Identification and Frequency Domain QML Estimation of Linearized DSGE Models*

Zhongjun Qu†  Denis Tkachenko†
Boston University  Boston University

August 1, 2010

Abstract

This paper considers issues related to identification, inference and computation in linearized Dynamic Stochastic General Equilibrium (DSGE) models. We first provide a necessary and sufficient condition for the local identification of the structural parameters based on the (first and) second order properties of the process. The condition allows for arbitrary relations between the number of observed endogenous variables and structural shocks and is simple to verify. The extensions including identification through a subset of frequencies, partial identification, conditional identification and identification under general nonlinear constraints are also studied. When lack of identification is detected, the method can be further used to trace out non-identification curves. For estimation, restricting our attention to nonsingular systems, we consider a frequency domain quasi-maximum likelihood (FQML) estimator and present its asymptotic properties. The limiting distribution of the estimator can be different from results in the related literature due to the structure of the DSGE model. Finally, we discuss a quasi-Bayesian procedure for estimation and inference. The procedure can be used to incorporate relevant prior distributions and is computationally attractive.

Keywords: structural identification, infinite dimensional mapping, non-identification curve, frequency domain estimation, MCMC.

*We thank Pierre Perron and Hsueh-Ling Huynh for useful comments and suggestions.
†Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (qu@bu.edu).
‡Department of Economics, Boston University, 270 Bay State Rd., Boston, MA, 02215 (tkatched@tcd.ie).
1 Introduction

The formal statistical analysis of DSGE models has become an important subject of modern macroeconomics. It is typically conducted in the time domain using a state space representation with the aid of Kalman or particle filtering, see An and Schorfheide (2007) and Fernández-Villaverde (2009) for reviews of related literature. This paper considers issues related to estimation, computation and inference from a spectral domain perspective. The goal is to present a unified framework for identifying and estimating linearized DSGE models based on the mean and the spectrum of the underlying process.

The identification of DSGE models is important for both calibration and formal statistical analysis, although the related literature is quite sparse. Substantial progress has been made recently, with an important contribution being Komunjer and Ng (2009). Other related works are Canova and Sala (2009), Consolo, Favero, and Paccagnini (2009) and Iskrev (2010). Komunjer and Ng (2009) documented that an inherent difficulty in the identification analysis is that the reduced form parameters (i.e., the ones appearing directly in the solution of the model) are in general not identifiable, thus the traditional approach of identifying structural parameters from the reduced form breaks down. Also, the solution system of a DSGE model can be singular (i.e., when the number of observed endogenous variables is greater than the number of exogenous shocks), which constitutes an additional layer of conceptual difficulty. They provided a necessary and sufficient condition for local identification of the dynamic parameters from the second order properties of the system by exploiting the dynamic structure of the model. Our identification analysis is inspired by theirs, although the approach we take and the results obtained all have distinctively different flavors. Our results, as well as the assumptions, are transparent and easily verifiable. Also, our approach permits to obtain results not contained in Komunjer and Ng (2009).

We first focus on the identification of the dynamic parameters from the spectrum. We treat the elements of the spectral density matrix as mappings from the structural parameter space to complex valued functions defined over $[-\pi, \pi]$ in a Banach space. Then, the parameters are locally identified if and only if the overall mapping is locally injective (that is, if any local change in parameter values leads to a different image). This leads to a necessary and sufficient rank condition for local identification, which depends on the first order derivative of the spectral density matrix with respect to the structural parameters of interest. Depending on the model at hand, the resulting condition can be easily evaluated analytically or numerically. The result is general because the assumptions mainly involve the uniqueness of the DSGE solution (i.e., determinacy) and the continuity and smoothness of the spectral density matrix. Unlike Komunjer and Ng (2009),
we do not impose restrictions on the relation between the number of observed endogenous variables and structural shocks. Note that although the identification condition is formulated in the spectral domain, it has a time domain interpretation as well. Specifically, under some regularity condition that ensures a one-to-one mapping between the spectral density matrix and the autocovariance functions, the condition is also necessary and sufficient for local identification through the complete set of autocovariances. Next, we incorporate the steady state parameters into the analysis and study identification through both the first and second order properties of the process. The result we obtain is analogous to the previous case with the addition of an extra term depending on the steady state parameters. When interpreted in the time domain, this condition is necessary and sufficient for local identification through the mean and the complete set of autocovariances.

We discuss various extensions of these two identification results. (1) We study identification through a subset of frequencies. This is relevant for situations where it is desirable to construct estimators based on a subset of frequencies to minimize the effect of unmodeled seasonality or measurement errors. (2) We consider partial identification, i.e., identifying a subset of parameters without making identification statements about the rest. (3) We give necessary and sufficient condition for conditional identification, i.e., the identification of a subset of parameters while holding the values of the other parameters fixed at some known value. (4) We also study identification under general nonlinear parameter constraints. For example, this allows to constrain some monetary shocks to have no long run effect on real variables, which can be easily formulated as a set of restrictions on the spectral density matrix at frequency zero. Note that the last three extensions are motivated by Komunjer and Ng (2009), although the assumptions used there are different.

Furthermore, when lack of identification is detected, our method can be used to trace out parameter values that yield processes with identical (first and) second order properties. We summarize the path of these values via non-identification curves and provide a simple algorithm to obtain them. We illustrate the method using a model considered by An and Schorfheide (2007) and document a serious concern about the identification of the parameters in the Taylor rule equation. The result shows that when varying parameters in this equation along a certain path the (mean and) spectrum of observables stay the same, thus it is impossible to uniquely pin down the parameter values even with an infinite sample. It also suggests that non-identification curves can be useful tools to characterize the nature of lack of identification, conveying richer information than a conventional zero-one identification analysis. It appears that our method is the first to deliver such curves.

As will be becoming clear, our results, as well as their proofs, are closely connected to Rothenberg (1971), who considered identification of parametric econometric models from the density func-
tions. However, there exists an important difference. Namely, the spectral density is a complex valued matrix that may be singular. This generates some conceptual and technical difficulties that do not arise in Rothenberg (1971). Consequently, our condition does not involve the Hessian matrix that played a central role in Rothenberg’s (1971) analysis.

An identification result is useful only if it corresponds to an estimator. This motivates formal consideration of the frequency domain quasi-maximum likelihood (FQML) estimation in this paper. The FQML approach was first proposed by Whittle (1951). Its statistical properties have been studied by, among others, Walker (1964), Hannan (1973), Dunsmuir and Hannan (1976), Dunsmuir (1979) and Hosoya and Taniguchi (1982) in the statistics literature. In the economics literature, Hansen and Sargent (1993) derived the FQML as an approximation to the time domain Gaussian QML and used it to understand the effect of seasonal adjustment in estimating rational expectations models. Diebold, Ohanian and Berkowitz (1998) laid out a general framework for estimation and model diagnostics based on a full second order comparison of the model and data dynamics. Their criterion function includes FQML as a special case.

The contribution of the current paper in the area of FQML estimation is threefold. First, we formally establish the link between the identification result and the property of the estimator, by showing that the rank condition derived is necessary and sufficient for the estimator to be asymptotically locally unique. Therefore, the identification result is empirically relevant. Second, we derive the limiting distribution of the estimator under mild conditions. We show that, due to the structure of the DSGE model, this distribution can be different from the ones obtained in the statistics literature, as, for example, in Dunsmuir and Hannan (1976), and Hosoya and Taniguchi (1982). Finally, we discuss a computationally attractive method to obtain the estimates, following the approach of Chernozhukov and Hong (2003). Besides the computational advantage, it allows us to impose priors on the parameters, thus having a (quasi) Bayesian interpretation.

Besides the above mentioned papers, there exists a small but growing literature that exploits the merits of estimation and diagnosis of econometric models in the spectral domain. Engle (1974) considered band spectrum regressions and demonstrated their value in dealing with errors in variables and seasonality. Altug (1989) applied FQML to estimate models with additive measurement errors. Watson (1993) suggested plotting the model and data spectra as one of the most informative diagnostics. Berkowitz (2001) considered the estimation of rational expectation models based on the spectral properties of the Euler residuals. Also, see Christiano, Eichenbaum and Marshall (1991) and Christiano and Vigfusson (2003) for applications of FQML to various problems. We believe that the identification, estimation and computational results obtained in this paper can be
useful for the further development of the literature in this field and for facilitating estimation and comparison of more sophisticated models.

The paper is organized as follows. The structure of the DSGE solution is discussed in Section 2. Section 3 considers the local identification of the structural parameters together with an algorithm to trace out non-identification curves and an illustrative example. The FQML estimator and its asymptotic properties are studied in Section 4. The discussion on interpretation of the estimates in misspecified models is also included. Section 5 presents a quasi-Bayesian approach for computation and inference. Section 6 concludes. All proofs are contained in the Appendix.

The following notation is used. $|z|$ is the modulus of $z$; the imaginary unit is denoted by $i$. $X^*$ stands for the conjugate transpose of a complex valued matrix $X$. For a random vector $x_t$, $x_{ta}$ denotes its $a$-th element. For a matrix $A$, $A_{ab}$ stands for its $(a,b)$-th entry. If $f_\theta \in R^k$ is a differentiable function of $\theta \in R^p$, then $\partial f_\theta / \partial \theta'$ is a $k$-by-$p$ matrix of partial derivatives evaluated at $\theta_0$. “$\rightarrow P$” and “$\rightarrow d$” signify convergence in probability and in distribution. And $O_p(\cdot)$ and $o_p(\cdot)$ are the usual symbols for stochastic orders of magnitude.

## 2 The model

Suppose a discrete time DSGE model has been solved and log-linearized around the steady state. Assume the solution is unique. Let $Y_t^d(\theta)$ be the log-deviations of endogenous variables from their steady states with $\theta$ being a finite dimensional structural parameter vector, i.e., containing the dynamic parameters. $Y_t^d(\theta)$ can be represented in various ways, and our method does not rely on a particular representation. To maintain generality, we only assume that they are representable as

$$Y_t^d(\theta) = \sum_{j=0}^{\infty} h_j(\theta) \epsilon_{t-j},$$

(1)

where $h_j(\theta)$ ($j = 0, ..., \infty$) are real valued matrices of constants and $\{\epsilon_t\}$ is a white noise process of unobserved structural shocks. The dimensions of the relevant variables and parameters are as follows:

$$Y_t^d(\theta) : n_Y \times 1, \quad \epsilon_t : n_\epsilon \times 1, \quad h_j(\theta) : n_Y \times n_\epsilon, \quad \theta : q \times 1.$$

Let $H(L; \theta)$ denote the matrix of lagged polynomials, i.e.,

$$H(L; \theta) = \sum_{j=0}^{\infty} h_j(\theta) L^j.$$

(2)

Then, $Y_t^d(\theta)$ can be written concisely as

$$Y_t^d(\theta) = H(L; \theta) \epsilon_t.$$  

(3)
Remark 1 We work directly with the vector moving average representation (3) without assuming invertibility, i.e., $\epsilon_t = \sum_{j=0}^{\infty} g_j(\theta) Y_{t-j}^d(\theta)$ for some $g_j(\theta)$. The latter is restrictive because it requires $n_Y \geq n_\epsilon$. Consequently, we allow for both $n_Y \geq n_\epsilon$ and $n_Y < n_\epsilon$. Note that the system is singular if $n_Y > n_\epsilon$.

Assumption 1. $\{\epsilon_t\}$ satisfies $E(\epsilon_t) = 0$, $E(\epsilon_t \epsilon_t') = \Sigma(\theta)$ with $\Sigma(\theta)$ being a nonsingular $n_\epsilon \times n_\epsilon$ matrix for all $\theta$ and $E(\epsilon_t \epsilon_s^{'}) = 0$ for all $t \neq s$. $\sum_{j=0}^{\infty} \text{tr}(h_j(\theta) \Sigma(\theta) h_j(\theta)') < \infty$.

Assumption 1 states that the shocks are serially uncorrelated with a finite covariance matrix. The assumption implies that $Y_t^d(\theta)$ is covariance stationary and has a spectral density matrix $f_\theta(\omega)$ that can be written as

$$f_\theta(\omega) = \frac{1}{2\pi} H(\exp(-i\omega); \theta) \Sigma(\theta) H(\exp(-i\omega); \theta)^*,$$

where $X^*$ denotes the conjugate transpose of a generic complex matrix $X$. To illustrate the flexibility of (1), we consider the following two examples.

Example 1 Consider a linear rational expectations system as in Sims (2002) (in this example and the next, we omit the dependence of the parameters on $\theta$ to simplify notation):

$$\Gamma_0 S_t = \Gamma_1 S_{t-1} + \Psi Z_t + \Pi \eta_t,$$  

where $S_t$ is a vector of model variables that includes the endogenous variables and the conditional expectation terms, $Z_t$ is an exogenously evolving, possibly serially correlated, random disturbance, and $\eta_t$ is an expectational error. Models with more lags or with lagged expectations can be accommodated by expanding the $S_t$ vector accordingly. Then, under some conditions (Sims, 2002, p. 12), the system can be represented as

$$S_t = \Theta_1 S_{t-1} + \Theta_0 Z_t + \Theta_S \sum_{j=1}^{\infty} \Theta_f^{j-1} \Theta_Z E_t Z_t+j,$$  

where $\Theta_0, \Theta_1, \Theta_S, \Theta_f$ and $\Theta_Z$ are functions of $\Gamma_0, \Gamma_1, \Psi$ and $\Pi$. Assuming $Z_t$ follows a vector linear process (for example, $Z_{t+1} = \Phi Z_t + \epsilon_{t+1}$), we then have $S_t = \Theta_1 S_{t-1} + B(L) \epsilon_t$ for some lag polynomial matrix $B(L)$, implying $S_t = (I - \Theta_1 L)^{-1} B(L) \epsilon_t$.

Let $A(L)$ be a matrix of finite order lag polynomials that specifies the observables such that

$$Y_t^d = A(L) S_t.$$

Then, we have

$$Y_t^d = A(L) (I - \Theta_1 L)^{-1} B(L) \epsilon_t.$$

Therefore, the spectral density of $Y_t^d$ is given by (4) with $H(L; \theta) = A(L) (I - \Theta_1 L)^{-1} B(L)$.
Remark 2 In the above example, the matrix $A(L)$ offers substantial flexibility since it allows us to study identification and estimation based on a subset of variables (equations) or a linear transformation of them. To see this, suppose $S_t$ includes two endogenous variables $x_t$ and $w_t$. Then $A(L)$ can be chosen such that $Y^d_t$ includes only $x_t$ but not $w_t$, or includes $x_t - x_{t-1}$ but not $x_t$. The fact that we do not impose restrictions on the relation between $n_Y$ and $n_x$ permits such analysis. This point is further illustrated in Section 3.2 through a concrete example.

Example 2 Another representation used in the literature by, among others, Uhlig (1999) and Komunjer and Ng (2009), is as follows:

$$
\begin{align*}
k_{t+1} &= Pk_t + Qz_t \\
w_t &= Rk_t + Sz_t \\
z_{t+1} &= \Psi z_t + \epsilon_{t+1},
\end{align*}
$$

where $k_t$ is a vector of observed endogenous (state) variables whose values are known at time $t$, $w_t$ is a vector of observed endogenous (jump) variables, $z_t$ has the same definition as in the previous example, and $P$, $Q$, $R$, $S$ and $\Psi$ are matrices of constants depending on the structural parameter $\theta$. Let

$$
Y^d_t = \begin{pmatrix} k_t \\ w_t \end{pmatrix}.
$$

Then, the spectral density of $Y^d_t$ is given by (4) with

$$
H(L; \theta) = \begin{pmatrix} L^{-1} [I - PL] & 0 \\ -R & I \end{pmatrix}^{-1} \begin{pmatrix} Q \\ S \end{pmatrix} [I - \Psi L]^{-1}.
$$

Again, one can study identification and estimation based on a subset of equations or a linear combination of them by picking an appropriate $A(L)$ and considering $Y^d_t = A(L)(k'_t, w'_t)'$ instead of (7), which corresponds to

$$
H(L; \theta) = A(L) \begin{pmatrix} L^{-1} [I - PL] & 0 \\ -R & I \end{pmatrix}^{-1} \begin{pmatrix} Q \\ S \end{pmatrix} [I - \Psi L]^{-1}.
$$

As will become clear later, if estimating the dynamic parameters is the main objective, then it is not necessary to specify the steady states of the DSGE solution. However, in some cases one may be interested in estimating the dynamic and steady state parameters jointly, for example, for conducting welfare analyses. Our framework permits this. As a matter of notation, let $\alpha$ denote
the structural parameters that affect only the steady states (parameters affecting both the steady states and log-deviations are treated as dynamic). Define the augmented parameter vector

\[ \bar{\theta} = (\theta', \alpha')' \]

and assume that the observables \( (Y_t) \) are related to the log-deviations \( (Y^d_t(\theta)) \) and the steady states \( (\mu(\bar{\theta})) \) via

\[ Y_t = \mu(\bar{\theta}) + Y^d_t(\theta). \]

In the remainder of the paper, we will examine the identification and estimation of \( \theta \) based on the properties of \( f_\theta(\omega) \) alone, and of \( \bar{\theta} \) based jointly on \( \mu(\bar{\theta}) \) and \( f_\theta(\omega) \).

3 Local identification of structural parameters

We first consider the identification of \( \theta \) at some \( \theta_0 \) and subsequently of \( \bar{\theta} \) at some \( \bar{\theta}_0 \). The next assumption imposes some restrictions on the parameter space.

**Assumption 2.** \( \theta \in \Theta_{\theta} \subset \mathbb{R}^q \) and \( \alpha \in \Theta_{\alpha} \subset \mathbb{R}^p \) with \( \Theta_{\theta} \) and \( \Theta_{\alpha} \) being compact and convex. \( \Theta = \Theta_{\theta} \times \Theta_{\alpha} \). Assume \( \theta_0 \) and \( \bar{\theta}_0 \) are interior points of \( \Theta_{\theta} \) and \( \Theta \) respectively.

Note that for identification analysis alone, we do not require the compactness and convexity assumptions on \( \Theta_{\theta} \) and \( \Theta_{\alpha} \). However, they are needed for studying the asymptotic properties of the parameter estimates.

The concept for location identification of the dynamic parameters is the same as in Komunjer and Ng (2009).

**Definition 1** The dynamic parameter vector \( \theta \) is said to be locally identifiable from the second order properties of \( \{Y_t\} \) at a point \( \theta_0 \) if there exists an open neighborhood of \( \theta_0 \) in which \( f_{\theta_1}(\omega) = f_{\theta_0}(\omega) \) for all \( \omega \in [-\pi, \pi] \) implies \( \theta_0 = \theta_1 \).

