Determinacy and Learnability in Monetary Policy Analysis: Additional Results

Bennett T. McCallum

Carnegie Mellon University

and

National Bureau of Economic Research

Preliminary

November 8, 2007

This paper is being prepared for the 2007 Annual Conference of the Central Bank of Chile, to be held in Santiago on November 15-16, 2007. I am indebted to Riccardo DiCecio and George Evans for helpful discussions on the topic of this paper.

1. Introduction

It is almost unnecessary to begin by emphasizing that recent research in monetary policy analysis has featured a great deal of work concerning conditions for determinacy—i.e., existence of a unique dynamically stable rational expectations equilibrium—under various specifications of policy behavior.[1] Indeed, there are a number of papers in which determinacy is the only criterion for a desirable monetary policy regime that is explicitly mentioned.[2]

By contrast, I have argued in a recent publications (McCallum, 2003a, 2007) that least-squares (LS) learnability is a compelling necessary condition for a rational expectations (RE) equilibrium to be considered plausible, since individuals must somehow learn about the exact nature of an economy from data generated by that economy itself, while the LS learning process is biased toward a finding of learnability. A similar position has also been expressed by Bullard (2006, p. 2004). From such a position it follows that in conditions in which there is more than one dynamically stable RE solution—i.e., indeterminacy—there may still be only one RE solution that is economically relevant, if the others are not LS learnable. In this sense, LS learnability is arguably a more important criterion than determinacy.

Substantively, my 2007 paper demonstrates that, in a very wide class of linear RE models, determinacy implies LS learnability (but not the converse) when individuals have knowledge of current conditions available for use in the learning process. This strong result does not pertain, however, if individuals have available, in the learning process, only information regarding previous values of endogenous variables.[3] One task of the present paper, accordingly, is to investigate the situation that obtains when only lagged information is available. In addition, the

[1] Prominent examples include Benhabib, et. al. (2001), Clarida, Gali, and Gertler (1999), Rotemberg and Woodford (1997), Sims (1994), and Woodford (2003). Discussion in a leading textbook is provided by Walsh (2003).
[2] See, for example, Carlstrom and Fuerst (2005).
[3] Another limitation of the analysis of McCallum (2007) is that it considers only solutions of a form that excludes "resonant frequency sunspot" solutions. That limitation, which is maintained here, is discussed briefly in Section 6.

paper will explore results that pertain when a further criterion of model plausibility, provisionally termed "well-formulated," characterizes the model's structure. In particular, it is shown that models that are well formulated, in the defined sense, invariably possess the property of E-stability and hence LS learnability if current-period information is available in the learning process, even if determinacy does not prevail. The situation in the case of lagged information is less favorable—i.e., learnability is assured only in special cases.

2. Model and Determinacy

It will be useful to begin with a summary of the formulation and results developed in McCallum (2007). Throughout we will work with a model of the form

(1) $y_t = A\, E_t y_{t+1} + C\, y_{t-1} + D\, u_t$,

where y_t is a m×1 vector of endogenous variables, A and C are m×m matrices of real numbers, D is m×n, and u_t is a n×1 vector of exogenous variables generated by a dynamically stable process

(2) $u_t = R\, u_{t-1} + \varepsilon_t$,

with ε_t a white noise vector. It will not be assumed, even initially, that A is invertible. This specification is useful in part because it is the one utilized in Section 10.3 of Evans and Honkapohja (2001), for which E-stability conditions are reported on their p. 238.[4] Furthermore, the specification is very broad; in particular, any model satisfying the formulations of King and Watson (1998) or Klein (2001), can be written in this form—which will accommodate any number of lags, expectational leads, and lags of leads. (See Appendix A.)

Following McCallum (1983, 1998), we consider solutions to (1)(2) of the form

(3) $y_t = \Omega\, y_{t-1} + \Gamma\, u_t$.

in which Ω is required to be real. Then we have that $E_t y_{t+1} = \Omega(\Omega y_{t-1} + \Gamma u_t) + \Gamma R u_t$ and

[4]Constant terms can be included in the equations of (1) by including an exogenous variable in u_t that is a random walk whose innovation has variance zero. In this case there is a borderline departure from process stability.

straightforward undetermined-coefficient reasoning shows that Ω and Γ must satisfy

(4) $A\Omega^2 - \Omega + C = 0$

(5) $\Gamma = A\Omega\Gamma + A\Gamma R + D.$

For any given Ω, (5) yields a unique Γ generically,[5] but there are many m×m matrices that solve (4) for Ω. Accordingly, the following analysis centers around (4). Since we do not assume that A is invertible, we write

