

# Winding Number Criterion for Existence and Uniqueness of Equilibrium in Linear Rational Expectations Models

A. Onatski  
Department of Economics  
Columbia University  
e-mail: ao2027@columbia.edu

August 25, 2003

## Abstract

This paper develops a simple geometric criterion for the existence and uniqueness of equilibrium in linear rational expectations (LRE) models. The class of models considered is very general and includes models with infinitely many lags and/or leads of endogenous variables. I show that the criterion generalizes Whiteman (1983) and is tightly related to the familiar Blanchard and Kahn (1980) conditions for the existence and uniqueness of equilibrium in LRE models. An application of the criterion to an analysis of the monetary policy robustness under the uncertainty about the standard New Keynesian model of the economy is given.

## 1. Introduction

This paper develops a simple geometric criterion for the existence and uniqueness of equilibrium in linear rational expectations (LRE) models of the form:

$$\sum_{j=-\infty}^{+\infty} A_j E_t x_{t-j} = z_t. \quad (1)$$

Here  $x$  and  $z$  are  $n$ -dimensional vectors of endogenous and exogenous variables respectively,  $A_j$  are  $n$ -dimensional square matrices of fixed coefficients and  $E_t$  represents mathematical expectation given information available at time  $t$  including the model's structure and all past and present realizations of the exogenous and endogenous processes.

Since the number of lags and leads of the endogenous variables in model (1) may be infinite, it does not, in general, reduce to any of the particular models

studied in the literature. Therefore, criteria of solvability developed in Blanchard and Kahn (1980), Hansen and Sargent (1980), Whiteman (1983), Anderson and Moore (1985), Broze, Gourieroux, and Szafarz (1995), Sims (1996), Klein (2000), Binder and Pesaran (1997), King and Watson (1998), McCallum (1998), Zdrozny (1998), and Uhlig (1999) cannot be applied. On the contrary, equations (1) nest a variety of models examined in the literature so that solvability of these models may be analyzed with the help of the criterion developed below.

There exist several reasons why studying models with an infinite number of lags and leads of endogenous variables may be interesting to economists. From a purely methodological perspective, analyzing more general models gives new insights about methods developed under more restrictive assumptions and allows their deeper interpretation. In the paper, I discuss in detail links between the new criterion and those of Whiteman (1983) and Blanchard and Kahn (1980). This discussion helps to look at the previously developed criteria from a new angle and to better see connections between them.

More substantively, models of form (1) may result from the first order conditions of optimization problems with non-parametric description of habit formation, adjustment costs, time discount schemes, and time-to-build constraints. To give a simple example, consider a monopolist deciding on how much capital to rent when production takes some time. In the spirit of Sargent (1987), we formalize this problem as

$$\max E_0 \sum_{t=1}^{\infty} \beta^t \left[ p_t y_t - r k_t - \frac{b}{2} (C(L) k_t)^2 \right]$$

subject to demand and production function constraints:

$$\begin{aligned} p_t &= 1 - \frac{a}{2} y_t + \varepsilon_t, \\ y_t &= D(L) k_t. \end{aligned}$$

Here  $y$  is the monopolist's production,  $p$  is the price that she is facing,  $k$  is the amount of capital employed,  $r$  is the rental price of capital assumed to be constant over time, and  $\varepsilon$  is the white noise demand shock. The term  $\frac{b}{2} (C(L) k_t)^2$  in the objective function corresponds to a generalized quadratic adjustment cost of capital. Capital is the only factor of production, and  $y$  is assumed to be equal to a distributed lag of  $k$  because the production takes time.

Unless we are modeling a production process for a specific industry using a particular technology, the number of lags that we must include into the lag polynomials  $C(L)$  and  $D(L)$  is disputable. To avoid ad hoc assumptions, it is reasonable to formulate the problem without restrictions on the functional form of  $C(L)$  and  $D(L)$ , so that  $C(L) = \sum_{i=0}^{\infty} c_i L^i$  and  $D(L) = \sum_{i=0}^{\infty} d_i L^i$ .

Assuming that the monopolist has been in business a long time, an econometrician wants to model the choice of capital by a stationary stochastic process. What are the restrictions on the process implied by the fact that the monopolist is maximizing her profit? It is shown in the Appendix that the restriction sought by the econometrician is

$$E_t \{ bC(L)C(L^{-1}) + aD(L)D(L^{-1}) \} \hat{k}_t = d_0 \varepsilon_t,$$

where  $\hat{k}_t$  denote deviations of capital from its steady-state amount. This equation can be cast in form (1) with  $x_t = \hat{k}_t$ ,  $z_t = d_0\varepsilon_t$ , and  $A_j = A_{-j} = \sum_{k=|j|}^{\infty} c_{k-|j|}c_k + \sum_{k=|j|}^{\infty} d_{k-|j|}d_k$ .

Typically, models such as this one are seen in the literature with  $C(L)$  and  $D(L)$  restricted to be either finite order lag polynomials or rational functions of  $L$ . In the former case, clearly, the first order condition of optimality uses only a finite number of lags and leads of endogenous variables. In the latter case, the first order condition can be cast in the finite lags and leads format by introducing auxiliary variables. Arguably, the main reason to restrict  $C(L)$  and  $D(L)$  to finite order polynomials or rational functions is not that the essential infinity of the lags or leads is too abstract to be useful in the economic research, but that the infinite models are believed to be too difficult to analyze rigorously so that the gain from avoiding the ad hoc assumptions about  $C$  and  $D$  may be not worth the extra complication.

This paper shows that, leaving aside degenerate models corresponding to zero measure events in the parameter space, analyzing the existence and uniqueness of equilibrium is no more difficult for the infinite LRE models than for the more conventional LRE models. In fact, when applied to the conventional models, the method proposed in the paper may be even simpler than Blanchard and Kahn's because it does not require representing a model in the canonical one-lag form which is straightforward but may be time consuming.

A young and growing field of economics where the infinite LRE models has been used for some time is the research on model uncertainty in macroeconomics and finance (see Hansen and Sargent (2002)).<sup>1</sup> This literature uses the methods of robust control to analyze optimal economic decisions when there is no single reliable model of the economy but a whole ball of potentially true alternatives in the model space. For this research, it is very convenient not to restrict the set of potentially true models to those that can be finitely parameterized.

To give an idea of why the infinite models become handy in the model uncertainty research let me consider a very stylized situation when a model of economic dynamics has the following AR(2) form:

$$x_t + a_1x_{t-1} + a_2x_{t-2} = \varepsilon_t. \quad (2)$$

The above equation can be interpreted as a reduced form of an underlying structural model of the economy including a description of a policy maker's decisions. We will assume that the policy decisions have a non-trivial effect on the equation's coefficients. Of course, the model, like any other theoretical model, is only an approximation to reality. In particular, it may even happen that although it implies stationarity of  $x_t$ , the actual dynamics of  $x_t$  is explosive.<sup>2</sup>

---

<sup>1</sup>For the papers using the infinite models see Onatski and Stock (2002), Tetlow and von zur Muehlen (2001), Onatski and Williams (in press).

<sup>2</sup>This is not just a theoretical possibility. Indeed, a potential explanation of the great inflation of the 70's is that the Fed adopted a wrong theory of the exploitable long-run trade-off between inflation and unemployment (see for example Sargent (1999)). The monetary policy based on the erroneous theory could have significantly added to the economic instability of that time.

Suppose that the policy maker is concerned that her policy may cause dynamic instability because of an unknown misspecification of the model. Then it would be helpful to know how large potentially true deviations from the AR(2) model must be before the policy implies dynamic instability. The larger the deviations necessary to “produce” the instability, the stronger the evidence in favor of the considered policy actions.

Formally, let us introduce a space of potentially true alternatives to the reference auto-regression (2) and choose a metric in this space. This can be done in many different ways. One possibility would be to consider the space of all AR(2) models of the form:

$$x_t + (a_1 + \delta_1)x_{t-1} + (a_2 + \delta_2)x_{t-2} = \varepsilon_t$$

and use as a distance between an alternative and the original models the Euclidean length of the vector  $(\delta_1, \delta_2)$ ,  $\sqrt{\delta_1^2 + \delta_2^2}$ . Another possibility would be to consider the space of all AR( $\infty$ ) models:

$$x_t + a_1x_{t-1} + a_2x_{t-2} + \Delta(L)x_t = \varepsilon_t,$$

where  $\Delta(L)$  is a potentially infinite lag polynomial with absolutely summable coefficients. In this case, the distance can be defined as the gain of the linear filter  $\Delta(L)$ , that is as  $\max_{\lambda \in [0, 2\pi]} |\Delta(e^{i\lambda})|$ .

In both cases, we are looking for the minimal size (in terms of the corresponding distance) of such deviation from the reference AR(2) model that brings at least one root of the corresponding characteristic equation inside the unit circle in the complex plane so that the instability results. The larger the minimal size of the “destabilizing” deviation, the more robust the policy associated with given coefficients  $a_1$  and  $a_2$  to potential misspecifications of the reference model (2).

It is not hard to see that the minimal size of the “destabilizing” deviation for the first description of potential alternative models is given by a solution to:

$$\begin{aligned} & \min \sqrt{\delta_1^2 + \delta_2^2}, \\ \text{s.t. } & \min \left\{ \left| \frac{-a_1 - \delta_1 \pm \sqrt{(a_1 + \delta_1)^2 - 4(a_2 + \delta_2)}}{2} \right| \right\} \leq 1. \end{aligned}$$

As is shown in the Appendix, for the second description, it is given by a solution to:

$$\min_{\lambda \in [0, 2\pi]} |1 + a_1e^{i\lambda} + a_2e^{i2\lambda}|. \quad (3)$$

Since in general it is easier to solve unconstrained minimization problems than constrained ones, restricting alternative models to the space of auto-regressions of the second order turns out to be less convenient analytically than considering alternatives of AR( $\infty$ ) form.

The criterion of the existence and uniqueness of equilibrium in the LRE models with an infinite number of lags and leads of endogenous variables is very useful

for addressing questions similar to the above one in the context of forward-looking reference models of the economy. In this paper, I consider an application of the criterion to an analysis of the robustness of monetary policy under uncertainty about a standard New Keynesian model of the economy, described by Woodford (1997) and Clarida, Gali and Gertler (1999) among others. The robustness of a particular policy is measured by the smallest size of a potential deviation from the New Keynesian model that results in dynamic instability or indeterminacy of equilibrium under the policy analyzed.

I find that simple policy rules that set the interest rate as a linear function of current inflation and the output gap are very robust to the model uncertainty for a wide range of the policy rules' coefficients. This stands in contrast to the robustness analysis of the policy rules under model uncertainty about Rudebusch and Svensson's (1999) purely backward-looking model of the economy performed by Onatski and Stock (2002). They showed that relatively more aggressive rules destabilize the economy for smaller deviations from Rudebusch and Svensson's model than relatively less aggressive rules. The difference in the results accords with the intuition that forward-looking models are in general more difficult to destabilize than models with adaptive expectations because rational agents react to would be instability by changing their actions so as to justify their beliefs in economic stability.

Before I turn to detailed discussion, let me informally describe the criterion developed in the paper. The criterion is based on the behavior of the graph of function:

$$A(\lambda) = \det \sum_{j=-\infty}^{+\infty} A_j e^{ij\lambda} \quad (4)$$

It is assumed that model (1) has no unit roots so that  $A(\lambda) \neq 0$  for any  $\lambda$ . Since  $e^{ij\lambda}$  are periodic functions with period  $2\pi$ , function  $A(\lambda)$  tracks a closed contour in the complex space when  $\lambda$  changes from 0 to  $2\pi$ . Under general conditions (see Proposition 2 below), if, moving along the contour,  $A(\lambda)$  rotates around zero counter-clockwise, then model (1) has no solutions; if it rotates around zero clockwise, then model (1) has multiple solutions; and if the contour does not encircle zero, then model (1) has a unique solution.

To illustrate the criterion let me consider a simple scalar LRE model:

$$ax_{t-1} + bx_t + cE_t x_{t+1} = z_t. \quad (5)$$

In the paper, we analyze the following two scenarios. First, both exogenous and endogenous processes  $\{z_t\}$  and  $\{x_t\}$  are assumed to be stationary and we are looking for a solution  $x_t$  that is representable as a square summable linear combination of past and present realizations of the process fundamental for  $\{z_t\}$ . Second, the exogenous and endogenous processes are assumed to satisfy asymptotic growth restrictions and the endogenous process must comply with initial conditions. The exogenous and endogenous processes are unrestricted otherwise and, in particular, may be non-stationary.

For the above example, we focus on the first scenario and look for a stationary solution to (5). The function  $A(\lambda)$  is equal to  $ae^{i\lambda} + b + ce^{-i\lambda}$ . Three different

Figure 1a:  $x_t + 2E x_{t+1} = z_t$

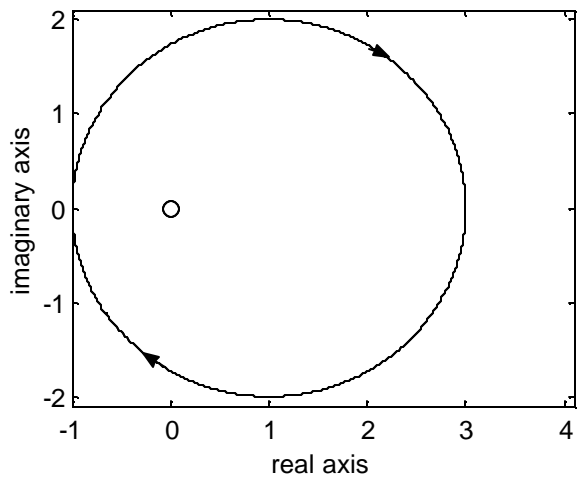


Figure 1b:  $x_{t-1} + 2x_t = z_t$

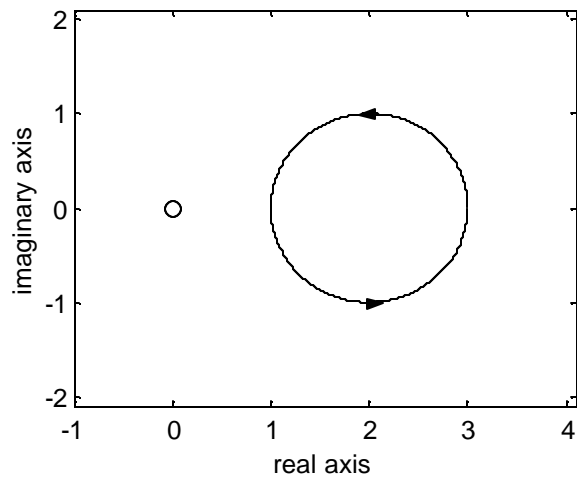


Figure 1c:  $2x_{t-1} + x_t = z_t$

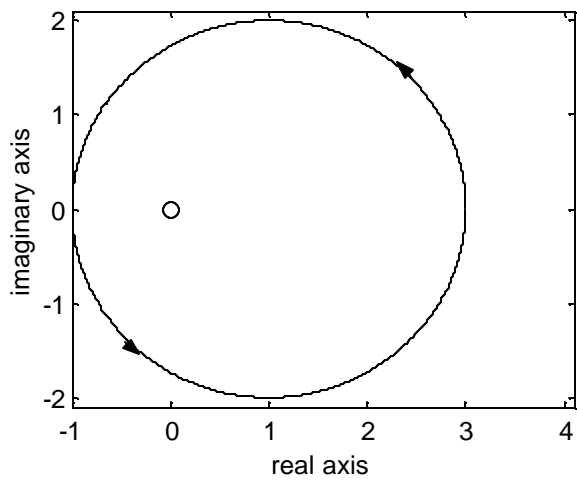
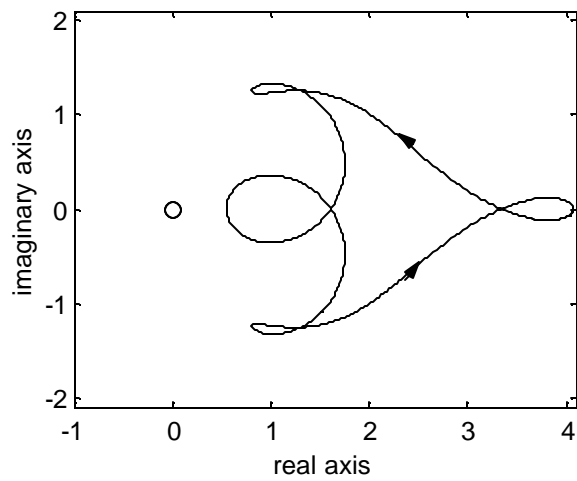


Figure 1d: Equation (4) in the text



choices of  $a, b$ , and  $c$ :  $(a, b, c) = (0, 1, 2)$ ,  $(a, b, c) = (1, 2, 0)$ , and  $(a, b, c) = (2, 1, 0)$ , correspond to situations when model (5) has, respectively, multiple, unique and no stationary solutions. The corresponding graphs of  $A(\lambda)$  are shown in figures 1a, 1b and 1c. In accordance with the criterion, the graph rotates around zero clockwise in figure 1a, does not encircle zero in figure 1b, and rotates around zero counter-clockwise in figure 1c.