The above concept is formulated in the frequency domain. However, there is an equivalent formulation in the time domain in terms of autocovariance functions. Specifically, suppose \( \{Y_t\} \) satisfy Assumption 2 with autocovariance function \( \Gamma(k) \) \( (k = 0, \pm 1, ...) \) satisfying \( \Gamma(k) = \Gamma(-k) \) and that \( f_\theta(\omega) \) is continuous in \( \omega \). Then, Theorem 1” in Hannan (1970) implies that there is a one-to-one mapping between \( \Gamma(k) \) \( (k = 0, \pm 1, ...) \) and \( f_\theta(\omega) \) \( (\omega \in [-\pi, \pi]) \) given by

\[ \Gamma(k) = \int_{-\pi}^{\pi} \exp(ik\omega) f_\theta(\omega) d\omega. \]
Therefore, $\theta$ is locally identifiable from $f_\theta(\omega)$ if and only if it is locally identifiable from the complete set of autocovariances $\{\Gamma(k)\}_{k=\infty}^{-\infty}$ of $Y_t$.

The spectral density matrix has $n_Y^2$ elements. Each element can be viewed as a map from $\Theta_\theta$ to complex valued functions defined over $[-\pi, \pi]$ in a Banach space. Therefore, the parameters are locally identified at $\theta_0$ if and only if the overall mapping is locally injective (i.e., any local change in parameter values will lead to a different image for some element). The mappings are infinite dimensional and difficult to analyze directly. However, it turns out the identification can be characterized by a finite dimensional matrix. To state this precisely, we start with the following assumption needed.

**Assumption 3.** The elements of $f_\theta(\omega)$ are continuous in $\omega$ and continuous and differentiable in $\theta$. The elements of the derivatives $\partial \vec{f}(\theta(\omega))/\partial \theta'$ are continuous in $\theta$ and $\omega$. Let

$$G(\theta) = \int_{-\pi}^{\pi} \left( \frac{\partial \vec{f}(\theta(\omega))}{\partial \theta'} \right)' \left( \frac{\partial \vec{f}(\theta(\omega))}{\partial \theta'} \right) d\omega$$

Assume there exists an open neighborhood of $\theta_0$ in which $G(\theta)$ has a constant rank.

This first part of the assumption requires the spectral density to be smooth with continuous first order derivatives. The second part requires $\theta_0$ to be a regular point of the matrix $G(\theta)$. These assumptions are quite mild. Note that in the definition of $G(\theta)$, "t" denote simple transposes rather than conjugate transposes.

**Remark 3** The dimension of $G(\theta)$ is always $q \times q$ and independent of $n_Y$ or $n_\epsilon$. Its $(j,k)$-th element is given by

$$G_{jk}(\theta) = \int_{-\pi}^{\pi} \text{tr} \left\{ \frac{\partial f_\theta(\omega)}{\partial \theta_j} \frac{\partial f_\theta(\omega)}{\partial \theta_k} \right\} d\omega.$$  

This representation is useful for computing $G(\theta)$ in practice. Lemma A.1 in the Appendix provides another representation that shows $G(\theta)$ is real, symmetric and positive semidefinite.

**Theorem 1** Let Assumptions 1-3 hold. Then, $\theta$ is locally identifiable from the second order properties of $\{Y_t\}$ at a point $\theta_0$ if and only if $G(\theta_0)$ is nonsingular.

The result allows for arbitrary relations between $n_Y$ and $n_\epsilon$. In particular, it permits $f_\theta(\omega)$ to be rank deficient, thus allowing for singular systems. The result, as well as its proof, is closely related to Theorem 1 in Rothenberg (1971), who considered identification in parametric models. In his case, $f_\theta(\omega)$ is replaced by the parametric density function and $G(\theta)$ is simply the information matrix. Since the information matrix describes the local curvature of the log-likelihood as a function
of \( \theta \), its rank naturally provides a measure for identification, for lack of identification is simply the lack of sufficient information to distinguish between alternative structures. In our case, the result is equally intuitive, since the parameters are locally identified if and only if any deviation of the parameters from \( \theta_0 \) leads to different mappings for \( f_{\theta}(\omega) \).

The main computational work in obtaining \( G(\theta_0) \) is to evaluate the first order derivatives and to compute the integral. This is typically straightforward using numerical methods. Because \( G(\theta) \) is real, symmetric and positive semidefinite, its eigen decomposition always exists. Therefore, the rank of \( G(\theta_0) \) can be evaluated using an algorithm for eigenvalue decomposition and counting the number of nonzero eigenvalues.

Given the insight conveyed by Theorem 1, it becomes straightforward to study the identification of \( \theta \) based on both the first and second order properties of the process.

**Definition 2**  The parameter vector \( \hat{\theta} \) is said to be locally identifiable from the first and second order properties of \( \{Y_t\} \) at a point \( \theta_0 \) if there exists an open neighborhood of \( \theta_0 \) in which \( \mu(\theta_1) = \mu(\theta_0) \) and \( f_{\theta_1}(\omega) = f_{\theta_0}(\omega) \) for all \( \omega \in [-\pi, \pi] \) implies \( \theta_0 = \theta_1 \).

**Assumption 4.** The elements of \( \mu(\hat{\theta}) \) are continuously differentiable with respect to \( \hat{\theta} \). Let

\[
\tilde{G}(\hat{\theta}) = \int_{-\pi}^{\pi} \left( \frac{\partial \text{vec}(f_{\theta}(\omega))}{\partial \hat{\theta}} \right) \left( \frac{\partial \text{vec}(f_{\theta}(\omega))}{\partial \hat{\theta}} \right)^\prime d\omega + \frac{\partial \mu(\hat{\theta})^\prime}{\partial \hat{\theta}} \frac{\partial \mu(\hat{\theta})}{\partial \hat{\theta}}.
\]

Assume there exists an open neighborhood of \( \theta_0 \) in which \( \tilde{G}(\hat{\theta}) \) has a constant rank.

**Remark 4**  \( \tilde{G}(\hat{\theta}) \) is a \((p + q)\)-by-\((p + q)\) matrix. The first term is a bordered matrix, consisting of \( G(\theta) \) with \( p \) rows and columns of zeros appended to it. Both terms are positive semidefinite, hence taking the sum cannot decrease the rank. Also, note that the \((j,k)\)-th element of \( \tilde{G}(\hat{\theta}) \) is given by

\[
\tilde{G}_{jk}(\hat{\theta}) = \int_{-\pi}^{\pi} \text{tr} \left\{ \frac{\partial f_{\hat{\theta}}(\omega)}{\partial \hat{\theta}_j} \frac{\partial f_{\hat{\theta}}(\omega)}{\partial \hat{\theta}_k} \right\} d\omega + \frac{\partial \mu(\hat{\theta})^\prime}{\partial \hat{\theta}_j} \frac{\partial \mu(\hat{\theta})}{\partial \hat{\theta}_k}.
\]

**Theorem 2**  Let Assumptions 1-4 hold. Then, \( \hat{\theta} \) is locally identifiable from the first and second order properties of \( \{Y_t\} \) at a point \( \tilde{\theta}_0 \) if and only if \( \tilde{G}(\tilde{\theta}_0) \) is nonsingular.

Iskrev (2010) is an interesting article that also studied identification from the first and second order properties. He suggested to identify the parameters from the mean and the first \( T \) auto-covariances of the observables. Because his result (Theorem 2) assumes \( T \) is finite, the resulting conditions are sufficient but not necessary, as pointed out by Komunjer and Ng (2009).

Theorems 1 and 2 can be further extended in various directions. In what follows we discuss four such extensions.
In practice, one may be interested in estimating the model using only a subset of frequencies, see Hansen and Sargent (1993), Diebold, Ohanian and Berkowitz (1998) and Berkowitz (2001). Such consideration can arise due to concerns about unmodeled seasonality or measurement errors. We now present an identification result for this case. Let $W(\omega)$ denote an indicator function defined on $[-\pi, \pi]$ that is symmetric around zero and equal to one over a finite number of closed intervals. Extend the definition of $W(\omega)$ to $\omega \in [\pi, 2\pi]$ by using $W(\omega) = W(2\pi - \omega)$.$^1$ Define the following two matrices

$$
G^W(\theta) = \left\{ \int_{-\pi}^\pi W(\omega) \left( \frac{\partial \text{vec}(f_\theta(\omega)')} {\partial \theta'} \right)' \left( \frac{\partial \text{vec}(f_\theta(\omega))} {\partial \theta} \right) d\omega \right\}, \\
\bar{G}^W(\tilde{\theta}) = \left\{ \int_{-\pi}^\pi W(\omega) \left( \frac{\partial \text{vec}(f_\tilde{\theta}(\omega)')} {\partial \tilde{\theta}'} \right)' \left( \frac{\partial \text{vec}(f_\tilde{\theta}(\omega))} {\partial \tilde{\theta}} \right) d\omega \right\} + \frac{\partial \mu(\theta)'} {\partial \theta} \frac{\partial \mu(\tilde{\theta})} {\partial \tilde{\theta}}
$$

**Corollary 1 (Identification from a subset of frequencies)**

1. Let Assumptions 1-3 hold, but with $G(\theta)$ replaced by $G^W(\theta)$. Then, $\theta$ is locally identifiable from the second order properties of $\{Y_t\}$ through the frequencies specified by $W(\omega)$ at a point $\theta_0$ if and only if $G^W(\theta_0)$ is nonsingular.

2. Let Assumptions 1-4 hold, but with $\bar{G}(\tilde{\theta})$ replaced by $\bar{G}^W(\tilde{\theta})$. Then, $\tilde{\theta}$ is locally identifiable from the first and second order properties of $\{Y_t\}$ through the frequencies specified by $W(\omega)$ at a point $\theta_0$ if and only if $\bar{G}^W(\theta_0)$ is nonsingular.

The proof is the same as for Theorems 1 and 2 because $W(\omega)$ is a nonnegative real valued function, therefore it is omitted.

The second extension concerns the identification of a subset of parameters without making identification statements about the rest (partial identification). Specifically, Let $\theta^s$ be a subset of parameters from $\theta$. We say it is locally identified from the second order properties of $\{Y_t\}$ if there exists an open neighborhood of $\theta_0$ in which $f_{\theta_1}(\omega) = f_{\theta_0}(\omega)$, for all $\omega \in [-\pi, \pi]$, implies $\theta^s_0 = \theta^s_1$. Note that, as in Rothenberg (1971, footnote p.586), the definition does not exclude there being two points satisfying $f_{\theta_1}(\omega) = f_{\theta_0}(\omega)$ and having the $\theta^s$ arbitrarily close in the sense of $||\theta^s_0 - \theta^s_1|| / ||\theta_0 - \theta_1||$ being arbitrarily small. Analogously, we can define the identification of a subset of $\tilde{\theta}$, say $\tilde{\theta}^s$, based on the first and second order properties. The following result is a consequence of Theorem 8 in Rothenberg (1971), which can be traced back to Wald (1950) and Fisher (1966).

---

$^1$This extension is needed for FQML estimation since the objective function involves summation over $\omega_j = 2\pi/T, ..., 2\pi(T-1)/T$, see (14).
Corollary 2 (Partial identification)

1. Let Assumptions 1-3 hold. Then, $\theta^s$ is locally identifiable from the second order properties of $\{Y_t\}$ at a point $\theta_0^s$ if and only if $G(\theta_0)$ and

$$G^a(\theta_0) = \begin{bmatrix} G(\theta_0) \\ \partial\theta_0^s/\partial\theta' \end{bmatrix}$$

have the same rank.

2. Let Assumptions 1-4 hold. Then, $\tilde{\theta}^s$ is locally identifiable from the first and second order properties of $\{Y_t\}$ at a point $\tilde{\theta}_0^s$ if and only if $\tilde{G}(\tilde{\theta}_0)$ and

$$\tilde{G}^a(\tilde{\theta}_0) = \begin{bmatrix} \tilde{G}(\tilde{\theta}_0) \\ \partial\tilde{\theta}_0^s/\partial\tilde{\theta}' \end{bmatrix}$$

have the same rank.

The proof is in the Appendix. Furthermore, one may be interested in studying the identification of a subset of parameters while keeping the values of the others fixed at $\theta_0$ (conditional identification). The result for this extension is formally stated below.

Corollary 3 (Conditional Identification).

1. Let Assumptions 1-3 hold. Then, a subvector of $\theta$, $\theta^s$, is conditionally locally identifiable from the second order properties of $\{Y_t\}$ at a point $\theta_0$ if and only if

$$G(\theta_0)^s = \int_{-\pi}^{\pi} \left( \frac{\partial \text{vec}(f_{\theta_0}(\omega))}{\partial \theta^s} \right)' \left( \frac{\partial \text{vec}(f_{\theta_0}(\omega))}{\partial \theta^s} \right) d\omega$$

is nonsingular.

2. Let Assumptions 1-4 hold. Then, a subvector of $\tilde{\theta}$, $\tilde{\theta}^s$, is conditionally locally identifiable from the first and second order properties of $\{Y_t\}$ at a point $\tilde{\theta}_0$ if and only if

$$\tilde{G}(\tilde{\theta}_0)^s = \int_{-\pi}^{\pi} \left( \frac{\partial \text{vec}(f_{\tilde{\theta}_0}(\omega))}{\partial \theta^s} \right)' \left( \frac{\partial \text{vec}(f_{\tilde{\theta}_0}(\omega))}{\partial \theta^s} \right) d\omega + \frac{\partial \mu(\tilde{\theta}_0)^s}{\partial \theta^s} \frac{\partial \mu(\tilde{\theta}_0)^s}{\partial \theta^s}$$

is nonsingular.
The proof is the same as for Theorems 1 and 2 because $G(\theta_0)^s$ and $\tilde{G}(\tilde{\theta}_0)^s$ have the same structure as $G(\theta_0)$ and $\tilde{G}(\tilde{\theta}_0)$ but with derivatives taken with respect to a subset of parameters. Therefore the detail is omitted. Corollaries 2 and 3 can be applied jointly in practice and convey complementary information.

Next, we consider identification under general constraints on the parameters. One potential example is that shocks to monetary variables have no long term effect on real variables, which can be formulated as a set of restrictions on the spectral density at frequency zero.

**Corollary 4 (Identification under general constraints)**

1. Let Assumptions 1-3 hold. Suppose $\theta_0$ satisfies $\psi(\theta_0) = 0$ with $\psi(\theta)$ a $k$-by-1 constraint vector continuously differentiable in $\theta$. Define the Jacobian matrix $\Psi(\theta)$ with the $(j,l)$-th element given by

   $$
   \Psi_{jl}(\theta) = \partial \psi_j(\theta) / \partial \theta_l.
   $$

   Suppose $\theta_0$ is a regular point of both $G(\theta)$ and $\Psi(\theta)$. Then, $\theta$ satisfying $\psi(\theta) = 0$ is locally identified from the second order properties of $\{Y_t\}$ at a point $\theta_0$ if and only if

   $$
   \begin{bmatrix}
   G(\theta_0) \\
   \Psi(\theta_0)
   \end{bmatrix}
   $$

   has full column rank equal to $q$.

2. Let Assumptions 1-4 hold and other conditions stated in part 1 of this corollary hold with $\theta$ replaced by $\bar{\theta}$. Then, $\bar{\theta}$ satisfying $\psi(\bar{\theta}) = 0$ is locally identified from the first and second order properties of $\{Y_t\}$ at a point $\theta_0$ if and only if

   $$
   \begin{bmatrix}
   G(\bar{\theta}_0) \\
   \Psi(\bar{\theta}_0)
   \end{bmatrix}
   $$

   has rank $(q + p)$.

Note that Corollary 4 can also be used to study conditional identification, because the latter is a special case of simple linear restrictions. However, Corollary 3 is simpler to apply, especially if the dimension of $\theta^s$ is much smaller compared to that of $\theta$. Clearly, Corollaries 2-4 can be applied in conjunction with Corollary 1 to study identification through a subset of frequencies. Note that the last three extensions are motivated by Komunjer and Ng (2009). Their conditions require $n_Y \geq n_\epsilon$ and are only for dynamic parameters.
3.1 Tracing out non-identification curves

In this section the discussion will focus on $\theta$ because for $\theta$ the procedure works in the same way. Suppose Theorem 1 or Corollary 1 shows that $\theta$ is locally unidentifiable.

First, consider the simple case where $G(\theta_0)$ has only one zero eigenvalue. Let $c(\theta_0)$ be a corresponding real eigenvector satisfying $\|c(\theta_0)\| = 1$. Then, $c(\theta_0)$ is unique up to multiplication by $-1$, and thus can be made unique by restricting its first nonzero element to be positive. This restriction is imposed in the subsequent analysis. Let $\delta(\theta_0)$ be an open neighborhood of $\theta_0$. Under Assumptions 1 to 3, $G(\theta)$ is continuous and has only 1 zero eigenvalue in $\delta(\theta_0)$, while $c(\theta)$ is continuous in $\delta(\theta_0)$. As in Rothenberg (1971), define a curve $\chi$ using the function $\theta(v)$ which solves the differential equation

$$\frac{\partial \theta(v)}{\partial v} = c(\theta),$$

$$\theta(0) = \theta_0,$$

where $v$ is a scalar that varies in a neighborhood of 0 such that $\theta(v) \in \delta(\theta_0)$. Then, along $\chi$, $\theta$ is not identified at $\theta_0$ because

$$\frac{\partial \text{vec} \left( f_{\theta(v)}(\omega) \right)}{\partial v} = \frac{\partial \text{vec} \left( f_{\theta(v)}(\omega) \right)}{\partial \theta(v)^t} c(\theta) = 0$$

for all $\omega \in [-\pi, \pi]$, where the last equality uses Assumption 3 and the fact that $c(\theta)$ is the eigenvector corresponding to the zero eigenvalue (c.f. (A.3) in the Appendix). We call $\chi$ the non-identification curve.

Clearly, this curve is continuous in $v$. It is also locally unique, in the sense that there does not exist another continuous curve containing $\theta_0$ and satisfying $f_{\theta_1}(\omega) = f_{\theta_0}(\omega)$ for all $\omega \in [-\pi, \pi]$. We state this result as a Corollary:

**Corollary 5** Let Assumptions 1-3 hold and $\text{rank}(G(\theta_0)) = q - 1$. Then, in a small neighborhood of $\theta_0$, there exists precisely one curve passing through $\theta_0$ that satisfies $f_{\theta_1}(\omega) = f_{\theta_0}(\omega)$ for all $\omega \in [-\pi, \pi]$.

Corollary 5 is not a trivial result because it involves infinite dimensional maps. The key idea in the proof is to reduce the problem to a finite dimensional one by considering projections of $f_\theta(.)$ associated with finite partitions of $[-\pi, \pi]$. Then, a standard constant rank theorem can be applied. The details of the proof are in the Appendix.
The non-identification curve can be evaluated numerically in various ways. The simplest example is the Euler method. First, obtain $c(\theta_0)$ as described above. Then, compute recursively

$$
\theta(v_{j+1}) \approx \theta(v_j) + c(\theta(v_j))(v_{j+1} - v_j), \quad v_{j+1} \geq v_j \geq 0, \quad j = 0, 1, \ldots
$$

where $|v_{j+1} - v_j|$ is the step size, which can be set to some small constant, say $h$. The associated approximation error in each step is of order $O(h^2)$ if $\theta(v)$ has bounded first and second derivatives. Therefore, the cumulative error over a finite interval is $O(h)$. It is important to note that because $\delta(\theta_0)$ is usually unknown, so is the domain of the curve. However, this is not a problem in practice, because we can first obtain a curve over a wide support, then resolve the model and compute the spectral density using points on this curve. The curve can then be truncated to exclude points yielding a different $f_\theta(\omega)$ from $f_{\theta_0}(\omega)$.

