Identifying asymmetric, multi-period Euler equations estimated by non-linear IV/GMM.

John Hunter*   Christos Ioannidis

Department of Economics, Brunel University, Uxbridge, Middlesex, UB8 3PH, England.

February 4, 2004

Abstract

In this article, the identification of instrumental variables and generalized method of moment (GMM) estimators with multi-period perceptions is discussed. The state space representation delivers a conventional first order condition that is solved for expectations when the Generalized Bézout Theorem holds. Here, it is shown that although weak instruments may be enough to identify the parameters of a linearized version of the Quasi-Reduced Form (Q-RF), their existence is not sufficient for the identification of the structural model. Necessary and sufficient conditions for local identification of the Quasi-Structural Form (Q-SF) derive from the product of the data moments and the Jacobian. Satisfaction of the moment condition alone is only necessary for local and global identification of the Q-SF parameters. While the conditions necessary and sufficient for local identification of the Q-SF parameters are only necessary to identify the expectational model that satisfies the regular solution. If the conditions required for the decomposition associated with the Generalized Bézout Theorem are not satisfied, then limited information estimates of the Q-SF are not consistent with the full solution. The Structural Form (SF) is not identified in the fundamental sense that the Q-SF parameters are not based on a forward looking expectational model. This suggests that expectations are derived from a forward looking model or survey data used to replace estimated expectations.

Keywords: Identification, Linear Operator Models, Order Condition, Rank Condition, Rational Expectations, Reduced Form, Structural Form

*Corresponding author, ecstjjh@brunel.ac.uk.

1 Introduction

This article considers identification of multivariate, multi-period rational expectations (RE) models discussed by Broze, Gourieroux and Szafarz (1995) and, Blinder and Pesaran (1997(BP97)). The article extends proposition 1 in BP97 via the Generalized Bézout Theorem to show, under more general conditions, when a regular solution to the RE problem exists. It extends the conditions for the identification of linear RE models in Pesaran (1987), to the multi-period expectations case. The conditions in Pesaran (1987) relate to linear models, which links identification to the question of valid instrument selection and validity. An issue paid particular attention in the literature on panel data models, where identification is generally accepted on satisfaction of tests of a number of over-identifying restrictions (Arellano, Hansen and Sentana (2000) and Phillips (2003)).

The question of identification becomes more complicated when RE models are derived from the solution to an objective problem. In addition to the existence of valid instruments there may be additional moment conditions that need to be satisfied, this occurs when the Generalised Method of Moments (GMM) estimator is used to correct for the presence of serial correlation (Hansen and Sargent (1982), West (1995) With panel data, a method devised by Bhargava and Sargan (1983), and extended by Arellano and Bond (1991) is commonly used to estimate dynamic euler equations. Again it is assumed that a test of over-identifying restrictions is sufficient to identify parameters of what are implicitly viewed as being forward looking expectations models. However, such tests of over-identifying restrictions are often accepted when there are weak instruments and when that is the case the tests are often unreliable (Stock and Wright (2000)). Satisfying a test of over-identifying restrictions from limited information estimates is not usually sufficient to identify expectations as their implicit estimates do not incorporate information about the forward solution to the RE problem.

In RE models further parametric restrictions apply to the mean and/or the variance equation, consequently higher order identification may rely on such non-linearity (Flôres and Szafarz (1994), Hunter (1992)). The impact of the discount rate is not discussed here, but apart from the case where there is a single equation the additional non-linearity introduced by the discount rate generally aids identification (Gregory et al (1993), Hunter (1989a,1992) and Sargan (1982)). Euler equations estimated using an errors in variables approach have a moving average error structure, which should either be accounted for directly or indirectly through the impact of the terminal condition on the time series properties or forecast performance of the model. The structure of the expectational model as well as the nature of the instrument set is important for identification. Identification may be lost when the estimator is not able to bind the parameters to the solution of the expectational problem and testing of the RE hypothesis in incomplete models may be compromised (Dufour (1997), Stock and Wright (2000)).

