very accurate, given the relatively short sample. Nevertheless, we see a significant improvement of the sampling distributions in particular for the two Whittle likelihood based estimators.

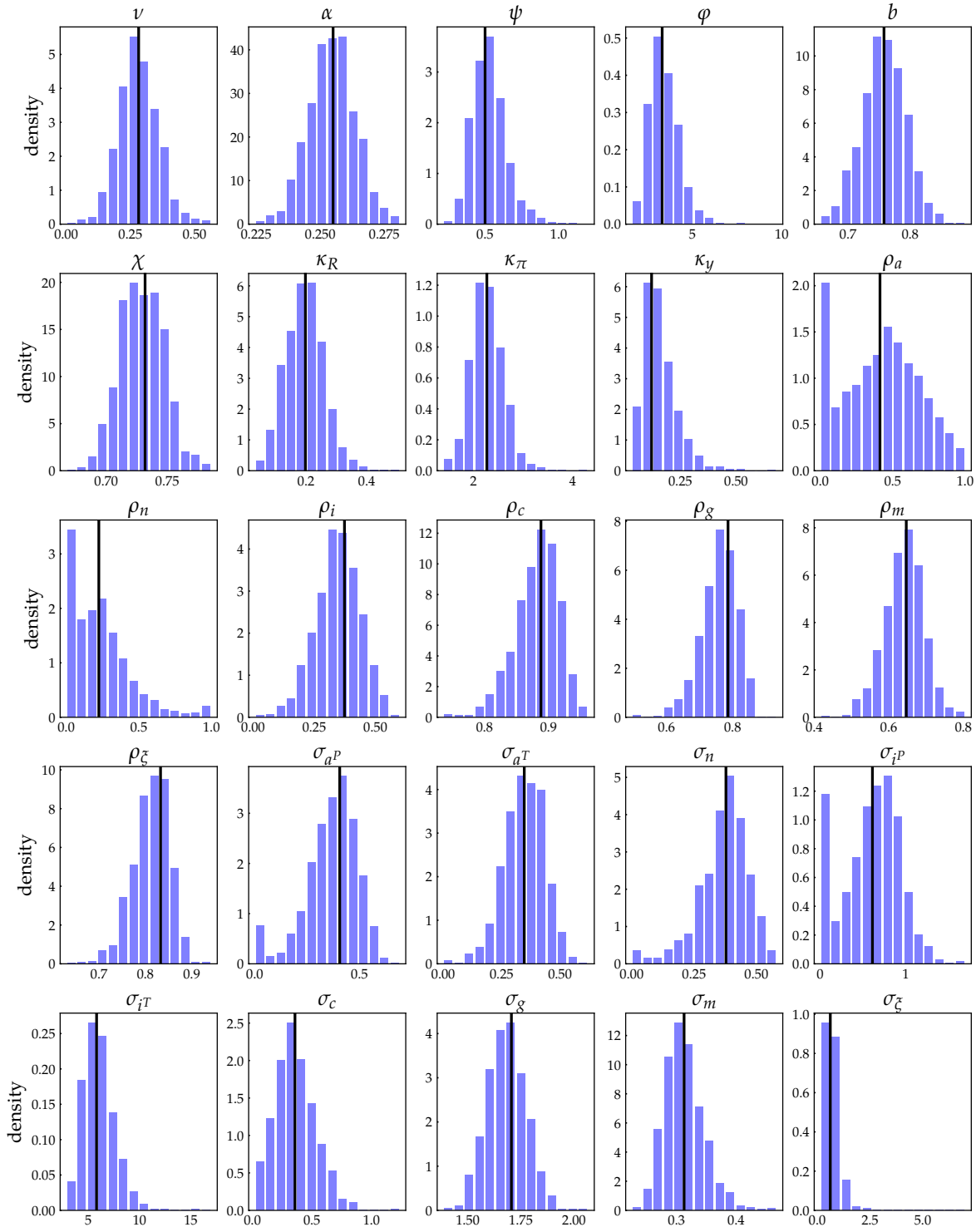

Figure 9: Alternative parametrization of the ACD model. Sampling distributions of the estimated parameters, time domain MLE. The black vertical lines indicate true values. The results are based on 1000 MC replications and sample size of $T = 192$.