Next, consider the case where $G(\theta_0)$ has multiple zero eigenvalues. Then, in general, there exists an infinite number of curves satisfying (9), because any linear combination of the eigenvectors points to a direction of non-identification. It is not useful to try reporting all such curves. To see this, suppose $\theta_0 = (\theta_0^1, \theta_0^2)'$ and that changing $\theta^1$ along a certain curve $\chi_1$ while keeping $\theta^2$ fixed at $\theta_0^2$ yields identical spectral densities. Also suppose the same property holds when we vary $\theta^2$ and fix $\theta^1$ at $\theta_0^1$, yielding a curve $\chi_2$. Suppose the rank of $G(\theta)$ stays constant in a local neighborhood of $\theta_0$. Then, changing $\theta^1$ and $\theta^2$ simultaneously can also generate new curves and there are infinitely many of them. In this example, $\chi_1$ and $\chi_2$ contain essentially all the information, as the rest of the curves are derived from them, and thus it suffices to report only two of them. Motivated by the above observation, we propose a simple four-step procedure that delivers a finite number of non-identification curves. The key idea underlying this procedure is to distinguish between separate sources of non-identification by using Corollary 3.

- **Step 1.** Apply Theorem 1 to see whether the full parameter vector is locally identified. Proceed to Step 2 if lack of identification is detected.

- **Step 2.** Consider all one-parameter subsets of $\theta_0$ and apply Corollary 3 to each of them. If a zero eigenvalue is reported, then it implies that the corresponding parameter is not locally conditionally identified and applying the procedure outlined above yields a non-identification curve (changing only this element and fixing others at $\theta_0$). Repeating this to all such subsets, we obtain a finite number of curves, with each curve being a scalar valued function of $v$.

- **Step 3.** Increase the number of parameters in the considered subsets of $\theta_0$ by 1. Single out
the subsets with the following two properties: (1) it does not include the subset detected in
previous steps as a proper subset, and (2) when applying Corollary 3, it reports only one
zero eigenvalue. Repeat the procedure outlined above for all such subsets to obtain non-
identification curves. Note that if the subset has $k$ elements, then the associated curve is a
$k$-by-1 vector valued function of $v$.

- **Step 4.** Continue Step 3 until all subsets are considered. Solve the model using parameter
values from the curves to determine the appropriate domain for $v$. Truncate the curves
accordingly.

Step 2 returns non-identification curves resulting from changing only one element in the para-
meter vector. In Step 3, the number of elements is increased sequentially. For each iteration, the
algorithm first singles out parameter subvectors whose elements are not separately identified. Then,
only subvectors satisfying the two properties outlined in Step 3 are further considered. The first
property is to rule out redundancy, because, if a $k$-element subset constitutes a non-identifi-
cation curve, including any additional element (fixing its value or varying it if itself is not conditional
identifiable) will by definition constitute another such curve, but conveying no additional infor-
mation. The second property serves the same purpose. Because if some subvector yields a $G(\theta)^s$
with multiple zero eigenvalues, then it must be a union of subvectors identified in previous steps
and containing fewer elements. To see that this is necessarily the case, suppose that for a given
subvector, two zero eigenvalues are reported. Then, there exists a linear combination of the two
corresponding eigenvectors that makes the first element of the resulting vector zero. Similarly,
there is a combination that makes the second element zero. The two resulting vectors are valid
eigenvectors, however, they correspond to lower dimensional subvectors of $\theta$. Now, apply Corollary
3 to these two subvectors. If single zero eigenvalues are reported, then it implies that they have
already been considered in the previous steps. Otherwise, the dimension of the subvectors can be
further reduced by using the same argument, eventually leading to the conclusion that they have
been previously considered. The general case with more than two zero eigenvalues can be analyzed
similarly.

This procedure will deliver a finite number of curves with the following two features. First, the
curves are minimal in the sense that, for each curve, all elements in the corresponding subvector
have to change to generate non-identification. Fixing the value of any element will shrink the
 corresponding curve to a single point. Second, the curves are sufficient in the sense that, for any
subvector that can generate a non-identification curve passing through $\theta_0$, it or one of its subsets
are already included. Finally, the procedure is simple to implement because it mainly involves
repeated applications of Corollary 3. This simplicity is achieved because we start with the lowest dimension, thus there is no need to directly handle the situation with multiple zero eigenvalues.

It should also be noted that, apart from evaluating the non-identification curves, the procedure is not computationally demanding. Once \( G(\theta) \) is computed in Step 1, the \( G(\theta)^s \) for any subvector considered can be obtained by simply picking out relevant elements of \( G(\theta) \) (c.f. Remark 3). Specifically, suppose we are interested in a particular \( k \)-element subvector of \( \theta \). If we number parameters inside \( \Phi \), and let \( \Phi = \{ \phi_1, \phi_2, ..., \phi_k \} \), then the \((i,j)\)-th element of \( G(\theta)^s \) is given by

\[
G(\theta)^s_{i,j} = G(\theta)_{\phi_i, \phi_j}, \quad i = 1, 2, ..., k; \quad j = 1, 2, ..., k
\]  

(11)

Also note that in case of Theorem 2, the same logic applies to the term \( \partial \mu(\theta_0)'/\partial \theta^s \) \( \partial \mu(\theta_0)/\partial \theta^s' \), i.e., having computed it once, one can repeatedly apply Corollary 3 by selecting relevant elements from it and \( \overline{G(\theta)^s} \) in the same fashion as in (11).

### 3.2 An illustrative example

To provide a frame of reference, we consider a DSGE model from An and Schorfheide (2007) whose identification was also studied by Komunjer and Ng (2009). The log-linearized solutions are given by:

\[
y_t = E_t y_{t+1} + g_t - E_t g_{t+1} - \frac{1}{\tau} (r_t - E_t \pi_{t+1} - E_t z_{t+1})
\]

\[
\pi_t = \beta E_t \pi_{t+1} + \frac{\tau (1 - \nu)}{\nu \pi^2 \bar{\phi}} (y_t - g_t)
\]

\[
c_t = y_t - g_t
\]

\[
r_t = \rho_r r_{t-1} + (1 - \rho_r) \psi_1 \pi_t + (1 - \rho_r) \psi_2 (y_t - g_t) + e_{rt}
\]

\[
g_t = \rho_g g_{t-1} + e_{gt}
\]

\[
z_t = \rho_z z_{t-1} + e_{zt}
\]

where \( e_{rt} = \epsilon_{rt}, \epsilon_{rt} \sim WN(0, \sigma_r^2) \), \( e_{gt} \sim WN(0, \sigma_g^2) \), and \( e_{zt} \sim WN(0, \sigma_z^2) \) are mutually uncorrelated shocks, \( \bar{\pi} \) is the steady state inflation rate. The vector of parameters to be identified is

\[
\theta = (\tau, \beta, \nu, \phi, \pi^2, \psi_1, \psi_2, \rho_r, \rho_g, \rho_z, \sigma_r^2, \sigma_g^2, \sigma_z^2).
\]

We can write the model as in (5) with

\[
S_t = (z_t, g_t, r_t, y_t, \pi_t, c_t, E_t(\pi_{t+1}), E_t(y_{t+1}))' \quad (12)
\]

16
The exact formulations of the matrices $\Gamma_0, \Gamma_1, \Psi$ and $\Pi$ are omitted here, as they directly follow from rearranging the model equations and corresponding program codes are available from An and Schorfheide (2007) (for GAUSS) and the web appendix of Komunjer and Ng (2009) (for MATLAB). We use the algorithm provided by Sims (2002) to obtain the model solution numerically in the form of (6), specifically

$$S_t = \Theta_1 S_{t-1} + \Theta_0 \epsilon_t,$$

where $\Theta_1$ and $\Theta_0$ are functions of $\theta$. The spectral density, as noted before, can then be computed using (4) with

$$H(L; \theta) = A(L)(I - \Theta_1 L)^{-1}\Theta_0.$$

Given the $S_t$ in (12) and $Y_t^d = (r_{t-1}, y_t, \pi_t, c_t)'$, the matrix $A(L)$ is given by

$$
\begin{pmatrix}
0 & 0 & L & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0
\end{pmatrix}.
$$

We choose

$$\theta_0 = (2, 0.9975, 0.1, 53.6797, 1.008^2, 1.5, 0.125, 0.75, 0.95, 0.9, 0.4, 0.36, 0.9),$$

as in Table 3 of An and Schorfheide (2007). The integral in $G(\theta_0)$ is approximated numerically by averaging over 10,000 Fourier frequencies from $-4999\pi/5000$ to $4999\pi/5000$ and multiplying by $2\pi$. The step size for the numerical differentiation is set to $10^{-7} \times \theta_0$. We obtain $\text{rank}(G(\theta_0)) = 10$. Because $q = 13$, this means that the entire parameter vector cannot be identified from the spectrum. This conclusion is in accordance with the findings of Komunjer and Ng (2009). The result was robust to varying the number of frequencies between 5,000 and 10,000 and the step size between $10^{-5} \times \theta_0$ and $10^{-7} \times \theta_0$. It should be pointed out that the choice of the step size for the numerical differentiation algorithm should not be too small (say $10^{-9} \times \theta_0$), otherwise the numerical error from the model solution algorithm will dominate and the conclusions will be misleading.

Since the model is not identified, we can follow the procedure outlined in Section 3.1 to pinpoint the sources of non-identification. In Step 2 we apply Corollary 3 to all one-element subsets of $\theta$ which, as noted above in (11), simply amounts to checking whether any diagonal elements of $G(\theta_0)$

$\footnote{Note that we scale the values for the variances ($\sigma_r^2, \sigma_y^2, \sigma_z^2$) from An and Schorfheide (2007) by $10^5$. This scaling is merely for ensuring numerical stability and does not affect any of our conclusions.}$

$\footnote{A simple two-point method is used. In our experience, using higher-order methods did not change the conclusions, but resulted in noticeably higher computational cost.}$
are zero. None are found, hence we continue to Step 3 and consider all two-element subvectors of \( \theta \). We find three subvectors that yield \( G^*(\theta_0) \) with one zero eigenvalue: \((\nu, \phi)\), \((\nu, \pi^2)\) and \((\phi, \pi^2)\). This finding is very intuitive, since all of these parameters enter the slope of the Phillips curve equation and thus are not separately identifiable, as noted by An and Schorfheide (2007). We do not report the non-identification curves for these cases, as they are trivial and can be eliminated by re-parameterizing the model with \( \kappa \equiv \tau(1 - \nu)/ (\nu \pi^2 \phi) \) as a new parameter instead. However, highlighting them does play a useful part in illustrating our procedure at work.

Before we continue, we exclude all three-parameter subvectors that contain either of the three non-identification sets identified above as proper subsets. This shrinks the remaining parameter space to further simplify our task. Considering all remaining three-element subvectors of \( \theta \) yields no new non-identification sets. However, there is one four-element subvector which has one zero eigenvalue:

\[ (\psi_1, \psi_2, \rho_r, \sigma_r^2). \]

Interestingly, all of these parameters enter the Taylor rule equation in the model. This finding is at odds with Komunjer and Ng (2009), who, after re-parameterizing the model using \( \kappa \), concluded that all dynamic parameters were identified. It is, however, consistent with Cochrane’s (2007) critical review on lack of parameter identification in Taylor rule equations.

Having excluded all subvectors containing the non-identification parameter sets above and repeating Step 4 with more parameters, we do not find any more sources of non-identification in this model. This is further confirmed when we repeat the exercise by considering a reparameterization of the model with \( \kappa \) as defined above. \( \theta \) is still not identified, and \( G(\theta_0) \) has only one zero eigenvalue. This means that there is only one direction of non-identification, which turns out to be, not surprisingly, along the \((\psi_1, \psi_2, \rho_r, \sigma_r^2)\) subvector.

We then proceed to approximate the non-identification curve, consisting of combinations of \( \psi_1, \psi_2, \rho_r, \) and \( \sigma_r^2 \), using the Euler method with step size \( h = 10^{-5} \) in a small neighborhood around \( \theta_0 \). The result is presented in Figure 1. The figure shows the non-identification curve pertaining to each parameter. The initial value is at 0, and the curve is extended \( 10^5 \) steps in each way using (10), which are marked on the graph by bold and dotted lines respectively.

(Figure 1 here)

Note that all four parameters have to change simultaneously to generate non-identification. To give a quantitative idea of the changes involved, we also present a sample of parameter values from
various points on the curve in Table 1. Specifically, ten points were taken at intervals of $10^4$ steps from $\theta_0$ in the positive and negative direction.

(Table 1 here)

Of course it is necessary to verify that the points on the curve result in identical spectral densities. We do this by computing the $f_\theta(\omega)$ at half of the Fourier frequencies used in the computation of $G(\theta_0)$ (i.e., 5000 frequencies between 0 and $\pi$)\(^4\) for each point on the curve and then compare it to the ones computed at $\theta_0$. Due to numerical error involved in solving the model, the computation of the $G()$ matrix and the approximation method for the differential equation, small discrepancies between the spectra computed at $\theta_0$ and the points on the curve should be expected. We therefore consider two different measures of the discrepancies (let $f_{\theta hl}(\omega)$ denote the ($h,l$)-th element of the spectral density matrix with parameter $\theta$ and $\Omega$ be the set that includes the 5000 frequencies between 0 and $\pi$):

Maximum absolute deviation :  \[ \max_{\omega_j \in \Omega} |f_{\theta hl}(\omega_j) - f_{\theta_0 hl}(\omega_j)| \]

Maximum relative deviation :  \[ \max_{\omega_j \in \Omega} \frac{|f_{\theta hl}(\omega_j) - f_{\theta_0 hl}(\omega_j)|}{|f_{\theta_0 hl}(\omega_j)|} \]

To save space, we only report results for the points in Table 1 as the rest are very similar.

(Tables 2 and 3 here)

Both Tables 2 and 3 show that even the largest observed deviations are quite modest (recall that the Euler method involves a cumulative approximation error that is of the same order as the step size, in this case $10^{-5}$). This confirms that the spectral density is constant along the curve.

Thus far in this example we have considered identification from the second order properties of the process. We now extend the analysis to incorporate the steady state parameters. Consider the measurement equations from An and Schorfheide (2007) relating the output growth, inflation, and the interest rate observed quarterly to the steady states and elements of $S_t$:

\[
\begin{align*}
YGR_t &= \gamma^{(Q)} + 100(y_t - y_{t-1} + z_t) \\
INFL_t &= \pi^{(A)} + 400\pi_t \\
INT_t &= \pi^{(A)} + r^{(A)} + 4\gamma^{(Q)} + 400r_t
\end{align*}
\]

where

\[
\gamma^{(Q)} = 100(\gamma - 1), \quad \pi^{(A)} = 400(\bar{\pi} - 1), \quad r^{(A)} = 400\left(\frac{1}{\beta} - 1\right)
\]

\(^4\)There is no need to consider $\omega \in [-\pi, 0]$ because $f_\theta(\omega) = f_\theta(-\omega)$.  

19
and $\gamma$ is a constant in the technological shock equation. The parameter vector becomes

$$\tilde{\theta} = (\tau, \beta, \nu, \phi, \pi, \psi_1, \psi_2, \rho_r, \rho_g, \rho_z, \sigma_\tau^2, \sigma_g^2, \sigma_z^2, \gamma^{(Q)})$$

where $\gamma^{(Q)}$ is the only non-dynamic parameter. Thus, we have

$$\mu(\tilde{\theta}) = \begin{pmatrix} \gamma^{(Q)} \\ 400(\pi - 1) \\ 400(\pi - 1) + 400(\frac{1}{\beta} - 1) + 4\gamma^{(Q)} \end{pmatrix}$$

and the $A(L)$ matrix in this case is

$$\begin{pmatrix} 100 & 0 & 0 & 100 - 100L & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 400 & 0 & 0 \\ 0 & 0 & 400 & 0 & 0 & 0 & 0 \end{pmatrix}.$$ 

Setting $\gamma^{(Q)} = 0.55$ as in An and Schorfheide (2007), we consider identification at

$$\tilde{\theta}_0 = (2, 0.9975, 0.1, 53.6797, 1.008, 1.5, 0.125, 0.75, 0.95, 0.9, 0.4, 0.36, 0.9, 0.55).$$

Note that $\mu(\tilde{\theta})$ can be easily differentiated analytically in this case.

Applying Theorem 2, we find $\text{rank}(C(\tilde{\theta}_0)) = 12$. Hence, $\tilde{\theta}_0$ is not identifiable from the first and second order properties of the observables either. After applying the procedure from Section 3.1, we find two subvectors, $(\nu, \phi)$ and $(\psi_1, \psi_2, \rho_r, \sigma_\tau^2)$, which account for non-identification. Intuitively, we no longer detect $(\nu, \pi)$ and $(\phi, \pi)$ as $\pi$ enters $\mu(\tilde{\theta})$ and hence is identifiable from the mean. Since the the two non-identification curves are exactly the same as in the dynamic parameter case, they are not reported here.

This example shows that the parameters in the model are not identifiable from the (first and) second order properties of observables. Our procedure attributes the source of non-identification to the parameters governing the Taylor rule equation. It is impossible to uniquely pin down $(\psi_1, \psi_2, \rho_r, \sigma_\tau^2)$ even having an infinite sample. This constitutes a serious concern for estimation in this and similar DSGE models.

**Remark 5** The results also have direct implications for Bayesian inference. Suppose we impose a tight prior on one of the four parameters, say $\psi_1$, while using flat priors on the rest. Then, the posterior distributions of $\psi_2, \rho_r$, and $\sigma_\tau^2$ will most often become concentrated due to their relation with $\psi_1$. Therefore, simply comparing the marginal priors and posteriors may give the false impression that the parameters are separately (or even strongly) identified and may overstate the informativeness of the data about the parameters.
4 FQML estimation

We first present a brief derivation of the FQML estimators and then study their asymptotic properties in both well and misspecified models. The subsequent analysis assumes that the system is nonsingular, i.e., $ny \leq ne$.