In addition to the type of parametric restriction considered above euler equations estimated using the errors in variables approach have a moving average error structure, which implies that the underlying models should either account for this or when such expectational errors are solved out then the terminal condition affect the time series properties and forecast performance of the model. Hence, the structure of the expectational model as well as the nature of the instrument set is important for identification. What happens to identification when the estimator is not able to bind the parameters to the solution of the expectational problem (i.e. the transversality condition discussed by Blinder and Pesaran (1995) or the difficulty in testing the RE hypothesis in incomplete models outlined in Dufour (1997)).

This article is structured as follows: in section 2 the solution to multivariate multi-period RE models is considered; in section 3 identification is viewed in terms of instrument validity and parametric restriction; in section 4 the impact on identification of the transversality condition and then conclusions are drawn.

2 Asymmetric Multi-period Expectations Models

The following model with future and past expectations has been considered by Broze and Szafarz (1991), Broze, Gourieroux and Szafarz (BGS95), and Blinder and Pesaran (BP95,BP97):

$$A_{00}y_t = \sum_{k=1}^{K} A_{ko}y_{t-k} + \sum_{k=1}^{K}\sum_{h=1}^{H} A_{kh}E(y_{t+h-k}|I_{t-k}) + u_t \tag{1}$$

where y_t and u_t are G vectors of decision and forcing variables. A_{kh}, $k = 0, 1, ...K$, $h = 0, 1, ...H$, are $G \times G$ dimensioned matrices of fixed coefficients and I_t represents a non-decreasing information set at time t, containing current and lagged values of y_t and u_t: $I_t = \{y_t,\ y_{t-1}, ...;\ u_t, u_{t-1}, ...\}$.Blinder and Pesaran (1995(BP95)) show that (1) has the following canonical form

$$\overline{x}_t = A\overline{x}_{t-1} + BE(\overline{x}_{t+1}|I_t) + w_t \tag{2}$$

where $\overline{x}_t = (x_t^{'}, x_{t-1}^{'}, ...x_{t-K+1}^{'})$, $x_t^{'} = (y_t^{'}, E(y_{t+1}^{'}|I_t), ...E(y_{t+H}^{'}|I_t))$, $A = -D_0^{-1}D_1$, $B = -D_0^{-1}D_{-1}$, $w_t = -D_0^{-1}\bar{\vartheta}_t$, $\bar{\vartheta}_t = (\vartheta_t^{'}, 0_{n\times1}^{'}, ...0_{n\times1}^{'})^{'}$, $\vartheta_t = (u_t^{'}, 0_{G\times1}^{'}, ...0_{G\times1}^{'})^{'}$; $\bar{\vartheta}_t$ and $\overline{x}_t$ are both $K(H+1)G \times 1$ matrices ($n = (H+1)G$), $0_{n\times1}^{'}$ is an $n \times 1$

vector of zeros, ϑ_t is of dimension $n \times 1$ and D_i for $i = -1, 0, 1$ are defined as:

$$
\begin{aligned}
D_{-1} &= \begin{bmatrix} \Gamma_{-1} & 0_n & \cdots & 0_n \\ 0_n & 0_n & \cdots & 0_n \\ & & \ddots & \\ 0_n & 0_n & \cdots & 0_n \end{bmatrix}, D_0 = \begin{bmatrix} \Gamma_0 & \Gamma_1 & \cdots & \Gamma_{K-1} \\ 0_n & I_n & \cdots & 0_n \\ & & \ddots & \\ 0_n & 0_n & \cdots & I_n \end{bmatrix} \\
D_1 &= \begin{bmatrix} 0_n & 0_n & \cdots & 0_n & \Gamma_K \\ I_n & 0_n & \cdots & 0_n & 0_n \\ & & \ddots & & \\ 0_n & 0_n & \cdots & I_n & 0_n \end{bmatrix} \text{ with } \Gamma_k, k = -1, 0, 1, ...K
\end{aligned}
$$