(6) $$\begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \Omega^2 \\ \Omega \end{bmatrix} = \begin{bmatrix} I & -C \\ I & 0 \end{bmatrix} \begin{bmatrix} \Omega \\ I \end{bmatrix},$$

in which the first row reproduces the matrix quadratic (4). Let the 2m×2m matrices on the left and right sides of (6) be denoted $\bar{A}$ and $\bar{C}$, respectively. Then instead of focusing on the eigenvalues of $\bar{A}^{-1}\bar{C}$, which does not exist when A is singular, we instead solve for the (generalized) eigenvalues of the matrix pencil $[\bar{C} - \lambda\bar{A}]$, alternatively termed the (generalized) eigenvalues of $\bar{C}$ with respect to $\bar{A}$ (e.g., Uhlig (1999)). Thus instead of diagonalizing $\bar{A}^{-1}\bar{C}$, as in Blanchard and Khan (1980), we use the Schur generalized decomposition, which serves the same purpose. Specifically, the Schur generalized decomposition theorem establishes that there exist unitary matrices Q and Z such that $Q\bar{C}Z = T$ and $Q\bar{A}Z = S$ with T and S triangular.[6] Then eigenvalues of the matrix pencil $(\bar{C} - \lambda\bar{A})$ are defined as t_{ii}/s_{ii}. Some of these eigenvalues may be "infinite," in the sense that some s_{ii} may equal zero. This will be the case, indeed, whenever A and therefore $\bar{A}$ are of less than full rank since then S is also singular. All of the foregoing is true for any ordering of the eigenvalues and associated columns of Z (and rows of Q). For the

---

[5] Generically, $I - R'\otimes[(I - A\Omega)^{-1}A]$ will be invertible, permitting solution of (5) for vec(Γ). Invertibility of $(I - A\Omega)$ is discussed below in Section 4.

[6] Provided only that there exists some λ for which $\det[\bar{C} - \lambda\bar{A}] \neq 0$. See Klein (2000) or Golub and Van Loan (1996, p. 377). Note that in McCallum (2007) the matrices $\bar{A}$ and A are denoted A and A_{11}, repectively.

present, let us focus on the arrangement that places the t_{ii}/s_{ii} in order of decreasing modulus.[7]

To begin the analysis, premultiply (6) by Q. Since $Q\bar{A} = SH$ and $Q\bar{C} = TH$, where $H \equiv Z^{-1}$, the resulting equation can be written as

$$(7) \qquad \begin{bmatrix} S_{11} & 0 \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \Omega^2 \\ \Omega \end{bmatrix} = \begin{bmatrix} T_{11} & 0 \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \Omega \\ I \end{bmatrix}.$$

The first row of (7) reduces to

$$(8) \qquad S_{11}(H_{11}\Omega + H_{12})\Omega = T_{11}(H_{11}\Omega + H_{12}).$$

Then if H_{11} is invertible the latter can be used to solve for Ω as

$$(9) \qquad \Omega = -H_{11}^{-1} H_{12} = -H_{11}^{-1}(-H_{11}Z_{12}Z_{22}^{-1}) = Z_{12}Z_{22}^{-1},$$

where the second equality comes from the upper right-hand submatrix of the identity $HZ = I$, provided that H_{11} is invertible, which we assume without significant loss of generality.[8] [9]

As mentioned above, there are many solutions Ω to (4). These correspond to different arrangements of the eigenvalues, which result in different groupings of the columns of Z and therefore different compositions of the submatrices Z_{12} and Z_{22}. Here, with the eigenvalues t_{ii}/s_{ii} arranged in order of decreasing modulus, the diagonal elements of S_{22} will all be non-zero provided that S has at least m non-zero eigenvalues, which we assume to be the case.[10] Clearly, for any solution under consideration to be dynamically stable, the eigenvalues of Ω must be smaller than 1.0 in modulus. In McCallum (2007) it is shown that

$$(10) \qquad \Omega = Z_{22}S_{22}^{-1}T_{22}Z_{22}^{-1}$$

[7] The discussion proceeds as if none of the t_{ii}/s_{ii} equals 1.0 exactly. If one does, the model can be adjusted, by multiplying some relevant coefficient by (e.g.) 0.9999.

[8] This invertibility condition, also required by King and Watson (1998) and Klein (2000), obtains except for degenerate special cases of (1) that can be solved by simpler methods than considered here. Note that the invertibility of H_{11} implies the invertibility of Z_{22}, given that Z and H are unitary.

[9] Note that it is not being claimed that all solutions are of the form (9).

[10] From its structure it is obvious that $\bar{A}$ has at least m nonzero eigenvalues so, since Q and Z are nonsingular, S must have rank of at least m. This necessary condition is not sufficient for S to have at least m nonzero eigenvalues, however; hence the assumption.

so Ω has the same eigenvalues as $S_{22}^{-1}T_{22}$. The latter is triangular, moreover, so the relevant eigenvalues are the m smallest of the 2m ratios t_{ii}/s_{ii} (given the decreasing-modulus ordering). For dynamic stability, the modulus of each of these ratios must then be less than 1. [In many cases, some of the m smallest moduli will equal zero.]