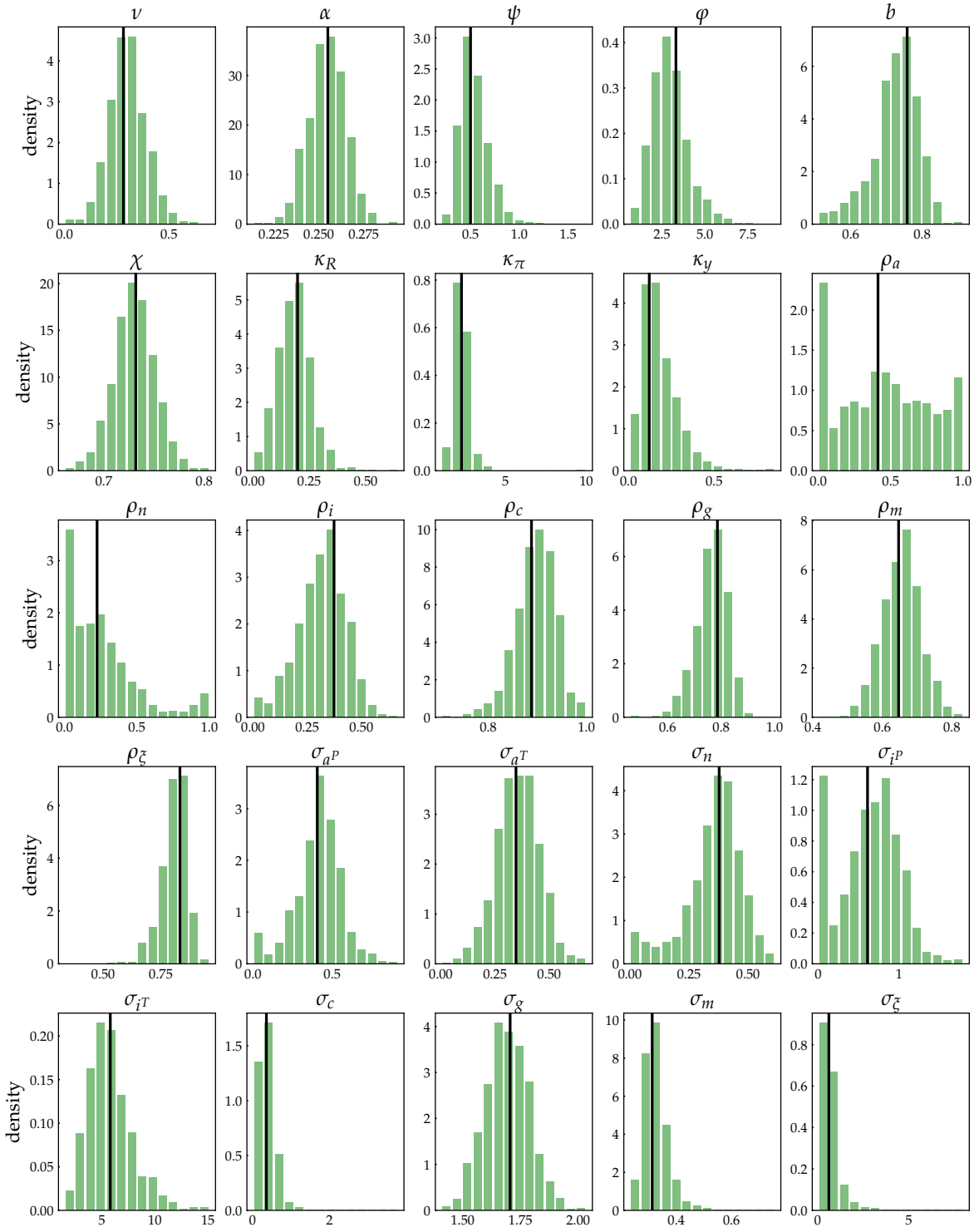

Figure 10: Alternative parametrization of the ACD model. Sampling distributions of the estimated parameters, frequency domain MLE using all frequencies. The black vertical lines indicate true values. The results are based on 1000 MC replications and sample size of $T = 192$.