$$
\begin{aligned}
\Gamma_{-1} &= \begin{bmatrix} 0_G & 0_G & \cdots & 0_G & 0_G \\ -I_G & 0_G & \cdots & 0_G & 0_G \\ & & \ddots & & \\ 0_G & 0_G & \cdots & -I_G & 0_G \end{bmatrix}, \Gamma_0 = \begin{bmatrix} I_G & -A_{01} & \cdots & -A_{0H} \\ 0_G & I_G & \cdots & 0_G \\ & & \ddots & \\ 0_G & 0_G & \cdots & I_G \end{bmatrix} \\
\Gamma_i &= \begin{bmatrix} -A_{i0} & -A_{i1} & \cdots & -A_{iH} \\ 0_G & 0_G & \cdots & 0_G \\ & & \ddots & \\ 0_G & 0_G & \cdots & 0_G \end{bmatrix} \text{ for } i = 1, ...K
\end{aligned}
$$

It follows from proposition 1 in BP97 that when λ_i defines a solution to the scalar problem:

$$\phi(\lambda_i) = \det(B\lambda_i^2 - \lambda_i I + A) = 0$$

there are finitely many Jordan matrices J_i for $i = 1, 2, ...l$ that solve $\phi(C) = 0$, where C=SJ_iS^{-1} for some non-singular matrix S. Any matrix J_i that solves $\phi(C) = 0$ also satisfies:

$$BS_i^2 - SJ_i + AS = 0. \tag{3}$$

Vectorizing (3):

$$((J_i^2)' \otimes B - J_i' \otimes I_m + I_m \otimes A)vecS = 0.$$

When $rank\{(J_i^2)' \otimes B - J_i' \otimes I_m + I_m \otimes A\} = m - 1$, then S is a non-singular matrix, and $J_i = \Lambda$ is a solution to $\phi(C) = 0$, which also solves:

$$P(C) = BC^2 - C + A = 0;$$

where $C = S\Lambda S^{-1}$. If $P(C) = 0$ then it follows from the Generalized Bézout Theorem (Gantmacher (1960)) that the polynomial in z-transfom, $P(z) =$($Bz^2 - zI + A$) has a left hand divisor of the form $(zI - C)$. Mapping $P(z) = F(z)(zI - C)$ onto the time domain:

$$P(L^{-1}) = F(L^{-1})(L^{-1}I - C)$$

where $F(L^{-1}) = (I-BC)(FL^{-1}-I)$. $P(C) = 0$ when (i) $(I-BC)^{-1}$ exists and (ii) $rank\{(J_j^2)' \otimes B - J_j' \otimes I_m + I_m \otimes A\} = m-1$. $P(L^{-1}) = (BL^{-1}\overline{x}_t - \overline{x}_t + AL\overline{x}_t)$ decomposes into backward and forward components:

$$-P(L^{-1})L\overline{x}_t = -(I-BC)(FL^{-1}-I)(L^{-1}I-C)L\overline{x}_t = \tag{4}$$

$$= (I-BC)(I-FL^{-1})(I-CL)\overline{x}_t = w_t \tag{5}$$

and $(I-FL^{-1})$ inverts to produce the regular solution in BP97 without the requirement that $AB = BA$:[1]

$$\overline{x}_t - C\overline{x}_{t-1} = \sum_{s=0}^{\infty} F^s (I_n - BC)^{-1} D_0^{-1} E(\bar{\vartheta}_{t+s}|I_t) \tag{6}$$

$$C = \begin{bmatrix} C_0 & \cdots & C_{K-2} & C_{K-1} \\ I_n & \cdots & 0_n & 0_n \\ & \ddots & & \\ 0_n & \cdots & I_n & 0_n \end{bmatrix}, \; C_i = \begin{bmatrix} C_{i0} & \cdots & C_{iH} \\ 0_G & \cdots & 0_G \\ & \ddots & \\ 0_G & \cdots & 0_G \end{bmatrix}, \; i = 0, ...K-1.$$