Let us henceforth refer to the solution under the decreasing-modulus ordering as the MOD solution. Now suppose that the MOD solution is stable. For it to be the only stable solution, there must be no other arrangement of the t_{ii}/s_{ii} that would result in a Ω matrix with all eigenvalues smaller in modulus than 1.0. Thus each of the t_{ii}/s_{ii} for i = 1,…, m must have modulus greater than 1.0, some perhaps infinite. Is there some m×m matrix whose eigenvalues relate cleanly to these ratios? Yes, it is the matrix $F \equiv (I - A\Omega)^{-1}A$, which appears frequently in the analysis of Binder and Pesaran (1995, 1997).[11] Regarding this F matrix, it is shown that, for any ordering such that H_{11} is invertible, including the MOD ordering, we have the equality

(11) $$H_{11} F H_{11}^{-1} = T_{11}^{-1}S_{11},$$

which implies that F has the same eigenvalues as $T_{11}^{-1}S_{11}$. In other words, it is the case that the eigenvalues of F are the same, for any given arrangement of the system's eigenvalues, as the <u>inverses</u> of the values of t_{ii}/s_{ii} for i = 1, …, m. Under the MOD ordering these are the inverses of the first (largest) m of the eigenvalues of the system's matrix pencil. Accordingly, for solution (9) to be the only stable solution, all the eigenvalues of the corresponding F must be smaller than 1.0 in modulus. This result, stated in different ways, is well known from Binder and Pesaran (1995), King and Watson (1998), and Klein (2000), and is an important generalization of one result of Blanchard and Khan (1980) for a model with nonsingular A.

---

[11] There is no general proof of invertibility of $[I - A\Omega]$, but if $A\Omega$ were by chance to have some eigenvalue exactly equal to 1.0, that condition could be eliminated by making some small adjustment to elements of A or C. Also, see Section 5 below.

Thus we have established notation for models of form (1)(2) and have reported results showing that the existence of a unique stable solution requires that all eigenvalues of the defined Ω matrix and the corresponding F must be less than 1.0 in modulus. It will be convenient to express that condition as follows: all $|\lambda_\Omega| < 1$ and all $|\lambda_F| < 1$.

3. E-Stability in Two Cases

We now turn to conditions for learnability under two different information assumptions. First we will review the main results from my JEDC paper, which assumes that agents have full information on current values of endogenous variables during the learning process, and then we will go on to the second assumption, namely, that only lagged values of endogenous variables are known during the learning process. The manner in which learning takes place in the Evans-Honkapohja (E&H) analysis is as follows. Agents are assumed to know the structure of the economy as specified in equations (1) and (2), in the sense that they know what variables are included, but do not know the numerical values of the parameters. What they need to know, to form expectations, is values of the parameters of the solution equations (3). In each period t, they form forecasts on the basis of least squares regression of the variables in y_{t-1} on previous values of y_{t-2} and any exogenous observables. Given those regression estimates, however, expectations of y_{t+1} may be calculated assuming knowledge of y_t or, alternatively, assuming that y_{t-1} is the most recent observation possessed by agents and thus usable in the forecasting process. In the former case, the conditions for E-stability reported by E&H (2001, p. 238) are that the following three matrices must have all eigenvalues with real parts less than 1.0:

(12a) $F \equiv (I - A\Omega)^{-1}A$

(12b) $[(I - A\Omega)^{-1}C]' \otimes F$

(12c) $R' \otimes F$.

In the second case, however, the analogous condition (E&H, 2001, p. 245) is that the following matrices must have all eigenvalues with real parts less than 1.0:

(13a) $A(I + \Omega)$

(13b) $\Omega' \otimes A + I \otimes A\Omega$

(13c) $R' \otimes A + I \otimes A\Omega$.

Except in the case that $\Omega = 0$, which will obtain when $C = 0$, these conditions are not equivalent to those in (12).

It is important to note that use of the first information assumption is not inconsistent with a model specification in which supply and demand decisions in period t are based on expectations formed in the past, such as $E_{t-1}y_{t+j}$ or $E_{t-2}y_{t+j}$. It might also be mentioned parenthetically that conditions (12) and (13) literally pertain to the E-stability of the model (1)(2) under the two information assumptions, not its learnability. Under quite broad conditions, however, E-stability is necessary and sufficient for LS learnability. This near-equivalence is referred to by E&H as the "E-stability principle" (E&H, 1999, p. 472; 2001, p. 41). Since E-stability is technically easier to verify, applied analysis typically focuses on it, rather than on direct exploration of learnability.

Given the foregoing discussion, it is a simple matter to verify that if a model of form (1)(2) is determinate, then it satisfies conditions (12). First, determinacy requires that all eigenvalues of F must have modulus less than 1.0, so their real parts must all be less than 1.0, thereby satisfying (12a). Second, from equation (4) it can be seen that $(I - A\Omega)^{-1}C = \Omega$. Therefore, matrix (12b) can be written as $\Omega' \otimes F$. Furthermore, it is a standard result (Magnus and Neudecker, 1988, p. 28) that the eigenvalues of a Kronecker product are the products of the eigenvalues of the relevant matrices (e.g., the eigenvalues of $\Omega' \otimes F$ are the products $\lambda_\Omega \lambda_F$).

Therefore, condition (12b) holds. Finally, since $|\lambda_F| < 1$, condition (12c) holds provided that all $|\lambda_R| \leq 1$, which we have assumed by specifying that (2) is dynamically stable.