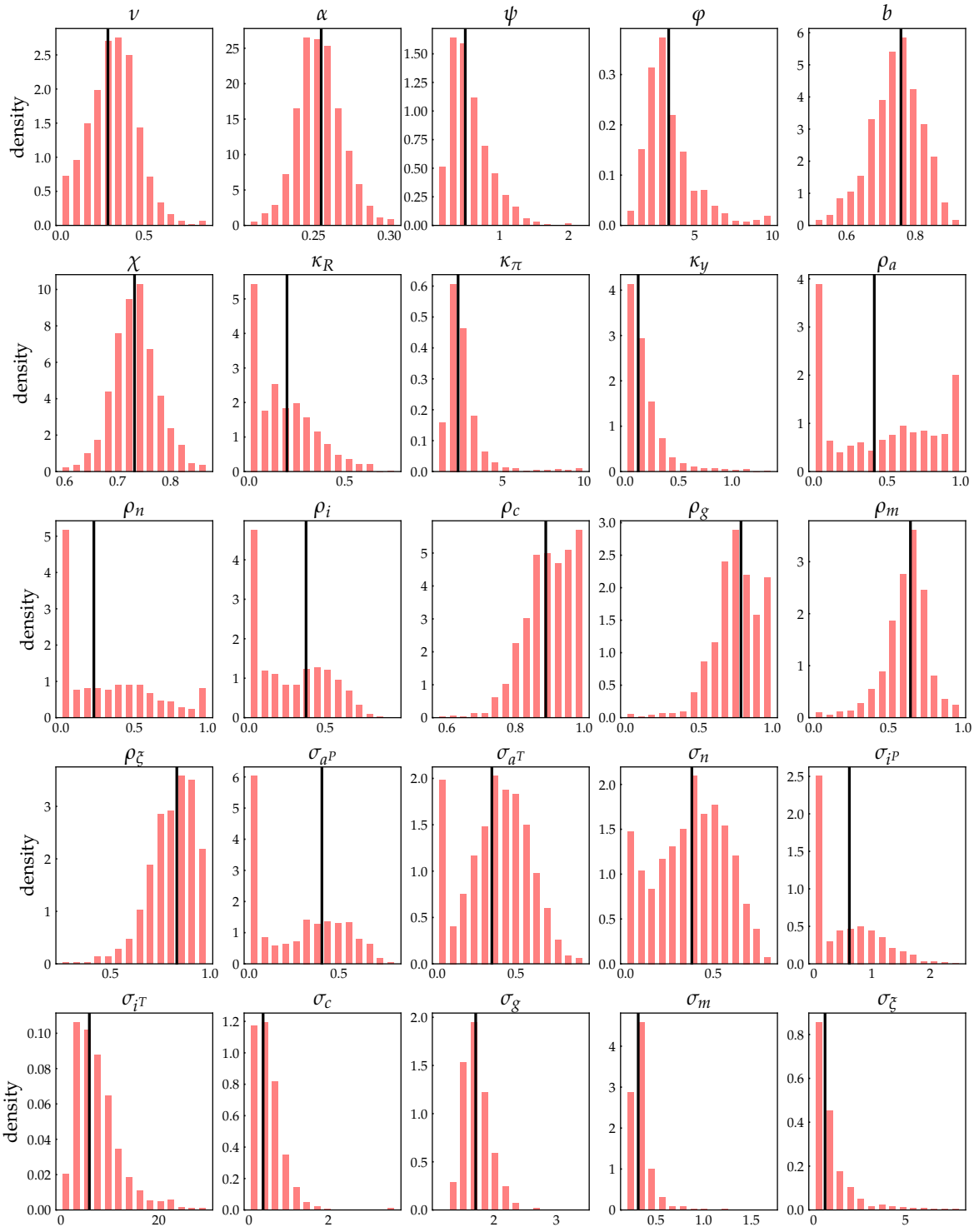

Figure 11: Alternative parametrization of the ACD model. Sampling distributions of the estimated parameters, frequency domain MLE using only business cycle frequencies. The black vertical lines indicate true values. The results are based on 1000 MC replications and sample size of $T = 192$.