If $u_t = \psi z_t + \varepsilon_t$ (BP95) then solving for expectations gives rise to a Vector Auto-Regressive (VAR) model (Pesaran (1987)):

$$\overline{x}_t - C\overline{x}_{t-1} = \Upsilon(L)\overline{z}_t + \overline{\varepsilon}_t \tag{7}$$

where $\Upsilon(L) = (\Upsilon_0 + \Upsilon_1 L + ...\Upsilon_{s-1}L^{s-1})$,

$$\Upsilon_{i(Kn \times Kn^*)} = \begin{bmatrix} \Phi_{0i} & 0_n & \cdots & 0_n \\ 0_n & 0_n & \cdots & 0_n \\ & & \ddots & \\ 0_n & 0_n & \cdots & 0_n \end{bmatrix}, \; \Phi_{0i} = \begin{bmatrix} \phi_{0i} & 0_G & \cdots & 0_G \\ 0_G & 0_G & \cdots & 0_G \\ & & \ddots & \\ 0_G & 0_G & \cdots & 0_G \end{bmatrix}, \; i = 1, ...s,$$

$\overline{z}_t' = [z_t^{*\prime}, 0, ..., 0]$ is a state vector containing exogenous process, $z_t^{*\prime} = [z_t', 0, ..., 0]$, $\overline{\varepsilon}_t' = [\varepsilon_t^{*\prime}, 0, ..., 0]$, is a state vector containing white noise residuals, $\varepsilon_t^{*\prime} =$

[1]The result presented replaces Proposition 2 in BP(97). It follows from the application of Frobenius Theorem or the appropriate matrices having Property P (Motzkin and Taussky (1954)), that $P(C) = BC^2 - C + A = 0$ has the following roots $\phi = \lambda_b\lambda^2 - \lambda + \lambda_a = 0$ when $AB = BA$. In (2), $AB = BA$ only occurs when the problem is first order and $B = \beta A$.

Alternatively, it is either required that $\lambda - \lambda_b\lambda^2 = \lambda_a$ and property P is satisfied or $P(C) = 0$ and $C - BC^2 = A$. It follows from the similarity of $C - BC^2$ and A that the roots (μ) of $C - BC^2$ equal λ_a.

Any matrix pair (C, B) with roots (λ, λ_b) is said to satisfy property P (see Motzkin and Taussky (1954)) when $F(C, B)$ has as its roots $f(\lambda, \lambda_b)$. More generally, Schneider (1955) shows that property P follows for every pair of matrices (A_i, A_j) in an ordered matrix polynomial when $(A_iA_j - A_jA_i)R_i = 0$.Hence, $C - BC^2$ has roots $\lambda - \lambda_b\lambda^2$ when $(BC - CB)R_i = 0$ for some matrix polynomial R_i. Property P will be satisfied when $BC = CB$ and $AB = BA$ Otherwise, when $\mu = \lambda_a$ and a Jordon form $C = SJ_iS^{-1}$ satisfies $P(C) = 0$, then $\mu = \lambda - \lambda_b\lambda^2$. Conditions (ii)-(iv) in BP(97) follow from these results without the requirement that A and B commute.

$[\varepsilon_t', 0, ..., 0]$ and $n^* = (H+1)G_z$. Now $\overline{x}_{t-1}$ contains perception variables $(E(y_{t-i}'|I_{t-j})$ for $j > i)$.[2].

Given measures for all elements in $\overline{x}_{t-1}$a solved form of the first order condition can be derived (Appendix A), the Quasi-Structural Form (Q-SF):

$$\begin{aligned} & (I + FC)\overline{x}_t - FE(\overline{x}_{t+1}|I_{t+1}) - C\overline{x}_{t-1} - (I_n - BC)^{-1}D_0^{-1}(\Psi\overline{z}_t + \overline{\varepsilon}_t) \\ = \; & F(\Psi R\bar{\zeta}_{t+1} + \overline{\varepsilon}_{t+1}') \end{aligned} \tag{8}$$

where $\bar{\vartheta}_t = \Psi\overline{z}_t + \overline{\varepsilon}_t$, $F = (I - BC)^{-1}B$, I is a Kn dimensional identity matrix and $R = \sum_{s=0}^{\infty} F^s(BC - I)^{-1}D_0^{-1}R_s$. The following linearization of (8) can be estimated consistently by Instrumental Variables (IV) when an optimal instrument set exists (Sargan (1983a)):