To see a less trivial example consider a scalar LRE model:

$$E_t x_{t+1} - (1 - L)^{\frac{1}{3}} x_t = z_t, \quad (6)$$

where we again focus on the stationary solution. For this model, function  $A(\lambda)$  is equal to  $e^{-i\lambda} - (1 - e^{i\lambda})^{\frac{1}{3}}$ . Its graph is shown in figure 1d. It rotates around zero clockwise so that, according to the criterion, the model has multiple solutions.

The remainder of this paper is organized as follows. In section 2, I restrict attention to univariate LRE models. I show that the newly proposed criterion can be viewed as a generalization of the conditions for the existence and uniqueness of equilibrium in LRE models developed by Whiteman (1983). Section 3 describes the stationary and non-stationary frameworks of the analysis performed in the paper and formulates the new criterion for multivariate LRE models. Section 4 studies connections between the new criterion and that of Blanchard and Kahn (1980). An application to an analysis of the robustness of monetary policy to the uncertainty about a standard New Keynesian model of the economy is considered in Section 5. Section 6 concludes. All proofs are given in the Appendix.

## 2. Connections to Whiteman's criterion

The rotation criterion introduced above may look like an extravagant way to check the existence and uniqueness of equilibrium in LRE models. However, it is very close in substance to many existent criteria. Perhaps the closest, even in form, to the new criterion is the criterion established by Whiteman (1983). To build intuition, let us explore this connection in some detail using a simple univariate framework. We will formulate and prove the new criterion for multivariate LRE models and explore the connections to the multivariate version of Whiteman's criterion and to the Blanchard and Kahn (1980) criterion in the next sections.

Whiteman (1983) studies univariate models of the form:

$$\sum_{j=-n}^m A_j E_t x_{t-j} = z_t, \quad (7)$$

where  $z_t$  is a zero-mean linearly regular<sup>3</sup> weakly stationary exogenous process with known Wold representation, and  $E_t$  denotes the operation of taking the best linear

---

<sup>3</sup>A zero-mean weakly stationary process is called linearly regular if it does not contain any component that can be predicted without an error by a linear function of the process' past values (see Rozanov (1967)).

forecast given information available at time  $t$ . In the spirit of Muth (1961), he restricts attention to those solutions to (7) that belong to the space spanned by time-independent square-summable causal linear combinations of the process fundamental for  $z_t$ . He writes a polynomial:

$$z^n f(z) = z^n \sum_{j=-n}^m A_j z^j = A_{-n} \prod_{j=1}^{n+m} (1 - \rho_j z)$$

and links the existence and uniqueness of equilibrium in (7) to the number of roots  $\rho_j$  outside the unit circle. If this number is equal to  $n$ , the number of essentially expectational terms in the model's equation, then the solution is unique. If it is less than  $n$ , there exist many solutions. Finally, if it is greater than  $n$ , there is no solution for general  $z_t$ .

Our geometric criterion applied to model (7) is also related to properties of function  $f(z)$ . Note that, for the model, the function  $A(\lambda)$  defined by (4) is equal to

$$A(\lambda) = f(e^{i\lambda}).$$

The new criterion says that there exists a unique equilibrium, multiple equilibria, or no equilibrium if the winding number of  $A(\lambda)$ , defined as the number of times the graph of  $A(\lambda)$  rotates around zero counter-clockwise when  $\lambda$  goes from 0 to  $2\pi$ , is zero, positive, or negative respectively.<sup>4</sup> A moment's reflection reveals that the two criteria are identical for models of form (7). Indeed, note that

$$\log A(\lambda) = \log A_{-n} + \sum_{j=1}^{n+m} \log(1 - \rho_j e^{i\lambda}) - n \log e^{i\lambda}. \quad (8)$$

The logarithm of an arbitrary complex number  $re^{i\lambda}$  is equal to the logarithm of its modulus,  $r$ , plus its argument,  $\lambda$ , multiplied by the imaginary unit  $i$ . Therefore, the winding number of  $A(\lambda)$ ,  $\text{wind } A(\lambda)$ , can alternatively be defined as  $\frac{1}{2\pi i} \log A(\lambda)|_{\lambda=0}^{2\pi}$ . Now, from (8) we have

$$\frac{1}{2\pi i} \log A(\lambda)|_{\lambda=0}^{2\pi} = \sum_{j=1}^{n+m} \frac{1}{2\pi i} \log(1 - \rho_j e^{i\lambda})|_{\lambda=0}^{2\pi} - n \frac{1}{2\pi i} \log e^{i\lambda}|_{\lambda=0}^{2\pi}$$

and hence

$$\text{wind } A(\lambda) = \sum_{j=1}^{n+m} \text{wind}(1 - \rho_j e^{i\lambda}) - n \text{wind}(e^{i\lambda}).$$

It is not hard to see that  $\text{wind}(e^{i\lambda}) = 1$  and  $\text{wind}(1 - \rho_j e^{i\lambda})$  is equal to zero if  $\rho_j$  is inside the unit circle and equal to 1 if  $\rho_j$  is outside the unit circle. Therefore, the winding number of  $A(\lambda)$  must be equal to the number of roots  $\rho_j$  outside the unit circle minus  $n$ , and the equivalence of the two criteria becomes obvious.

---

<sup>4</sup>A clockwise rotation of the graph around zero is counted with the negative sign.

In contrast to Whiteman's criterion, ours remains valid for models with an infinite number of lags and leads of the endogenous variables. Clearly, when the number of lags and leads is infinite, it becomes impossible to compare the number of roots outside the unit circle to  $n$  because the  $n$  and possibly the number of roots outside the unit circle become infinite. However, at least as long as the coefficients of infinite LRE models remain absolutely summable, the winding number of  $A(\lambda)$  will be well defined, finite, and related to the existence and uniqueness of equilibrium in the same way as for conventional LRE models.

In mathematics, the winding number is a basic concept of the complex analysis. Its modifications have proven to be extremely useful in topology and operator theory. In engineering, a criterion of stability of linear dynamical systems based on the winding numbers, which is very close in form to the one developed in this paper, is called the Nyquist stability criterion (for a textbook treatment see Driels (1996)).

Although the winding number is less visible in economics, its generalization, called the degree of a function, lies in the foundation of the theory of general equilibrium. To see this, let me give you a glimpse of the idea of how the winding number relates to Mas-Colell's discussion (1985, pp. 45-50) of the foundations of Brouwer's fixed-point theorem. A fixed point of a function  $g(x)$  can be viewed as a zero of the vector field  $f(x) = x - g(x)$ . The fact that any continuous function from an  $n$ -dimensional ball to itself has a fixed point (the Brouwer theorem) follows from a special case of the Poincare-Hopf theorem, which says that the sum of indices of the vector field  $f(x)$  at different zeros of this field must be equal to 1. Now, the index of the vector field  $f$  at its zero,  $x_0$ , measures the number of times the vector field "turns around" near  $x_0$ . This notion can be formalized by considering a mapping  $f(x_0 + \varepsilon v) / |f(x_0 + \varepsilon v)|$  of a circle around  $x_0$  (of small radius  $\varepsilon$ ) to the unit circle. The index of the vector field is a degree of this mapping which measures how many times the circle around  $x_0$  wraps around the unit circle. It is a direct generalization of the winding number of a function.

This hasty dive to the foundations of the general equilibrium theory leaves a question unanswered: What is the deep reason why Whiteman's criterion can be reformulated in terms of winding numbers? Why do clockwise rotations correspond to multiple equilibria and counter-clockwise ones correspond to no equilibrium? I will try to address these questions below by providing an intuitive explanation for both criteria.<sup>5</sup> For the greater part of this section, I will remain in Whiteman's stationary framework of the analysis focusing on the conventional, finite lags and leads, univariate model (7).

The main principle that I employ below is to reduce model (7) to a much simpler model that will have the same (zero, one, or infinite) number of solutions as the original one. This simpler model will be primitive so that its solvability properties will be obvious and no extra analysis will be needed. For the following discussion it is convenient to represent model (7) in the form:

$$E_t A(L)x_t = z_t, \tag{9}$$

---

<sup>5</sup>To the best of my knowledge, the explanation of Whiteman's criterion proposed here is new.

where  $L$  is the lag operator and  $A(L) = \sum_{j=-n}^m A_j L^j$ .

One way to obtain a model equivalent to (9) in terms of the number of solutions is to redefine the endogenous variable. Consider a new endogenous variable  $y_t = A_+(L)x_t$ , where  $A_+(L)$  is a polynomial in the lag operator, such that the moving average representation,  $y_t = A_+(L)x_t$ , is invertible. Then there exists a one-to-one correspondence between solutions to (9) in terms of  $x_t$  and in terms of  $y_t$ . Therefore, if  $A(L)$  can be factorized as  $A_1(L)A_+(L)$ , we can reduce (9) to:

$$E_t A_1(L) y_t = z_t.$$

For the finite lags and leads LRE model that am focusing on,  $L^n A(L)$  is a finite order lag polynomial and, hence,  $A(L)$  can be factorized as

$$A(L) = A_{-n} L^{-n} \prod_{j=1}^{n+m} (1 - \rho_j L),$$

and we can take  $A_+(L)$  and  $A_1(L)$  to be equal to:

$$A_+(L) = \prod_{j: |\rho_j| < 1} (1 - \rho_j L), \quad A_1(L) = A_{-n} L^{-n} \prod_{j: |\rho_j| > 1} (1 - \rho_j L). \quad (10)$$

Although the corresponding reduced model will be of lower order than the original one, it is not reduced enough to have obvious solvability properties. Let me, therefore, describe an additional method of reducing the model.

This time, I will aim for “eliminating” those factors of  $A(L)$  that have the form of a polynomial in the inverse of the lag operator,  $L^{-1}$ . Let  $A_-(L)$  be such a polynomial. I will require that the inverse of  $A_-(L)$  be representable as a series in non-positive powers of  $L$ . For the model under study we can take  $A_-(L)$  to be equal to:

$$A_-(L) = A_{-n} \prod_{j: |\rho_j| > 1} (L^{-1} - \rho_j),$$

so that  $A(L)$  is factorized as

$$A(L) = A_-(L) A_0(L) A_+(L), \quad (11)$$

where  $A_0(L) = L^{-n+r}$ , and  $r$  is the number of roots  $\rho_j$  outside the unit circle.

Note that model (9) can be rewritten in the form:

$$A(L)x_t = z_t + \zeta_t \quad \text{for some } \zeta_t : E_t \zeta_t = 0$$

where we have replaced the projection of  $A(L)x_t$  by its value less an arbitrary element  $\zeta_t$  with zero projection. Multiplying both sides of the above equation by  $A_-^{-1}(L)$  and projecting on the space spanned by the history of  $z_t$ , we have:

$$E_t A_0(L) A_+(L) x_t = E_t A_-^{-1}(L) z_t, \quad (12)$$

where we omit term  $E_t A_-^{-1}(L)\zeta_t$  from the right hand side because  $A_-^{-1}(L)\zeta_t$  can be represented as a linear combination of the future and present realizations of  $\zeta_t$  (recall that  $A_-^{-1}(L)$  is a series in non-positive powers of  $L$ ), and therefore its projection is equal to zero by the law of iterated expectations.

Since, as has just been explained, (9) implies (12), solutions to (9) are included in the set of solutions to (12). It is perhaps surprising that the converse is also true. To see this, note that we can rewrite (12) as

$$A_0(L)A_+(L)x_t = E_t A_-^{-1}(L)z_t + \xi_t \quad \text{for some } \xi_t : E_t \xi_t = 0$$

Multiplying both sides of the above equality by  $A_-(L)$  and taking the projection, we obtain:

$$E_t A(L)x_t = E_t A_-(L)E_t A_-^{-1}(L)z_t,$$

where we omit term  $E_t A_-(L)\xi_t$  from the right hand side because  $A_-(L)\xi_t$  can be represented as a linear combination of the future and present realizations of  $\xi_t$  and therefore its projection is equal to zero by the law of iterated expectations. Let us now represent  $A_-^{-1}(L)z_t$  as a sum of two terms: a term that belongs to the space spanned by the history of  $z_t$ , say  $a_t$ , and a term orthogonal to this space, say  $b_t$ . Then, by the definition of  $a_t$ ,  $E_t A_-^{-1}(L)z_t = a_t$ , and we can rewrite the above equality as

$$E_t A(L)x_t = E_t A_-(L)a_t.$$

Now, let us substitute  $a_t = A_-^{-1}(L)z_t - b_t$  in the above equality. We have:

$$E_t A(L)x_t = E_t A_-(L) (A_-^{-1}(L)z_t - b_t) = z_t - E_t A_-(L)b_t.$$

But  $E_t A_-(L)b_t = 0$  by the same token as  $E_t A_-^{-1}(L)\zeta_t$  and  $E_t A_-(L)\xi_t$  are equal to zero. Hence,

$$E_t A(L)x_t = z_t,$$

and therefore (12) implies (9).

Substituting  $y_{t+n-r} = A_0(L)A_+(L)x_t$  into (12), we get the model

$$E_t y_{t+n-r} = E_t A_-^{-1}(L)z_t, \tag{13}$$

which is equivalent to (9) in terms of the number of solutions. This new model has transparent solvability properties. Indeed, if  $n$ , the number of essentially expectational terms in the original model's equation, is less than  $r$ , the number of roots outside the unit circle, (13) in general has no solutions because it requires a past realization of an endogenous variable to contain information about a present realization of the exogenous variable.<sup>6</sup> If  $n$  is larger than  $r$ , then there exist multiple solutions, for example, of the form:

$$y_t = E_{t-n+r} A_-^{-1}(L)z_{t-n+r} + k\varepsilon_t,$$

---

<sup>6</sup>To see this, choose  $z_t$  to be, for example, a white noise process so that  $E_t A_-^{-1}(L)z_t$  is proportional to  $z_t$ .

where  $k$  is an arbitrary scalar and  $\varepsilon_t$  is the innovation for the exogenous process  $z_t$ . Finally, if  $n = r$  then there exist a unique solution:

$$y_t = E_t A_-^{-1}(L) z_t,$$

and, in terms of the original endogenous variable,  $x_t$ ,

$$x_t = A_+^{-1}(L) E_t A_-^{-1}(L) z_t.$$

Of course, we have just derived Whiteman's criterion for the existence and uniqueness of equilibrium in models (7). Now it should become clearer why the criterion can be reformulated in terms of the winding numbers. The winding number of  $A(\lambda)$  is simply the degree of the "time shift",  $-n + r$ , of the new endogenous variable  $y_t$  in the primitive model (13) corresponding to the original one. Indeed, taking into account the factorization (11) of  $A(L)$  we can write:

$$\begin{aligned} \text{wind } A(\lambda) &= \text{wind } A_-(\lambda) + \text{wind } A_0(\lambda) + \text{wind } A_+(\lambda) \\ &= \sum_{j: |\rho_j| > 1} \text{wind}(e^{-i\lambda} - \rho_j) + \text{wind } e^{i\lambda(-n+r)} + \sum_{j: |\rho_j| < 1} \text{wind}(1 - \rho_j e^{i\lambda}). \end{aligned}$$

It is not hard to see that the first and the last terms in the above sum are equal to zero, whereas the winding number of  $e^{i\lambda(-n+r)}$  is equal to the time shift,  $-n + r$ . Therefore, the winding number of  $A(\lambda)$  is positive, zero, or negative if and only if the time shift is, respectively, into the past, zero, or into the future, and hence if and only if the original model has, respectively, no solutions, a unique solution, or multiple solutions.

The described strategy of reducing the original model to a primitive one can be generalized to infinite LRE models (1) that have absolutely summable coefficients. Below, we briefly and informally discuss ideas behind such a generalization, postponing any rigorous proofs and formulations until the next section. Note that we can rewrite model (1) as (9), where now  $A(L) = \sum_{j=-\infty}^{\infty} A_j L^j$  is represented by a series of integer powers of  $L$ . To apply the reduction strategy we need to find a factorization of  $A(L)$  into the product of  $A_-(L)$ ,  $A_0(L)$  and  $A_+(L)$ , where  $A_-(L)$  and  $A_-^{-1}(L)$  have the form of a series of non-positive powers of  $L$  with absolutely summable coefficients,  $A_+(L)$  and  $A_+^{-1}(L)$  have the form of a series of non-negative powers of  $L$  with absolutely summable coefficients,<sup>7</sup> and  $A_0(L)$  is equal to  $L^p$  where  $p$  is some integer. If such a factorization is found, then the original model is equivalent in terms of the number of solutions to primitive model (13).