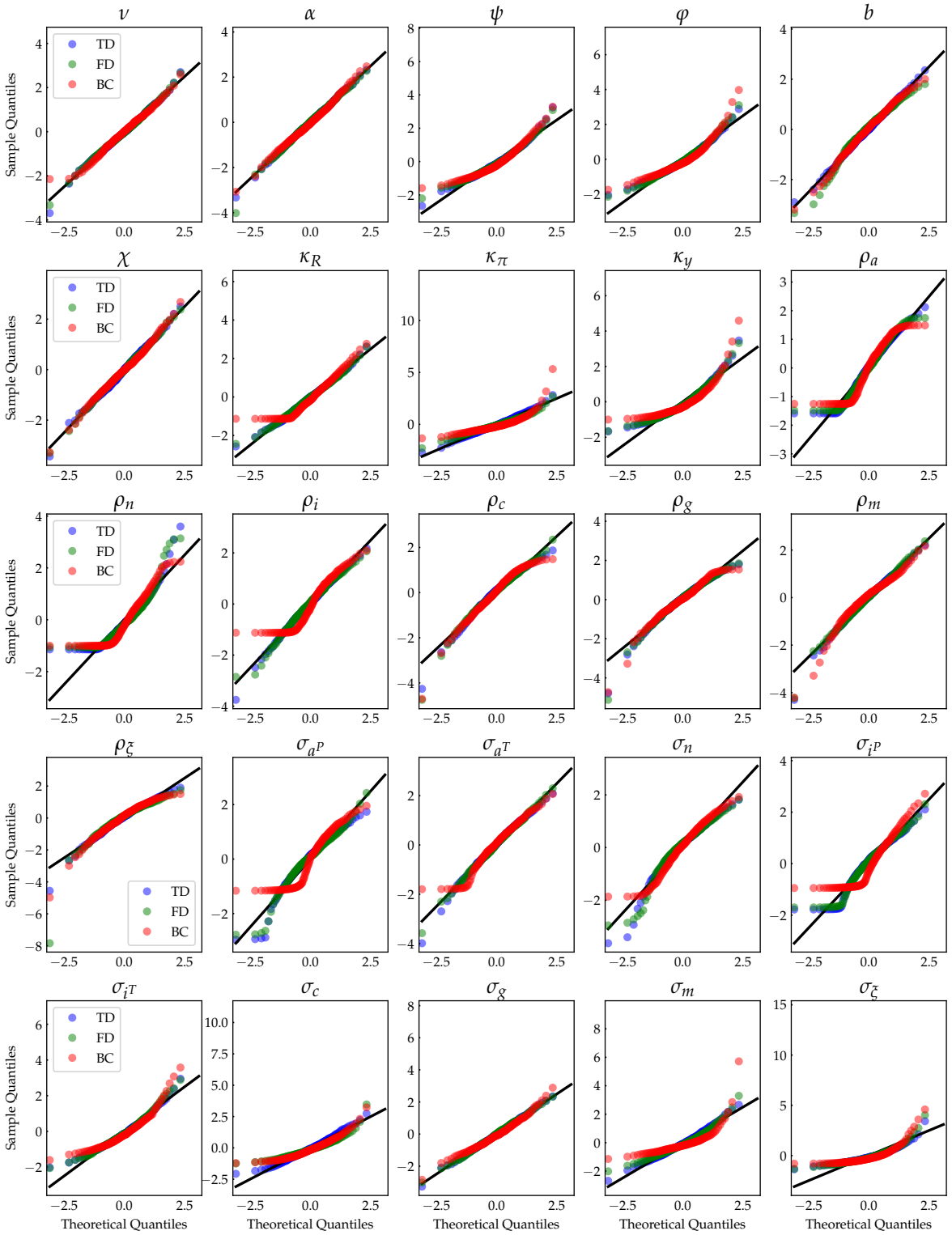

Figure 12: Q-Q plots for the alternative parametrization of the ACD model.

5 Concluding Comments

The purpose of this study was to shed some light on the performance of the band-spectral estimation approach applied to business cycle models. Band-spectral methods are widely used to study frequency-dependent relationships among variables. In business cycle research, this approach permits the estimation of structural models on the basis of the frequencies they are best suited to represent, such as the business cycle frequencies. In particular, the frequency domain approximation of the likelihood function (the Whittle likelihood) can be used to estimate the parameters of fully-specified models on the basis of a targeted band of frequencies. Using the medium-scale model of Angeletos et al. (2018) as a data-generating process, I performed a Monte Carlo study to evaluate the finite-sample properties of the band-spectral maximum likelihood estimator (MLE) and to compare them to those of the full spectrum and the exact time-domain MLE. The results show that using the band-spectral estimator leads to considerable biases and efficiency losses for most estimated parameters. In fact, the performance of both Whittle likelihood-based estimators is found to be seriously deficient in terms of bias and accuracy, in contrast to that of the time domain estimator, which successfully recovers all model parameters. I showed how these findings can be explained with the theoretical properties of the underlying model, and describe simple-to-use tools and diagnostics that can be used to detect potential problems in band-spectral estimation for a wide class of macroeconomic models.

References

Angeletos, G.-M., F. Collard, and H. Dellas (2018): "Quantifying confidence," *Econometrica*, 86, 1689–1726.

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

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

Engle, R. F. (1974): "Band Spectrum Regression," *International Economic Review*, 15, 1–11.

Hannan, E. J. (1963): "Regression for time series with errors of measurement," *Biometrika*, 50, 293–302.

Hansen, L. P. and T. J. Sargent (1993): "Seasonality and approximation errors in rational expectations models," *Journal of Econometrics*, 55, 21–55.

Qu, Z. and D. Tkachenko (2012): "Identification and frequency domain quasi-maximum likelihood estimation of linearized dynamic stochastic general equilibrium models," *Quantitative Economics*, 3, 95–132.

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

Uhlig, H. (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, pp. 125–139.