$$Q_o\overline{x}_t - Q_1\overline{x}_{t+1} - Q_2\overline{x}_{t-1} - \Pi\overline{z}_t = \varsigma_{t+1} \tag{9}$$

where ς_{t+1}= $(I_n - BC)^{-1}D_0^{-1}\overline{\varepsilon}_t - F\overline{\varepsilon}_{t+1}' - F\Psi R\bar{\zeta}_{t+1}$, an MA(1) error in state space form, $Qo = (I - BC)^{-1}$, $Q_1 = (I - BC)^{-1}B$, $Q_2 = C$ and $\Pi = (BC - I)^{-1}D_0^{-1}\Psi$. Multiplying through by $(I - BC)$ yields the Q-RF

$$\overline{x}_t - P_1\overline{x}_{t+1} - P_2\overline{x}_{t-1} - P_3\overline{z}_t = \varsigma_{t+1}^* \tag{10}$$

Where $\varsigma_{t+1}^* = \overline{\varepsilon}_t^+ - FD_0\overline{\varepsilon}_{t+1}^+ - F\Psi R\bar{\zeta}_{t+1}$, $\overline{\varepsilon}_t^+ = D_0^{-1}\overline{\varepsilon}_t$, $P_1 = B = D_0^{-1}D_{-1}$, $P_2 = C - BC^2$ and $P_3 = D_0^{-1}\Psi$.

3 Identification of Expectations Models

The conditions presented in Pesaran (1987) can be extended to identify linearized versions of (10) and (9). Identification of SF parameters follows from the existence of a well defined RF or Q-RF and identification of (7) from the existence of sufficient lagged information.[3]

If Q_o is non-singular, $\overline{x}_t = C\overline{x}_{t-1} + \Upsilon(L)\overline{z}_t + \overline{\varepsilon}_t$ then $E(\overline{x}_t|I_t)$ depends on $\overline{x}_{t-1}$ and $\overline{z}_{t-i}$, for $i = 0, 1, 2..s - 1$. As (9) is a state space analogue of the model in Pesaran (1987) and $b = [Q_0, Q_1, \Pi, C]$, is identified when UQ below has appropriate rank:[4]

$$UQ = \begin{bmatrix} b\Phi & O & O & ... & O \\ O & I_{Kn} & O & ... & O \\ O & O & I_{Kn^*} & ... & O \\ O & C^2 & \Xi_0 & ... & \Xi_{S-1} \end{bmatrix},$$

[2]Such variables are either directly measurable at time t, can be calculated from VAR or solved using the Quadratic Determinantal Equation Method discussed in BP(95, 97).

[3]Such conditions derive from Rothenberg(1971) and relate to the instrument matrix having sufficient rank.

[4]Implicit in the notion of a Q-RF is the idea that either $B_0 = I$ or $C = I$.

$$U=\left[\begin{array}{cccc} Q_0 & -Q_1 & -\Pi & -Q_2 \\ O & I_{Kn} & O & O \\ O & O & I_{Kn^*} & O \\ O & O & O & I_{Kn} \end{array}\right] \text{ and } Q=\left[\begin{array}{ccccc} \Phi_i & C & \Upsilon_0 & ... & \Upsilon_{S-1} \\ O & I_{Kn} & O & ... & O \\ O & O & I_{Kn^*} & ... & O \\ O & C^2 & \Xi_0 & ... & \Xi_{S-1} \end{array}\right];.$$

where $Q_0C - Q_1 - Q_2C^2 = 0$, $Q_0\Upsilon_0 - \Pi - Q_2\Xi_0 = 0$, $Q_0\Upsilon_i - Q_2\Xi_i = 0$ *for* $i = 2, 3, ...s-1$ and without further homogenous linear restrictions $\Phi = [\Phi_i : O]$. This simplifies as $rank(UQ) = Kn + Kn^* + rank(UQ^*)$, where:

$$UQ^* = \left[\begin{array}{cccc} b\Phi & O & ... & O \\ O & \Xi_1 & ... & \Xi_{S-1} \end{array}\right].$$