Determinacy does not imply learnability, however, under the second information assumption. This point, which is developed by E&H (2001, pp. 174-181), can be illustrated by means of a bivariate example.[12] Let the y_t vector in (1) include two variables, y_{1t} and y_{2t}, related by the dynamic model that follows:

$$(14) \quad \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} -0.01 & 0.01 \\ 0.99 & -0.01 \end{bmatrix} \begin{bmatrix} E_t y_{1t+1} \\ E_t y_{2t+1} \end{bmatrix} + \begin{bmatrix} 0.02 & 1.10 \\ 0.01 & 0.06 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}.$$

Then for the MOD solution we have

$$(15) \quad A\Omega = \begin{bmatrix} -0.01 & 0.01 \\ 0.99 & -0.01 \end{bmatrix} \begin{bmatrix} 0.0218 & 1.1133 \\ -0.095 & -0.774 \end{bmatrix} = \begin{bmatrix} -0.0012 & -0.0189 \\ 0.0225 & 1.1099 \end{bmatrix},$$

with eigenvalues of Ω being –0.148 and –0.604, while $F = \begin{bmatrix} 0.1604 & 0.00831 \\ -9.040 & 0.0893 \end{bmatrix}$, which has (complex) eigenvalues $0.1249 \pm 0.2717\, i$. Inspection of these shows that this solution is determinate, and that conditions (12a) and (12b), relevant for E-stability in the case in which current information is available during learning, are satisfied. Let us assume R = 0, i.e., white noise disturbances, for simplicity. Then the determinate RE solution is E-stable and learnable under the first information assumption.

But for the case with only lagged information during learning, we need to consider the eigenvalues of the matrices shown in expressions (13). For (13a), the matrix $A(I + \Omega)$ is $\begin{bmatrix} -0.0112 & -0.0089 \\ 1.0125 & 1.0999 \end{bmatrix}$ whose eigenvalues are –0.0030 and 1.0918. The last of these violates the

[12] Its specification is close numerically to the qualitative version of the E&H example that is used in McCallum (2007, pp. 1386-1388).

condition for (13a), however, so under the lagged-information assumption the relevant E-stability condition is not satisfied and the determinate RE equilibrium is not LS learnable.

This result exemplifies the fact that determinacy is not generally sufficient for learnability of RE solutions, although it is sufficient under the first information assumption. Of equal importance, in my opinion, is the fact that determinacy is not necessary for learnability. In particular, the MOD solution can be learnable, and be the only learnable solution, in cases in which indeterminacy prevails. One such example is given in my JEDC paper on p. 1386. In such cases, the position that learnability is necessary for a solution to be plausible would suggest that there may be no problem implied by the absence of determinacy.[13]

4. Well-Formulated Models

In an unpublished working paper (McCallum, 2003b), I have suggested that there is a distinct and neglected property that dynamic models should possess to be considered "well-formulated" and plausible for the purpose of economic analysis. To begin the discussion, consider first the single-variable case of specification (1),

(16) $y_t = aE_t y_{t+1} + cy_{t-1} + u_t,$

with $u_t = (1-\rho)\eta + \rho u_{t-1} + w_t$ with $|\rho| < 1$ and w_t white noise. Thus u_t is an exogenous forcing variable with an unconditional mean of η (assumed nonzero) and units have been chosen so that there is no constant term. Applying the unconditional expectation operator to (16) yields

(17) $E\, y_t = aEy_{t+1} + cEy_{t-1} + \eta.$

In this case y_t will be covariance stationary, and we have

(18) $E\, y_t = \eta / [1-(a+c)].$

But from the latter, it is clear that as a + c approaches 1.0 from above, the unconditional mean of

[13] Disregarding, that is, "sunspot" solutions not of form (3).

y_t approaches $-\infty$ (assuming without loss of generality that $\eta > 0$), whereas if $a + c$ approaches 1.0 from below, the unconditional mean approaches $+\infty$. Thus there is an infinite discontinuity at $a + c = 1.0$. This implies that a tiny change in $a + c$ could alter the average (i.e., steady state) value Ey_t from an arbitrarily large positive number to an arbitrarily large negative number. Such a property seems highly implausible and therefore unacceptable for a well-formulated model.[14] The substantive problem is not eliminated, obviously, by adoption of the zero-measure exclusion $a + c \neq 1$.

In light of the foregoing observation, it is my contention that, to be considered well formulated (WF), the model at hand needs to include a restriction on its admissible parameter values; a restriction that rules out $a + c = 1$, and yet admits a large interval of values that includes $(a, c) = (0, 0)$. In the case at hand, the appropriate restriction is $a + c < 1$. Of course, $a + c > 1$ would serve just as well mathematically to avoid the infinite discontinuity, but it seems clear that $a + c < 1$ is vastly more appropriate from an economic perspective since it includes the values (0, 0).[15] Since we want this condition to apply to $a + c$ sums between zero and that value that pertains to the model at hand, our requirement for WF is that a and c satisfy $1 - \varepsilon(a + c) > 0$ for all $0 \leq \varepsilon \leq 1$. [It should be clear, in addition, that the foregoing argument could be easily modified to apply to y_t processes that are trend stationary, rather than strictly (covariance) stationary.] It is shown in McCallum (2003b) that under this requirement, plus a second one to be discussed shortly, the univariate model (16) is invariably E-stable.[16]

Next, for the bivariate case of model (1), extension of the foregoing WF property requires

---

[14] The model could be formulated with the exogenous variable also written in terms of percent or fractional deviations from the reference level η, e.g., $\hat{u}_t = u_t - \eta$. But that would not alter the relationship between Ey_t and η, which can be extremely different for tiny changes in $a + c$.