Note that there exists a one-to-one correspondence between the lag operator functions  $A(L) = \sum_{j=-\infty}^{\infty} A_j L^j$  and functions of a real variable  $A(\lambda) = \sum_{j=-\infty}^{\infty} A_j e^{ij\lambda}$ . This correspondence preserves the factorizations of  $A(L)$  and  $A(\lambda)$ . For example, if

---

<sup>7</sup>The requirement of the absolute summability of the coefficients in the series representations of  $A(L)$  and the factors  $A_+(L)$  and  $A_-(L)$  is essential. This requirement guarantees that, when applied to an arbitrary weakly stationary process, these operator functions will produce a well defined weakly stationary process.

$A(\lambda)$  can be factorized as  $B(\lambda)C(\lambda)$ , then  $A(L)$  is factorizable as  $B(L)C(L)$  where  $B(L)$  and  $C(L)$  are constructed from the Fourier series of  $B(\lambda)$  and  $C(\lambda)$  by replacing powers of  $e^{i\lambda}$  with the corresponding powers of  $L$ . It turns out that any function  $A(\lambda) = \sum_{j=-\infty}^{\infty} A_j e^{ij\lambda}$  which does not take on zero values and which is such that  $\sum_{j=-\infty}^{\infty} |A_j| < \infty$  admits factorization:

$$A(\lambda) = A_-(\lambda)A_0(\lambda)A_+(\lambda),$$

such that  $A_-(\lambda)$  and  $A_-^{-1}(\lambda)$  have absolutely summable Fourier expansions with zero coefficients on the positive powers of  $e^{i\lambda}$ ;  $A_+(\lambda)$  and  $A_+^{-1}(\lambda)$  have absolutely summable Fourier expansions with zero coefficients on the negative powers of  $e^{i\lambda}$ ; and  $A_0(z) = z^p$ , where  $p$  is an integer number.<sup>8</sup> Such a factorization resembles the canonical spectral factorization (see for example Priestly (1981) pp.733-736) of the spectral density function of a stationary process which is a basic ingredient of Kolmogorov and Wiener's solutions to the problem of predicting a future value of the process. For the spectral factorization,  $A_-(\lambda)$  is the complex conjugate of  $A_+(\lambda)$  and  $A_0(z) = z^0 = 1$ . However, the spectral factorization is only valid for functions that, besides satisfying other requirements, must be real-valued, whereas  $A(\lambda)$  will typically be complex for LRE models.<sup>9</sup>

Of course, for the infinite models, the relationship between the time shift,  $p$ , in the primitive model and the roots of  $A(\lambda)$  are lost. However, the time shift remains equal to the winding number of  $A(\lambda)$ . The reason for this is that both wind  $A_-(\lambda)$  and wind  $A_+(\lambda)$  are equal to zero. Indeed, consider for example wind  $A_+(\lambda)$ . Note that the function  $A_+(z)$ , defined by replacing  $e^{i\lambda}$  in the Fourier expansion of  $A_+(\lambda)$  by a complex variable  $z$ , is analytic and has no zeros inside the unit circle on the complex plane. Hence,  $A_+(z)$  maps the region inside the unit circle to a compact region on the complex plain that does not contain zero. The unit circle itself is mapped to the boundary of this compact region,  $A_+(\lambda)$ ,  $\lambda \in [0, 2\pi]$ , which encircles only the region's inner points and, therefore, does not encircle zero. Similar arguments can be used to show that wind  $A_-(\lambda) = 0$ .

In this section, we have described the strong connections between the newly proposed criterion and Whiteman's criterion. We have used a simple univariate framework to build up intuition and to illustrate the fundamental ideas on which the new criterion is based. Since the majority of LRE models that arise in practice are multivariate, in the next section we depart from the univariate framework and formally state and prove the new criterion for multivariate models (1). We will treat both the case when solutions are required to be stationary processes with no initial conditions stated, and the case when solutions are allowed to be non-stationary and the initial conditions are specified. After the formal proof of the new criterion, we

---

<sup>8</sup>In mathematical literature the factorization is called the Wiener-Hopf factorization (see, for example, Bottcher and Silbermann 1999)

<sup>9</sup>I thank an anonymous referee for pointing out the relationship between the Wiener-Hopf factorization and the canonical spectral factorization. The existence of the Wiener-Hopf factorization can be proven with the help of techniques used by Whittle (1963, p. 26) to show the existence of the canonical spectral factorization.

will discuss its close connections to Blanchard and Kahn's (1980) criterion for the solvability of LRE models.

### 3. The new criterion for multivariate LRE models

In this section we formulate and prove the winding number criterion for multivariate LRE models of form (1), which we restate here for convenience:

$$\sum_{j=-\infty}^{+\infty} A_j E_t x_{t-j} = z_t.$$

We assume that the matrix coefficients  $A_j$  are element-by-element absolutely summable, that is

$$\sum_{j=-\infty}^{+\infty} |A_j| < \infty, \tag{14}$$

where  $|A_j|$  denote matrices with entries equal to absolute values of the corresponding entries of  $A_j$ . We will consider two different frameworks: a stationary framework where the exogenous and endogenous processes are assumed to be weakly stationary and no initial conditions for the endogenous process are given, and a non-stationary framework where initial conditions on the endogenous process are specified and it is assumed that the exogenous and endogenous processes must satisfy asymptotic growth rate restrictions but are unrestricted otherwise. For both frameworks, we will say that the model has a unique solution if for any process  $\{z_t\}$  which is admissible, that is satisfying the framework's restrictions, there exists one and only one solution to the model's equation. We will say that the model has no solutions if there exists an admissible process  $\{z_t\}$  such that the model's equation does not have solutions. We will say that there are multiple solutions to the model if it has neither no solutions nor a unique solution.

Similar to the univariate case studied in the previous section, the analysis of solvability of (1) in both frameworks is based on a factorization of function:

$$A(\lambda) = \sum_{j=-\infty}^{\infty} A_j e^{ij\lambda},$$

which is assumed not to vanish:

$$A(\lambda) \neq 0 \text{ for any } \lambda \in [0, 2\pi].$$

The factorization has the form:

$$A(\lambda) = A_-(\lambda)A_0(\lambda)A_+(\lambda), \tag{15}$$

where  $A_+(\lambda)$  and  $A_+^{-1}(\lambda)$  have absolutely summable Fourier expansions with zero coefficients on the negative powers of  $e^{i\lambda}$ ;  $A_-(\lambda)$  and  $A_-^{-1}(\lambda)$  have absolutely summable

Fourier expansions with zero coefficients on the positive powers of  $e^{i\lambda}$ ; and  $A_0(\lambda)$  has a diagonal matrix form:

$$A_0(\lambda) = \text{diag}(e^{i\lambda k_1}, e^{i\lambda k_2}, \dots, e^{i\lambda k_n}),$$

where the integers  $k_1 \leq k_2 \leq \dots \leq k_n$  are uniquely determined and called the partial indices of  $A(\lambda)$ .

Similar to the univariate case, the factorization of a matrix function  $A(\lambda)$  is closely related to the canonical spectral factorization of the spectral density of a weakly stationary linearly regular  $n$ -dimensional process. Of course, the canonical spectral factorization only holds for positive definite matrix functions, whereas factorization (15) holds for all complex-valued matrix functions with absolutely summable coefficients of their element-by-element Fourier expansions.<sup>10</sup> The existence of such a factorization was established by Gohberg and Krein (1960). A full proof of the existence can also be found in Gohberg and Feldman (1974), Theorem 2.2, Chapter VIII.

In light of the discussion in the previous section, it should not be surprising that the sum of the partial indices of  $A(\lambda)$  must be equal to the winding number of  $\det A(\lambda)$ . To see this, note that

$$\begin{aligned} \det A(\lambda) &= \det A_-(\lambda) \det A_0(\lambda) \det A_+(\lambda) \\ &= \det A_-(\lambda) e^{i\lambda(k_1+k_2+\dots+k_n)} \det A_+(\lambda) \end{aligned}$$

and, in full analogy to the univariate case,

$$\text{wind} [\det A(\lambda)] = \text{wind} [e^{i\lambda(k_1+k_2+\dots+k_n)}] = k_1 + k_2 + \dots + k_n.$$

Let us now turn to the analysis of the existence and uniqueness of equilibrium in model (1). We start from our stationary framework.

### 3.1. Stationary framework

We assume that the exogenous process  $\{z_t\}$  is a weakly stationary process with known coefficients of Wold's representation and  $E_t$  denotes the operation of taking the best linear forecast based on information available at time  $t$  that consists of all past and present realizations (up to and including time  $t$ ) of the exogenous and endogenous processes. Following Whiteman (1983), we restrict attention to those solutions  $\{x_t\}$  to the model that can be represented as time-invariant square summable linear combination of past and present realizations of the process fundamental for  $\{z_t\}$ , where the square summability of the representation's matrix coefficients is understood in the element-by-element sense. The convergence of the infinite sum in the model's equation is understood in the mean square sense.

A proof of the following proposition can be found in the Appendix.

---

<sup>10</sup>Note, however, that in contrast to factorization (15), the existence of the canonical spectral factorization does not require the positive definite matrix  $A(\lambda)$  to have an absolutely summable Fourier expansion.

**Proposition 1.** *Model (1) has a unique solution if all partial indices of  $A(\lambda)$  are equal to zero. It has no solutions if at least one of the partial indices of  $A(\lambda)$  is positive. The model has many solutions if all partial indices of  $A(\lambda)$  are non-positive and at least one of them is negative.*

In the situation where there exists a unique solution to (1), this solution can be analytically expressed through the factors in the factorization of  $A(\lambda)$ . Let a Wold decomposition of the exogenous process  $z_t$  and the spectral representation of an  $m$ -dimensional process  $\{\varepsilon_t\}$  fundamental for  $z_t$  be  $z_t = \sum_{i=0}^{\infty} C_i \varepsilon_{t-i}$  and  $\varepsilon_t = \int_{-\pi}^{\pi} e^{it\lambda} \Phi(d\lambda)$  respectively.<sup>11</sup> Let us also define an annihilation operator,  $[\cdot]_+$ , acting on matrix functions  $F(\lambda)$  which are element-by-element square integrable over  $\lambda \in [-\pi, \pi]$  as  $[F(\lambda)]_+ = \sum_{j=0}^{\infty} F_j e^{ij\lambda}$ , where  $F(\lambda) = \sum_{j=-\infty}^{\infty} F_j e^{ij\lambda}$  is the Fourier expansion of  $F(\lambda)$ . Then, as is shown in the Appendix, the unique solution to (1) has the form:

$$x_t = \int_{-\pi}^{\pi} e^{it\lambda} A_+^{-1}(\lambda) [A_-^{-1}(\lambda) C(\lambda)]_+ \Phi(d\lambda),$$

where  $C(\lambda) = \sum_{j=0}^{\infty} C_j e^{ij\lambda}$ .

In practice, to be able to apply the above formula we need to construct factorization (15) of  $A(\lambda)$ . It is relatively easy to do this for rational functions (see Clancey and Gohberg 1981) which correspond to conventional LRE models with a finite number of lags and leads of endogenous variables. Hence, the majority of special models studied in the literature can be solved with the help of the factorization technique described above. Unfortunately, construction of the factorizations for functions  $A(\lambda)$  that correspond to “infinite” LRE models is a highly non-trivial task. Note, however, that to analyze the existence and uniqueness of solutions to models (1) we do not need to construct the factorizations explicitly. All we need to know is whether all partial indices of  $A(\lambda)$  are zero, non-positive, or neither of the above. As we show below, the partial indices of  $A(\lambda)$  can be effectively analyzed with the help of the winding number of  $A(\lambda)$ .

Had all partial indices of  $A(\lambda)$  been of the same sign or zero we would have been able to distinguish between the three cases described in Proposition 1 by computing  $\text{wind det } A(\lambda)$ . Indeed, in such a case, a zero winding number implies that all indices are zero, a positive winding number indicates that at least one partial index is positive, and a negative winding number corresponds to all partial indices being non-positive and at least one of them being negative.

Luckily, it turns out that almost all functions  $A(\lambda)$  have partial indices of the same sign or zero. More precisely, let us denote a normed space of  $n \times n$  matrix functions over  $\lambda \in [-\pi, \pi]$  with an absolutely summable Fourier series and the maximum of the element by element supremums norm as  $W_{n \times n}$ . Then, according to Theorem 2.1 and Corollary 2.2 in Chapter 10 of Clancy and Gohberg (1981), the set

---

<sup>11</sup>We allow the dimensionality of the fundamental process,  $m$ , to be less than the dimensionality of the exogenous process,  $n$ . In particular, some of the equations from the  $n$  equations described by (1) may be identities not subject to any random shocks. For good expositions of the spectral analysis of stationary processes see Rozanov (1963) and Priestley (1981).

of functions  $A(\lambda)$  whose partial indices are of the same sign or zero includes an open and everywhere dense set in  $W_{n \times n}$ . Hence, if we have a LRE model such that some of the corresponding partial indices are positive and some are negative then we can slightly perturb parameters  $A_j$  to get a new model with the corresponding partial indices of the same sign or zero. Moreover, partial indices associated with this new model will be stable under small perturbations of its parameters.

We will, therefore, define a generic LRE model as a model (1) that corresponds to  $A(\lambda)$  whose partial indices are of the same sign or zero. We have just established the following winding number criterion for the existence and uniqueness of equilibrium in multivariate LRE models:

**Proposition 2.** *A generic LRE model (1) has a unique solution, multiple solutions, or, in general, no solutions if and only if the winding number of  $\det A(\lambda)$  is equal to zero, less than zero, or greater than zero respectively.*

Note that since for the univariate models considered in the previous section,  $A(\lambda)$  has only one partial index, the univariate models are always generic. Unfortunately, there is no guarantee that non-generic models will not be encountered in practice. In fact, Dotsey and Boyd (1990) give examples of simple economic models that are not generic. In the next section, I show that the genericity of LRE models is equivalent to relatively non-restrictive familiar regulatory conditions for some special models studied in the literature.

It is interesting to compare the above criterion to the multivariate version of Whiteman's (1983) criterion. In chapter 4 of his book, Whiteman studies multivariate LRE models:

$$\sum_{j=-l}^m A_j E_t x_{t-j} = z_t.$$

He claims that if  $A_{-l}$  is of full rank, the roots of  $\det \left[ z^l \sum_{j=-l}^m A_j z^j \right]$  are distinct, and  $r * n$  of these roots are inside the unit circle while the others lie outside the unit circle, then (i) if  $r < l$ , there are many solutions to the model, (ii) if  $r = l$ , there is one solution to the model, and (iii) if  $r > l$ , there is no solution to the model.

Similar to the univariate case analyzed in the previous section, the winding number of  $A(\lambda)$  is equal to the number of roots of  $\det \left[ z^l \sum_{j=-l}^m A_j z^j \right]$  inside the unit circle minus  $l * n$  (where  $n$  is the dimensionality of the vector processes  $\{x_t\}$  and  $\{z_t\}$ ). Hence, the statements (i), (ii), and (iii) of the Whiteman criterion can be reformulated as: (i) if  $\text{wind} \det A(\lambda) < 0$ , there are many solutions to the model, (ii) if  $\text{wind} \det A(\lambda) = 0$ , there is one solution to the model, and (iii) if  $\text{wind} \det A(\lambda) > 0$ , there is no solution to the model.

Furthermore, Whiteman's conditions, that  $A_{-l}$  is of full rank and all roots of  $\det \left[ z^l \sum_{j=-l}^m A_j z^j \right]$  are distinct, imply that the LRE model is generic. Suppose not, so that some of the partial indices of  $A(\lambda)$  are negative and some are positive. As we show below, this assumption leads to a contradiction. Indeed, from Proposition 1 we know that in such a case the model has no solutions. Therefore, for Whiteman's

criterion to hold, it must be the case that  $r > l$ . The following example shows that this is not always true.

Consider a LRE model:

$$\begin{pmatrix} -\frac{1}{4} & 0 \\ 0 & 1 \end{pmatrix} E_t x_{t+1} + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} x_{t-1} = z_t.$$

Note that  $A_{-1}$  is of full rank and that  $\det \left[ z \sum_{j=-1}^1 A_j z^j \right] = z^2 - \frac{1}{4}$  has two distinct roots (hence,  $r = 1$ ), both of which are inside the unit circle, so that the model satisfies Whiteman's conditions. The factorization of  $A(\lambda)$  corresponding to this model is

$$A(\lambda) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e^{i\lambda} & 0 \\ 0 & e^{-i\lambda} \end{pmatrix} \begin{pmatrix} 1 - \frac{1}{4}e^{-i2\lambda} & 0 \\ 0 & 1 \end{pmatrix}.$$

The two partial indices are 1 and -1. Therefore, the model is non-generic, but  $r = 1$ . Hence, Whiteman's conditions must imply genericity.