If UQ is dimensioned $2Kn \times (r + Kn^*(s-1))$ with r conventional restrictions on B and $rank(\Xi_i = C\Upsilon_i) \leq \min(Kn, Kn^*)$, then a unique solution to b, to a normalization implies $rank(UQ^*) = 2Kn - 1$. A necessary condition for identification is:

$$2K(1+H)G - 1 < r + K(1+H)K_z(s-1).$$

Exact identification requires r restrictions and enough pre-determined information via lags on the VAR (s) and/or exogenous variables (K_z).[5] Given enough a priori restrictions, rank($\left[\begin{array}{ccc} \Xi_1 & ... & \Xi_{S-1} \end{array}\right]$) $= Kn$ is all that is required. But for a process in x truncated then C is rank deficient and irrespective of s identification is lost. Technically, identification will be accepted when there are enough non-zero lags on the exogenous variables in the RF, but empirically the RF is well estimated and the parameters distinguished from zero when these lags are significant. It may be possible that rank($\left[\begin{array}{ccc} \Xi_1 & ... & \Xi_{S-1} \end{array}\right]$) $= Kn$ when a long enough dynamic process is estimated, but identification though technically accepted may be invalidated by weak instruments. The local conditions also globally identify the linearized Q-SF parameters and when $Q_0 = I$ the Q-RF.

Sargan (1983) shows that when it is possible to solve the underlying parameters from well defined estimates of a RF, the conditions due to Hunter (1992) are sufficient for global identification. Given that rank($\left[\begin{array}{ccc} \Xi_1 & ... & \Xi_{S-1} \end{array}\right]$) $= Kn$, a necessary condition for identification of the RF and Q-RF parameters, then the Q-SF parameters can be solved in the following way using their definitions (10) and the imposition of some additional linear restrictions. Therefore, D_0 may be identified from homogenous linear restrictions on the parameters of the SF (Sargan(1988)) and cross equation restrictions that derive from $D_{-1} - D_0P_1 = 0$,[6] C follows from the solution to $P_2 - C + P_1C^2 = 0$ and $\Psi = D_0P_3$. When

[5] When $s = 2$ and $G = G_z$ then this; order condition has a more usual form:

$$r > K(1+H)G - 1.$$

[6] Be aware that not all these restrictions should be effective as D_{-1} is a fixed matrix for the case considered by BP97 and consequently P_1 is non-invertible.

$P_2 = A = C - P_1 C^2$ then (2) and (10) are equivalent but full equivalence of these two forms only occurs when the transversality conditions hold and $P_2 - C + BC^2 = 0$ defines a valid solution to the forward looking expectational problem (BP95 and BP97). If $rank(P_3) = Kn$, then D_0^{-1} exists and the parameter matrices D_{-1}, D_0 and D_1 along with their associated sub-matrices are also identified. The existence of such a solution is sufficient for global identification. of the Q-RF and Q-SF parameters.

Flôres and Szafarz (1994) derive conditions, which depend on the rank of the Jacobian Matrix considered without expectations by Rothenberg (1971). When compared with Flôres and Szafarz, the following condition developed from Sargan (1983a) depends on the product of the Jacobian and the moment matrix of the data for models that include perception variables ($K > 1$).The Sargan approach to local identification is considered next.

When (10) is stacked across the sample the system can be written:

$$V(\theta)X^{*\prime} = E' \tag{11}$$

where $V(\theta) = [D_0 \vdots\text{-}\Psi : -D_{-1} : -D_0 C + D_{-1}C^2\ ]$, $X^* = [X\ Z\ X_{+1}\ X_{-1}\ ]$, X and is an $N \times Kn$, matrix of observations on y_t, Z is an $N \times Kn^*$ stacked matrix of observations on $\overline{z}_t$, $E'$is an $N \times Kn$ stacked matrix of observations on ς^*_{t+1}, N is the number of time observations, matrices subscripted by $_{+1}$ relate to observations for the period $t+1$ and $_{-1}$ to period $t-1$.