4.1 The estimators

For the sole purpose of deriving the quasi-likelihood function, assume that the process $\{Y_t\}$ is Gaussian. Let $\omega_j$ denote the Fourier frequencies, i.e., $\omega_j = 2\pi j/T \ (j = 1, 2, ..., T - 1)$. The discrete Fourier transforms are given by

$$w_T(\omega_j) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} Y_t \exp(-i\omega_j t), \ j = 1, 2, ..., T - 1.$$ 

Note that replacing $Y_t$ by $Y_t - \mu(\theta)$ does not affect the value of $w_T(\omega_j)$ at these frequencies. $w_T(\omega_j)$ have a complex valued multivariate normal distribution, and for large $T$ are approximately independent, each with the probability density function (see Hannan (1970), p. 223-225)

$$\frac{1}{\pi^{ny} \det(f_{\theta}(\omega_j))} \exp \left[-\text{tr} \left\{ f_{\theta}^{-1}(\omega_j) w_T(\omega_j) w_T^*(\omega_j) \right\} \right], \ j = 1, 2, ..., T - 1.$$ 

Therefore, an approximate log-likelihood function of $\theta$ based on observations $Y_1, ..., Y_T$ is given, up to constant multiplication, by

$$-\sum_{j=1}^{T-1} \left[ \log \det (f_{\theta}(\omega_j)) + \text{tr} \left\{ f_{\theta}^{-1}(\omega_j) I_T(\omega_j) \right\} \right], \quad (13)$$

where $I_T(\omega_j) = w_T(\omega_j) w_T^*(\omega_j)$ denotes the periodogram. Letting $W(\omega_j)$ be an indicator function as defined in the previous section, we consider the following generalized version of (13):

$$L_T(\theta) = -\sum_{j=1}^{T-1} W(\omega_j) \left[ \log \det (f_{\theta}(\omega_j)) + \text{tr} \left\{ f_{\theta}^{-1}(\omega_j) I_T(\omega_j) \right\} \right], \quad (14)$$

Then, the FQML estimator for $\theta$ is given by

$$\hat{\theta}_T = \arg \min_{\theta \in \Theta_0} L_T(\theta). \quad (15)$$

Thus, the above procedure allows us to estimate the dynamic parameters based on the second order properties of $\{Y_t\}$ without any reference to the steady state parameters. Compared with the time domain QML, the estimate here can be obtained without demeaning the data.
It is also simple to estimate both dynamic and steady state parameters jointly. Let
\[ w_{\bar{\theta},T}(0) = \frac{1}{\sqrt{2\pi T}} \sum_{t=1}^{T} Y_t - \mu(\bar{\theta}) \] and
\[ I_{\bar{\theta},T}(0) = w_{\bar{\theta},T}(0) w_{\bar{\theta},T}(0)' . \]
Noticing that \( w_{\bar{\theta},T}(0) \) has a multivariate normal distribution with asymptotic variance \( f_{\bar{\theta}}(0) \) and is asymptotically independent of \( w_T(\omega_j) \) for \( j = 1, 2, ..., T - 1 \), we arrive at the following approximate log-likelihood function of \( \bar{\theta} \):
\[ \bar{L}_T(\bar{\theta}) = L_T(\theta) - \left[ \log \det (f_{\theta}(0)) + \text{tr} \left\{ f_{\theta}^{-1}(0)I_{\bar{\theta},T}(0) \right\} \right] . \]
Then, the FQML estimator for \( \bar{\theta} \) is given by
\[ \hat{\theta}_T = \arg \min_{\bar{\theta} \in \Theta} \bar{L}_T(\bar{\theta}) . \] (16)

### 4.2 Asymptotic properties of the FQML estimators

The asymptotic properties of the estimator (15), with \( W(\omega_j) = 1 \) for all \( \omega_j \), have been studied under various data generating processes in the statistics literature, see, for example, Dunsmuir and Hannan (1976) and Dunsmuir (1979). The estimator (16) received less attention. One exception is Hansen and Sargent (1993), who formally established that \( T^{-1}\bar{L}_T(\bar{\theta}) \) converges to the same limit as the time domain Gaussian quasi-maximum likelihood function for \( \bar{\theta} \) uniformly in \( \bar{\theta} \in \Theta \). Their result allows for non-Gaussianity and model misspecification. This section can be viewed as a further development of their work in the following sense. First, we formally establish the relationship between the identification condition and the asymptotic properties of the estimator. Second, we explicitly derive the limiting distribution of the estimator, which is important for inference and model comparison.

We will gradually tighten the assumptions to obtain increasingly stronger results. To analyze the first issue, the following assumptions are imposed on the second and fourth order properties of the observed process \( \{Y_t\} \).

**Assumption 5.**

(i) \( \{Y_t\} \) is generated by
\[ Y_t = \mu(\theta_0) + Y^d_t(\theta_0) \]
with \( Y^d_t(\theta) \) satisfying (1). (ii) \( f_{\theta}(\omega) \) is positive definite with eigenvalues bounded away from 0 and \( \infty \) uniformly in \( \omega \) for all \( \theta \in \Theta_{\theta} \). The elements of \( \partial \vec{f}_{\theta}(\omega)/\partial \theta' \) are bounded away from \( \infty \) uniformly in \( \omega \) for all \( \theta \in \Theta_{\theta} \). The elements of \( f_{\theta}(\omega) \) belong to \( \text{Lip}(\beta) \) with respect to \( \omega \), the Lipschitz class of degree \( \beta, \beta > 1/2 \).
Assumption 6. $\epsilon_t$ is fourth-order stationary. Let $Q_{h,l,g,k}(j_1,j_2,j_3)$ be the joint cumulant of $\epsilon_{t+h}, \epsilon_{(t+j_1)l}, \epsilon_{(t+j_2)g}$ and $\epsilon_{(t+j_3)k}$. Assume $\sum_{j_1,j_2,j_3=-\infty}^{\infty} |Q_{h,l,g,k}(j_1,j_2,j_3)| < \infty$ for any $1 \leq h, l, g, k \leq n_e$.

The first part of Assumption 5 states that the model is correctly specified. This will be relaxed in Section 4.3. The second part strengthens the first condition in Assumption 3. It is satisfied by stationary finite order VARMA processes with finite error covariance matrices, which are the forms that the solutions to linearized DSGE models typically take. In Assumption 6, the summability of the fourth cumulant is weaker than the independence assumption, a sufficient condition is provided in Andrews (1991, Lemma 1).

We now define the concept of a locally unique maximizer.

Definition 3 Let $L(\phi)$ be some generic criterion function. We say $\varphi_0$ is a locally unique maximizer of $L(\varphi)$ if there exists an open neighborhood of $\varphi_0$ such that $L(\varphi) < L(\varphi_0)$ for all $\varphi$ different from $\varphi_0$ in this neighborhood.

Define the following quantities as the limits of $T^{-1}L_T(\theta)$ and $T^{-1}\bar{L}_T(\theta)$:

\[
L_\infty(\theta) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) \left[ \log \det(\phi(\omega)) + \text{tr} \left\{ \phi^{-1}(\omega) \phi_0(\omega) \right\} \right] d\omega, \\
\bar{L}_\infty(\bar{\theta}) = L_\infty(\theta) - \frac{1}{2\pi} (\mu(\bar{\theta}_0) - \mu(\bar{\theta}))' \phi^{-1}(0) (\mu(\bar{\theta}_0) - \mu(\bar{\theta})).
\]

Lemma 1 Let Assumptions 1-3, 5 and 6 hold. Then,

1. $T^{-1}L_T(\theta) \to^p L_\infty(\theta)$ uniformly over $\theta \in \Theta_\theta$.

2. $\theta_0$ is a locally unique maximizer of $L_\infty(\theta)$ if and only if it is locally identified. Furthermore, if $\theta_0$ is globally identified,\footnote{The parameter vector $\theta$ is said to be globally identifiable from the second order properties of $\{Y_t\}$ at a point $\theta_0$ if for any $\theta_1 \in \Theta_\theta$, $f_{\theta_1}(\omega) = f_{\theta_0}(\omega)$ for all $\omega \in [-\pi, \pi]$ implies $\theta_0 = \theta_1$.} then it is the unique maximizer of $L_\infty(\theta)$.

3. $\hat{\theta}_T \to^p \theta_0$ if one of the following two conditions is satisfied: 1) $\theta_0$ is globally identified, or 2) $\theta_0$ is locally identified and the maximization is carried over the corresponding small neighborhood of identification, say $\delta(\theta_0)$, instead of $\Theta_\theta$.

4. Let Assumptions 1-6 hold. Then, Properties 1-3 hold when $\theta, \theta_0, \hat{\theta}_T, L_T(\theta)$ and $L_\infty(\theta)$ are replaced by $\bar{\theta}, \bar{\theta}_0, \hat{\theta}_T, \bar{L}_T(\bar{\theta})$ and $\bar{L}_\infty(\bar{\theta})$, respectively.
The first result is essentially due to Lemma A.3.3(1) in Hosoya and Taniguchi (1982). Their result is pointwise in $\theta$ and is established with $W(\omega) = 1$. Our result strengthens theirs to uniform convergence, which is important for showing Property 3. The second result formally establishes the close link between the identification conditions and the asymptotic properties of the FQML estimator. The result is quite intuitive ex post, however, it is worth documenting given that the identification property is derived without explicitly referring to the likelihood function. The first two results lead directly to Property 3 by a uniform weak law of large numbers. Property 4 holds based on the same arguments.

To derive the limiting distribution of the estimators, the assumptions on $\{\epsilon_t\}$ need to be further strengthened.

**Assumption 7.** (i) $\{\epsilon_t\}$ is a vector of martingale difference sequences with respect to the $\sigma-$field generated by $\epsilon_t: s \leq t$. $E(\epsilon_ta\epsilon_tb|\mathcal{F}_{t-\tau}) = \xi_{ab}$, $E(\epsilon_ta\epsilon_tb\epsilon_tc|\mathcal{F}_{t-\tau}) = \zeta_{abcd}$ a.s. with $\Sigma_{aa} > 0$ and $\zeta_{aadd} > 0$ for all $1 \leq a, b, c, d \leq n_{\epsilon}$. (ii) Let $c(t,r) = \epsilon_t\epsilon_t'_{i+r} - E(\epsilon_t\epsilon_t'_{i+r})$. Assume $\lim_{T \to \infty} T^{-1} \sum_{t=0}^{L} \sum_{t=1}^{T} E \left[ c_{ab}(t,r)^2 \mathbf{1} \{c_{ab}(t,r)^2 > \varepsilon T\} \right] < \varepsilon$ holds for any $\varepsilon > 0$, $L < \infty$ and all $1 \leq a, b \leq n_{\epsilon}$.

Part (i) of Assumption 7 imposes restrictions on the conditional moments up to order 4. $\Sigma_{aa} > 0$ and $\zeta_{aadd} > 0$ are the usual positive variance conditions. It is essentially the same as Assumption C2.3 in Dunsmuir (1979). This part can be further relaxed to allow some conditional heteroskedasticity at the cost of some technical and notational complications, see Theorem 3.1 in Hosoya and Taniguchi (1982). Part (ii) is a Lindeberg-type condition. It ensures that the sample autocovariances $T^{-1/2} \sum_{t=1}^{T-r} c(t,r)$ ($r = 0, 1, ..., L$) satisfy a central limit theorem for any finite fixed $L$. It can be replaced by other sufficient conditions that serve the same purpose. The next result states the limiting distributions of $\hat{\theta}_T$ and $\tilde{\theta}_T$.

**Theorem 3** Suppose $\theta_0$ and $\tilde{\theta}_0$ are globally identified or the maximizations (15) and (16) are over convex compact sets in which they are locally identified and are interior points.

1. Let Assumptions 1-3 and 5-7 hold. Then,

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \to^d N(0, M^{-1}VM^{-1}),$$
where \( M \) and \( V \) are \( q \)-by-\( q \) matrices, with the \((j,l)\)-th element given by

\[
M_{jl} = \int_{-\pi}^{\pi} W(\omega) \text{tr} \left\{ f_{\theta_0}(\omega) \frac{\partial f_{\theta_0}^{-1}(\omega)}{\partial \theta_j} f_{\theta_0}(\omega) \frac{\partial f_{\theta_0}^{-1}(\omega)}{\partial \theta_l} \right\} d\omega
\]

\[
V_{jl} = 4\pi M_{jl} + \sum_{a,b,c,d=1}^{n_c} \kappa_{abcd} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H^*(\omega) \frac{\partial H^{-1}(\omega)}{\partial \theta_j} H(\omega) d\omega \right]_{ab} \times \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H^*(\omega) \frac{\partial H^{-1}(\omega)}{\partial \theta_l} H(\omega) d\omega \right]_{cd},
\]

where \([.,.]_{ab}\) denotes the \((a,b)\)-th element of the matrix, \( \kappa_{abcd} \) is the fourth cross cumulant of \( \epsilon_{ta}, \epsilon_{tb}, \epsilon_{tc} \) and \( \epsilon_{td} \), \( H(\omega) = H(\exp(-i\omega) ; \theta_0) = \sum_{j=0}^{\infty} h_j(\theta_0) \exp(-i\omega j) \) (c.f. (3)) and \( H^*(\omega) \) is its conjugate transpose.

2. Let Assumptions 1-7 hold. Then, \( \sqrt{T}(\hat{\theta}_T - \theta_0) \to^d N(0, \bar{M}^{-1}V \bar{M}^{-1}) \), where \( \bar{M} \) and \( \bar{V} \) are \((q+p)\)-by-\((q+p)\) matrices, with the \((j,l)\)-th element given by

\[
\bar{M}_{jl} = \int_{-\pi}^{\pi} W(\omega) \text{tr} \left\{ f_{\theta_0}(\omega) \frac{\partial f_{\theta_0}^{-1}(\omega)}{\partial \theta_j} f_{\theta_0}(\omega) \frac{\partial f_{\theta_0}^{-1}(\omega)}{\partial \theta_l} \right\} d\omega + 2 \frac{\partial \mu(\bar{\theta}_0)'}{\partial \theta_j} f_{\theta_0}^{-1}(0) \frac{\partial \mu(\bar{\theta}_0)}{\partial \theta_l}
\]

\[
\bar{V}_{jl} = 4\pi \bar{M}_{jl} + \sum_{a,b,c,d=1}^{n_c} \kappa_{abcd} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H^*(\omega) \frac{\partial H^{-1}(\omega)}{\partial \theta_j} H(\omega) d\omega \right]_{ab} \times \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H^*(\omega) \frac{\partial H^{-1}(\omega)}{\partial \theta_l} H(\omega) d\omega \right]_{cd} + A_{jl} + A_{lj}
\]

with \( A_{jl} = 2 \sum_{a,b,c=1}^{n_c} \xi_{abc} \left[ \int_{-\pi}^{\pi} W(\omega) \left[ H^*(\omega) \frac{\partial H^{-1}(\omega)}{\partial \theta_j} H(\omega) \right]_{ab} d\omega \right] \times \left[ \frac{\partial H(\omega)}{\partial \theta_l} f_{\theta_0}^{-1}(0) H(0) \right]_{ce} \) and \( \xi_{abc} = E(\epsilon_{ta} \epsilon_{tb} \epsilon_{tc}) \).

When \( W(\omega) = 1 \), the first result reduces to Corollary 2.2 in Dunsmuir (1979, p. 497) and Proposition 3.1 in Hosoya and Taniguchi (1982). The limiting distribution depends on the fourth order properties of the process. Intuitively, this is because in DSGE models the same set of parameters affects both the conditional mean and the conditional covariance of the process. Technically, the term \( h_0(\theta) \) is in general not an identity matrix, but rather depends on unknown parameters. This causes the second term in \( V_{jl} \) to be in general nonzero. However, in the important special case where \( \epsilon_t \) are Gaussian with diagonal covariance matrix, \( \kappa_{abcd} = 0 \) and the limiting distribution depends only on the second order property of the process. This holds for different specifications of \( W(\omega) \).

The second result is new in the literature even for the case with \( W(\omega) = 1 \). The inclusion of the steady state parameter makes the limiting distribution dependent on the third order properties of
Y_t, namely \( \xi_{abc} \). Again, in the important special case with Gaussianity and a diagonal covariance matrix, \( \xi_{abc} = 0 \) and only the second order property matters.

To construct the confidence interval, \( f_{\theta_0}^*(\omega) \), \( H(\omega) \) and \( H^*(\omega) \) \((\omega \in [-\pi, \pi])\) can be consistently estimated by replacing \( \theta_0 \) and \( \tilde{\theta}_0 \) with \( \hat{\theta}_T \) and \( \tilde{\theta}_T \) and applying (2) and (4). The derivatives and the integrals can be evaluated numerically. The cumulants \( \xi_{abc} \) and \( \kappa_{abcd} \) can be replaced by their sample counterparts.

4.3 Misspecified models

We consider the interpretation of the parameter estimates when the DSGE models are viewed as approximations. The next assumption allows the true data generating process to be different from that implied by the DSGE solution.

**Assumption MI.** The observations \( \{Y_t\}_{t=1}^T \) follow a covariance stationary process given by \( Y_t = \mu_0 + \sum_{j=0}^{\infty} h_{0j} \varepsilon_{t-j} \), whose mean \( \mu_0 \) and spectral density \( f_0(\omega) \) are possibly different from \( \mu(\tilde{\theta}_0) \) and \( f_{\theta_0}(\omega) \). Also, \( Y_t \) satisfies Assumptions 5(ii) with \( f_{\theta}(\omega) \) replaced by \( f_0(\omega) \) and Assumptions 6 and 7 with \( \varepsilon_t \) replaced by \( \varepsilon_t \).

Suppose the estimates \( \hat{\theta}_T \) and \( \tilde{\theta}_T \) are constructed in the same way as before and define the following pseudo-true values

\[
\theta_0^m = \arg \min_{\theta \in \Theta_\theta} L_\infty^m (\theta) \quad \text{and} \quad \tilde{\theta}_0^m = \arg \min_{\tilde{\theta} \in \Theta} L_\infty^m (\tilde{\theta}),
\]

where

\[
L_\infty^m (\theta) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) \left[ \log \det (f_{\theta}(\omega)) + \text{tr} \left\{ f_{\theta}^{-1}(\omega) f_0(\omega) \right\} \right] d\omega,
\]

\[
L_\infty^m (\tilde{\theta}) = L_\infty^m (\theta) - \frac{1}{2\pi} \left( \mu_0 - \mu(\tilde{\theta}) \right)' f_{\theta}^{-1}(0) \left( \mu_0 - \mu(\tilde{\theta}) \right).
\]

Suppose \( \theta_0^m \) and \( \tilde{\theta}_0^m \) lie in the interior of \( \Theta_\theta \) and \( \Theta \).

**Corollary 6** Suppose \( \theta_0^m \) and \( \tilde{\theta}_0^m \) are globally identified or the maximizations (15) and (16) are over convex compact sets in which they are locally identified and are interior points. Let Assumption MI hold.