The converse is not true. For example, there exist plenty of conventional finite lags and leads LRE models that are generic, but do not satisfy Whiteman's condition that  $A_{-l}$  be non-singular. One of such models is

$$\begin{pmatrix} -\frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix} E_t x_{t+1} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} x_t + \begin{pmatrix} 0 & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} x_{t-1} = z_t.$$

The above discussion shows that Whiteman's criterion for existence and uniqueness of equilibrium in multivariate LRE models is a special case of the winding number criterion. First, our condition that the LRE model is generic is weaker than Whiteman's conditions that  $A_{-l}$  be non-singular and that all roots of the determinant polynomial  $\det \left[ z^l \sum_{j=-l}^m A_j z^j \right]$  be distinct.<sup>12</sup> Hence, our criterion generalizes Whiteman's for conventional finite lags and leads LRE models. Second, the winding number criterion remains valid for the infinite lags and leads LRE models, which is not the case for Whiteman's criterion.

In practice, the winding number criterion can be checked numerically in a straightforward manner. Computation of the winding number of a complex function  $A(\lambda)$  amounts to counting the number of times that  $\text{Re } A(\lambda)$  changes its sign given that  $\text{Im } A(\lambda)$  is positive. The latter can be directly programmed as a function of the roots of  $\text{Re } A(\lambda)$  and  $\text{Im } A(\lambda)$  and the value of  $A(\lambda)$  at zero frequency. In many relatively less complex cases, the criterion can be checked visually by inspecting the graph of  $\det A(\lambda)$ .

The criterion sheds some light on the validity of using finite lag/lead approximations to study the existence and uniqueness of equilibrium in (1). Restricting attention to generic models, the existence and uniqueness properties of equilibrium for the approximation will not change as long as the corresponding winding number is the same as that for the approximated model. Note that since the winding number

---

<sup>12</sup>Whiteman's conditions are very useful since they provide a simple way to check the genericity of a model. The genericity can also be verified directly. An example of such a direct verification is given in the Appendix, in the proof of Proposition 3.

wind  $\det A(\lambda)$  is a continuous function of  $A_j$  for any  $A(\lambda) \in W_{n \times n}$ , all sufficiently accurate finite lag/lead approximations will have the same winding number as the approximated model.

It can be shown that the winding number criterion remains true for LRE models:

$$\sum_{j=-\infty}^{+\infty} \sum_{s=0}^r A_{j,s} E_{t-s} x_{t-j} = z_t \quad (16)$$

that may include lagged expectations.<sup>13</sup> For such models, we define function  $A(\lambda)$  as

$$A(\lambda) = \sum_{j=-\infty}^{+\infty} \sum_{s=0}^r A_{j,s} e^{i(j-r)\lambda}.$$

Note that, in contrast to the finite lag/lead case, models (16) cannot be reduced to (1) by introduction of a finite number of additional endogenous variables.

The stationary framework adopted above allows the use of powerful spectral techniques for the analysis of solvability of LRE models (1). However, most of the economic processes are not stationary. Although sometimes detrending and other normalizing transformations can be used to reformulate a question of interest in terms of stationary processes, the transformation methods used are often ad hoc and a non-transformed formulation of the question seems relatively more appealing. Moreover, often an appropriate transformation does not exist in principle. For example, if a researcher wants to analyze an economy in transition after a one-time event such as a cardinal policy regime shift or a big structural change, then the very nature of the problem precludes using the stationarity assumption.

For these and other reasons, the majority of authors writing about LRE models relax the stationarity assumption. Some of them, for example Broze, Gourieroux, and Szafarz (1995) and Binder and Pesaran (1997) do not put any restrictions on the exogenous and endogenous processes beyond requiring the existence of the first moments. Others, including Blanchard and Kahn (1980), require the endogenous and exogenous processes to satisfy asymptotic growth restrictions and a predetermined part of the endogenous process to comply with initial conditions. In the next subsection of this paper, we establish the winding number criterion for a non-stationary framework similar in spirit to that of Blanchard and Kahn (1980).

### 3.2. Non-stationary framework

In contrast to the stationarity case, we will interpret  $E_t$  not as a linear projector but as the mathematical expectation given the information available at time  $t$ . We now assume that the information,  $\Omega_t$ , available at time  $t$  includes but is not limited to all past and present realizations of the endogenous  $\{x_t\}$  and exogenous  $\{z_t\}$  processes. Precisely, the “extra” information in  $\Omega_t$  may include some future values of the exogenous process  $\{z_t\}$  and some values, past, present or future, of an “additional”

---

<sup>13</sup>A proof of such an extension can be obtained from the author upon request.

exogenous process  $\{\mu_t\}$ , for example, the intensity of the sun's radiation or a position of the Polar star in the sky during time  $t$ . As is standard in the literature, we require that the information set does not decrease over time, that is, that  $\Omega_t \subseteq \Omega_{t+1}$ . We will denote the exclusively exogenous information, that is, the information about  $\{z_t\}$  and  $\{\mu_t\}$ , as  $\bar{\Omega}_t$  and the operation of taking conditional expectation given  $\bar{\Omega}_t$  as  $\bar{E}_t$ .

Following Blanchard and Kahn (1980), Sims (1995), Klein (1998), and Zadrozny (1998), among others, we restrict attention to the exogenous and endogenous processes that satisfy asymptotic growth restrictions. Precisely, we require that there exists  $0 < \delta \leq 1$  such that for any  $t$  there exist positive  $n$ -dimensional random vectors  $X_t$  and  $Z_t$  with finite expectation,  $\mathbb{E}X_t, \mathbb{E}Z_t < \infty$ , such that:

$$\delta^i |E_t x_{t+i}| \leq X_t, \quad \delta^i |\bar{E}_t z_{t+i}| \leq Z_t \quad \forall i \geq 0. \quad (17)$$

These restrictions are relatively weak. In particular, they do not require processes  $\{x_t\}$  and  $\{z_t\}$  to have second moments, only the first moments must exist. In the literature, the growth restriction on the exogenous process  $z_t$  is typically formulated in terms of  $E_t z_{t+i}$ , whereas we formulate it in terms of  $\bar{E}_t z_{t+i}$ . Our restriction is weaker since the information we are conditioning on excludes the endogenous part of  $\Omega_t$ .

The main reason that we formulate the growth restriction in terms of  $\bar{E}_t$  is technical. Our proof of the winding number criterion uses such a formulation non-trivially. Essentially, conditioning on the endogenous information makes the set of exogenous processes that satisfy the growth restriction depend on the potential solution. Therefore, it becomes unclear what the definition of solvability of the LRE model is. Indeed, we would like to say that the LRE is solvable if for any exogenous process satisfying the growth restrictions there exists a solution satisfying its own growth restriction. However, whether an exogenous process satisfies the growth restriction depends, at least formally, on the potential solution, which makes the definition of solvability self-referring.

In this context, it is interesting to note that conditioning on the endogenous information creates some complications in Broze, Gourieroux and Szafarz's (1995) treatment of the Cagan multivariate model:

$$x_t = AE(x_{t+1}|I_t) + z_t.$$

The authors mention that if  $A$  is nilpotent, then there exists a unique, unrestricted except for the requirement of existence of the first moments, solution to the model.

For example, if  $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  then the unique solution is

$$x_t = AE(z_{t+1}|I_t) + z_t,$$

or if we define  $z_t = (z_{1t}, z_{2t})'$  and  $x_t = (x_{1t}, x_{2t})'$ , we have:

$$\begin{aligned} x_{1t} &= E(z_{2,t+1}|I_t) + z_{1t} \\ x_{2t} &= z_{2t}. \end{aligned}$$

Strangely enough, because of the endogeneity of  $I_t$ , the above formula may describe many solutions rather than one. Take, for example,  $z_{1t} = 0$  for all  $t$  and  $z_{2t} = 0$  for all  $t \neq \tau$ . Assume that  $z_{2\tau}$  is distributed uniformly on segment  $[0,1]$ . Define  $I_t$  as consisting of all past and current realizations of  $\{x_t\}$  and  $\{z_t\}$ . Then, one solution described by the above equations is

$$\begin{aligned} x_{1t} &= z_{2,t+1} \\ x_{2t} &= z_{2t} \end{aligned}$$

and another solution is

$$\begin{aligned} x_{1t} &= 0, \text{ if } t \neq \tau - 1 \\ x_{1,\tau-1} &= \frac{1}{2} \\ x_{2t} &= z_{2t}. \end{aligned}$$

The additional source of the non-uniqueness of solutions to LRE models resulting from the conditioning on the endogenous information seems to be overlooked by Broze, Gourieroux and Szafarz (1995).

Turning back to the formulation of our non-stationary framework, we assume that  $\{x_t\}$  satisfies initial conditions. Precisely, denote the  $i$ -th component of the  $n$ -dimensional vector  $x_t$  as  $x_t^{(i)}$ . Then we assume that the values  $x_t^{(i)}$  are fixed for all those  $i$  and  $t < 0$ , such that the  $i$ -th column of the matrix  $A_{-t}$  is not equal to zero,  $A_{-t}^{(i)} \neq 0$ , and the model's equations (1) hold for all  $t \geq 0$ :

$$\sum_{j=-\infty}^{+\infty} A_j E_t x_{t-j} = z_t, \quad t = 0, 1, 2, \dots \quad (18)$$

We require that the initial history of  $x_t$  satisfies:

$$\sup_{t < 0, A_{-t}^{(i)} \neq 0} \delta^t |x_t^{(i)}| < \infty, \quad i = 1, \dots, n. \quad (19)$$

In case  $\delta < 1$ , this means that values of  $x_t, t < 0$  decay exponentially at least at rate  $\delta$  when  $t$  goes further and further into the past. To us, this assumption seems to be technical rather than substantial. Indeed, among other possible histories, the assumption allows such  $x_t, t < 0$  that have arbitrarily many, but a finite number, of potentially very large non-zero records. Arguably, such histories are the most interesting to consider. When  $\delta = 1$ , assumption (19) says that the initial history of  $\{x_t\}$  must be bounded. The qualification  $A_{-t}^{(i)} \neq 0$  is needed to accommodate the conventional "finite" LRE models that are typically formulated with initial conditions that do not go infinitely into the past. We will call the initial conditions on  $x_t^{(i)}$ , such that  $A_{-t}^{(i)} \neq 0$ , relevant and initial conditions, if any, on  $x_t^{(i)}$ , such that  $A_{-t}^{(i)} = 0$ , irrelevant.

The matrix coefficients  $A_j$  in (1) are assumed to satisfy:

$$\sum_{j=-\infty}^{+\infty} \delta^j |A_j| < \infty,$$

and the convergence on the right hand side of (1) is understood as the almost sure convergence. Note that we can focus on the case  $\delta = 1$ , because, by introducing new endogenous variables, exogenous variables, and coefficients of the model:

$$\tilde{x}_t = \delta^t x_t, \tilde{z}_t = \delta^t z_t, \tilde{A}_j = \delta^j A_j,$$

we can reduce the general setup to a special one with  $\delta = 1$ .

Although the described framework can accommodate many different LRE models that do not assume the stationarity of the exogenous and endogenous processes, there exist important limitations. Perhaps most important, we require that  $A(\lambda) \neq 0$ , that is  $\sum_{j=-\infty}^{\infty} A_j L^j$  is assumed not to have unit roots. This substantially reduces the generality of the framework.<sup>14</sup> We cannot, for example, fit in a purely backward-looking model of the form:

$$x_t = x_{t-1} + z_t, \quad t = 1, 2, 3, \dots \quad x_{t=0} = x_0.$$

Not unrelated to the above limitation is the fact that we restrict attention either to processes that may explode not faster than at a certain rate, if  $\delta < 1$ , or to “bounded processes”, if  $\delta = 1$ . Hence, we exclude the case when processes can be growing polynomially. Such a case is included, for example, in the treatment of the LRE models by Blanchard and Kahn (1980).

Finally, the framework does not fit models with some relevant initial conditions missing. For example, the following model:

$$x_t = \frac{1}{2}x_{t-1}, \quad t = 1, 2, 3, \dots$$

does not fit into our framework. Of course, the above model implies:

$$E_t x_{t+1} = \frac{1}{2}x_t, \quad t = 0, 1, 2, \dots,$$

which is accommodated by the framework. But the above equation is only equivalent to the original one if information available at time  $t$  includes  $x_{t+1}$ , a situation that we forbid. Here we should stress that the fact that we require  $I_t$  to include only past and current realizations of  $x_t$  does not preclude us from considering predetermined variables in the sense of Blanchard and Kahn (1980). If there are some variables that have values known in advance, we can simply redefine the time index on such variables so that no variables with indices larger than  $t$  belong to  $I_t$ . We will say more on this in the next section. The case of missing initial conditions requires a special consideration and is outside the scope of this paper.<sup>15</sup>

<sup>14</sup>The unit root case is problematic for many approaches to the LRE models. For example, Klein (1998) excludes unit roots from the consideration in his paper.

<sup>15</sup>It is perhaps not usually appreciated that the familiar Blanchard and Kahn (1980) framework also cannot fit some of the models with missing initial values. For example, the model

$$x_t - 2x_{t-1} - \frac{1}{4}x_{t-2} + \frac{1}{2}x_{t-3} = 1, \quad x_{t=-1} = 0, x_{t=-3} = 0$$

Our results on the solvability properties of the LRE models that fit into the above non-stationary framework can be summarized as follows. As is shown in the Appendix to this paper, Proposition 1 remains true for the non-stationary framework if  $\delta = 1$ . Since Proposition 2 of the previous subsection was obtained from Proposition 1 without using the stationarity assumption, it holds for the non-stationary framework with  $\delta = 1$  too. For the cases when  $\delta < 1$ , Propositions 1 and 2 hold, given a modified definition of  $A(\lambda)$  :

$$A(\lambda) = \sum_{j=-\infty}^{+\infty} \delta^j A_j e^{ij\lambda}.$$

#### 4. Connections to Blanchard and Kahn's criterion

The criterion of the existence and uniqueness of equilibrium in LRE models given in Proposition 1 is intimately related to many other criteria proposed in the literature. We have already studied the connections of the new criterion to that of Whiteman (1983). Now, after we have formally established the criterion in the multivariate non-stationary framework, we will discuss its connections to the well-known Blanchard and Kahn (1980) conditions for the existence and uniqueness of equilibrium.

In their seminal article, Blanchard and Kahn (1980) study models of form:

$$\begin{pmatrix} y_{t+1} \\ {}_t p_{t+1} \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} y_t \\ p_t \end{pmatrix} + z_t, \quad t = 0, 1, 2, \dots, \quad y_{t=0} = y_0 \quad (20)$$

where  $y$  is a  $k \times 1$  vector of variables predetermined at  $t$ ,  $p$  is a  $m \times 1$  vector of variables non-predetermined at  $t$ ,  $z$  is a  $(k + m) \times 1$  vector of exogenous variables, and  ${}_t p_{t+1}$  is the agent's expectations of  $p_{t+1}$  held at  $t$ . In other words,

$${}_t p_{t+1} = E(p_{t+1} | \Omega_t),$$

where  $E(\cdot)$  is the mathematical expectation operator;  $\Omega_t$  is the information set at  $t$ ;  $\Omega_t \subseteq \Omega_{t+1}$ ; and  $\Omega_t$  includes at least past and current values of  $y, p, z$  (it may include other exogenous variables than  $z$ ; it may include future values of exogenous variables). Finally, Blanchard and Kahn require that:

$$\forall t, \exists \bar{z}_t \in R^{k+m}, \theta_t \in R \text{ such that}$$

$$-(1+i)^{\theta_t} \bar{z}_t \leq E(z_{t+i} | \Omega_t) \leq (1+i)^{\theta_t} \bar{z}_t, \quad \forall i \geq 0.$$

---

cannot be represented in Blanchard and Kahn's canonical form. Here, the initial value  $x_{t=-2}$  is missing. If, for example, we define a predetermined variable  $y_t = (x_{t-1}, x_{t-3})'$  and a non-predetermined variable  $p_t = x_{t-2}$ , then the information set will not satisfy Blanchard and Kahn's requirement that  $\Omega_t \subseteq \Omega_{t+1}$ . Moreover, if we ignore this violation and apply the Blanchard and Kahn criterion for the existence and uniqueness of equilibrium in LRE models, then we will find that the model has a unique bounded solution, whereas, in fact, it has no such solutions. For a detailed discussion of the connections between the Blanchard and Kahn criterion and ours see next section.

A predetermined variable is defined as a variable that is a function only of variables known at time  $t$ , that is of variables in  $\Omega_t$ , so that  $y_{t+1} =_t y_{t+1}$ , and such that its value is given at time  $t = 0$ . A non-predetermined variable  $p_{t+1}$  can be a function of any variable in  $\Omega_{t+1}$ .