[15] In models of the linear form (16), one would expect coefficients a and c typically to represent elasticities and often to be numerically small relative to 1.

[16] That paper's analysis of multivariate systems is, however, unsatisfactory.

that A and C are such that det[I – ε(A + C)] is positive for all $0 \leq \varepsilon \leq 1$; otherwise the steady-state values of the variables may possess infinite discontinuities. But there are other requirements as well. Let ac_{ij} temporarily denote the ijth element of A + C. Then the model with $y_1 = Ey_{1t}$, $y_2 = Ey_{2t}$, $\eta_1 = Eu_{1t}$, and $\eta_2 = Eu_{2t}$ implies

$$(19) \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} ac_{11} & ac_{12} \\ ac_{21} & ac_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$$

so that $Ey = [I - (A+C)]^{-1}\eta$ can be written as

$$(20) \qquad \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} 1-ac_{22} & ac_{12} \\ ac_{21} & 1-ac_{11} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$$

where $\Delta = \det[I - (A + C)] = (1-ac_{11})(1-ac_{22}) - ac_{12}ac_{21}$. Then the counterpart of the univariate requirement that $1 - (a+c) > 0$ includes the condition $\Delta > 0$. We must rule out, however, the case in which $\Delta > 0$ results from $1-ac_{11}$ and $1-ac_{22}$ both being negative.[17] The condition on Δ should be extended, therefore, to also require $1-ac_{11} > 0$ and $1-ac_{22} > 0$. And, furthermore, it should be the case that $1-ac_{22}$ is larger in absolute magnitude than ac_{12}, with a similar requirement for y_2. Otherwise, (20) could imply that the sign of y_i and η_i would be different. But under the conditions just stated, that unattractive anomaly would not occur.

How are these WF requirements extended to pertain to cases with more than two variables? One would naturally require that [I – (A+C)] must be a P-matrix, which has all its principal minors positive, and implies that $[I - (A+C)]^{-1}$ is also a P-matrix .[18] However, it transpires that that condition is necessary but not sufficient to establish the argument with respect to learnability that will be of concern below. Accordingly, suppose that we follow the suggestion above, that we require that the off-diagonal elements of $[I - (A+C)]^{-1}$ do not

[17] This is clear for the case in which A + C is a diagonal matrix.

[18] On the topic of P-matrices, see Horn and Johnson (1991) and Gale and Nikaido (1965).

outweigh those on the diagonal. That possibility can be ruled out by requiring that [I – (A+C)] is (positive) dominant-diagonal (PDD). This condition is somewhat stronger than is literally required by our objective of ruling out specifications in which leading implications of the model are hyper-sensitive to parameter values, but the PDD requirement is sufficient for our purpose and is one that has an important tradition stemming from the literature on multimarket stability analysis.

As a brief but relevantdigression, one example of a matrix that is a P-matrix and yet is not positive dominant-diagonal is as follows:

$$(21) \qquad \begin{bmatrix} 0.08 & -0.92 & 0.90 \\ 0.92 & 0.07 & -0.03 \\ -0.72 & 0.30 & 0.04 \end{bmatrix}.$$

Clearly, the entries in any row show immediately that this matrix is not positive dominant diagonal (PDD). But its determinant is 0.3087 and the three second-order minors are 0.0118, 0.651, and 0.852. Since the diagonal elements are also all positive, the matrix is a P-matrix. For future reference, we note that its eigenvalues are –0.0067 + 1.2319i, –0.0067 – 1.2319i, and 0.2034. Thus the example illustrates the fact that, although a P-matrix cannot have a negative real eigenvalue, it can have a complex eigenvalue pair with negative real parts.[19]

Returning now to the main line of argument, there is a second type of discontinuity that should also be eliminated for a model to be viewed as WF, namely, infinite discontinuities in its impulse response functions. In model (1)(2) with solution (3), the impulse response to the shock vector u_t (3) involves the matrix Γ, which is given by

$$(22) \qquad \Gamma = A\Omega\Gamma + A\Gamma R + D.$$

Thus $(I - A\Omega)\Gamma = A\Gamma R + D$ so using $F = (I - A\Omega)^{-1}A$, equation (22) can be written as

[19] See Horn and Johnson (1991, p. 123).

(23) $\Gamma = F\Gamma R + (I - A\Omega)^{-1}D$.

Then using the well-known identity that, for any conformable matrix product ABC, it is true that vec ABC = $(C' \otimes A)$ vec B,[20] we have

(24) vec $\Gamma = (R' \otimes F)$ vec Γ + vec $[(I - A\Omega)^{-1}D]$

implying

(25) vec $\Gamma = [I - (R' \otimes F)]^{-1}$ vec $[(I - A\Omega)^{-1}D]$.