1. Assume the DSGE solution \( Y_t^d(\theta) \) satisfies Assumptions 1-3 and 5(ii). Then,

\[
\sqrt{T} (\hat{\theta}_T - \theta_0^m) \overset{d}{\rightarrow} N(0, \Omega^{-1} \Pi \Omega^{-1})
\]
with

\[ \Omega = \int_{-\pi}^{\pi} W(\omega) \left[ \frac{\partial^2}{\partial \theta \partial \theta'} \log \det (f_{\theta_0^n}(\omega)) + \frac{\partial^2}{\partial \theta \partial \theta'} \text{tr} \left\{ f_{\theta_0^n}^{-1}(\omega) f_0(\omega) \right\} \right] d\omega \]

\[ \Pi_{jl} = 4\pi \int_{-\pi}^{\pi} W(\omega) \text{tr} \left\{ f_0(\omega) \frac{\partial f_{\theta_0^n}^{-1}(\omega)}{\partial \theta_j} f_0(\omega) \frac{\partial f_{\theta_0^n}^{-1}(\omega)}{\partial \theta_l} \right\} d\omega 
+ \sum_{a,b,c,d=1} \kappa_{abcd} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H_0^*(\omega) \frac{\partial f_{\theta_0^n}^{-1}(\omega)}{\partial \theta_j} H_0(\omega) dw \right]_{ab} 
\times \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H_0^*(\omega) \frac{\partial f_{\theta_0^n}^{-1}(\omega)}{\partial \theta_l} H_0(\omega) dw \right]_{cd}, \]

where \( \kappa_{abcd} \) is the fourth cross cumulant of \( \varepsilon_{ta}, \varepsilon_{tb}, \varepsilon_{tc} \) and \( \varepsilon_{td} \), \( H_0(\omega) = \sum_{j=0}^\infty h_{0j} \exp(-i\omega j) \).

2. Assume the DSGE solution is given by \( \mu(\theta) + Y_t^d(\theta) \) and satisfies Assumptions 1-4 and 5(ii). Then, \( \sqrt{\bar{T}}(\hat{\theta}_T - \theta_0^n) \rightarrow^{d} N(0, \Omega^{-1}\Pi^{-1}) \) with

\[ \Omega = \int_{-\pi}^{\pi} W(\omega) \left[ \frac{\partial^2}{\partial \theta \partial \theta'} \log \det (f_{\theta_0^n}(\omega)) + \frac{\partial^2}{\partial \theta \partial \theta'} \text{tr} \left\{ f_{\theta_0^n}^{-1}(\omega) f_0(\omega) \right\} \right] d\omega 
+ 2 \frac{\partial \mu(\theta_0^n)}{\partial \theta} f_{\theta_0^n}^{-1}(0) \frac{\partial \mu(\theta_0^n)}{\partial \theta'} 
+ \sum_{a,b,c,d=1} \kappa_{abcd} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H_0^*(\omega) \frac{\partial f_{\theta_0^n}^{-1}(\omega)}{\partial \theta_j} H_0(\omega) dw \right]_{ab} 
\times \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H_0^*(\omega) \frac{\partial f_{\theta_0^n}^{-1}(\omega)}{\partial \theta_l} H_0(\omega) dw \right]_{cd}, \]

\[ \Pi_{jl} = 4\pi \left\{ \int_{-\pi}^{\pi} W(\omega) \text{tr} \left\{ f_0(\omega) \frac{\partial f_{\theta_0^n}^{-1}(\omega)}{\partial \theta_j} f_0(\omega) \frac{\partial f_{\theta_0^n}^{-1}(\omega)}{\partial \theta_l} \right\} d\omega 
+ 2 \frac{\partial \mu(\theta_0^n)}{\partial \theta_j} f_{\theta_0^n}^{-1}(0) \frac{\partial \mu(\theta_0^n)}{\partial \theta_l} \right\} 
+ \sum_{a,b,c,d=1} \kappa_{abcd} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H_0^*(\omega) \frac{\partial f_{\theta_0^n}^{-1}(\omega)}{\partial \theta_j} H_0(\omega) dw \right]_{ab} \right. 
\times \left. \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H_0^*(\omega) \frac{\partial f_{\theta_0^n}^{-1}(\omega)}{\partial \theta_l} H_0(\omega) dw \right]_{cd} \right), \]

with \( A_{jl} = 2 \sum_{a,b,c,d=1} \xi_{abc} \left\{ \int_{-\pi}^{\pi} W(\omega) \left[ H_0^*(\omega) \frac{\partial f_{\theta_0^n}^{-1}(\omega)}{\partial \theta_j} H_0(\omega) \right]_{ab} \right\} d\omega \times \left[ \frac{\partial \mu(\theta_0^n)}{\partial \theta_j} f_{\theta_0^n}^{-1}(0) H_0(0) \right]_{c} \]
and \( \xi_{abc} = E(\varepsilon_{ta} \varepsilon_{tb} \varepsilon_{tc}) \).

Misspecification in general affects both the mean and the variance of the estimate. Note that when only estimating the dynamic parameters, misspecifying \( \mu(\theta) \) has no effect on the estimate \( \hat{\theta}_T \).

5 Quasi-Bayesian inference

This section extends the above framework to incorporate prior distributions on the DSGE parameters. It also discusses a computationally attractive procedure to obtain parameter estimates. The analysis is motivated by Chernozhukov and Hong (2003). We will focus on \( \theta_0 \) because the procedure is identical for \( \theta_0 \).
Consider the function
\[ p_T(\theta) = \frac{\pi(\theta) \exp(L_T(\theta))}{\int_\Theta \pi(\theta) \exp(L_T(\theta)) d\theta}, \] where \( L_T(\theta) \) is the same as in (14), and \( \pi(\theta) \) can be a proper prior probability density or, more generally, a weight function that is strictly positive and continuous over \( \Theta_\theta \). Because \( \exp(L_T(\theta)) \) is a more general criterion function than the likelihood, \( p_T(\theta) \) is in general not a true posterior in the Bayesian sense. However, it is a proper distribution density over the parameters of interest, and is termed quasi-posterior in Chernozhukov and Hong (2003).

The estimate for \( \theta_0 \) can be taken to be the quasi-posterior mean:
\[ \hat{\theta}_T = \int_{\Theta_\theta} \theta p_T(\theta) d\theta. \]

To compute the estimator, we can use Markov chain Monte Carlo (MCMC) methods, such as the Metropolis–Hastings algorithm, to draw a Markov chain
\[ S = (\theta^{(1)}, \theta^{(2)}, ... \theta^{(B)}) \]
whose marginal density is approximately given by \( p_T(\theta) \), and \( \hat{\theta}_T \) can be computed as
\[ \hat{\theta}_T = \frac{1}{B} \sum_{j=1}^{B} \theta^{(j)}. \]

Meanwhile, for a given continuously differentiable function \( g: \Theta_\theta \rightarrow \mathbb{R} \), for example, an impulse response at a given horizon, its estimate can be obtained via
\[ g(\hat{\theta}_T) = \frac{1}{B} \sum_{j=1}^{B} g(\theta^{(j)}). \]

Here we omit the details on the construction of the Markov Chains, since they follow standard procedures. One may refer to Chernozhukov and Hong (2003, Section 5) or An and Schorfheide (2007) for more details.

The next result provides an asymptotic justification for the estimator under correct model specification.

**Theorem 4** Suppose \( \theta_0 (\bar{\theta}_0) \) is globally identified or \( \pi(\theta) (\pi(\bar{\theta})) \) is strictly positive only over a compact convex neighborhood of \( \theta_0 (\bar{\theta}_0) \) in which they are locally identified and are interior points, then \( \hat{\theta}_T (\bar{\theta}_T) \) has the same limiting distribution as in Theorem 3 under the corresponding assumptions stated there.
Consider the construction of confidence intervals for the elements of \( \theta_0 \) or, more generally, of \( g(\theta_0) \). In the important special case of Gaussianity with \( \Sigma(\theta) \) being diagonal, the confidence intervals can be obtained directly from the the quantiles of the MCMC sequence \( (\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(B)}) \). Such intervals are asymptotically valid because \( \kappa_{abcd} = 0 \) and therefore \( M = V \). The same result holds for \( \bar{\theta}_0 \) because \( \xi_{abc} = 0 \), thus \( \bar{M} = \bar{V} \). In the general case, because \( \exp(L_T(\theta)) \) is a more general criterion function, implying \( M \neq V \), such an interval is not necessarily asymptotically valid, as clearly demonstrated in Chernozhukov and Hong (2003). However, valid large sample inference can still be easily carried out using the Delta method, as suggested in Chernozhukov and Hong (2003, Theorem 4). Specifically, let \( \bar{M}^{-1} \) be \( T \) times the variance-covariance matrix of the MCMC sequence \( (\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(B)}) \). Let \( \bar{V} \) be an estimator for \( V \), which can be obtained using the formula in Theorem 3 by replacing \( H(\omega), \kappa_{abcd} \) and \( \partial f_{\theta_0^{-1}}(\omega)/\partial \theta_j \) (\( j=1, \ldots, q \)) with their consistent estimates. Then, a valid \( (1 - \alpha) \) percent confidence interval for \( g(\theta_0) \) is given by

\[
[c_{g,T}(\alpha/2), c_{g,T}(1 - \alpha/2)],
\]

where

\[
c_{g,T}(\alpha) = g(\bar{\theta}_T) + q_\alpha T^{-1/2} \sqrt{\frac{\partial g(\bar{\theta}_T)}{\partial \theta'} \bar{M}^{-1} \bar{V} \bar{M}^{-1} \frac{\partial g(\bar{\theta}_T)}{\partial \theta}}
\]

with \( q_\alpha \) being the \( \alpha \)-quantile of the standard normal distribution. Analogous argument can be applied to construct confidence intervals for \( g(\bar{\theta}_0) \). The asymptotic validity of such intervals can be verified using the same argument as in Chernozhukov and Hong (2003, Theorem 4). Therefore, the details are omitted here.

Under misspecification, a result analogous to Theorem 4 can be obtained, with the true value replaced by the pseudo-true values and the covariance matrix modified accordingly.

The key computational difference between the above method and the time domain quasi-Bayesian inference is in computing the Kalman filter versus the spectral density at the different parameter values. Therefore, the computation costs are similar. The spectral domain approach has some additional advantages. First, one can exclude some frequencies by specifying an appropriate \( W(\omega) \), which is not easy to achieve in the time domain. Second, if the sole interest is in estimating the dynamic parameters, it is not necessary to specify \( \mu(\bar{\theta}) \) or to demean the data. Third, although not pursued in the current paper, the spectral domain approach can be extended to handle models without requiring log-linearizations. The idea is that as long as the spectral density can be computed, analytically or by simulation, a similar criterion function as (13) can be constructed to obtain parameter estimates. Such an idea has been mentioned elsewhere, for example in Diebold, Ohanian and Berkowitz (1998), but has not been formally studied. Finally, it provides a platform
for conducting hypothesis testing and model diagnosis from the spectral domain, as emphasized by Watson (1993). For example, one can readily obtain estimates and confidence interval for components of the spectral density matrix and contrast them with the observed data. Also, it is simple to construct tests for restrictions imposed on a given frequency component, such as the zero frequency. We plan to explore such developments in future work.

6 Conclusion

We have provided a unified treatment of issues related to identification, inference and computation in linearized DSGE models in the frequency domain. Besides presenting a necessary and sufficient condition for local identification of the structural parameters, we also proposed a method to trace out non-identification curves when lack of identification is detected. For estimation, we considered a frequency domain quasi-maximum likelihood (FQML) estimator and showed that it permits to incorporate relevant prior distributions and is computationally attractive. Future works include extending the framework to handle models without log-linearizations and to conduct hypothesis testing and model diagnosis in the spectral domain.
References


Appendix

The spectral density matrix $f_\theta(\omega)$ is a Hermitian matrix satisfying $f_\theta(\omega)^* = f_\theta(\omega)$. It is in general not symmetric. The following correspondence is useful for understanding and proving the identification results:

$$f_\theta(\omega) \leftrightarrow f_\theta(\omega)^R \text{ with } f_\theta(\omega)^R = \begin{bmatrix} \text{Re}(f_\theta(\omega)) & \text{Im}(f_\theta(\omega)) \\ -\text{Im}(f_\theta(\omega)) & \text{Re}(f_\theta(\omega)) \end{bmatrix}, \tag{A.1}$$

where $\text{Re}()$ and $\text{Im}()$ denote the real and the imaginary part of a complex matrix, i.e., if $C = A + Bi$, then $\text{Re}(C) = A$ and $\text{Im}(C) = B$. Because $f_\theta(\omega)$ is Hermitian, $f_\theta(\omega)^R$ is real and symmetric (see Lemma 3.7.1(v) in Brillinger, 2001). To simplify notation, let

$$R(\omega; \theta) = \text{vec}(f_\theta(\omega)^R).$$

The following lemma is crucial for proving the subsequent results.

**Lemma A.1** We have the following identity:

$$\left( \frac{\partial \text{vec}(f_\theta(\omega)')}{\partial \theta'} \right)' \left( \frac{\partial \text{vec}(f_\theta(\omega))}{\partial \theta} \right) = \frac{1}{2} \left( \frac{\partial R(\omega; \theta)}{\partial \theta'} \right)' \left( \frac{\partial R(\omega; \theta)}{\partial \theta} \right). \tag{A.2}$$

**Proof of Lemma A.1.** The $(j,k)$-th element of the term on the left hand side is equal to

$$\begin{align*}
\left( \frac{\partial \text{vec}(f_\theta(\omega)')}{\partial \theta_j} \right)' \left( \frac{\partial \text{vec}(f_\theta(\omega))}{\partial \theta_k} \right) &= \text{tr} \left\{ \frac{\partial f_\theta(\omega)}{\partial \theta_j} \frac{\partial f_\theta(\omega)}{\partial \theta_k} \right\} \\
 &= \text{tr} \left\{ \frac{\partial f_\theta(\omega)}{\partial \theta_j} \frac{\partial f_\theta(\omega)}{\partial \theta_k} \right\} \\
 &= \frac{1}{2} \text{tr} \left\{ \left( \frac{\partial f_\theta(\omega)}{\partial \theta_j} \right)^R \left( \frac{\partial f_\theta(\omega)}{\partial \theta_k} \right)^R \right\} \\
 &= \frac{1}{2} \text{tr} \left\{ \left( \frac{\partial f_\theta(\omega)^R}{\partial \theta_j} \right) \left( \frac{\partial f_\theta(\omega)^R}{\partial \theta_k} \right) \right\},
\end{align*}$$

where the first equality is because of the identity $\text{vec}(A')' \text{vec}(B) = \text{tr}(AB)$ for generic matrices $A$ and $B$, the second is because $f_\theta(\omega)$ is Hermitian, thus this term is real valued, the third equality is because of the definition (A.1), the fourth is because, for generic complex matrices, if $Z = XY$, then $Z^R = X^R Y^R$ (see Lemma 3.7.1(ii) in Brillinger, 2001), and the fifth is because $f_\theta(\omega)^R$ is real and symmetric. The last term in the display is simply the $(j,k)$-th element of the right hand side term in (A.2). This completes the proof.

**Proof of Theorem 1.** Lemma A.1 implies that $G(\theta)$ defined by (8) is real, symmetric, positive semidefinite and is equal to

$$\frac{1}{2} \int_{-\pi}^{\pi} \left( \frac{\partial R(\omega; \theta_0)}{\partial \theta'} \right)' \left( \frac{\partial R(\omega; \theta_0)}{\partial \theta} \right) d\omega.$$
This allows us to adopt the arguments in Theorem 1 in Rothenberg (1971) to prove the result.

Suppose \( \theta_0 \) is not locally identified. Then, there exists an infinite sequence of vectors \( \{ \theta_k \}_{k=1}^{\infty} \) approaching \( \theta_0 \) such that, for each \( k \),

\[
R(\omega; \theta_0) = R(\omega; \theta_k) \quad \text{for all } \omega \in [-\pi, \pi].
\]

For an arbitrary \( \omega \in [-\pi, \pi] \), by the mean value theorem and the differentiability of \( f_\theta(\omega) \) in \( \theta \),

\[
0 = R_j(\omega; \theta_k) - R_j(\omega; \theta_0) = \frac{\partial R_j(\omega; \theta_0)}{\partial \theta^j}(\theta_k - \theta_0),
\]

where the subscript \( j \) denotes the \( j \)-th element of the vector and \( \theta_0 \) lies between \( \theta_k \) and \( \theta_0 \) and in general depends on both \( \omega \) and \( j \). Let

\[
d_k = \frac{\theta_k - \theta_0}{\| \theta_k - \theta_0 \|},
\]

then

\[
\frac{\partial R_j(\omega; \theta_0)}{\partial \theta^j} d_k = 0 \quad \text{for every } k.
\]

The sequence \( \{d_k\} \) is an infinite sequence on the unit sphere and therefore there exists a limit point \( d \) (note that \( d \) does not depend on \( j \) or \( \omega \)). As \( \theta_k \to \theta_0 \), \( d_k \) approaches \( d \) and we have

\[
\lim_{k \to \infty} \frac{\partial R_j(\omega; \theta_0)}{\partial \theta^j} d_k = \frac{\partial R_j(\omega; \theta_0)}{\partial \theta^j} d = 0,
\]

where the convergence result holds because \( f_\theta(\omega) \) is continuously differentiable in \( \theta \) (Assumption 3). Because this holds for an arbitrary \( j \), it holds for the full vector \( R(\omega; \theta_0) \). Therefore

\[
\frac{\partial R(\omega; \theta_0)}{\partial \theta^j} d = 0,
\]

which implies

\[
d \left( \frac{\partial R(\omega; \theta_0)}{\partial \theta^j} \right) \left( \frac{\partial R(\omega; \theta_0)}{\partial \theta^j} \right)' d = 0.
\]

Because the above result holds for an arbitrary \( \omega \in [-\pi, \pi] \), it also holds when integrating over \([-\pi, \pi]\). Thus

\[
d \left\{ \int_{-\pi}^{\pi} \left( \frac{\partial R(\omega; \theta_0)}{\partial \theta^j} \right) \left( \frac{\partial R(\omega; \theta_0)}{\partial \theta^j} \right)' d\omega \right\} d = 0.
\]

Applying Lemma A1, because \( d \neq 0 \), \( G(\theta_0) \) is singular.

To show the converse, suppose that \( G(\theta) \) has constant rank \( \rho < q \) in a neighborhood of \( \theta_0 \) denoted by \( \delta(\theta_0) \). Then, consider the characteristic vector \( c(\theta) \) associated with one of the zero roots of \( G(\theta) \). Because

\[
\int_{-\pi}^{\pi} \left( \frac{\partial R(\omega; \theta)}{\partial \theta^j} \right) \left( \frac{\partial R(\omega; \theta)}{\partial \theta^j} \right)' d\omega \times c(\theta) = 0,
\]

A-2
we have
\[ \int_{-\pi}^{\pi} \left( \frac{\partial R(\omega; \theta)}{\partial \theta'} c(\theta) \right)' \left( \frac{\partial R(\omega; \theta)}{\partial \theta'} c(\theta) \right) d\omega = 0. \]
Since the integrand is continuous in \( \omega \) and always non-negative, we must have
\[ \left( \frac{\partial R(\omega; \theta)}{\partial \theta'} c(\theta) \right)' \left( \frac{\partial R(\omega; \theta)}{\partial \theta'} c(\theta) \right) = 0 \]
for all \( \omega \in [-\pi, \pi] \) and all \( \theta \in \delta(\theta_0) \). Furthermore, this implies
\[ \frac{\partial R(\omega; \theta)}{\partial \theta'} c(\theta) = 0 \tag{A.3} \]
for all \( \omega \in [-\pi, \pi] \) and all \( \theta \in \delta(\theta_0) \). Because \( G(\theta) \) is continuous and has constant rank in \( \delta(\theta_0) \), the vector \( c(\theta) \) is continuous in \( \delta(\theta_0) \). Consider the curve \( \chi \) defined by the function \( \theta(v) \) which solves for \( 0 \leq v \leq \bar{v} \) the differential equation
\[ \frac{\partial \theta(v)}{\partial v} = c(\theta), \]
\[ \theta(0) = \theta_0. \]
Then,
\[ \frac{\partial R(\omega; \theta(v))}{\partial v} = \frac{\partial R(\omega; \theta(v))}{\partial \theta(v)'} \frac{\partial \theta(v)}{\partial v} = \frac{\partial R(\omega; \theta(v))}{\partial \theta(v)'} c(\theta) = 0 \]
for all \( \omega \in [-\pi, \pi] \) and \( 0 \leq v \leq \bar{v} \), where the last equality uses (A.3). Thus, \( R(\omega; \theta) \) is constant on the curve \( \chi \). This implies that \( f_\theta(\omega) \) is constant on the same curve and that \( \theta_0 \) is unidentifiable. This completes the proof.