Matrix  $M$  from (20) has Jordan decomposition  $M = B \times \text{diag}(J_1, J_2) \times B^{-1}$  where  $B$  is the matrix of the right eigenvectors of matrix  $M$ ;  $J_1$  is the  $\bar{k} \times \bar{k}$  matrix of all those Jordan blocks that correspond to the eigenvalues of  $M$  on or inside the unit circle; and  $J_2$  is the  $\bar{m} \times \bar{m}$  matrix of the remaining Jordan blocks.

Blanchard and Kahn formulate the following conditions for the existence and uniqueness of equilibrium in (20). Assume that a sub-matrix,  $B_{11}$ , corresponding to the first  $k$  rows and the first  $\bar{k}$  columns of the matrix of the right eigenvectors,  $B$ , has full rank. Then if the number of the eigenvalues of  $M$  outside the unit circle,  $\bar{m}$ , is equal to the number of the non-predetermined variables,  $m$ , then there exists a unique solution to (20). If the number of the eigenvalues outside the unit circle is greater than the number of the non-predetermined variables, then, in general, there is no solution. Finally, if the number of the eigenvalues outside the unit circle is less than the number of the non-predetermined variables, then there are many solutions.

As we will soon see, the Blanchard and Kahn (1980) framework substantially overlaps with the non-stationary framework described in the previous section. However, there exist some important differences. First, we do not require a LRE model to be in the canonical, one lead, form (20). Although the canonical representation is straightforward for conventional “finite” LRE models, it may take substantial time to derive. Besides, for the “infinite” LRE models studied in this paper the canonical representation of the model is not possible in principle without making the number of equations in the model infinite. Strictly speaking, the canonical representation would be impossible in our framework even for the conventional “finite” models because we require that all predetermined variables be reindexed so that the information available at time  $t$  includes only those realizations of the endogenous variables that have an index less than or equal to  $t$ .

Second, Blanchard and Kahn focus on the non-exploding processes, whereas we allow the processes to explode. Their framework, however, can be easily extended to the exploding case. Third, as we have already mentioned, Blanchard and Kahn allow the processes to grow polynomially. We exclude such cases from consideration. In contrast to our framework, that of Blanchard and Kahn can fit models with unit roots. Fourth, as was discussed in detail above, we formulate the growth restriction on the exogenous process in terms of  $\bar{E}_t z_{t+i}$ , whereas they use  $E_t z_{t+i}$ .

Finally, Blanchard and Kahn allow more flexibility in defining the information set  $\Omega_t$  than we do. In particular, as I mentioned before, they allow  $\Omega_t$  to contain future realizations of both predetermined and non-predetermined variables. Of course, unless  $\Omega_t$  includes all realizations of the endogenous variables infinitely far away in the future, we can reindex the variables so that  $\Omega_t$  contains no variables with indices greater than  $t$ . However, when we reindex the variables, say, one period back, we may lose some relevant initial conditions. As was discussed before, the case of the missing relevant initial conditions is not considered in this paper.

Here I would like to illustrate the importance of the difference between our and Blanchard and Kahn's formulation of the information set  $\Omega_t$  using an intentionally simple example. Consider a purely backward-looking model:

$$x_t - \frac{1}{2}x_{t-1} = 0, \quad t = 0, 1, 2, \dots \quad (21)$$

Assume that no initial conditions on  $x_t, t < 0$  are stated and that  $\Omega_t$  contains only the past and present realizations of  $x_t$ . As was discussed in the previous section, such a missing initial condition case is not treated by our framework. However, Blanchard and Kahn can simply define a non-predetermined variable  $p_{t+1} = x_t$ , so that now  $\Omega_t$  contains the "future" realization of  $p$ ,  $p_{t+1}$ , and write:

$${}_{t}p_{t+1} - \frac{1}{2}p_t = 0, \quad t = 0, 1, 2, \dots \quad (22)$$

Blanchard and Kahn's criterion applied to the above model will of course say that the model has an infinite number of solutions. Note that our criterion, formally applied to model (22), will also say that the model has an infinite number of solutions. However, we did not prove the criterion for the situation when  $\Omega_t$  may contain  $p_{t+1}$ , and hence, the formal application of our criterion to the above model would not be justified. Of course, as was discussed in the previous section, without requiring that  $\Omega_t$  contains  $p_{t+1}$ , model (22) is not equivalent to (21).

Let us now assume that an LRE model described by Blanchard and Kahn's canonical representation (20) does not have unit roots, and that we are interested in those solutions to the model that have bounded expectations conditional on the information available at  $t$ , and consider only those exogenous processes that have bounded expectations conditional on the exogenous information available at  $t$ . We also assume that the endogenous part of the information available at time  $t$  includes only  $y_\tau, \tau \leq t + 1$  and  $p_s, s \leq t$ , that is, no future values of the non-predetermined variables and two-or-more-periods-into-the-future values of the predetermined variables belong to  $\Omega_t$ . Then, we can cast the problem in the format described in the previous section.

Define  $x_t$  as  $(y'_{t+1}, p'_t)'$ . Note that the first part of the vector of the endogenous variables  $x_t$  is defined as the next period's realization of  $y$  and now  $\Omega_t$  includes only the "past and current" realizations of  $x_t$ . With these notations, model (20) can be rewritten in form (1) as follows:

$$A_{-1}E_t x_{t+1} + A_0 x_t + A_1 x_{t-1} = z_t, \quad t = 0, 1, 2, 3, \dots, x_{-1}^{(i)} = y_0^{(i)}, i = 1, \dots, k \quad (23)$$

where  $A_{-1} = \begin{pmatrix} 0 & 0 \\ 0 & I_m \end{pmatrix}$ ,  $A_0 = \begin{pmatrix} I_k & -M_{12} \\ 0 & -M_{22} \end{pmatrix}$ ,  $A_1 = \begin{pmatrix} -M_{11} & 0 \\ -M_{21} & 0 \end{pmatrix}$ . The corresponding function  $A(\lambda)$  has the following form:

$$A(\lambda) = \begin{pmatrix} I_k - M_{11}e^{i\lambda} & -M_{12} \\ -M_{21}e^{i\lambda} & I_m e^{-i\lambda} - M_{22} \end{pmatrix}.$$

**Proposition 3.** *Matrix  $B_{11}$  has full rank if and only if all partial indices of  $A(\lambda)$  are of the same sign or zero.*

The proof of this proposition is given in the Appendix. Recall that an LRE model is generic, in the sense described in the previous section, if and only if the function  $A(\lambda)$  associated with the model has all partial indices of the same sign or zero. Hence, Proposition 3 implies that Blanchard and Kahn's (1980) requirement that  $B_{11}$  is of full rank is equivalent to the requirement that model (23) is generic.

Now let me show that the Blanchard-Kahn conditions on the number of the unstable eigenvalues vs. the number of the non-predetermined variables can be reformulated in terms of the behavior of the graph of  $\det A(\lambda)$ . Note that  $A(\lambda)$  can be represented as the following product:

$$A(\lambda) = B \left[ \begin{pmatrix} I_k e^{-i\lambda} & 0 \\ 0 & I_m e^{-i\lambda} \end{pmatrix} - \begin{pmatrix} J_1 & 0 \\ 0 & J_2 \end{pmatrix} \right] B^{-1} \begin{pmatrix} I_k e^{i\lambda} & 0 \\ 0 & I_m \end{pmatrix}.$$

Therefore,

$$\det A(\lambda) = e^{ik\lambda} \prod_{|\rho_j| < 1} (e^{-i\lambda} - \rho_j) \prod_{|\rho_j| > 1} (e^{-i\lambda} - \rho_j),$$

where  $\rho_j$  are the eigenvalues of  $M$ . As we showed in the previous sections, such a decomposition implies that:

$$\text{wind } \det A(\lambda) = \text{wind } e^{ik\lambda} + \sum_{|\rho_j| < 1} \text{wind } (e^{-i\lambda} - \rho_j) + \sum_{|\rho_j| > 1} \text{wind } (e^{-i\lambda} - \rho_j).$$

Clearly,  $\text{wind } e^{ik\lambda} = k$  and the winding number of  $(e^{-i\lambda} - \rho_j)$  is equal to -1 if  $|\rho_j| < 1$ , and it is equal to zero if  $|\rho_j| > 1$ . Hence, the winding number of  $\det A(\lambda)$  is equal to the number of the predetermined variables minus the number of the eigenvalues of  $M$  that lie outside the unit circle. Therefore, the graph of  $\det A$  does not rotate around zero, rotates around zero clockwise, or rotates around zero counterclockwise if and only if the number of the stable eigenvalues of  $M$  coincides, is larger than, or is less than the number of the predetermined variables respectively. This is the same as to say that  $\text{wind } \det A(\lambda)$  is zero, negative, or positive if and only if the number of the unstable eigenvalues of  $M$  coincides, is less than, or is larger than the number of the non-predetermined variables respectively, which constitutes the Blanchard and Kahn conditions for the model to have a unique, multiple or no solutions respectively.

To summarize, the non-stationary framework described in the previous section significantly overlaps with that of Blanchard and Kahn. However, neither of the frameworks is a partial case of the other. Our approach allows the infinite-dimensional LRE models and exploding processes, whereas that of Blanchard and Kahn allows unit roots and polynomially growing processes and has greater flexibility in formulation of the structure of the information set  $\Omega_t$ . For the LRE models that can be cast in both frameworks, the two criteria for the existence and uniqueness of equilibrium are equivalent.

## 5. Analysis of monetary policy under model uncertainty

In this section we will apply the criterion developed above to analyze the robustness of simple Taylor-type rules to the model uncertainty about the standard new Keynesian model of the economy described, for example, in Woodford (1997) and Clarida, Gali, and Gertler (1999). The model consists of two equations

$$\pi_t = \beta E_t \pi_{t+1} + \alpha y_t + \varepsilon_t \quad (24)$$

$$y_t = E_t y_{t+1} - \varphi(i_t - E_t \pi_{t+1}) + \eta_t \quad (25)$$

Here  $\pi_t$  is inflation at an annual percentage rate,  $y_t$  is the output gap, measured as 100 times the log ratio of actual real output to potential output, and  $i_t$  is the federal funds rate at an annual rate. Parameters  $\beta$ ,  $\alpha$ , and  $\varphi$  represent the time discount factor of the private sector, a measure of the speed of price adjustment, and the intertemporal elasticity of substitution respectively. In the numerical analysis below I assume  $\alpha = 0.024$ ,  $\varphi = 6.37$ , and  $\beta = 0.99$  as in Woodford (1997).

We will assume that the above model is a reference model for a monetary authority, say Fed, that is able to control the interest rate. The Fed views the model as only an approximation of the reality and wants to make sure that its policy works well not only for the reference model, but also for some deviations from it. Denote the right hand side of the reference model equations (24) and (25) as  $ref_\pi$  and  $ref_y$  respectively. Following Onatski and Stock (2002), we model the potential deviations as:

$$\begin{aligned} \pi_t &= ref_\pi + w_1 E_t \Delta_1(L)(\pi_{t+1} - \pi_t) + w_2 E_t \Delta_2(L)y_t \\ y_t &= ref_y + w_3 E_t \Delta_3(L)y_{t+1} + w_4 E_t \Delta_4(L)(i_t - E_t \pi_{t+1}), \end{aligned}$$

where  $\Delta_1, \Delta_2, \Delta_3$ , and  $\Delta_4$  are potentially infinite series in the positive and negative powers of the lag operator,  $L$ , with absolutely summable coefficients. The deltas are meant to represent the economic dynamics that was left unmodeled by the reference model. The weights  $w_i$  are supposed to reflect the relative importance of the uncertainties.

The Fed is assumed to restrict attention only to those deviations from the reference model that have some empirical or theoretical value. We will model this restriction by considering only those deviations that correspond to relatively small  $\Delta_i$ . The size of  $\Delta_i$  can be defined as

$$\|\Delta_i\| = \sup_{\lambda \in [0, 2\pi]} |\Delta_i(e^{i\lambda})|.$$

We can, for example, focus only on the deviations with  $\|\Delta_i\| < r, i = 1, \dots, 4$ , where  $r$  is some real number that describes the size of the model uncertainty. For convenience of notation, we introduce  $\Delta = \{\Delta_1, \dots, \Delta_4\}$  and define  $\|\Delta\|$  as the maximum of  $\|\Delta_i\|, i = 1, \dots, 4$ .

We take  $w_1 = 0.6, w_2 = 0.01, w_3 = 0.01, w_4 = 3$ . The weight  $w_1$  was chosen so that  $\Delta_1 = L$ , which has size 1, introduces an empirically plausible amount of

backward-lookingness into the Phillips curve (see Rudebusch 2002 for a discussion). The weights  $w_2$  and  $w_4$  were chosen to be half as large as the corresponding structural parameters,  $\alpha$  and  $\varphi$ . Finally,  $w_3$  was chosen so that “small” deviations from the unit root in the output gap’s own dynamics in the IS curve are allowed. Such a choice of weights is arbitrary. For a more careful calibration of the model uncertainty see Onatski and Williams (in press).

We assume that the Fed wants to study the robustness to the model uncertainty of policy rules:

$$i_t = g_\pi \pi_t + g_y y_t.$$

The above type rules were introduced by J. Taylor (1993) and since then were subject to active research. The famous Taylor rule corresponds to  $g_\pi = 1.5$ ,  $g_y = 0.5$ . Such rules fit the data quite well, at least since the late 80’s. Using the timeless perspective paradigm suggested by Woodford (1999), we assume that the quality of a rule is assessed as if the rule were always in operation. Such an approach allows us to assume the stationarity of the economic processes and avoids the complications related to the non-stationarity of transition between old and new policy regimes.

Onatski and Stock (2002) analyze the model uncertainty about purely backward-looking models of the economy. They describe the robustness of a given policy rule by a minimal size of the model uncertainty, which includes potential deviations from the reference model such that, controlled by the policy rule, they lead to the dynamic instability of the economic processes. If such a minimal size is relatively large, the rule does not lead to the instability under a variety of deviations from the reference model and, therefore, is relatively robust. This measure of robustness can be generalized to models that include rational expectations by requiring that relatively more robust rules should lead neither to the dynamic instability nor to the indeterminacy of equilibrium for model deviations of relatively large size. In the rest of this section, we, therefore, will be interested in the question: How, for a given policy rule, to find the minimal size of the model uncertainty that includes LRE models that either have no stationary solutions or have multiple solutions?

Let us assume that the reference model controlled by a given policy rule has a unique solution. If this is not so, then the rule is not robust to an arbitrarily small amount of the model uncertainty and should be rejected. According to our criterion for the existence and uniqueness of equilibrium in LRE models, the graph of the function  $\det A(\lambda)$  corresponding to the reference model controlled by the rule does not encircle zero. Let the function  $A_\Delta(\lambda)$  correspond to the deviation from the reference model characterized by  $\Delta$ . The graph of  $\det A_\Delta(\lambda)$  is a continuous function of  $\Delta$ . Therefore, as long as the size of  $\Delta$  is small, and the deviation remains a generic LRE, the graph of  $\det A_\Delta(\lambda)$  will not encircle zero and the deviation will still have a unique solution. The uniqueness or existence of equilibrium will be lost when the graph “hits” zero or when the deviation becomes a non-generic model.

The potential non-genericity of the deviation creates a problem for our analysis of robustness because we cannot detect the non-genericity using the behavior of the graph of  $\det A(\lambda)$ . However, as we said above, the non-genericity of a model is not stable to small model perturbations. This means that the non-existence or multiplicity

Figure 1:

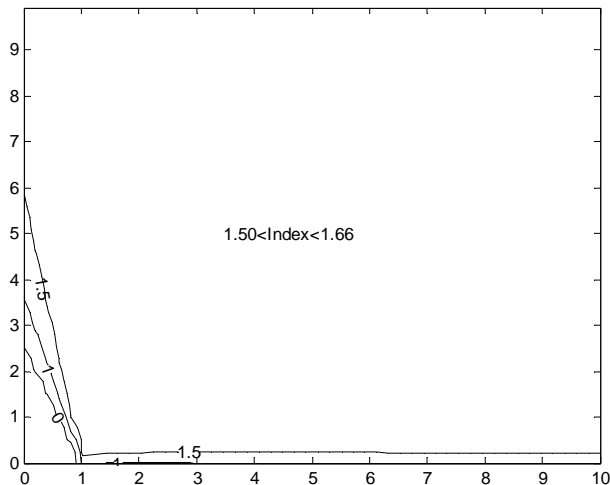


Figure 2: The index of robustness for Taylor-type rules to the model uncertainty about the standard New Keynesian model of the economy.

ity problem created by such a non-generic deviation from the reference model does not persist under small perturbations. On the contrary, if the graph of  $\det A_\Delta(\lambda)$  “hits” zero for  $\Delta = \Delta_0$ , and thus the non-uniqueness or multiplicity problem is created for some  $\|\Delta\| > \|\Delta_0\|$ , the problem becomes insensitive to small fluctuations. We will, therefore, ignore non-generic deviations and measure the robustness of a given policy rule by the smallest size of the model uncertainty that includes a generic deviation from the reference model that either has no solutions or has many solutions. In other words, the proposed measure of the robustness, that we will call the index of robustness, is equal to the minimum of the uncertainty size  $\|\Delta\|$  such that  $\det A_\Delta(\lambda) = 0$  for some  $\lambda \in [0, 2\pi]$ .