Accordingly, our second WF requirement is for $[I - (R' \otimes F)]$ and $(I - A\Omega)$ to be well behaved in the same manner as $I - (A + C)$, i.e., that they are PDD matrices

5. E-Stability in WF Models

In this section, I begin by showing that if a model of form (1) is well-formulated, in the sense specified above, then the solution provided by the MOD ordering is, in all cases, E-stable and therefore LS learnable under the first information assumption. The WF property stipulates that the matrices $(I - A\Omega)$ and $[I - (R' \otimes F)]$ are well behaved in the sense of being PDD matrices. But this implies that their eigenvalues all have positive real parts—see Horn and Johnson (1985, p. 349). Thus we see from the second of these expressions that the E&H criterion (12c) is met.[21] But, furthermore, one of the eigenvalues of R will be 1.0, since we will include among the system's disturbances a random walk with zero variance, in order to be able to include constant terms in the specification. Then, with 1.0 as one of the eigenvalues of R, it will be the case that M of the eigenvalues of $(R' \otimes F)$ will be the eigenvalues of F. Then use of the property in footnote 21 shows that criterion (12a) is met. Then how about the remaining condition (12b)? Here we recognize that, by rearrangement of (4), $(I - A\Omega)^{-1}C = \Omega$.

[20] See, for example, E&H (2001, p. 117) or Magnus and Neudecker (1988, p. 28).

[21] Here, and often in what follows, I use the fact that the eigenvalues of a matrix of form (I – B) satisfy $\lambda_{I-B} = 1 - \lambda_B$.

Accordingly, (12b) becomes $\Omega' \otimes F$. But then note that with the MOD ordering it is the case that all $|\lambda_\Omega| < 1/|\lambda_F|$ so all $|\lambda_\Omega|\,|\lambda_F| < 1$. But $|\lambda_\Omega|\,|\lambda_F| = |\lambda_\Omega\lambda_F| \geq \text{Re}(\lambda_\Omega\lambda_F)$ so (12b) is invariably satisfied. Accordingly, we see that, with current information available during the learning process, the MOD solution to all well-formulated models of form (1)(2) is E-stable and thus LS learnable. This result does not require or imply determinacy; there can be a multiplicity of (dynamically) stable solutions. Also, it should be said explicitly, WF is not a necessary condition for either E-stability or determinacy. The setup in equations (14)-(15), for example, is not a well formulated model.

Given the importance of the matrices Ω, AΩ, and F, it is interesting that they are related to A and C by the following identity,

(26) $(I - A\Omega)(I - F)(I - \Omega) = I - (A + C)$,

which is mentioned by B&P (1995, fn. 34).[22] From this equation we see that that non-singularity of $I - (A + C)$ implies that the three matrices $(I - A\Omega)$, $(I - F)$, and $(I - \Omega)$ are all nonsingular. Accordingly, the WF requirement that $\det[I - \varepsilon(A + C)]$ is positive for all $0 \leq \varepsilon \leq 1$ also implies that the real eigenvalues of Ω, AΩ, and F are all less than 1.0 in value. Unfortunately, this does not imply that there is not some complex eigenvalue of F or AΩ with real part greater than 1.0, which is what is needed to satisfy the E&H conditions (12) for learnability.

Next we consider learnability for WF models under the second information assumption, for which the relevant conditions are that all eigenvalues of the matrices in (13a)-(13c) have real parts less than 1.0. First consider (13a), which implies that $I - A(I + \Omega)$ must have all eigenvalues with real parts that are positive. Using the definition of F, we can write

(27) $(I - A\Omega)(I - F) = (I - A\Omega)[I - (I - A\Omega)^{-1}A] = (I - A\Omega) - A = I - A(I + \Omega)$.

[22] It can be verified by writing out F in left side of (26), multiplying, cancelling, and inserting C for $\Omega - A\Omega^2$.

Now, our discussion above indicates that $I - A\Omega$ and $I - F$ will both have eigenvalues with all real parts positive under the WF assumption, so (27) might seem to suggest that this property would carry over to $I - A(I + \Omega)$. Unfortunately, however, this property seems not to be implied, although it would be if all the relevant eigenvalues were real.

Indeed, I have not been able to find any general results pertaining to conditions (13), but we can consider a couple of special cases that are of some interest. First, consider the case in which $C = 0$, so there are no predetermined variables in the solution, which implies that $\Omega = 0$. Then we have $F = (I - A\Omega)^{-1}A = A$ and thus (13a) becomes the same as (12a). Furthermore, (13b) is irrelevant with $\Omega = 0$ and (13c) becomes $(R' \otimes A)$, which is the same as in (12c). So in this case, the two information assumptions yield the same E-stability conditions and the WF restrictions imply that E-stability obtains in the case at hand. Second, suppose that $C \neq 0$, but that the exogenous variables are white noise, i.e., $R = 0$. Then (13c) becomes $(I \otimes A\Omega)$ and the result based on $(I - A\Omega)^{-1}$ shows that this condition will be satisfied if the WF conditions are relevant. But conditions pertaining to (13a) and (13b) are not necessarily satisfied. Of course, one can examine specific cases numerically.