**Proof of Theorem 2.** Using Lemma A.1 again, \( G(\tilde{\theta}) \) can be equivalently represented as
\[ G(\tilde{\theta}) = \frac{1}{2} \int_{-\pi}^{\pi} \left( \frac{\partial R(\omega; \theta)}{\partial \theta'} \right)' \left( \frac{\partial R(\omega; \theta)}{\partial \theta'} \right) d\omega + \left( \frac{\partial \mu(\tilde{\theta})}{\partial \theta'} \right)' \frac{\partial \mu(\tilde{\theta})}{\partial \theta'} \]
with both terms on the right hand side being real, symmetric and positive semidefinite. Let
\[ \bar{R}(\omega; \tilde{\theta}) = \begin{bmatrix} R(\omega; \theta) \\ \frac{1}{\sqrt{\pi}} \mu(\tilde{\theta}) \end{bmatrix}, \]
then
\[ \tilde{G}(\tilde{\theta}) = \frac{1}{2} \int_{-\pi}^{\pi} \left( \frac{\partial \bar{R}(\omega; \tilde{\theta})}{\partial \theta'} \right)' \left( \frac{\partial \bar{R}(\omega; \tilde{\theta})}{\partial \theta'} \right) d\omega. \]
Using this representation, the proof proceeds in the same way as in Theorem 1, with \( \theta \) replaced by \( \tilde{\theta} \) and \( R(\omega; \theta) \) by \( \bar{R}(\omega; \tilde{\theta}) \). The detail is omitted.

**Proof of Corollary 2.** We only prove the first result, as the second can be proven analogously using the formulation in the proof of Theorem 2.
Suppose the subvector $\theta_0^s$ is not locally identified. Write $\theta = (\theta^a, \theta^r)'$. There exists an infinite sequence of vectors $\{\theta_k\}_{k=1}^{\infty}$ approaching $\theta_0$ such that

$$R(\omega; \theta_0) = R(\omega; \theta_k) \text{ for all } \omega \in [-\pi, \pi] \text{ and each } k.$$ 

By the definition of partial identification, $\{\theta_k^s\}$ can be chosen such that $\|\theta_k^s - \theta_0^s\| / \|\theta_k - \theta_0\| > \varepsilon$ with $\varepsilon$ being some arbitrarily small positive number. The values of $\theta_k^s$ can either change or stay fixed in this sequence; no restriction is imposed on them besides those in the preceding display. As in the proof of Theorem 1, in the limit, we have

$$\frac{\partial R(\omega; \theta_0)}{\partial \theta^s} d = 0,$$

with $d^s \neq 0$ (where $d^s$ is comprised of the elements in $d$ that correspond to $\theta^s$). Therefore, on one hand,

$$G(\theta_0)d = 0,$$

on the other hand, because $d^s \neq 0$ and by definition $\partial \theta_0^s / \partial \theta^r = [I_{\dim(\theta^s)}, 0_{\dim(\theta^r)}]$, we have

$$\frac{\partial \theta_0^s}{\partial \theta^r} d = d^s \neq 0,$$

which implies

$$G^a(\theta_0)d \neq 0.$$

Thus, we have identified a vector that falls into the orthogonal column space of $G(\theta_0)$ but not of $G^a(\theta_0)$. Because the former orthogonal space always includes the latter as a subspace, this result implies $G^a(\theta_0)$ has a higher column rank than $G(\theta_0)$.

To show the converse, suppose that $G(\theta)$ and $G^a(\theta)$ have constant ranks in a neighborhood of $\theta_0$ denoted by $\delta(\theta_0)$. Because the rank of $G(\theta)$ is lower than that of $G^a(\theta)$, there exists a vector $c(\theta)$ such that

$$G(\theta)c(\theta) = 0 \text{ but } G^a(\theta)c(\theta) \neq 0,$$

which implies for all $\omega \in [-\pi, \pi]$ and all $\theta \in \delta(\theta_0)$ (c.f. arguments leading to (A.3)):

$$\frac{\partial R(\omega; \theta)}{\partial \theta^r} c(\theta) = 0$$

but

$$\begin{bmatrix} \partial R(\omega; \theta) / \partial \theta^s \\ \partial \theta^s / \partial \theta^r \end{bmatrix} c(\theta) = \begin{bmatrix} 0 \\ c^s(\theta) \end{bmatrix} \neq 0,$$

where $c^s(\theta)$ denotes the elements in $c(\theta)$ that correspond to $\theta^s$. Because $G(\theta)$ is continuous and has constant rank in $\delta(\theta_0)$, the vector $c(\theta)$ is continuous in $\delta(\theta_0)$. As in Theorem 1, consider the curve $\chi$ defined by the function $\theta(v)$ which solves for $0 \leq v \leq \bar{v}$ the differential equation

$$\frac{\partial \theta(v)}{\partial v} = c(\theta), \ \theta(0) = \theta_0.$$
On one hand, because $c^*(\theta) \neq 0$ and $c^*(\theta)$ is continuous in $\theta$, points on this curve correspond to different $\theta^s$. On the other hand,

\[ \frac{\partial R(\omega; \theta(v))}{\partial v} = \frac{\partial R(\omega; \theta(v))}{\partial \theta(v)'} \frac{\partial \theta(v)}{\partial v} = \frac{\partial R(\omega; \theta(v))}{\partial \theta(v)'} c(\theta) = 0 \]

for all $\omega \in [-\pi, \pi]$ and $0 \leq v \leq \bar{v}$, implying $f_\theta(\omega)$ is constant on the same curve. Therefore, $\theta^s_0$ is not locally identifiable.

**Proof of Corollary 4.** The proof is essentially the same as in Rothenberg (1971, Theorem 2), and is included for the matter of completeness. Suppose $\Psi(\theta)$ has rank $s$ for all $\theta$ in a neighborhood of $\theta_0$. Then, by the implicit function theorem, there exists a partition of $\theta$ into $\theta^1 \in R^s$ and $\theta^2 \in R^{q-s}$ such that

\[ \theta^1 = q(\theta^2) \]

for all solutions of $\psi(\theta) = 0$ in a neighborhood of $\theta_0$ with $\theta^2_0$ being an interior point of that neighborhood. Consequently, the spectral density can be rewritten as

\[ f_\theta(\omega) = f_{q(\theta^2),\theta^2}(\omega), \]

which involves only $q - s$ parameters. Let

\[ Q(\theta^2) = \frac{\partial q(\theta^2)}{\partial \theta^2} \quad \text{and} \quad \tilde{G}(\theta) = \begin{bmatrix} Q(\theta^2)' & I \end{bmatrix} G(\theta) \begin{bmatrix} Q(\theta^2) \\ I \end{bmatrix}. \]

Then, by Theorem 1, $\theta_0$ is identified if and only if $\tilde{G}(\theta_0)$ has full rank.

Suppose there exists a vector $d \in R^{q-s}$ such that

\[ \tilde{G}(\theta_0) d = 0. \tag{A.4} \]

Then, the structure of $G(\theta)$ (c.f. Lemma A1) implies that (A.4) holds if and only if

\[ G(\theta_0) \begin{bmatrix} Q(\theta^2_0) \\ I \end{bmatrix} d = 0. \]

Let

\[ c = \begin{bmatrix} Q(\theta^2_0) \\ I \end{bmatrix} d, \]

then we have: (1) $c \neq 0$ if and only $d \neq 0$, and (2)

\[ \begin{bmatrix} G(\theta_0) \\ \Psi(\theta_0) \end{bmatrix} c = 0 \]

if and only if (A.4) holds, where $\Psi(\theta_0)c = 0$ always holds because $\theta_0$ satisfies the constraint $\psi(\theta) = 0$. Thus, the preceding matrix has full rank if and only if $\theta_0$ is identified under the constraints. This completes the proof.
Proof of Corollary 5. Without loss of generality, assume $n_Y = 1$. Otherwise, the proof can be carried out by analyzing $R(\omega; \theta)$. The map $\theta \mapsto f_\theta$ is infinite dimensional. The proof therefore involves two steps. The first is to reduce it to a finite dimensional problem. The second is to apply a constant rank theorem (a generalization of the implicit function theorem).

Consider a positive integer $N$ and a partition of the interval $[-\pi, \pi]$ by $\omega_j = (2\pi j/2^N) - \pi$, with $j = 0, 1, \ldots, 2^N$. Then, the following map is finite dimensional:

$$\theta \mapsto (f_\theta(\omega_0), \ldots, f_\theta(\omega_{2^N})). \quad (A.5)$$

To simplify notation, let $f_{\theta,N} = (f_\theta(\omega_0), \ldots, f_\theta(\omega_{2^N}))'$. Conventionally, the rank of the above map is defined as the rank of the Jacobian matrix $\partial f_{\theta,N}/\partial \theta'$, which is of dimension $(2^N + 1)$-by-$q$ with rank no greater than $q - 1$ at $\theta_0$, because if the rank equals to $q$, then $\theta_0$ becomes locally identified, contradicting the assumption in the corollary. Note that, for a given $N$, its rank can be strictly less than $q - 1$.

We now show that there exists a finite $N$ such that $\partial f_{\theta,N}/\partial \theta'$ has rank $q - 1$ at $\theta_0$. Suppose such an $N$ does not exist. Then, the rank of $\partial f_{\theta,N}/\partial \theta'$ is at most $q - 2$ for arbitrarily large $N$. This implies that the rank of

$$G_N(\theta_0) = \frac{2\pi}{2^N + 1} \sum_{j=0}^{2^N} \left( \frac{\partial f_{\theta_0}(\omega_j)}{\partial \theta'} \right)' \left( \frac{\partial f_{\theta_0}(\omega_j)}{\partial \theta'} \right)$$

is at most $q - 2$ for arbitrarily large $N$, because vectors orthogonal to $\partial f_{\theta,N}/\partial \theta'$ are also orthogonal to $G_N(\theta)$ by construction. Let $\lambda_{N,j}$ ($j = 1, \ldots, q$) be the eigenvalues of $G_N(\theta_0)$ sorted in an increasing order. Then, for any finite $N$,

$$\lambda_{N,1} = \lambda_{N,2} = 0.$$

On the other hand, because $G_N(\theta_0) \to G(\theta_0)$, so do its eigenvalues. Thus, for any $\varepsilon > 0$, there exists a finite $N$ such that $|\lambda_2 - \lambda_{N,2}| < \varepsilon$, where $\lambda_2$ is the second smallest eigenvalue of $G(\theta_0)$. Choosing $\varepsilon = \lambda_2/2$ leads to

$$|\lambda_{N,2}| > \lambda_2/2.$$

Since rank($G(\theta_0)$) = $q - 1$ by Assumption, $\lambda_2$ is strictly positive. Thus, we reach a contradiction. Because the convergence of $G_N(\theta) \to G(\theta)$ is uniform in an open neighborhood of $\theta_0$, say $\delta(\theta_0)$, the above analysis also implies there exists an $N$ such that $\partial f_{\theta,N}/\partial \theta'$ has constant rank $q - 1$ in that neighborhood.

Use such an $N$ and consider again the map $\theta \mapsto f_{\theta,N}$, which is finite dimensional, continuously differentiable and has constant rank $q - 1$ in $\delta(\theta_0)$. Define the level set

$$\{ \theta \in \delta(\theta_0) : f_{\theta,N} = f_{\theta_0,N} \}.$$

Then, the rank theorem (Krantz and Parks 2002, Theorem 3.5.1 and the discussion on page 56) implies that the level set constitutes a smooth, parameterized one dimensional manifold. Thus, there exists a unique level curve passing through $\theta_0$ satisfying $f_{\theta,N} = f_{\theta_0,N}$.
Therefore, we have established the result for a particular finite $N$. Further increasing $N$ leads to finer partitions of $[-\pi, \pi]$. This cannot decrease the rank of the map (A.5), therefore cannot increase the number of level curves passing through $\theta_0$. Thus, in the limit, there is at most one level curve passing through $\theta_0$. The existence of such a curve for the infinite dimensional case has already been shown in the main text, given by (9). This completes the proof.

**Proof of Lemma 1.** Applying Lemma A.3.3 (1) in Hosoya and Taniguchi (1982), for a given $\theta \in \Theta_\theta$, we have

$$\plim_{T \to \infty} \frac{1}{T} \sum_{j=1}^{T-1} \tr \{ W(\omega_j) f_\theta^{-1}(\omega_j) I_T(\omega_j) \} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \tr \{ W(\omega) f_\theta^{-1}(\omega) f_{\theta_0}(\omega) \} \, d\omega.$$ 

To prove stochastic equicontinuity, consider for any $\theta_1, \theta_2 \in \Theta_\theta$

$$\frac{1}{T} \sum_{j=1}^{T-1} \tr \{ W(\omega_j) \left( f_{\theta_1}^{-1}(\omega_j) - f_{\theta_2}^{-1}(\omega_j) \right) I_T(\omega_j) \}.$$

Apply a first order Taylor expansion,

$$\frac{1}{T} \sum_{j=1}^{T-1} \tr \{ W(\omega_j) \left( f_{\theta_1}^{-1}(\omega_j) - f_{\theta_2}^{-1}(\omega_j) \right) I_T(\omega_j) \}$$

$$= \frac{1}{T} \sum_{j=1}^{T-1} \frac{\partial \tr \{ W(\omega_j) f_\theta^{-1}(\omega_j) I_T(\omega_j) \} }{\partial \theta'} (\theta_1 - \theta_2)$$

$$= -\frac{1}{T} \sum_{j=1}^{T-1} W(\omega_j) \vec{\tr} \{ I_T(\omega_j)' \} \{ f_\theta^{-1}(\omega_j)' \otimes f_\theta^{-1}(\omega_j) \} \frac{\partial \vec{\tr} (f_\theta(\omega_j))}{\partial \theta'} (\theta_1 - \theta_2), \tag{A.6}$$

where $\tilde{\theta}$ lies between $\theta_1$ and $\theta_2$. The norm of (A.6) is bounded by

$$\frac{1}{T} \sum_{j=1}^{T-1} \left\| \vec{\tr} (I_T(\omega_j)') \right\| \left\| \{ f_\theta^{-1}(\omega_j)' \otimes f_\theta^{-1}(\omega_j) \} \frac{\partial \vec{\tr} (f_\theta(\omega_j))}{\partial \theta'} \right\| \| \theta_1 - \theta_2 \|.$$

The quantity

$$\left\| (f_\theta^{-1}(\omega_j)' \otimes f_\theta^{-1}(\omega_j)) \frac{\partial \vec{\tr} (f_\theta(\omega_j))}{\partial \theta'} \right\|$$

is uniformly bounded by Assumption 5(ii). The term $T^{-1} \sum_{j=1}^{T-1} \| \vec{\tr} (I_T(\omega_j)') \|$ only depends on $\theta_0$ and is $O_p(1)$ because the diagonal elements of $T^{-1} \sum_{j=1}^{T-1} I_T(\omega_j)$ are positive and satisfy a law of large numbers (Hosoya and Taniguchi, 1982, Lemma A.3.3 (1)), and the norm of the off-diagonal elements can be bounded by the diagonal elements using the Cauchy–Schwarz inequality. Therefore, the term (A.6) can be made uniformly small by choosing a small $\| \theta_1 - \theta_2 \|$. Meanwhile,

$$\frac{1}{T} \sum_{j=1}^{T-1} W(\omega_j) \log \det f_\theta(\omega_j) \to \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) \log \det f_\theta(\omega) \, d\omega$$

A-7
uniformly in $\theta \in \Theta$. Thus, the first result holds.

For the second result, we first show that $\theta_0$ maximizes $L_\infty (\theta)$. Apply the same argument as in Hosoya and Taniguchi (1982, p.149). For every $\omega \in [-\pi, \pi]$

$$W(\omega) \left[ \log \det f_0(\omega) + tr \left\{ f_0^{-1}(\omega)f_{\theta_0}(\omega) \right\} \right] = W(\omega) \log \det f_{\theta_0}(\omega) + W(\omega) \left[ tr \left\{ f_0^{-1}(\omega)f_{\theta_0}(\omega) \right\} - \log \det \left\{ f_0^{-1}(\omega)f_{\theta_0}(\omega) \right\} \right]$$

$$= W(\omega) \log \det f_{\theta_0}(\omega) + W(\omega) \sum_{j=1}^{n_Y} \lambda_j(\omega) - \log \lambda_j(\omega) - 1 + W(\omega)n_Y,$$

where $\lambda_j(\omega)$ is the j-th eigenvalue of $f_0^{-1}(\omega)f_{\theta_0}(\omega)$. Because $\lambda_j(\omega) - \log \lambda_j(\omega) - 1 \geq 0$ and the equality holds if and only if $\lambda_j(\omega) = 1, j = 1, \ldots, n_Y$. This implies

$$L_\infty (\theta) \leq -\frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) \left( \log \det f_{\theta_0}(\omega) + n_Y \right) d\omega,$$

which holds with equality if and only if $\lambda_j(\omega) = 1$ for all $\omega \in [-\pi, \pi]$ (j = 1,...,n_Y). However, $\lambda_j(\omega) = 1$ implies $f_{\theta_0}(\omega) = f_\theta(\omega)$ because $f_0^{-1}(\omega)f_{\theta_0}(\omega)$ is Hermitian, therefore having orthonormal eigenvectors. Hence, $\theta_0$ is a global maximizer.

The above result implies that any other parameter vector, say $\theta_1$, is a maximizer if and only if $f_{\theta_1}(\omega) = f_{\theta_0}(\omega)$ for all $\omega \in [-\pi, \pi]$. Now, suppose the parameters are locally identified. Then, there are no parameter values close to $\theta_0$ satisfying this equality. Thus, $\theta_0$ is the locally unique maximizer. To see the converse, suppose $\theta_0$ is the locally unique maximizer, then, there cannot be any parameter close to $\theta_0$ satisfying $f_{\theta_0}(\omega) = f_\theta(\omega)$ for all $\omega$. Thus, by definition, we have local identification. The argument to establish the result for the global identification proceeds in the same way.