There exist effective computational methods that estimate  $\min \{\|\Delta\| : \det A_\Delta(\lambda) = 0\}$  from below and from above. The methods are described in detail, for example, in Packard and Doyle (1993). The commercially available mu-analysis and synthesis toolbox for MATLAB and the LMI control toolbox for MATLAB contain the relevant computer codes. We compute the upper and lower bounds on the index of robustness for Taylor-type policy rules with parameters  $g_\pi \in [0, 10]$  (grid of 0.2),  $g_y \in [0, 10]$  (grid of 0.25). The upper and lower bounds turn out to be very close to each other. Figure 2 reports the contour plot of the upper bound on the index.

We see that the rules with a smaller than 1 reaction to inflation and a relatively weak response to the output gap perform very poorly. Such rules destabilize even the reference model of the economy. The other rules have about the same index

of robustness. This situation is drastically different from that for purely backward-looking models. Onatski and Stock (2002) showed that the index of robustness is relatively high for the rules with a low (but higher than 1) reaction to inflation and a relatively high (about 2) reaction to the output gap. More aggressive rules are relatively less robust and may even destabilize the reference model (the Rudebusch and Svensson model (1999) for that study).

The fact that the robustness of the policy rules in the forward-looking new Keynesian model does not deteriorate with a rise in the aggressiveness of the rules agrees with the intuition that forward-looking models are much more difficult to destabilize than backward-looking models. Note, however, that in forward-looking models, there exists an additional potential source of non-robustness corresponding to the multiplicity of equilibria. It turns out that the potential of the multiplicity of equilibria does not add enough to the variability of the index of robustness to differentiate between more and less aggressive rules on our grid.

In the above analysis, our criterion for the existence and uniqueness of equilibrium in “infinite” LRE models played an instrumental role. Since the deviations from the reference model were formulated in terms of  $\Delta(L)$ , which were allowed to be infinite series in the positive and negative powers of  $L$ , we could not analyze the existence and uniqueness of equilibrium using conventional criteria. Of course, we could have restricted attention to  $\Delta(L)$ , which are sums of a finite number of positive and negative powers of  $L$ . However, as was explained in the introduction, restricting model deviations to such a parametric class would have created difficulties for the analysis of robustness. Precisely, given the current state of knowledge, it would become much more difficult, if not impossible, to compute accurate bounds on the minimal size of the deviations from the restricted class, which result in the dynamic instability or indeterminacy of equilibrium for a particular policy rule.

## 6. Conclusion

This paper develops the following simple geometric criterion for the existence and uniqueness of equilibrium in LRE models of form:

$$\sum_{j=-\infty}^{+\infty} A_j E_t x_{t-j} = z_t.$$

Assuming that  $\sum_{j=-\infty}^{+\infty} |A_j| < \infty$  and that the model is generic, the model has a unique equilibrium, multiple equilibria or no equilibrium at all if and only if the graph of  $\det A(\lambda)$ , where  $A(\lambda) = \sum_{j=-\infty}^{+\infty} A_j e^{ij\lambda}$ , does not rotate around zero, rotates around zero clockwise, or rotates around zero counter-clockwise, respectively.

We consider two different frameworks of the analysis: a stationary and non-stationary one. For the stationary framework, it is assumed that the exogenous process  $\{z_t\}$  is linearly regular weakly stationary and that the solutions are sought among time-invariant, square summable linear combinations of the past and current realizations of the process fundamental for  $\{z_t\}$ . For the non-stationary framework,

it is assumed that the model equations hold for  $t \geq 0$ , the relevant initial conditions  $x_t, t < 0$  are given and bounded, both the exogenous and the endogenous processes have bounded conditional expectations of the future values, and information available at time  $t$  consists of the current and past values of the endogenous process  $\{x_t\}$ , the current, past, and possibly future values of the exogenous process  $\{z_t\}$ , and the current, past, and possible future values of an additional exogenous process  $\{\mu_t\}$ , that, for example, may represent sun spots.

We show that Whiteman's (1983) criterion for the existence and uniqueness of equilibrium in LRE models is a special case of the above criterion. For the conventional finite lags and leads models, our requirement of the models' genericity is weaker than the conditions stated in the formulation of Whiteman's criterion. In addition, in contrast to Whiteman's criterion, ours remains valid for models that contain an infinite number of lags and leads of the endogenous variable.

We explore the connections between the new criterion and that of Blanchard and Kahn (1980). We show that our non-stationary framework of the analysis and Blanchard and Kahn's framework overlap considerably. However, there exist important differences. Our approach allows infinite-dimensional LRE models and exploding processes, whereas theirs allows unit roots and polynomially growing processes and allows for greater flexibility in formulation of the structure of the information set  $\Omega_t$ . For models that can be accommodated by both the present and Blanchard and Kahn's frameworks, the two criteria are equivalent.

This paper applies the winding number criterion to a study of robustness of monetary policy rules under the model uncertainty about the standard New Keynesian model of the economy. I show that, in contrast to the uncertainty about purely backward-looking models of the economy, the index of robustness of the policy rules is relatively insensitive to the aggressiveness of the rule.

## References

- [1] Anderson, Gary and G. Moore (1985). "A Linear Algebraic Procedure for Solving Linear Perfect Foresight Models", *Economics Letters* 17, pp. 247-52.
- [2] Binder, Michael and M Hashem Pesaran (1997). Multivariate Linear Rational Expectations Models: Characterization of the Nature of the Solutions and Their Fully Recursive Computation. *Econometric Theory*, vol. 13, no. 6, pp.877-88
- [3] Blanchard O. J. and C. M. Kahn (1980) The Solution of Linear Difference Models under Rational Expectations. *Econometrica*, Vol. 48, No. 5, pp.1305-1311.
- [4] Bottcher, Albrecht and B. Silbermann (1998), Introduction to large truncated Toeplitz matrices. Springer.
- [5] Broze, Laurence, C. Gourieroux and A. Szafarz (1995), "Solutions of Multivariate Rational Expectations Models", *Econometric Theory*, vol. 11, pp.229-57.
- [6] Clancey, Kevin, I. Gohberg (1981), Factorization of matrix functions and singular integral operators. Birkhauser Verlag.
- [7] Clarida, R., Gali, J. and M. Gertler (1999) The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature*. Vol. 37 (4). pp. 1661-1707.
- [8] Dotsey M. and John H. Boyd, III (1990) Interest Rate Rules and Nominal Determinacy. Working Paper #90-1 Federal Reserve Bank of Richmond.
- [9] Driels, Morris (1996) Linear Control Systems Engineering. McGraw-Hill.
- [10] Gohberg, I. C. and I. A. Fel'dman (1974) Convolution Equations and Projection Methods for their Solution. Translations of Mathematical Monographs, Vol. 41. American Mathematical Society, Providence, Rhode Island.
- [11] Gohberg I., L. Lerer and L. Rodman (1978) Factorization Indices for Matrix Polynomials. *Bulletin of the American Mathematical Society*, 84, pp. 275-77.
- [12] Gohberg, I. and M.G. Krein (1960). Systems of integral equations on a half-line with kernel depending upon the difference of the arguments. *American Mathematical Society Translations* (2) 14, pp. 217-287.
- [13] Hansen, L. P. and T. J. Sargent, (1980), Formulating and Estimating Dynamic Linear Rational Expectations Models. *Journal of Economic Dynamics and Control* 2, 7-46.
- [14] Hansen, L. P. and T. J. Sargent, (2002), Robust Control and Model Uncertainty in Macroeconomics. Unpublished book manuscript.

- [15] King, Robert G., M. Watson (1998). The Solution of Singular Linear Difference Systems under Rational Expectations. *International Economics Review* 39(4), pp. 1015-1026.
- [16] Klein, Paul (2000). Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model. *Journal of Economic Dynamics and Control*, vol. 24, no. 10, pp. 1405-23
- [17] Kolmogorov, A. N. and S. V. Fomin (1970) Introductory Real Analysis. Englewood Cliffs, N.J., Prentice-Hall
- [18] Luenberger, D. G. (1969), Optimization by Vector Space Methods. John Wiley and Sons, Inc.
- [19] Mas-Colell, Andreu (1985). The theory of general economic equilibrium: a differentiable approach. Cambridge University Press.
- [20] McCallum, Bennett T 1998, "Solutions to Linear Rational Expectations Models: A Compact Exposition", *Economics Letters*, vol. 61, pp. 143-47.
- [21] Muth, John F. (1961). "Rational Expectations and the Theory of Price Movements", *Econometrica* 29, pp.315-35.
- [22] Onatski A., Stock J. (2002), Robust Monetary Policy under Model Uncertainty in a Small Model of the US Economy. *Macroeconomic Dynamics*, Vol 6, No. 1.
- [23] Onatski A. and N. Williams (in press), Modeling Model Uncertainty. *Journal of European Economic Association*.
- [24] Packard A., Doyle J. (1993) The Complex Structured Singular Value, *Automatica*, Vol. 29, No. 1, pp. 71-109.
- [25] Priestly, M. B. (1981), Spectral Analysis and Time Series. Academic Press.
- [26] Rozanov, Yu. A. (1967) Stationary Random Processes. Holden-Day. San Francisco, Cambridge, London, Amsterdam.
- [27] Rudebusch, G.D. and L.E.O. Svensson (1999) Policy Rules for Inflation Targeting. In J. Taylor (ed.), *Monetary Policy Rules*, pp. 203-246. Chicago: University of Chicago Press for the NBER.
- [28] Rudebusch, G.D. (2002) Assessing Nominal Income Rules for Monetary Policy with Model and Data Uncertainty. *The Economic Journal*, 112, pp. 402-432.
- [29] Sargent, Thomas J. (1987). Dynamic macroeconomic Theory. Cambridge, Mass.: Harvard University Press.
- [30] Sargent, Thomas J. (1999). The Conquest of American Inflation. Princeton University Press, Princeton, N.J.

- [31] Sims, Christopher A. (1996). Solving Linear Rational Expectations Models. Seminar paper.
- [32] Taylor, J. (1977). "Conditions for Unique Solution in Stochastic Macroeconomic Models with Rational Expectations", *Econometrica* 45, pp. 1277-1385.
- [33] Taylor, John B. (1993) Discretion versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Policy* 39, 195-214.
- [34] Tetlow R. and P. von zur Muehlen (2001) Robust Monetary Policy with Misspecified Models: Does Model Uncertainty Always Call for Attenuated Policy? *Journal of Economic Dynamics and Control*, 25, pp. 911-49.
- [35] Uhlig, Harald (1999). A toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily. User's Guide.
- [36] Whiteman, Charles H. (1983), *Linear Rational Expectations Models; A User's Guide*. University of Minnesota Press.
- [37] Whittle, P. (1963), *Prediction and Regulation by Linear Least-Square Methods*. The English Universities Press LTD.
- [38] Woodford, M. (1997), Control of the Public Debt: a Requirement for Price Stability? in Calvo, G., and M. King, *The Debt Burden and Monetary Policy*, London: Macmillan
- [39] Woodford, M. (1999), "Commentary on: How should Monetary Policy be Conducted in an Era of Price Stability?", In *New Challenges for Monetary Policy*, A Symposium Sponsored by the Federal Reserve Bank of Kansas City.
- [40] Zdrozny, Peter A. (1998) An eigenvalue method of undetermined coefficients for solving linear rational expectations models. *Journal of Economic Dynamics and Control*, 22, pp.1353-73.

## 7. Appendix

**The first order condition for the optimization problem described in the introduction:**

Assuming that the coefficients in  $C(L)$  and  $D(L)$  are absolutely summable and  $C(1) > r$ , the steady-state values of the output and capital are:

$$\begin{aligned}\bar{y} &= D(1)\bar{k}, \\ \bar{k} &= \frac{C(1) - r}{aC(1)^2 + bD(1)^2}.\end{aligned}$$

We will interpret the assumption that the monopolist has been in business a long time by saying that the “time zero” is infinitely far away into the past and, therefore, the deviations of capital and output from the steady-state values,  $\hat{k}_t$  and  $\hat{y}_t$ , are purely stochastic, zero-mean stationary processes. We can rewrite the problem in terms of the deviations as

$$\max E \left[ \left( -\frac{a}{2}\hat{y}_t + \varepsilon_t \right) \hat{y}_t - \frac{b}{2} (C(L)\hat{k}_t)^2 \right],$$

or, substituting  $\hat{y}_t = D(L)\hat{k}_t$  and shifting by  $-\frac{E\varepsilon_t^2}{2a}$ , which does not affect the solution, as

$$\max E \left[ -\frac{a}{2} \left( D(L)\hat{k}_t - \frac{\varepsilon_t}{a} \right)^2 - \frac{b}{2} (C(L)\hat{k}_t)^2 \right].$$

The discount factor  $\beta$  does not play role here because of the stationarity of the deviations. We will be looking for a solution to the above problem in the Hilbert space of zero mean stationary processes having Wold representation with innovations  $\varepsilon_t$ . The functional  $f(\hat{k}_t) = E \left[ -\frac{a}{2} \left( D(L)\hat{k}_t - \frac{\varepsilon_t}{a} \right)^2 - \frac{b}{2} (C(L)\hat{k}_t)^2 \right]$  that needs to be maximized is everywhere Frechet differentiable (for a definition of Frechet derivative and an excellent discussion of optimization of functionals in vector spaces see Luenberger (1969)) with Frechet derivative  $f'(\hat{k}_t)$ , equal to a linear functional over the space of stationary processes  $h_t$ , defined as

$$f'(\hat{k}_t) \circ h_t = E \left\{ \left( -aD(L^{-1})D(L)\hat{k}_t - bC(L^{-1})C(L)\hat{k}_t + d_0\varepsilon_t \right) h_t \right\}.$$

At the maximum, the derivative must be equal to zero functional, that is

$$-aD(L^{-1})D(L)\hat{k}_t - bC(L^{-1})C(L)\hat{k}_t + d_0\varepsilon_t = 0.$$

### **Proof of formula (3):**

For the characteristic equation corresponding to the model:

$$x_t + a_1x_{t-1} + a_2x_{t-2} + \Delta(L)x_t = z_t$$

to have a zero at the unit circle, there must exist  $\lambda$  such that

$$1 + a_1e^{i\lambda} + a_2e^{i2\lambda} = -\Delta(e^{i\lambda}).$$

Therefore, the size of the minimal destabilizing deviation must be no smaller than the solution to (3).

On the other hand, suppose that the minimum in (3), call it  $r$ , is attained at  $\lambda_0$  and let  $1 + a_1 e^{i\lambda_0} + a_2 e^{i2\lambda_0}$  be represented as  $r e^{i\theta}$ . Consider  $\Delta(L) = a \frac{L-b}{1-bL}$ , where  $a = -r$  and  $b = \sin \frac{\theta-\lambda_0}{2} / \sin \frac{\theta+\lambda_0}{2}$  if  $\sin^2 \frac{\theta-\lambda_0}{2} \leq \sin^2 \frac{\theta+\lambda_0}{2}$ , and  $a = r$  and  $b = \cos \frac{\theta-\lambda_0}{2} / \cos \frac{\theta+\lambda_0}{2}$  otherwise. It is not hard to see that  $\|\Delta\| = r$  and  $|b| \leq 1$ , so that  $\Delta(L)$  can be represented as a series in the non-negative powers of  $L$  with absolutely summable coefficients. It is straightforward to check that  $\Delta(e^{i\lambda_0}) = -r e^{i\theta}$ , so that the characteristic equation has a root at  $e^{i\lambda_0}$ . Hence, the described  $\Delta(L)$  is destabilizing and the size of the minimal destabilizing deviation must be no larger than the solution to (3).