6. General Issues

A number of possible objections to the foregoing argument need to be addressed. Probably the most prominent among researchers in the area would be the fact that our analysis has been concerned only with solutions of form (3), which excludes sunspot solutions of the "resonant frequency" type. It is my position, however, that the learning process pertaining to solutions of this type is much less plausible than for solutions of form (3). In particular, the solutions are not of the standard vector-autoregression (VAR) form. Therefore, an agent who experimented with many different specifications of VAR models, using the economy's generated time series data, would still not be led to such a solution. Indeed, it seems to me that arguments suggesting that that type of learning could exist in actual economies are utterly implausible. Of course, literally speaking, RE itself is implausible—as early critics emphasized. Nevertheless RE is rightly regarded by mainstream researchers as the appropriate assumption for economic analysis, especially policy analysis. That is the case because RE is fundamentally the assumption that agents optimize with respect to their expectational behavior, just as they do (according to neoclassical economic analysis) with respect to other basic economic activities such as selection of consumption bundles, selection of quantities produced and inputs utilized, etc.—for a necessary condition for optimization is that individuals eliminate any systematically erroneous component of their expectational behavior. Also, RE is doubly attractive (to researchers) from a policy perspective, for it assures that a researcher does not propose policy rules that rely upon policy behavior that is designed to exploit consistent patterns of suboptimal expectational behavior by individuals.

Another issue is the possible use of learning behavior not as a device for assessing the plausibility of rational expectations, but as a replacement for the latter. This type of approach is discussed by E&H (2001, Ch. 14) and has been prominent in the work of Orphanides and Williams (2005), among others. Use of decreasing-gain learning (E&H, 2001, pp. 338-341) provides a sensible alternative to the constant-gain learning implicit in the LS learning/E-stability literature. I do not believe, however, that this approach solves the "startup" problem, i.e., the issue of how the economy will behave in the first several periods following the adoption of a new policy rule or the occurrence of some other structural change. I doubt that economies move promptly to new RE equilibria following such a change, and I would doubt that they move promptly to a modelled learning path. In both cases, I share the opinion voiced by Lucas (1980), to the effect that after a structural change (including policy regime changes), reliable analysis should pertain to the economy's behavior after it has had time to settle into a new dynamic stochastic equilibrium.

7. Conclusion

Let us now conclude with a very brief review of the points developed above. First, the paper reviews a previous result to the effect that, under the information assumption that agents possess knowledge of current endogenous variables in the learning process, determinacy of a RE equilibrium is sufficient but not necessary for least-squares learnability of that equilibrium. Thus, since learnability is an attractive necessary condition for plausibility of any equilibrium, there may exist a single plausible RE solution even in cases of indeterminacy. The paper proposes and outlines a distinct criterion that models should possess, termed "well formulated," that rules out infinite discontinuities in the model's implied steady-state values of endogenous variables and in its impulse response functions. The paper then demonstrates that under the first information assumption, the "natural" decreasing-modulus solution is, in all well-formulated models, learnable—even in the absence of determinacy. Under the second information assumption, the situation is less favorable in the sense that learnability can be guaranteed only under special assumptions.

Appendix A

To demonstrate that a very wide variety of linear RE models can be written in form (1)(2), consider the formulation of King and Watson (1998) or Klein (2001), as exposited by McCallum (1998), as follows:

$$\text{(A-1)}\quad \begin{bmatrix} A_{11}^{*} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} E_t x_{t+1} \\ k_{t+1} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} x_t \\ k_t \end{bmatrix} + \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \left[v_t \right].$$

Here v_t is an AR(1) vector of exogenous variables (including shocks) with stable AR matrix R while x_t and k_t are $m_1 \times 1$ and $m_2 \times 1$ vectors of non-predetermined and predetermined endogenous variables, respectively. We assume without significant loss of generality that B_{11} is invertible[23] and that $G_2 = 0$.[24] Then we define $y_t = [x_t' \; k_t' \; x_{t-1}' \; k_{t-1}']'$ and write the system in form (1) with $u_t = v_t$ and the matrices given as follows:

$$\text{(A-2)}\quad A_{11} = \begin{bmatrix} B_{11}^{-1}A_{11}^{*} & 0 & 0 & -B_{11}^{-1}B_{12} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 0 & 0 & 0 & 0 \\ B_{21} & B_{22} & 0 & 0 \\ I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} -B_{11}^{-1}G_1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

This representation is important because it is well known that the system (A-1) permits, via use of auxiliary variables, any finite number of lags, expectational leads, and lags of expectational leads for the basic endogenous variables. Also, any higher-order AR process for the exogenous variables can be written in AR(1) form.[25] Thus we have shown that the Evans and Honkapohja (2001) formulation in their Section 10.3 is in fact rather general, although it does not pertain to asymmetric information models.

---

[23] For the system (A-1) to be cogent, each of the m_1 non-predetermined variables must appear in at least one of the m_1 equations of the first matrix row. Then the diagonal elements of B_{11} will all be non-zero and to avoid inconsistencies the rows of B_{11} must be linearly independent. This implies invertibility.

[24] If it is desired to include a direct effect of v_t on k_{t+1}, this can be accomplished by definition of an auxiliary variable (equal to v_{t-1}) in x_t (in which case v_t remains in the information set for period t). Also, auxiliary variables can be used to include expectations of future values of exogenous variables.

[25] Binder and Pesaran (1995) show that virtually any linear model can be put in form (1), but in doing so admit a more general specification than (2) for the process generating the exogenous variables.

References

Binder, M. and M. H. Pesaran, 1995. Multivariate Rational Expectations Models and Macro-econometric Modeling: A Review and Some New Results. Handbook of Applied Econometrics, edited by M. H. Pesaran and M. Wickens. Basil Blackwell Publishers.