The third result follows directly from the uniform weak law of large numbers.

**Proof of Theorem 3.** We only prove the second result which includes the first is a special case.

The first order condition gives

$$2\pi T^{-1/2} \sum_{j=0}^{T-1} W(\omega_j) \frac{\partial \text{vec} \left( f_{\tilde{\theta}_T}(\omega_j) \right)' \left( f_{\tilde{\theta}_T}^{-1}(\omega_j)' \otimes f_{\tilde{\theta}_T}^{-1}(\omega_j) \right)}{\partial \theta} \text{vec} (I_T(\omega_j) - f_{\tilde{\theta}_T}(\omega_j))$$

$$+ 2T^{-1/2} \sum_{t=1}^{T} \frac{\partial \mu(\tilde{\theta}_T)' f_{\tilde{\theta}_T}^{-1}(0) (Y_t - \mu(\tilde{\theta}_T))}{\partial \theta} = 0.$$
which holds because $\widehat{\theta}_T \to_p \tilde{\theta}_0$, $f_{\tilde{\theta}_0}(\omega_j)$ and $\mu(\tilde{\theta}_0)$ are continuously differentiable, and that $f_{\tilde{\theta}_0}^{-1}(\omega_j)$ have bounded eigenvalues. Apply a first order Taylor expansion around $\tilde{\theta}_0$, then the left hand side of the preceding display is equal to

$$2\pi T^{-1/2} \sum_{j=0}^{T-1} W(\omega_j) \frac{\partial \text{vec}(f_{\tilde{\theta}_0}(\omega_j))'}{\partial \theta} \left( f_{\tilde{\theta}_0}^{-1}(\omega_j)' \otimes f_{\tilde{\theta}_0}^{-1}(\omega_j) \right) \text{vec} \left( I_T(\omega_j) - f_{\tilde{\theta}_0}(\omega_j) \right) \quad (I) \quad (A.7)$$

$$+ 2T^{-1/2} \sum_{t=1}^{T} \frac{\partial \mu(\tilde{\theta}_0)'}{\partial \theta} f_{\tilde{\theta}_0}^{-1}(0) \left( y_t - \mu(\tilde{\theta}_0) \right) \quad (II)$$

$$- 2\pi T^{-1} \sum_{j=0}^{T-1} W(\omega_j) \frac{\partial \text{vec}(f_{\tilde{\theta}_0}(\omega_j))'}{\partial \theta} \left( f_{\tilde{\theta}_0}^{-1}(\omega_j)' \otimes f_{\tilde{\theta}_0}^{-1}(\omega_j) \right) \frac{\partial \text{vec}(f_{\tilde{\theta}_0}(\omega_j))}{\partial \theta'} T^{1/2}(\widehat{\theta} - \tilde{\theta}_0) \quad (III)$$

$$- 2 \frac{\partial \mu(\tilde{\theta}_0)'}{\partial \theta} f_{\tilde{\theta}_0}^{-1}(0) \frac{\partial \mu(\tilde{\theta}_0)}{\partial \theta'} T^{1/2}(\widehat{\theta} - \tilde{\theta}_0) \quad (IV)$$

$$+ o_p(1).$$

First consider term (III), the quantity in front of $T^{1/2}(\widehat{\theta} - \tilde{\theta}_0)$ converges to

$$\int_{-\pi}^{\pi} W(\omega) \frac{\partial \text{vec}(f_{\tilde{\theta}_0}(\omega))'}{\partial \theta} \left( f_{\tilde{\theta}_0}^{-1}(\omega)' \otimes f_{\tilde{\theta}_0}^{-1}(\omega) \right) \frac{\partial \text{vec}(f_{\tilde{\theta}_0}(\omega))}{\partial \theta'} d\omega,$$

whose $(h,k)$-th element is given by

$$\int_{-\pi}^{\pi} \text{tr} \left\{ W(\omega) f_{\tilde{\theta}_0}(\omega) \frac{\partial f_{\tilde{\theta}_0}^{-1}(\omega)}{\partial \theta_h} f_{\tilde{\theta}_0}(\omega) \frac{\partial f_{\tilde{\theta}_0}^{-1}(\omega)}{\partial \theta_k} \right\} d\omega.$$

Therefore, the above expansion implies (c.f. Theorem 3 for the definition of $\tilde{M}$)

$$T^{1/2}(\widehat{\theta} - \tilde{\theta}_0) = \tilde{M}^{-1} * (I) + \tilde{M}^{-1} * (II) + o_p(1).$$

Term (I) satisfies a central limit theorem, whose covariance matrix has the $(h,k)$-th element given by (see Theorem 3.1 and Proposition 3.1 in Hosoya and Taniguchi, 1982, in particular, their formula for $U_{ij}$)

$$4\pi \int_{-\pi}^{\pi} W(\omega) \text{tr} \left\{ f_{\tilde{\theta}_0}(\omega) \frac{\partial f_{\tilde{\theta}_0}^{-1}(\omega)}{\partial \theta_h} f_{\tilde{\theta}_0}(\omega) \frac{\partial f_{\tilde{\theta}_0}^{-1}(\omega)}{\partial \theta_k} \right\} d\omega$$

$$+ \sum_{a,b,c,d=1}^{n_\epsilon} \kappa_{abcd} \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H^*(\omega) \frac{\partial f_{\tilde{\theta}_0}^{-1}(\omega)}{\partial \theta_h} H(\omega) d\omega \right]_{ab} \times \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega) H^*(\omega) \frac{\partial f_{\tilde{\theta}_0}^{-1}(\omega)}{\partial \theta_k} H(\omega) d\omega \right]_{cd},$$

Term (II) also satisfies CLT, with covariance matrix given by

$$8\pi \frac{\partial \mu(\tilde{\theta}_0)'}{\partial \theta} f_{\tilde{\theta}_0}^{-1}(0) \frac{\partial \mu(\tilde{\theta}_0)}{\partial \theta'}.$$

To complete the proof, we only need to verify the covariance matrix between (I) and (II). Let

$$A = \text{Cov}((I),(II))$$
and consider its \((h,k)\)-th element:

\[
A_{hk} = 4\pi \text{Cov} \left\{ \operatorname{tr} \left( \frac{1}{\sqrt{T}} \sum_{j=0}^{T-1} W(\omega_j) \frac{\partial f_{\theta_0}^{-1}(\omega_j)}{\partial \theta_h} (I_T(\omega_j) - f_{\theta_0}(\omega)) \right), \left( \frac{1}{\sqrt{T}} \frac{\partial \mu(\tilde{\theta}_0)}{\partial \theta_k} f_{\theta_0}^{-1}(0) \sum_{t=1}^{T} (Y_t - \mu(\tilde{\theta}_0)) \right) \right\}
\]

Define

\[
\phi^h(\omega_j) = \frac{\partial f_{\theta_0}^{-1}(\omega_j)}{\partial \theta_h} \quad \text{and} \quad \psi^k(0) = \frac{\partial \mu(\tilde{\theta}_0)}{\partial \theta_k} f_{\theta_0}^{-1}(0).
\]

Then,

\[
A_{hk} = 4\pi \text{Cov} \left\{ \operatorname{tr} \left( \frac{1}{\sqrt{T}} \sum_{j=0}^{T-1} W(\omega_j) \phi^h(\omega_j) (I_T(\omega_j) - f_{\theta_0}(\omega)) \right), \left( \frac{1}{\sqrt{T}} \psi^k(0) \sum_{t=1}^{T} (Y_t - \mu(\tilde{\theta}_0)) \right) \right\}
\]

\[
= 4\pi \text{Cov} \left\{ \frac{1}{\sqrt{T}} \sum_{j=0}^{T-1} W(\omega_j) \sum_{a,b=1}^{n_y} \phi^h_{ab}(\omega_j) (I_{T_{ba}}(\omega_j) - f_{\theta_0 ba}(\omega)), \frac{1}{\sqrt{T}} \sum_{c=1}^{n_y} \psi^k_{c}(0) \sum_{t=1}^{T} (Y_{tc} - \mu_c(\tilde{\theta}_0)) \right\}
\]

\[
= 4\pi \sum_{a,b,c=1}^{n_y} \text{Cov} \left\{ \frac{1}{\sqrt{T}} \sum_{j=0}^{T-1} W(\omega_j) \phi^h_{ab}(\omega_j) (I_{T_{ba}}(\omega_j) - f_{\theta_0 ba}(\omega)), \frac{1}{\sqrt{T}} \psi^k_{c}(0) \sum_{t=1}^{T} (Y_{tc} - \mu_c(\tilde{\theta}_0)) \right\},
\]

where \(I_{T_{ba}}(\omega_j)\) is the \((b,a)\)-th element of \(I_T(\omega_j)\) and other quantities are defined analogously. Consider the two terms inside the curly brackets separately. Applying the same argument as in Theorem 10.8.5 in Brockwell and Davis (1991), we have

\[
\frac{1}{\sqrt{T}} \sum_{j=0}^{T-1} W(\omega_j) \phi^h_{ab}(\omega_j) (I_{T_{ba}}(\omega_j) - f_{\theta_0 ba}(\omega))
\]

\[
= \frac{1}{\sqrt{T}} \sum_{j=0}^{T-1} \sum_{f,g=1}^{n_x} W(\omega_j) \phi^h_{ab}(\omega_j) H_{bf}(\omega_j) \left( I_{T_{fg}}(\omega_j) - E I_{T_{fg}}(\omega_j) \right) H^*_g(\omega_j) + o_p(1),
\]

where and \(I_{T_{fg}}(\omega_j)\) denote the \((f,g)\)-th element of the periodogram of \(\epsilon_t\). Apply Theorem 10.3.1 in Brockwell and Davis (1991), we have

\[
\frac{1}{\sqrt{T}} \psi^k_{c}(0) \sum_{t=1}^{T} (Y_{tc} - \mu_c(\tilde{\theta}_0)) = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \sum_{l=1}^{T} \psi^k_{c}(0) H_{cl}(0) \epsilon_{lt} + o_p(1),
\]

where \(H(0) = \sum_{j=0}^{\infty} h_j(\theta_0)\) (c.f. (3)). Therefore, their covariance is equal to

\[
\frac{1}{T} \sum_{t=1}^{T} \sum_{j=0}^{T-1} \sum_{f,g,l=1}^{n_x} W(\omega_j) \phi^h_{ab}(\omega_j) H_{bf}(\omega_j) H^*_g(\omega_j) \psi^k_{c}(0) H_{cl}(0) E \left\{ \left( I_{T_{fg}}(\omega_j) - E I_{T_{fg}}(\omega_j) \right) \epsilon_{lt} \right\} + o_p(1)
\]

\[
= \frac{1}{T} \sum_{t=1}^{T} \sum_{f,g,l=1}^{n_x} W(\omega_j) \phi^h_{ab}(\omega_j) H_{bf}(\omega_j) H^*_g(\omega_j) \psi^k_{c}(0) H_{cl}(0) \xi_{fgl} + o_p(1)
\]

\[
= \frac{1}{2\pi} \sum_{f,g,l=1}^{n_x} \left\{ \int_{-\pi}^{\pi} W(\omega) H^*(\omega) \phi^h_{ab}(\omega) H_{bf}(\omega) d\omega \right\} \times \xi_{fgl} \times \left\{ \psi^k_{c}(0) H_{cl}(0) \right\} + o_p(1).
\]
Some algebra shows that
\[ A_{hk} = 2 \sum_{f,g,l=1}^{n_k} \left[ \int_{-\pi}^{\pi} W(\omega) H(\omega) \frac{\partial f_{\theta_0}^{-1}(\omega)}{\partial \theta_k} H(\omega) d\omega \right]_{gf} \times \xi_{fl} \times \left[ \frac{\partial \mu(\theta_0)^{f'}}{\partial \theta_k} f_{\theta_0}^{-1}(0) H(0) \right]_l. \]

**Proof of Corollary 6.** We will prove the second result. Because the argument is very similar to Theorem 3 and Taniguchi (1979, Theorem 2), we will only provide an outline. The estimate \( \hat{\theta} \) solves
\[ \frac{\partial L_T(\hat{\theta})}{\partial \theta} = 0 \]  
and the pseudo true value \( \hat{\theta}_0^m \) satisfies
\[ \frac{\partial L_\infty(\hat{\theta}_0^m)}{\partial \theta} = 0. \]  
Consider a Taylor expansion of (A.8) around \( \hat{\theta}_0^m \):
\[ T^{1/2} (\hat{\theta} - \hat{\theta}_0^m) = \left[ -\frac{2\pi \partial^2 L_T(\hat{\theta})}{T} \right]^{-1} \left( 2\pi T^{-1/2} \frac{\partial L_T(\hat{\theta}_0^m)}{\partial \theta} - 2\pi T^{-1/2} \frac{\partial L_\infty(\hat{\theta}_0^m)}{\partial \theta} \right). \]

Further,
\[ -\frac{2\pi \partial^2 L_T(\hat{\theta})}{T} \rightarrow \int_{-\pi}^{\pi} W(\omega) \left[ \frac{\partial^2}{\partial \theta \partial \theta'} \log \det f_{\theta_0}^{-1}(\omega) \right] + \frac{\partial^2}{\partial \theta \partial \theta'} \text{tr} \left\{ f_{\theta_0}^{-1}(\omega) f_0(\omega) \right\} + 2 \frac{\partial \mu(\hat{\theta}_0^m)^{f'}}{\partial \theta} f_{\theta_0}^{-1}(0) \frac{\partial \mu(\hat{\theta}_0^m)}{\partial \theta'} \]
because \( \hat{\theta} \rightarrow^p \hat{\theta}_0^m \) and the continuity of integrand. Also,
\[
2\pi T^{-1/2} \frac{\partial L_T(\hat{\theta}_0^m)}{\partial \theta} - 2\pi T^{1/2} \frac{\partial L_\infty(\hat{\theta}_0^m)}{\partial \theta} \\
= -2\pi T^{-1/2} \sum_{j=1}^{T-1} W(\omega_j) \frac{\partial}{\partial \theta} \left\{ f_{\theta_0}^{-1}(\omega_j) (I_T(\omega_j) - f_0(\omega)) \right\} + 2T^{-1/2} \sum_{t=1}^{T} \frac{\partial \mu(\hat{\theta}_0^m)^{f'}}{\partial \theta} f_{\theta_0}^{-1}(0) (Y_t - \mu_0) + o_p(1)
\]
\[ = (M1) + (M2) + o_p(1). \]

Terms (M1) and (M2) satisfy a central limit theorem and can be analyzed in the same way as terms (I) and (II) in (A.7). The limiting covariance matrix can be verified accordingly. The detail is omitted.
Proof of Theorem 4: It suffices to verify that Assumptions 1 to 4 in Chernozhukov and Hong (2003) hold under our set of conditions. Relabel these assumptions as CH1 to CH4. CH1 and CH2 are trivial. CH3 is implied by Lemma 1(1), 1(2) and 1(4). To verify CH4, applying a second order Taylor expansion of $L_T(\theta)$ around $\theta_0$ (c.f. CH4(i)):

$$L_T(\theta) - L_T(\theta_0) = (\theta - \theta_0)' \frac{\partial L_T(\theta_0)}{\partial \theta} + \frac{1}{2} (\theta - \theta_0)' \frac{\partial^2 L_T(\theta_0)}{\partial \theta \partial \theta'} (\theta - \theta_0) + R_T(\theta)$$

with

$$R_T(\theta) = (\theta - \theta_0)' \left\{ \frac{\partial^2 L_T(\tilde{\theta}_T)}{\partial \theta \partial \theta'} - \frac{\partial^2 L_T(\theta_0)}{\partial \theta \partial \theta'} \right\} (\theta - \theta_0),$$

where $\tilde{\theta}_T$ lies between $\theta$ and $\theta_0$. Now,

$$T^{-1/2} \frac{\partial L_T(\theta_0)}{\partial \theta} \overset{d}{\rightarrow} N(0, V),$$

therefore CH4(ii) is satisfied ($V$ corresponds to $\Omega_n$ in CH4). For CH4(iii), note that $V$ is nonrandom and positive definite, and that

$$-T^{-1} \frac{\partial^2 L_T(\theta_0)}{\partial \theta \partial \theta'}$$

$$= T^{-1} \sum_{j=1}^{T-1} W(\omega_j) \left( \frac{\partial \text{vec}(f_{\theta_0}(\omega_j)')}{\partial \theta'} \right)' \left\{ f_{\theta_0}^{-1}(\omega_j)' \otimes f_{\hat{\theta}_0}^{-1}(\omega_j) \right\} \left( \frac{\partial \text{vec}(f_{\theta_0}(\omega_j))}{\partial \theta'} \right)$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} W(\omega_j) \left( \frac{\partial \text{vec}(f_{\theta_0}(\omega)')}{\partial \theta'} \right)' \left( \frac{\partial \text{vec}(f_{\hat{\theta}_0}(\omega))}{\partial \theta'} \right) d\omega + o(1)(A.10)$$

where the leading term on the right hand side is a nonrandom and positive definite because $f_{\theta_0}^{-1}(\omega)$ and

$$\int_{-\pi}^{\pi} W(\omega_j) \left( \frac{\partial \text{vec}(f_{\theta_0}(\omega)')}{\partial \theta'} \right)' \left( \frac{\partial \text{vec}(f_{\hat{\theta}_0}(\omega))}{\partial \theta'} \right) d\omega$$

are positive definite by Assumption 5 and local identification. It is $O(1)$ because the integrand is bounded, see Assumption 5. Therefore CH4(iii) is satisfied. CH4(iv.a) holds because

$$|R_T(\theta)| \leq \left\| T^{1/2}(\theta - \theta_0) \right\|^2 \left\| T^{-1} \frac{\partial^2 L_T(\tilde{\theta}_T)}{\partial \theta \partial \theta'} - T^{-1} \frac{\partial^2 L_T(\theta_0)}{\partial \theta \partial \theta'} \right\|,$$

where the second term can be made arbitrarily small by choosing $\|\theta - \theta_0\|$ small because of (A.10) and the boundedness and continuity of $\frac{\partial \text{vec}(f_\theta(\omega))}{\partial \theta'}$ and $f_\theta^{-1}(\omega)$ in $\theta$ (Assumptions 3 and 5(ii)). CH4(iv.b) holds because of the preceding argument and the fact that $\|T^{1/2}(\theta - \theta_0)\|^2 = O(1)$.