Define Z^*as $\left[\hat{X}\ Z\ \hat{X}_{+1}\ X_{-1}\ \right]$ where $\hat{X}$ and $\hat{X}_{+1}$ are matrices of predictions or forecasts of X and X_{+1}. Post multiplying (11) by the instrument matrix (Z^*):

$$V(\theta)X^{*\prime}Z^* = E'Z^* \tag{12}$$

Satisfaction of the following criterion, the probability limit of (12) is required to estimate the parameters of (9) consistently (Sargan (1983a))[7]:

$$V(\theta)\frac{plim(X^{*\prime}Z^*)}{N} = 0 \tag{13}$$

The criterion is made operational by replacing $\hat{X}$ and $\hat{X}_{+1}$by their instruments $Z^+ =$$[X_{-1}, Z, Z_{-1}, Z_{-2}, ...Z_{-s}]$ so the moment matrix of the data can be written:

$$p\lim(\frac{Z^{+\prime}X^*}{N}) = M = plim\left[\frac{Z^{+\prime}X}{N} : \frac{Z^{+\prime}Z}{N} : \frac{Z^{+\prime}X_{+1}}{N} : \frac{Z^{+\prime}X_{-1}}{N}\right].$$

Vectorizing ($V(\theta)M$):

$$vec(V(\theta)M') = (M \otimes I_{Kn})vec(V(\theta)) = 0. \tag{14}$$

Following Sargan(1983a) necessary and sufficient condition for the local identification of dynamic autoregressive models estimated by IV can be derived from

[7]As is explained by Arellano (2002) GMM has generalized this the type of NLIV criterion to consider broader forms of non-linearity, but at the penalty of losing Sargan's original motivation to handle measurement error, the expectational case considered here. Also dynamic information from the mean and variance equations is often discarded when GMM is applied.

the first derivative of (12). Differentiating (14) with respect to the parameter vector:

$$\frac{dvec(V(\theta)M')}{d\theta} = (M \otimes I_{Kn})\frac{dvec(V(\theta))}{d\theta}, \tag{15}$$

gives rise to the rank condition:

$$rank\{(M \otimes I_{Kn})\frac{dvec(V(\theta))}{d\theta}\} < m = 2Kn + Kn^*. \tag{16}$$

If $\theta = [vec(D_0)' : vec(\Psi)' : vec(C)']$, then:

$$\frac{dvec(V(\theta))}{d\theta} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \\ -(C' \otimes I) & 0 & -(I \otimes D_0) + ((I \otimes D_0C) + (C' \otimes D_0) \end{bmatrix} \tag{17}$$

and conditions that are essentially the same as those due to Flôres and Szafarz (1994) follow from the rank of Jacobian of the transformation (17). But Sargan (1983a) shows for a class of dynamic model that the condition on the rank of the Jacobian is only sufficient for identification.[8] Pre-multiplying $\frac{dvec(V(\theta))}{d\theta}$ by $(M \otimes I_{Kn})$:

$$\begin{aligned}
(M \otimes I_{Kn})\frac{dvec(V(\theta))}{d\theta} &= [(M_0 \otimes I_{Kn}) - (M_3 \otimes I_{Kn})(C' \otimes I_{Kn}) : \\
-(M_1 \otimes I_{Kn}) &: (M_3 \otimes I_{Kn})(-(I \otimes D_0) + ((I \otimes D_0C) + (C' \otimes D_0))] \\
&= [(M_0 \otimes I_{Kn}) - (M_3C' \otimes I_{Kn}) : -(M_1 \otimes I_{Kn}) : \\
&\quad (M_3 \otimes I_{Kn})((C' - I) \otimes D_0 + (I \otimes D_0C))] \\
&= [((M_0 - M_3C') \otimes I) : -(M_1 \otimes I_{Kn}) : \\
&\quad (M_3(C' - I) \otimes D_0) + (M_3 \otimes D_0C))]
\end{aligned} \tag{18}$$

where $M = [M_0 : M_1 : M_2 : M_3]$. A necessary condition for identification is that rank$(M) < (2Kn + Kn^*)$. Given, the definition of the model in section 2, D_1 is a fixed matrix and is identified a priori. The remaining parameters are not identified when the above rank condition fails or:

$$\begin{aligned}
rank(M_0 - M_3C') &< Kn && \text{(I)} \\
rank(M_1) &< Kn^* && \text{(II)} \\
rank(M_3(C' - I) \otimes D_0) + (M_3 \otimes D_0C)) &< Kn && \text{(III)}
\end{aligned}$$