____________________________, 1997. Multivariate Linear Rational Expectations Models: Characterization of the Nature of the Solutions and their Fully Recursive Computation. Econometric Theory 13, 877-888.

Benhabib, J., S. Schmitt-Grohe, and M. Uribe, 2001. The Perils of Taylor Rules. Journal of Economic Theory 96, 40-69.

Bernanke, B.S., and M. Woodford, 1997. Inflation Forecasts and Monetary Policy, Journal of Money, Credit, & Banking 24, 653-684.

Blanchard, O. and C. M. Kahn, 1980. The Solution of Linear Difference Models Under Rational Expectations. Econometrica 48, 1305-1311.

Bullard, J.B., 2006. The Learnability Criterion and Monetary Policy, Federal Reserve Bank of St. Louis Review 88/3, 203-217.

Bullard, J.B., and K. Mitra, 2002. Learning About Monetary Policy Rules, Journal of Monetary Economics 49, 1105-1129.

Carlstrom, C.T., and T.S. Fuerst, 2005. Investment and Interest Rate Policy. Journal of Economic Theory 123, 4-20.

Cho, S., and A. Moreno, 2007. The Forward Solution for Linear Rational Expectations Models, Working Paper, Yonsei University, Seoul, Korea. July.

Clarida, R., J. Gali, and M. Gertler, 1999. The Science of Monetary Policy: A New Keynesian Perspective," Journal of Economic Literature 37, 1661-1707.

Evans, G. W., 1986. Selection Criteria for Models with Non-Uniqueness. Journal of Monetary Economics 18, 147-157.

Evans, G. W. and S. Honkapohja, 1999. Learning Dynamics. In Handbook of Macroeconomics, J.B. Taylor and M. Woodford, eds. North-Holland.

______________ and __________, 2001. Learning and Expectations in Macroeconomics. Princeton Univ. Press.

______________ and ____________, 2003. Expectations and the Stability Problem for Optimal Monetary Policies. Review of Economic Studies 70, 807-824.

Gale, D., and H. Nikaido, 1965. The Jacobian Matrix and Global Univalence of Mappings, Mathematische Annalen 159, 81-93.

Golub, G. H, and C. F. Van Loan, 1996. Matrix Computations, 3rd ed. Johns Hopkins University Press.

Horn, R.A., and C.R. Johnson, 1985. Matrix Analysis. Cambridge Univ. Press.

________________________, 1991. Topics in Matrix Analysis. Cambridge Univ. Press.

King, R. G., 2000. The New IS-LM Model: Language, Logic, and Limits, Federal Reserve Bank of Richmond Economic Quarterly 86 (Summer), 45-103.

King, R. G., and M. W. Watson, 1998. The Solution of Singular Linear Difference Systems Under Rational Expectations. International Economic Review 39, 1015-1026.

Klein, P., 2000. Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model. Journal of Economic Dynamics and Control 24, 1405-1423.

Lucas, R.E., Jr., 1980. "Rules, discretion, and the role of the economic advisor," in S. Fischer, ed. Rational Expectations and Economic Policy. University of Chicago Press for NBER.

Magnus, J. R., and H. Neudecker, 1988. Matrix Differential Calculus with Applications In Statistics and Econometrics. John Wiley and Sons.

McCallum, B. T., 1983. On Non-Uniqueness in Rational Expectations Models: An Attempt at Perspective. Journal of Monetary Economics 11, 139-168.

_________________, 1998. Solutions to Linear Rational Expectations Models: A Compact Exposition. Economics Letters 61, 143-147.

_____________, 2003a. Multiple-Solution Indeterminacies in Monetary Policy Analysis. Journal of Monetary Economics 50, 1153-1175.

_____________, 2003b. The Unique Minimum State Variable RE Solution is E-Stable in all Well Formulated Linear Models. NBER Working Paper 9960, September.

_____________, 2004. On the Relationship Between Determinate and MSV Solutions in Linear RE Models. Economics Letters 84, 55-60.

_____________, 2007. E-stability vis-a-vis Determinacy Results for a Broad Class of Linear Rational Expectations Models. Journal of Economic Dynamics and Control 31, 1376-91.

Orphanides, A., and J.C. Williams, 2005. Imperfect Knowledge, Inflation Expectations, and Monetary Policy, in The Inflation-Targeting Debate, B.S. Bernanke and M. Woodford, eds. University of Chicago Press for NBER.

Rotemberg, J.J., and M. Woodford, 1999.

Sims, C. A., 1994. A Simple Model for the Study of the Determination of the Price Level and the Interaction of Monetary and Fiscal Policy, Economic Theory 4, 381-399.

Uhlig, H., 1999. A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily. In Computational Methods for the Study of Dynamic Economies, R. Marimon and A. Scott, eds. Oxford University Press.

Walsh, C., 2003. Monetary Theory and Policy, 2nd ed. MIT Press.

Woodford, M., 1994. Nonstandard Indicators for Monetary Policy: Can Their Usefulness Be Judged From Forecasting Regressions? In Monetary Policy, N.G. Mankiw, ed. University of Chicago Press for NBER.

__________, 2003. Interest and Prices, Princeton University Press.