The proof for $\tilde{\theta}_T$ involves the same argument and is therefore omitted.
Figure 1. The Non-identification curve \( (\psi_1, \psi_2, \rho_r, \sigma_r^2) \)
Table 1: Parameter values along the non-identification curve

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( \psi_1 )</th>
<th>( \psi_2 )</th>
<th>( \rho )</th>
<th>( \sigma^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_0 )</td>
<td>1.500000</td>
<td>0.125000</td>
<td>0.750000</td>
<td>0.400000</td>
</tr>
<tr>
<td>( \theta_1 )</td>
<td>1.549721</td>
<td>0.038640</td>
<td>0.744275</td>
<td>0.393917</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>1.599427</td>
<td>-0.047694</td>
<td>0.738284</td>
<td>0.387601</td>
</tr>
<tr>
<td>( \theta_3 )</td>
<td>1.649117</td>
<td>-0.133998</td>
<td>0.732008</td>
<td>0.381039</td>
</tr>
<tr>
<td>( \theta_4 )</td>
<td>1.698787</td>
<td>-0.220271</td>
<td>0.725425</td>
<td>0.374217</td>
</tr>
<tr>
<td>( \theta_5 )</td>
<td>1.748437</td>
<td>-0.306507</td>
<td>0.718514</td>
<td>0.367121</td>
</tr>
<tr>
<td>( \theta_6 )</td>
<td>1.798064</td>
<td>-0.392704</td>
<td>0.711250</td>
<td>0.359735</td>
</tr>
<tr>
<td>( \theta_7 )</td>
<td>1.847666</td>
<td>-0.478855</td>
<td>0.703605</td>
<td>0.352043</td>
</tr>
<tr>
<td>( \theta_8 )</td>
<td>1.897238</td>
<td>-0.564956</td>
<td>0.695549</td>
<td>0.344028</td>
</tr>
<tr>
<td>( \theta_9 )</td>
<td>1.946777</td>
<td>-0.651001</td>
<td>0.687048</td>
<td>0.335670</td>
</tr>
<tr>
<td>( \theta_{10} )</td>
<td>1.996280</td>
<td>-0.736982</td>
<td>0.678066</td>
<td>0.326951</td>
</tr>
</tbody>
</table>

Panel (a). Direction 1

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( \psi_1 )</th>
<th>( \psi_2 )</th>
<th>( \rho )</th>
<th>( \sigma^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_{-1} )</td>
<td>1.450265</td>
<td>0.211383</td>
<td>0.755476</td>
<td>0.405862</td>
</tr>
<tr>
<td>( \theta_{-2} )</td>
<td>1.400518</td>
<td>0.297788</td>
<td>0.760718</td>
<td>0.411514</td>
</tr>
<tr>
<td>( \theta_{-3} )</td>
<td>1.350760</td>
<td>0.384213</td>
<td>0.765741</td>
<td>0.416966</td>
</tr>
<tr>
<td>( \theta_{-4} )</td>
<td>1.300992</td>
<td>0.470655</td>
<td>0.770558</td>
<td>0.422299</td>
</tr>
<tr>
<td>( \theta_{-5} )</td>
<td>1.251214</td>
<td>0.557112</td>
<td>0.775182</td>
<td>0.427312</td>
</tr>
<tr>
<td>( \theta_{-6} )</td>
<td>1.201428</td>
<td>0.643585</td>
<td>0.779625</td>
<td>0.432224</td>
</tr>
<tr>
<td>( \theta_{-7} )</td>
<td>1.151635</td>
<td>0.730071</td>
<td>0.783895</td>
<td>0.436972</td>
</tr>
<tr>
<td>( \theta_{-8} )</td>
<td>1.101834</td>
<td>0.816569</td>
<td>0.788004</td>
<td>0.441565</td>
</tr>
<tr>
<td>( \theta_{-9} )</td>
<td>1.052027</td>
<td>0.903078</td>
<td>0.791960</td>
<td>0.446010</td>
</tr>
<tr>
<td>( \theta_{-10} )</td>
<td>1.002214</td>
<td>0.989598</td>
<td>0.795772</td>
<td>0.450313</td>
</tr>
</tbody>
</table>

Panel (b). Direction 2

Note. \( \theta_j \) are points on the curve that are \( |j| \times 10^4 \) steps away from \( \theta_0 \) in respective directions. Results are rounded to the nearest sixth digit to the right of decimal.
Table 2: Deviations of spectra across frequencies (direction 1)

<table>
<thead>
<tr>
<th>Spectral density matrix element number</th>
<th>(1,1)</th>
<th>(2,1)</th>
<th>(3,1)</th>
<th>(4,1)</th>
<th>(2,2)</th>
<th>(3,2)</th>
<th>(4,2)</th>
<th>(3,3)</th>
<th>(4,3)</th>
<th>(4,4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum absolute deviations across frequencies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\theta_1)</td>
<td>1.2E-06</td>
<td>2.3E-07</td>
<td>9.0E-07</td>
<td>2.3E-07</td>
<td>1.2E-07</td>
<td>1.3E-07</td>
<td>1.2E-07</td>
<td>6.0E-07</td>
<td>1.3E-07</td>
<td>1.2E-07</td>
</tr>
<tr>
<td>(\theta_2)</td>
<td>3.8E-07</td>
<td>1.6E-06</td>
<td>1.5E-06</td>
<td>1.6E-06</td>
<td>5.1E-07</td>
<td>1.0E-06</td>
<td>5.1E-07</td>
<td>1.5E-06</td>
<td>1.0E-06</td>
<td>5.1E-07</td>
</tr>
<tr>
<td>(\theta_3)</td>
<td>8.8E-07</td>
<td>1.8E-06</td>
<td>2.4E-06</td>
<td>1.8E-06</td>
<td>4.3E-07</td>
<td>1.1E-06</td>
<td>4.3E-07</td>
<td>2.1E-06</td>
<td>1.1E-06</td>
<td>4.3E-07</td>
</tr>
<tr>
<td>(\theta_4)</td>
<td>2.1E-06</td>
<td>2.0E-06</td>
<td>3.3E-06</td>
<td>2.0E-06</td>
<td>3.2E-07</td>
<td>1.2E-06</td>
<td>3.2E-07</td>
<td>2.6E-06</td>
<td>1.2E-06</td>
<td>3.2E-07</td>
</tr>
<tr>
<td>(\theta_5)</td>
<td>3.4E-06</td>
<td>2.2E-06</td>
<td>4.2E-06</td>
<td>2.2E-06</td>
<td>2.8E-07</td>
<td>1.3E-06</td>
<td>2.8E-07</td>
<td>3.2E-06</td>
<td>1.3E-06</td>
<td>2.8E-07</td>
</tr>
<tr>
<td>(\theta_6)</td>
<td>4.8E-06</td>
<td>2.4E-06</td>
<td>5.1E-06</td>
<td>2.4E-06</td>
<td>4.3E-07</td>
<td>1.4E-06</td>
<td>4.3E-07</td>
<td>3.7E-06</td>
<td>1.4E-06</td>
<td>4.3E-07</td>
</tr>
<tr>
<td>(\theta_7)</td>
<td>6.1E-06</td>
<td>2.6E-06</td>
<td>6.0E-06</td>
<td>2.6E-06</td>
<td>6.1E-07</td>
<td>1.4E-06</td>
<td>6.1E-07</td>
<td>4.3E-06</td>
<td>1.4E-06</td>
<td>6.1E-07</td>
</tr>
<tr>
<td>(\theta_8)</td>
<td>7.5E-06</td>
<td>2.7E-06</td>
<td>6.9E-06</td>
<td>2.7E-06</td>
<td>7.9E-07</td>
<td>1.4E-06</td>
<td>7.9E-07</td>
<td>4.8E-06</td>
<td>1.4E-06</td>
<td>7.9E-07</td>
</tr>
<tr>
<td>(\theta_9)</td>
<td>8.8E-06</td>
<td>2.8E-06</td>
<td>7.8E-06</td>
<td>2.8E-06</td>
<td>1.0E-06</td>
<td>1.4E-06</td>
<td>1.0E-06</td>
<td>5.3E-06</td>
<td>1.4E-06</td>
<td>1.0E-06</td>
</tr>
<tr>
<td>(\theta_{10})</td>
<td>1.0E-05</td>
<td>2.9E-06</td>
<td>8.6E-06</td>
<td>2.9E-06</td>
<td>1.2E-06</td>
<td>1.4E-06</td>
<td>1.2E-06</td>
<td>5.7E-06</td>
<td>1.4E-06</td>
<td>1.2E-06</td>
</tr>
</tbody>
</table>

| Maximum relative deviations across frequencies |
| \(\theta_1\) | 5.4E-08 | 2.9E-08 | 6.4E-08 | 2.9E-08 | 7.1E-09 | 2.4E-08 | 6.4E-08 | 6.6E-08 | 2.4E-08 | 6.4E-08 |
| \(\theta_2\) | 1.7E-08 | 2.0E-07 | 1.1E-07 | 2.0E-07 | 2.2E-09 | 1.9E-07 | 1.5E-07 | 1.6E-07 | 1.9E-07 | 1.5E-07 |
| \(\theta_3\) | 3.9E-08 | 2.3E-07 | 1.7E-07 | 2.3E-07 | 1.8E-09 | 2.1E-07 | 1.3E-07 | 2.3E-07 | 2.1E-07 | 1.3E-07 |
| \(\theta_4\) | 9.6E-08 | 2.6E-07 | 2.3E-07 | 2.6E-07 | 1.4E-09 | 2.3E-07 | 9.8E-08 | 2.9E-07 | 2.3E-07 | 9.8E-08 |
| \(\theta_5\) | 1.5E-07 | 2.8E-07 | 3.0E-07 | 2.8E-07 | 3.6E-08 | 2.4E-07 | 1.9E-07 | 3.5E-07 | 2.4E-07 | 1.9E-07 |
| \(\theta_6\) | 2.1E-07 | 3.0E-07 | 3.6E-07 | 3.0E-07 | 4.5E-08 | 2.5E-07 | 2.7E-07 | 4.1E-07 | 2.5E-07 | 2.7E-07 |
| \(\theta_7\) | 2.7E-07 | 3.2E-07 | 4.3E-07 | 3.2E-07 | 5.4E-08 | 2.6E-07 | 3.7E-07 | 4.7E-07 | 2.6E-07 | 3.7E-07 |
| \(\theta_8\) | 3.3E-07 | 3.4E-07 | 4.9E-07 | 3.4E-07 | 6.2E-08 | 2.6E-07 | 4.7E-07 | 5.2E-07 | 2.6E-07 | 4.7E-07 |
| \(\theta_9\) | 4.0E-07 | 3.6E-07 | 5.5E-07 | 3.6E-07 | 7.1E-08 | 2.6E-07 | 5.7E-07 | 5.8E-07 | 2.6E-07 | 5.7E-07 |
| \(\theta_{10}\) | 4.6E-07 | 3.7E-07 | 6.2E-07 | 3.7E-07 | 8.0E-08 | 2.6E-07 | 6.8E-07 | 6.3E-07 | 2.6E-07 | 6.8E-07 |

Note. \(\theta_1\) to \(\theta_{10}\) are as defined in Table 1. The upper triangular elements are omitted due to symmetry.
Table 3: Deviations of spectra across frequencies (direction 2)

<table>
<thead>
<tr>
<th>Spectral density matrix element number</th>
<th>(1,1)</th>
<th>(2,1)</th>
<th>(3,1)</th>
<th>(4,1)</th>
<th>(2,2)</th>
<th>(3,2)</th>
<th>(4,2)</th>
<th>(3,3)</th>
<th>(4,3)</th>
<th>(4,4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum absolute deviations across frequencies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\theta_{-1}$</td>
<td>8.3E-07</td>
<td>3.6E-08</td>
<td>4.9E-07</td>
<td>3.6E-08</td>
<td>1.4E-07</td>
<td>9.7E-08</td>
<td>1.4E-07</td>
<td>2.5E-07</td>
<td>9.7E-08</td>
<td>1.4E-07</td>
</tr>
<tr>
<td>$\theta_{-2}$</td>
<td>1.7E-06</td>
<td>8.7E-08</td>
<td>9.9E-07</td>
<td>8.7E-08</td>
<td>2.7E-07</td>
<td>1.8E-07</td>
<td>2.7E-07</td>
<td>5.2E-07</td>
<td>1.8E-07</td>
<td>2.7E-07</td>
</tr>
<tr>
<td>$\theta_{-3}$</td>
<td>4.0E-06</td>
<td>2.3E-07</td>
<td>2.0E-06</td>
<td>2.3E-07</td>
<td>5.0E-07</td>
<td>3.3E-07</td>
<td>5.0E-07</td>
<td>1.1E-06</td>
<td>3.3E-07</td>
<td>5.0E-07</td>
</tr>
<tr>
<td>$\theta_{-4}$</td>
<td>1.7E-06</td>
<td>8.7E-08</td>
<td>9.9E-07</td>
<td>8.7E-08</td>
<td>2.7E-07</td>
<td>1.8E-07</td>
<td>2.7E-07</td>
<td>5.2E-07</td>
<td>1.8E-07</td>
<td>2.7E-07</td>
</tr>
<tr>
<td>$\theta_{-5}$</td>
<td>4.0E-06</td>
<td>2.3E-07</td>
<td>2.0E-06</td>
<td>2.3E-07</td>
<td>5.0E-07</td>
<td>3.3E-07</td>
<td>5.0E-07</td>
<td>1.1E-06</td>
<td>3.3E-07</td>
<td>5.0E-07</td>
</tr>
<tr>
<td>$\theta_{-6}$</td>
<td>4.9E-06</td>
<td>4.2E-07</td>
<td>3.1E-06</td>
<td>4.2E-07</td>
<td>6.8E-07</td>
<td>4.4E-07</td>
<td>6.8E-07</td>
<td>1.8E-06</td>
<td>4.4E-07</td>
<td>6.8E-07</td>
</tr>
<tr>
<td>$\theta_{-7}$</td>
<td>5.7E-06</td>
<td>5.3E-07</td>
<td>3.7E-06</td>
<td>5.3E-07</td>
<td>7.6E-07</td>
<td>4.9E-07</td>
<td>7.6E-07</td>
<td>2.1E-06</td>
<td>4.9E-07</td>
<td>7.6E-07</td>
</tr>
<tr>
<td>$\theta_{-8}$</td>
<td>6.5E-06</td>
<td>6.6E-07</td>
<td>4.2E-06</td>
<td>6.6E-07</td>
<td>8.3E-07</td>
<td>5.3E-07</td>
<td>8.3E-07</td>
<td>2.5E-06</td>
<td>5.3E-07</td>
<td>8.3E-07</td>
</tr>
<tr>
<td>$\theta_{-9}$</td>
<td>7.3E-06</td>
<td>7.9E-07</td>
<td>4.8E-06</td>
<td>7.9E-07</td>
<td>9.0E-07</td>
<td>5.7E-07</td>
<td>9.0E-07</td>
<td>2.9E-06</td>
<td>5.7E-07</td>
<td>9.0E-07</td>
</tr>
<tr>
<td>$\theta_{-10}$</td>
<td>8.1E-06</td>
<td>9.4E-07</td>
<td>5.4E-06</td>
<td>9.4E-07</td>
<td>9.6E-07</td>
<td>6.0E-07</td>
<td>9.6E-07</td>
<td>3.3E-06</td>
<td>6.0E-07</td>
<td>9.6E-07</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Maximum relative deviations across frequencies</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_{-1}$</td>
<td>3.7E-08</td>
<td>4.5E-09</td>
<td>3.5E-08</td>
<td>4.5E-09</td>
<td>3.6E-09</td>
<td>4.9E-08</td>
<td>6.2E-08</td>
<td>2.7E-08</td>
<td>4.9E-08</td>
<td>6.2E-08</td>
</tr>
<tr>
<td>$\theta_{-2}$</td>
<td>7.4E-08</td>
<td>1.1E-08</td>
<td>7.1E-08</td>
<td>1.1E-08</td>
<td>7.5E-09</td>
<td>9.4E-08</td>
<td>1.2E-07</td>
<td>5.7E-08</td>
<td>9.4E-08</td>
<td>1.2E-07</td>
</tr>
<tr>
<td>$\theta_{-3}$</td>
<td>1.1E-07</td>
<td>1.9E-08</td>
<td>1.1E-07</td>
<td>1.9E-08</td>
<td>1.2E-08</td>
<td>1.4E-07</td>
<td>1.8E-07</td>
<td>8.7E-08</td>
<td>1.4E-07</td>
<td>1.8E-07</td>
</tr>
<tr>
<td>$\theta_{-4}$</td>
<td>1.5E-07</td>
<td>2.8E-08</td>
<td>1.4E-07</td>
<td>2.8E-08</td>
<td>1.6E-08</td>
<td>1.8E-07</td>
<td>2.3E-07</td>
<td>1.2E-07</td>
<td>1.8E-07</td>
<td>2.3E-07</td>
</tr>
<tr>
<td>$\theta_{-5}$</td>
<td>1.8E-07</td>
<td>4.0E-08</td>
<td>1.8E-07</td>
<td>4.0E-08</td>
<td>2.0E-08</td>
<td>2.1E-07</td>
<td>2.8E-07</td>
<td>1.6E-07</td>
<td>2.1E-07</td>
<td>2.8E-07</td>
</tr>
<tr>
<td>$\theta_{-6}$</td>
<td>2.2E-07</td>
<td>5.2E-08</td>
<td>2.2E-07</td>
<td>5.2E-08</td>
<td>2.5E-08</td>
<td>2.5E-07</td>
<td>3.3E-07</td>
<td>1.9E-07</td>
<td>2.5E-07</td>
<td>3.3E-07</td>
</tr>
<tr>
<td>$\theta_{-7}$</td>
<td>2.5E-07</td>
<td>6.7E-08</td>
<td>2.6E-07</td>
<td>6.7E-08</td>
<td>2.9E-08</td>
<td>2.8E-07</td>
<td>3.7E-07</td>
<td>2.3E-07</td>
<td>2.8E-07</td>
<td>3.7E-07</td>
</tr>
<tr>
<td>$\theta_{-8}$</td>
<td>2.9E-07</td>
<td>8.2E-08</td>
<td>3.0E-07</td>
<td>8.2E-08</td>
<td>3.4E-08</td>
<td>3.1E-07</td>
<td>4.1E-07</td>
<td>2.7E-07</td>
<td>3.1E-07</td>
<td>4.1E-07</td>
</tr>
<tr>
<td>$\theta_{-9}$</td>
<td>3.3E-07</td>
<td>1.0E-07</td>
<td>3.4E-07</td>
<td>1.0E-07</td>
<td>3.9E-08</td>
<td>3.3E-07</td>
<td>4.5E-07</td>
<td>3.2E-07</td>
<td>3.3E-07</td>
<td>4.5E-07</td>
</tr>
<tr>
<td>$\theta_{-10}$</td>
<td>3.6E-07</td>
<td>1.2E-07</td>
<td>3.8E-07</td>
<td>1.2E-07</td>
<td>4.3E-08</td>
<td>3.6E-07</td>
<td>4.8E-07</td>
<td>3.6E-07</td>
<td>3.6E-07</td>
<td>4.8E-07</td>
</tr>
</tbody>
</table>

Note. $\theta_{-1}$ to $\theta_{-10}$ are as defined in Table 1. The upper triangular elements are omitted due to symmetry.