**Proof of Proposition 1 for the stationary framework:**

First, note that we can interchange the order of the linear projection  $E_t$  and the infinite sum in (1). Indeed, since the series of matrices  $\{A_j\}$  is element by element absolutely summable and solutions  $\{x_t\}$  are sought among the weakly stationary processes, the partial sums  $\sum_{j=-n}^n A_j E_t x_{t-j}$  and  $\sum_{j=-n}^n A_j x_{t-j}$  converge in the Hilbert space formed by the completion in mean square norm of a linear space spanned by all finite linear combinations of all components of vector process  $\{z_t\}$ . Validity of the change in the order of the infinite sum and the projection follows from the fact that the projection operator  $E_t$  is continuous, so that

$$\lim_{n \rightarrow \infty} E_t \sum_{j=-n}^n A_j x_{t-j} = E_t \lim_{n \rightarrow \infty} \sum_{j=-n}^n A_j x_{t-j}.$$

Now, consider a Wold's decomposition of the exogenous process  $z_t$

$$z_t = \sum_{k=0}^{\infty} C_k \varepsilon_{t-k}, \tag{26}$$

where  $\{\varepsilon_t\}$  is a fundamental process for  $\{z_t\}$  and  $C_k$  are the  $n \times m$  matrix coefficients of the decomposition.<sup>16</sup> Let

$$\varepsilon_t = \int_{-\pi}^{\pi} e^{it\lambda} \Phi(d\lambda) \tag{27}$$

be the spectral representation of the fundamental process' values. Since  $\{\varepsilon_t\}$  is uncorrelated stationary process with unit variance, the vector random measure  $\Phi(d\lambda)$  satisfies the following conditions:

$$E\Phi(\Delta)\Phi^*(\tilde{\Delta}) = 0, \tag{28}$$

$$E\Phi(\Delta)\Phi^*(\Delta) = I_m \Lambda(\Delta) \tag{29}$$

for any Borel sets  $\Delta, \tilde{\Delta} \in [-\pi, \pi]$  that do not intersect. Here  $(.)^*$  denotes the operation of complex conjugation,  $I_m$  is the  $m \times m$  identity matrix,  $\Lambda(.)$  is the Lebesgue

---

<sup>16</sup>Different Wold's decompositions may be obtained by orthogonal transformations of the vector of fundamental shocks,  $\varepsilon$ .

measure on the real line, and  $E$  denotes the operation of taking the mathematical expectation.

Equations (26) and (27) imply the following representation for  $z_t$  :

$$z_t = \int_{-\pi}^{\pi} e^{it\lambda} C(\lambda) \Phi(d\lambda),$$

where

$$C(\lambda) = \sum_{k=0}^{\infty} C_k e^{ik\lambda}.$$

Note that since the matrix coefficients of the Wold representation are element by element square summable, all elements of the matrix function  $C(\lambda)$  belong to  $L^2[-\pi, \pi]$ . In what follows, all properties stated for matrices must be understood element by element for simplicity of notation. So, for example, we simply say that  $C(\lambda) \in L^2[-\pi, \pi]$ .

Recall that we are looking for solutions representable in the form:

$$x_t = \sum_{k=0}^{\infty} D_k \varepsilon_{t-k},$$

where  $\{D_k\}$  is a square summable series of  $n \times m$  matrices. Let us define

$$D(\lambda) = \sum_{k=0}^{\infty} D_k e^{ik\lambda}.$$

Then a potential solution must have the following form:

$$x_t = \int_{-\pi}^{\pi} e^{it\lambda} D(\lambda) \Phi(d\lambda). \quad (30)$$

Furthermore, since  $\{A_j\}$  is absolutely summable, the linear transformation,  $\sum_{j=-\infty}^{\infty} A_j x_{t-j}$ , must have form:

$$\sum_{j=-\infty}^{\infty} A_j x_{t-j} = \int_{-\pi}^{\pi} e^{it\lambda} A(\lambda) D(\lambda) \Phi(d\lambda),$$

where

$$A(\lambda) = \sum_{j=-\infty}^{\infty} A_j e^{ij\lambda}.$$

Hence, we can equivalently rewrite (1) as

$$E_t \int_{-\pi}^{\pi} e^{it\lambda} A(\lambda) D(\lambda) \Phi(d\lambda) = \int_{-\pi}^{\pi} e^{it\lambda} C(\lambda) \Phi(d\lambda). \quad (31)$$

The projection operator  $E_t$  transforms any two-sided linear combination of the fundamental process' values in the one-sided linear combination by annihilation of

all terms in the original linear combination corresponding to the future values of  $\varepsilon_t$  (from the point of view of time  $t$ ). Therefore, (31) is equivalent to

$$\int_{-\pi}^{\pi} [e^{it\lambda}A(\lambda)D(\lambda)]_+ \Phi(d\lambda) = \int_{-\pi}^{\pi} e^{it\lambda}C(\lambda)\Phi(d\lambda), \quad (32)$$

where the annihilation operator,  $[\cdot]_+$ , on matrix functions from  $L^2[-\pi, \pi]$  is defined as follows. For any  $F(\lambda)$  with the Fourier expansion  $F(\lambda) = \sum_{j=-\infty}^{\infty} F_j e^{ij\lambda}$  we have:

$$[F(\lambda)]_+ = \sum_{j=0}^{\infty} F_j e^{ij\lambda}.$$

Let us denote the difference between two integrands in (32) as  $Z(\lambda)$ , that is,

$$Z(\lambda) = [e^{it\lambda}A(\lambda)D(\lambda)]_+ - e^{it\lambda}C(\lambda).$$

It follows from properties (28) and (29) of the vector random measure  $\Phi(d\lambda)$  and from equation (32) that

$$\int_{-\pi}^{\pi} Z(\lambda)Z^*(\lambda)d\lambda = 0,$$

and therefore  $[e^{it\lambda}A(\lambda)D(\lambda)]_+ = e^{it\lambda}C(\lambda)$ , and

$$[A(\lambda)D(\lambda)]_+ = C(\lambda) \quad (33)$$

almost everywhere with respect to the Lebesgue measure on  $[-\pi, \pi]$ .<sup>17</sup> Note that solving model (1) is equivalent to finding  $D(\lambda)$  satisfying (33). Therefore, we will call (33) the spectral equivalent of (1), or simply the spectral model.

Substituting the Wiener-Hopf factorization of  $A(\lambda)$  (15) into the spectral model, we have:

$$[A_-(\lambda)A_0(\lambda)A_+(\lambda)D(\lambda)]_+ = C(\lambda).$$

The above equality is equivalent to the existence of a function  $F(\lambda) \in L^2[-\pi, \pi]$  that has a one-sided Fourier expansion

$$F(\lambda) = \sum_{j=-\infty}^{-1} F_j e^{ij\lambda}, \quad (34)$$

such that

$$A_-(\lambda)A_0(\lambda)A_+(\lambda)D(\lambda) = C(\lambda) + F(\lambda).$$

Multiplying both sides of the above equation by  $A_-^{-1}(\lambda)$ , we have:

$$A_0(\lambda)A_+(\lambda)D(\lambda) = A_-^{-1}(\lambda)C(\lambda) + A_-^{-1}(\lambda)F(\lambda).$$

---

<sup>17</sup>All equations stated below must be understood as true almost everywhere with respect to the Lebesgue measure. We will not mention this qualification in what follows to save the space.

Let us denote  $A_+(\lambda)D(\lambda)$  as  $\tilde{D}(\lambda)$ . Of course, since  $A_+(\lambda) \neq 0$  on  $[-\pi, \pi]$ , finding  $D(\lambda)$  that solves the spectral model is equivalent to finding  $\tilde{D}(\lambda)$  that satisfies

$$A_0(\lambda)\tilde{D}(\lambda) = A_-^{-1}(\lambda)C(\lambda) + A_-^{-1}(\lambda)F(\lambda) \quad (35)$$

for some  $F(\lambda) \in L^2[-\pi, \pi]$  of form (34).

Since  $A_+(\lambda)$  and  $D(\lambda)$  both have one-sided Fourier expansions with only non-zero coefficients on  $e^{ij\lambda}$  with  $j \geq 0$ , their product,  $\tilde{D}(\lambda)$ , also must have non-zero coefficients only on the non-negative powers of  $e^{i\lambda}$ . Note that  $\tilde{D}(\lambda) \in L^2[-\pi, \pi]$  because  $D(\lambda) \in L^2[-\pi, \pi]$  and  $A_+(\lambda)$  is bounded (recall that it has absolutely summable Fourier coefficients). Now, if the partial indices of  $A(\lambda)$  are equal to  $k_1 \leq k_2 \leq \dots \leq k_n$  we have:

$$A_0(\lambda) = \text{diag} (e^{ik_1\lambda}, e^{ik_2\lambda}, \dots, e^{ik_n\lambda}),$$

and therefore the coefficients on  $e^{ij\lambda}$  in the Fourier expansion of the elements in the  $r$ -th row of matrix  $A_0(\lambda)\tilde{D}(\lambda)$  are zero for all  $j < k_r$ .

Suppose that the largest partial index of  $A(\lambda)$ ,  $k_n$ , is positive. Then, in general, there is no solution to the spectral model and therefore to the original model (1). Indeed, we will show that there exists  $C(\lambda) \in L^2[-\pi, \pi]$  such that for any  $F(\lambda)$  of form (34), at least one element in the  $n$ -th row of the matrix on the right hand side of equation (35) has the Fourier expansion with a non-zero coefficient on the constant  $1 = e^{ij0}$ . On the other hand, since  $k_n > 0$ , any element in the  $n$ -th row of the matrix on the left hand side of (35) has zero Fourier coefficient corresponding to the constant.

To see that some of the elements of the  $n$ -th row of the right hand side of (35) have a non-zero Fourier coefficient on the constant, denote the Fourier coefficients in the expansion of  $A_-(\lambda)$  as  $R_j$  and those in the expansion of  $A_-^{-1}(\lambda)$  as  $K_j$ , so that

$$A_-(\lambda) = \sum_{j=-\infty}^0 R_j e^{ij\lambda}, A_-^{-1}(\lambda) = \sum_{j=-\infty}^0 K_j e^{ij\lambda}.$$

Clearly, the product  $A_-(\lambda)A_-^{-1}(\lambda)$  must have the Fourier coefficient on the constant equal to  $R_0K_0$ . But the product is the identity matrix, hence  $R_0K_0 = I$ ,  $K_0$  is non-singular, and at least one element on the  $n$ -th row of  $K_0$  must be different from zero. Without loss of generality, let such a non-zero coefficient be the first one in the row. By choosing  $C(\lambda)$  equal to a constant matrix with the first  $m$  rows constituting the  $I_m$  matrix, and the rest  $n - m$  rows zero, we therefore guarantee that the element in the  $n$ -th row and the first column of  $A_-^{-1}(\lambda)C(\lambda)$  has a non-zero coefficient on the constant in its Fourier expansion. Finally, note that  $A_-^{-1}(\lambda)F(\lambda)$  has zero Fourier coefficients on  $e^{ij\lambda}$ ,  $\lambda \geq 0$  and therefore the Fourier coefficients on the constant for the right hand side of (35) coincide with those of  $A_-^{-1}(\lambda)C(\lambda)$ .

Suppose now that all partial indices of  $A(\lambda)$  are non-positive, and some are negative. Then there exists an infinite number of solutions to the spectral model and, hence, an infinite number of solutions to the original model (1). Indeed, let us define

$$[A_-^{-1}(\lambda)C(\lambda)]_- = A_-^{-1}(\lambda)C(\lambda) - [A_-^{-1}(\lambda)C(\lambda)]_+,$$

and a  $n \times m$  matrix function  $M(\lambda)$  such that all elements of this matrix except the one in the first row and the first column are equal to zero. We will define the upper left element of  $M(\lambda)$  as

$$M_{11}(\lambda) = e^{-i(k_1+1)\lambda},$$

where  $k_1$  is the smallest (and hence negative) partial index of  $A(\lambda)$ . Now, for any real number  $k$  the following is a solution to the spectral model.

$$\tilde{D}(\lambda) = A_0^{-1}(\lambda) [A_-^{-1}(\lambda)C(\lambda)]_+ + kM(\lambda).$$

To check this, substitute the above formula in (35) and choose  $F(\lambda)$  as

$$F(\lambda) = A_-(\lambda) \left( kA_0(\lambda)M(\lambda) - [A_-^{-1}(\lambda)C(\lambda)]_- \right).$$

Finally, if all partial indices of  $A(\lambda)$  are zero, then  $A_0(\lambda)$  is the  $n \times n$  identity matrix and  $\tilde{D}(\lambda)$  must be equal to

$$\tilde{D}(\lambda) = [A_-^{-1}(\lambda)C(\lambda) + A_-^{-1}(\lambda)F(\lambda)]_+.$$

But for any  $F(\lambda)$  satisfying (34), all Fourier coefficients of  $A_-^{-1}(\lambda)F(\lambda)$  on the non-negative powers of  $e^{i\lambda}$  must be equal to zero. Therefore,

$$\tilde{D}(\lambda) = [A_-^{-1}(\lambda)C(\lambda)]_+$$

is the unique solution to the spectral model and thus the original model has a unique solution. Recalling that

$$D(\lambda) = A_+^{-1}(\lambda)\tilde{D}(\lambda)$$

and substituting to (30) we have an explicit form for this solution:

$$x_t = \int_{-\pi}^{\pi} e^{it\lambda} A_+^{-1}(\lambda) [A_-^{-1}(\lambda)C(\lambda)]_+ \Phi(d\lambda).$$

**Proof of Proposition 1 for the non-stationary framework:**

We will first rewrite (18) so that values of  $x_t$  that form the initial history appear on the right hand side of the equation:

$$\sum_{j=-\infty}^t A_j E_t x_{t-j} = z_t - \sum_{j=t+1}^{\infty} A_j E_t x_{t-j}, \quad t = 0, 1, 2, \dots,$$

or, changing the index of summation and denoting the left hand side of the above equation as  $\zeta_t$ ,

$$\sum_{k=0}^{\infty} A_{t-k} E_t x_k = \zeta_t, \quad t = 0, 1, 2, \dots \quad (36)$$

Note that growth restrictions (14) and (19) imply that  $\{\zeta_t\}$  satisfies the asymptotic growth conditions:

$$|\bar{E}_t \zeta_{t+i}| \leq \tilde{B}_t \quad \forall i \geq 0 \quad (37)$$

for some random variables  $\tilde{B}_t$  with a finite absolute moment,  $E|\tilde{B}_t| < \infty$ .

Applying the conditional expectation operation  $E_s$  to both sides of equation (36) we have:

$$E_s \sum_{k=0}^{\infty} A_{t-k} E_t x_k = E_s \zeta_t, \quad t = 0, 1, 2, \dots$$

and therefore:

$$\sum_{k=0}^{\infty} A_{t-k} E_s x_k = E_s \zeta_t, \quad 0 \leq s \leq t. \quad (38)$$

The latter equations follow from the former ones because of the law of iterated expectations and the fact that the order of the conditional expectation and the infinite sum on the left hand side of the former equation is interchangeable. The order is interchangeable because condition (17) and the Lebesgue Convergence Theorem (see for example Kolmogorov and Fomin 1970) imply that the partial sums of the infinite sum converge not only almost surely but also in mean. The interchangeability follows then from the fact that the operation of taking the conditional expectation is continuous with respect to the convergence in mean.

Let us now denote  $E_s x_k, 0 \leq s \leq k$  as  $\xi_{sk}$ . For  $k < s$ , using the law of iterated expectations, we have:  $E_s x_k = E_k x_k = \xi_{kk}$ . Similarly, denote  $E_s \zeta_t, 0 \leq s \leq t$  as  $\eta_{st}$ . We will interpret equations (38) as an infinite system of equations to determine a triangular array of random variables  $\xi_{sk}, 0 \leq s \leq k$ , given a triangular array of random variables  $\eta_{st}, 0 \leq s \leq t$ . We will analyze the solvability properties of this system and then will reinterpret our results in terms of the existence and uniqueness of a process  $\{x_k\}$  satisfying (38) and, hence, satisfying the original equations (18).

We will first rewrite (38) in terms of  $\xi_{sk}$  and  $\eta_{st}$  for  $s = 0$  only. We have

$$\sum_{k=0}^{\infty} A_{t-k} \xi_{0k} = \eta_{0t}, \quad 1 \leq t. \quad (39)$$

If series  $\eta_{st}$  to be interpreted as  $E_s \zeta_t$ , the growth conditions (37) imply that  $\{\eta_{0t}\}$  belong to the space  $l_n^\infty$  of bounded sequences of  $n$ -dimensional vectors almost surely. Similarly, if  $\xi_{sk}$  to be interpreted as  $E_s x_k$ , the growth conditions (17) imply that  $\{\xi_{0k}\} \in l_n^\infty$  almost surely. In what follows, we will omit the qualification “almost surely” to save the space and will simply write  $\{\xi_{0k}\} \in l_n^\infty$  or  $\{\eta_{0t}\} \in l_n^\infty$ .

According to Theorem 4.2, Chapter VIII of Gohberg and Feldman (1974), there exists a unique sequence  $\{\xi_{0k}\} \in l_n^\infty$  satisfying equations (39) if and only if all partial indices of  $A(\lambda)$  are equal to zero. There exists no solution to (39) at least for some sequences  $\{\eta_{0t}\} \in l_n^\infty$  if and only if at least one of the partial indices of  $A(\lambda)$  is positive. Finally, there exist infinitely many solutions to (39) if and only if all partial indices of  $A(\lambda)$  are non-positive and at least one of them is negative.

Now let us consider equations (38) reformulated in terms of  $\xi_{sk}$  and  $\eta_{st}$  for  $s = 1$  only. We have

$$A_t \xi_{00} + \sum_{k=1}^{\infty} A_{t-k} \xi_{1k} = \eta_{1t}, \quad 1 \leq t.$$

Rearrange this to get

$$\sum_{k=1}^{\infty} A_{t-k} \xi_{1k} = \eta_{1t} - A_t \xi_{00}, \quad 1 \leq t, \quad (40)$$

where we assume that  $\xi_{00}$  is determined by the solution to (39). Using the theorem from Goldberg and Feldman (1974) one more time, we can show that the solvability of (40) considered as a system of equations in  $\{\xi_{1k}\}, k = 1, 2, 3, \dots$  depends on the partial indices of  $A(\lambda)$  in the same way the solvability of (39) does.

In general, we can show that equations

$$\sum_{k=i}^{\infty} A_{t-k} \xi_{ik} = \eta_{it} - \sum_{k=0}^{i-1} A_{t-k} \xi_{kk}, \quad 0 \leq i \leq t, \quad (41)$$

considered as the equations to determine a triangular array of random variables  $\{\xi_{ik}\} \in l_n^\infty, 0 \leq i \leq t$ , have a unique solution for any triangular array of random variables  $\{\eta_{st}\} \in l_n^\infty, 0 \leq s \leq t$  if all partial indices of  $A(\lambda)$  are zero, have no solutions at least for some  $\{\eta_{st}\} \in l_n^\infty$  if at least one partial index of  $A(\lambda)$  is positive, and have multiple solutions for any  $\{\eta_{st}\} \in l_n^\infty$  if all partial indices of  $A(\lambda)$  are non-positive and at least one of them is negative.

Let us now reformulate the above solvability conditions in terms of the existence and uniqueness of stochastic process  $\{x_t\}$  solving (38). We will show that: 1) if (41) has no solution, then (38) has no solution, 2) if (41) has a unique solution, then (38) has a unique solution, and 3) if (41) has many solutions, then (38) has many solutions. In other words, (41) and (38) have the same solvability properties. Let me turn to the proof of the above three statements.

1) Suppose that (41) have no solution for some array  $\{\eta_{st}\} \in l_n^\infty$ . Then at least one of the partial indices of  $A(\lambda)$  is positive and we can choose the array  $\{\eta_{st}\}$  so that (39) have no solutions for  $\{\eta_{0t}\}$ . We can further choose  $\eta_{0t}$  to be constants almost surely, and  $\eta_{st} = \eta_{0t}$  for any  $0 \leq s \leq t$ . Indeed, equations (39) have no solutions if there are no solutions to

$$\sum_{k=0}^{\infty} A_{t-k} \xi_{0k}(\omega) = \eta_{0t}(\omega), \quad 0 \leq t,$$

where  $\omega$  is arbitrary elementary event in the underlying probability space  $\Omega$ . Therefore, we can choose  $\eta_{0t} \equiv \eta_{0t}(\omega)$  so that  $\eta_{0t}$  are constants almost surely. Equations (41) will have no solution if (39) have no solution. Therefore we can choose  $\{\eta_{st}\} \in l_n^\infty, s > 0$  arbitrarily, in particular, we can choose them so that  $\eta_{st} = \eta_{0t}$  for any  $0 \leq s \leq t$ .

Assume now that (38) have solutions for any exogenous process  $\{\zeta_t\}$  satisfying the asymptotic growth restrictions. And let  $\{x_t\}$  be a solution to (38) for the exogenous process  $\zeta_t = \eta_{tt}$ , which, clearly, satisfies the growth restrictions. Consider  $\tilde{\xi}_{sk} = E_s x_k$ ,  $\tilde{\eta}_{st} = E_s \zeta_t$ . Then  $\{\tilde{\xi}_{sk}\}$  is a solution to (41) for  $\{\tilde{\eta}_{st}\}$ . But  $\{\tilde{\eta}_{st}\} = \{\eta_{st}\}$ . Indeed, we have:

$$\tilde{\eta}_{st} = E_s \zeta_t = E_s \eta_{tt} = E_s \eta_{0t} = \eta_{0t} = \eta_{st}.$$

Hence, our assumption that (38) has a solution is wrong and our first statement has been proven.

2) Suppose that (41) have a unique solution  $\{\xi_{sk}\} \in l_n^\infty$  for any  $\{\eta_{st}\} \in l_n^\infty$ . We will choose  $\{\eta_{st}\}$  so that

$$\eta_{st} = \bar{E}_s \zeta_t, 0 \leq s \leq t. \quad (42)$$

The above cited theorem from Gohberg and Feldman (1974) implies that, for any  $0 \leq s \leq k$ , solution  $\xi_{sk}$  is a linear combination of  $\eta_{it}, i \leq s, t \geq i$  with absolutely summable coefficients. Therefore, we can also write:

$$\eta_{st} = E(\zeta_t | \zeta_s, \zeta_{s-1}, \dots, \zeta_0, \xi_{ss}, \dots, \xi_{00}, \dots), 0 \leq s \leq t.$$

Consider a stochastic process  $\{x_t\}$  defined as  $x_t = \xi_{tt}, 0 \leq t$ . We will prove that  $E_s x_t = \xi_{st}, 0 \leq s \leq t$  so that  $\{x_t\}$  is a solution to (38).

Recall that  $E_s$  is defined as the operation of mathematical expectation conditional on the information available at time  $s$ , including  $\zeta_s, \zeta_{s-1}, \dots, \zeta_0$  and  $x_s, x_{s-1}, x_{s-2}, \dots$ . Hence, for  $x_t$  just defined,  $E_s$  is equivalent to the operation of mathematical expectation conditional on  $\zeta_s, \zeta_{s-1}, \dots, \zeta_0, \xi_{ss}, \dots, \xi_{00}$  and any additional exogenous information available at  $t$ , and, therefore, by the law of iterated expectations,  $E_s \eta_{it} = \eta_{st}$  for any  $s \leq i$ .

Let us apply  $E_s$  to both sides of (41). We have:<sup>18</sup>

$$\sum_{k=i}^{\infty} A_{t-k} E_s \xi_{ik} = E_s \eta_{it} - \sum_{k=0}^{i-1} A_{t-k} E_s \xi_{kk}, 0 \leq i \leq t.$$

For  $0 \leq s \leq i$  the above equation simplifies to

$$\sum_{k=i}^{\infty} A_{t-k} E_s \xi_{ik} = \eta_{st} - \sum_{k=0}^{i-1} A_{t-k} E_s \xi_{kk}, 0 \leq i \leq t.$$

But we assumed that equations (41) have unique solution. Therefore,

$$E_s \xi_{ik} = \xi_{sk}$$

for any  $0 \leq s \leq i \leq k$ . In particular

$$E_s x_i = E_s \xi_{ii} = \xi_{si}$$

and  $\{x_t\}$  is a solution to (38).

Suppose that (38) have many solutions, then (38) have a nontrivial solution  $\{\varphi_t\}$  for  $\zeta_t \equiv 0$ . Solution  $\{\varphi_t\}$  generates a non-trivial solution  $\phi_{sk} = E_s \varphi_k$  to (41) for  $\eta_{sk} \equiv 0$ . Now if  $\xi_{sk}$  is any solution to (41) with arbitrary  $\eta_{st}$ , then  $\xi_{ks} + \gamma \phi_{sk}$  is also a solution, where  $\gamma$  is any real number. But we assumed that (41) has only one

---

<sup>18</sup>The above discussion of the validity of interchanging the order of operations of conditional expectation and infinite sum applies here.

solution. Therefore, (38) has only one solution and our second statement has been proven.

3) Suppose that (41) have many solutions  $\{\xi_{sk}\} \in l_n^\infty$  for any  $\{\eta_{st}\} \in l_n^\infty$ . As we showed above,  $A(\lambda)$  must have non-positive partial indices some of which are negative. This is also the necessary and sufficient condition for the existence of an infinite number of solutions to (39). Therefore, there exists a solution  $\{\phi_{sk}\}, 0 \leq s \leq k$  to (41) with  $\eta_{st} \equiv 0$ , such that, for some  $k, \phi_{0k} \neq 0$ . The random variables  $\phi_{sk}$  can be chosen to be equal to constants almost everywhere. Define  $\varphi_t = \phi_{0t}$ . Clearly,  $\{\varphi_t\}$  is a non-trivial solution to (38) for  $\zeta_t \equiv 0$ . Hence, if for arbitrary  $\{\zeta_t\}$  satisfying growth restrictions (37) there exists at least one solution  $\{x_t\}$  to (38) that satisfy the growth restrictions (17), then all processes  $x_t + \gamma\varphi_t$  are solutions to (38). That is, the number of solutions to (38) is infinite.

It remains to be proven that there is at least one solution to (38) for arbitrary  $\zeta_t$  satisfying the growth restrictions. Let  $\{\eta_{st}\}$  be as in (42) and  $\{\xi_{sk}\}$  is a solution to (41) such that  $\xi_{sk}$  is a linear combination of  $\eta_{it}, i \leq s, t \geq i$  with absolutely summable coefficients (existence of such a solution is implied by Gohberg-Feldman theorem referred to above). We now define an alternative solution  $\{\tilde{\xi}_{sk}\}$  and a process  $\{x_t\}$  that will be a solution to (38). Our definition is recursive:

$$\begin{aligned} \tilde{\xi}_{0k} &= \xi_{0k}, k \geq 0, & x_0 &= \tilde{\xi}_{00}, \\ \tilde{\xi}_{1k} &= \xi_{1k} - \left[ E(\xi_{1k} | \zeta_0, x_0, \dots) - \tilde{\xi}_{0k} \right], k \geq 1, & x_1 &= \tilde{\xi}_{11}, \\ &\dots & & \\ \tilde{\xi}_{ik} &= \xi_{ik} - \left[ E(\xi_{ik} | \zeta_{i-1}, \dots, \zeta_0, x_{i-1}, \dots, x_0, \dots) - \tilde{\xi}_{i-1,k} \right], k \geq i, & x_i &= \tilde{\xi}_{ii}, \\ &\dots & & \end{aligned}$$

As before, the conditioning sets include, beside  $\zeta_{i-1}, \dots, \zeta_0, x_{i-1}, \dots, x_0$ , all “extra” exogenous information available at time  $i - 1$ .

First, note that  $\{\tilde{\xi}_{sk}\}$  belongs to  $l_n^\infty$ . This follows from the fact that  $\xi_{sk}$  is a linear combination of  $\eta_{it}, i \leq s, t \geq i$  with absolutely summable coefficients, from the definition of  $\eta_{it}$  (42), and from growth restrictions (37) satisfied by  $\{\zeta_t\}$ . The fact that  $\{\tilde{\xi}_{sk}\}$  is a solution to (41) is trivially true for  $s = 0$  and can be proven for general  $s$  by the mathematical induction.

Let us show that  $\{x_t\} = \{\tilde{\xi}_{tt}\}$  is a solution to (38). It is enough to prove that  $E_s x_t = E_s \tilde{\xi}_{tt} = \tilde{\xi}_{st}$  because if this is true, then equations (41) with  $i = s$  imply (38). We have:

$$\begin{aligned} E_{t-1} \tilde{\xi}_{tt} &= E_{t-1} \xi_{tt} - E(\xi_{tt} | \zeta_{t-1}, \dots, \zeta_0, x_{t-1}, \dots, x_0, \dots) + \tilde{\xi}_{t-1,t} \\ &= E_{t-1} \xi_{tt} - E_{t-1} \xi_{tt} + \tilde{\xi}_{t-1,t} = \tilde{\xi}_{t-1,t}. \end{aligned}$$

Noting that  $E_s \tilde{\xi}_{tt} = E_s E_{s+1} \dots E_{t-1} \tilde{\xi}_{tt}$ , and applying the above equality many times, we get the desired result.

### Proof of Proposition 3:

Let us consider a function

$$L(\lambda) = e^{i\lambda}A(\lambda) = B \left[ I_{k+m} - \begin{pmatrix} J_1 e^{i\lambda} & 0 \\ 0 & J_2 e^{i\lambda} \end{pmatrix} \right] B^{-1} \begin{pmatrix} I_k e^{i\lambda} & 0 \\ 0 & I_m \end{pmatrix}.$$

According to Gohberg, Lerer, and Rodman (1978), the following equalities hold for the partial indices  $l_1 \leq l_2 \leq \dots \leq l_{k+m}$  of  $L(\lambda)$  :

$$l_i = \# \{j | k + m + r_{j-1} - r_j \leq i - 1, j = 1, 2\}, \quad i = 1, 2, \dots, k + m, \quad (43)$$

where

$$r_1 = \text{rank} \begin{bmatrix} C_{-1} \\ C_{-2} \end{bmatrix}, r_2 = \text{rank} \begin{bmatrix} C_{-1} & C_{-2} \\ C_{-2} & C_{-3} \end{bmatrix},$$

and

$$C_{-j} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ij\lambda} L^{-1}(\lambda) d\lambda.$$

From the definition of  $L(\lambda)$ , it is clear that the partial indices of  $A(\lambda)$  are equal to those of  $L(\lambda)$  minus one. Therefore, the partial indices of  $A(\lambda)$  are not all of the same sign or zero if and only if  $l_1 - 1 < 0$  and  $l_{k+m} - 1 > 0$ . Since  $0 \leq l_i \leq 2$ , the former inequalities are equivalent to  $l_1 = 0$ ,  $l_{k+m} = 2$  and (43) implies

$$1 \leq r_1 < k + m, \quad 1 \leq r_2 - r_1 < k + m. \quad (44)$$

Below, we will show that these inequalities hold if and only if  $B_{11}$  does not have full rank, which will prove the proposition.

Indeed, using the above formula for  $C_{-j}$ , we get

$$C_{-1} = \begin{pmatrix} B_{11} & 0 \\ 0 & B_{22}J_2^{-1} \end{pmatrix}, C_{-2} = \begin{pmatrix} 0 & B_{12}J_2^{-1} \\ 0 & B_{22}J_2^{-2} \end{pmatrix}, C_{-3} = \begin{pmatrix} 0 & B_{12}J_2^{-2} \\ 0 & B_{22}J_2^{-3} \end{pmatrix}.$$

Therefore,

$$\begin{aligned} r_1 &= \text{rank} \begin{bmatrix} B_{11} & 0 \\ 0 & B_{22}J_2^{-1} \\ 0 & B_{12}J_2^{-1} \\ 0 & B_{22}J_2^{-2} \end{bmatrix} = \text{rank } B_{11} + \text{rank} \begin{bmatrix} B_{12} \\ B_{22} \end{bmatrix} = \text{rank } B_{11} + \bar{m}, \\ r_2 &= \text{rank} \begin{bmatrix} B_{11} & 0 & 0 & B_{12}J_2^{-1} \\ 0 & B_{22}J_2^{-1} & 0 & B_{22}J_2^{-2} \\ 0 & B_{12}J_2^{-1} & 0 & B_{12}J_2^{-2} \\ 0 & B_{22}J_2^{-2} & 0 & B_{22}J_2^{-3} \end{bmatrix} = \text{rank} \begin{bmatrix} B_{11} & 0 & B_{12}J_2^{-1} \\ 0 & B_{22}J_2^{-1} & 0 \\ 0 & B_{12}J_2^{-1} & 0 \\ 0 & B_{22}J_2^{-2} & 0 \end{bmatrix} \\ &= \text{rank} [ B_{11} \quad B_{12} ] + \text{rank} \begin{bmatrix} B_{12} \\ B_{22} \end{bmatrix} = k + \bar{m}. \end{aligned}$$

Hence, inequalities (44) can be rewritten as:

$$1 \leq \text{rank } B_{11} + \bar{m} < k + m, \quad 1 \leq k - \text{rank } B_{11} < k + m,$$

which is the same as

$$-\bar{m} < \text{rank } B_{11} < \bar{k}, \quad -m < \text{rank } B_{11} < k,$$

or, finally,

$$-\min \{m, \bar{m}\} < \text{rank } B_{11} < \min \{k, \bar{k}\}.$$

The right-most inequality is equivalent to the statement that  $B_{11}$  does not have full rank. We, however, have an “extra” inequality  $-\min \{m, \bar{m}\} < \text{rank } B_{11}$  that must be satisfied for some partial indices of  $A(\lambda)$  to have different signs. This inequality is only non-trivial if  $\min \{m, \bar{m}\} = 0$ . But in such a case,  $B_{11}$  either contains at least one full row of  $B$  if  $\bar{m} = 0$ , or contains at least one full column of  $B$  if  $m = 0$ . Therefore, since  $B$  cannot be singular,  $\text{rank } B_{11}$  must be non-zero and the inequality  $-\min \{m, \bar{m}\} < \text{rank } B_{11}$  always holds and does not bring any information in addition to the requirement that  $B_{11}$  is not of full rank.