It follows that D_0 is not-identified when $M_0 = M_3C'$ or certain rows and columns in M_0 and M_3C' are subject to some cancellation. A special case of this

[8] The results presented here do not consider conditions on the variance-covariance matrix, but the authors of this article believe that a key distinction between this paper and that of Flores and Safarz follows form the interaction of the moment conditions with the Jacobian. The conditions on the variance-covariance matrix induce further non-linearities, which only make the identification story more complicated.

occurs when there are unit roots in the process driving the exogenous variables. Failure of (II) occurs when M_1 is rank deficient and Ψ is not identified; there are insufficient instruments or z is cointegrating exogenous for a subset of the deep parameters (Hunter (1989)). Cointegrating exogeneity implies that the long-run processes forcing the y do not force z. Failure of (III) occurs when either D_0 or M_3 are rank deficient. Otherwise, $C = 0$ when $M_3(C'-I)\otimes D_0) = -(M_3 \otimes D_0 C)$ or C is not fully identified when there are dependencies between $M_3(C'-I)\otimes D_0)$ and $(M_3 \otimes D_0 C)$. There may also be rank dependencies across the columns of (18) associated with cointegration between the x variables alone This will lead to further restrictions, which imply that not all the parameters in Ψ are identified.[9]

Non-identification occurs when conditions on the product of the Jacobian matrix and the Moment matrix of the data are rejected, but as Phillips (2003) suggests, the literature often considers it enough to satisfy the moment conditions or an empirical test of the type of rank condition found in Pesaran (1987). However, Dufour (1997) and Stock and Wright (2000) question the performance of such tests of moments. Dufour is critical of the power of the underlying tests based on what are limited information estimators, while Stock and Wright consider the question of Weak Instruments. Furthermore, Stock Wright and Yogo (2002) explain that the IV and GMM sample distributions are non-normal under weak instruments and as a result point estimates and hypothesis tests are unreliable. This should be seen in the context of a broader observation made by Sargan (1983) that a loss of identification may be associated with estimators that have fat-tailed distributions. The poor size of tests of instrument validity may be associated with any loss of identification due to failure of (I)-(III). Both weak instruments and/or an empirical incapacity to detect the supposed RE structure.

The latter observation leads to the final proposition of this article, that identification thus far described relates to the parametric identification of the Q-RF and Q-SF parameters. alone and not to the forward looking representation of the model. That is estimates of parameters based on (2) (6) and (7) need to be observationally equivalent to those derived from (8), (9) and (10).for full identification of the structure As the (Q-RF) and (Q-SF) do not impose the solution, then equivalence with (2) only occurs when the instruments well approximate the true expectations, but the true expectations are unobserved. Although the observation of weak instruments may lead to a valid rejection of the expectations model, as Stock and Wright (2000) suggest, poor instrument performance may be the cause of the test to falsely reject either the null or the alternative using conventional criterion.

[9] As compared with Flôres and Szafarz, the rank condition considered here is affected by the existence of unit roots and cointegration. As this is associated with rank deficiencies amongst the moment matrices of the exogenous and endogenous variables, then identification of some parameters is lost.

4 The impact on Identification of the Transversality condition

The previous section considers identification of SF and RF form parameters in relation to models that do not fully exploit the solution to the RE model. Therefore the local conditions developed in the previous section are necessary and sufficient to identify the parameters of both (10) and (9), but they are only necessary for the models (10) and (2) to be isomorphic.

Two models are isomorphic when the two sets of parameters are observationally equivalent. However, the Q-RF (10) is a model with an MA error structure that does not use estimates of expectations that bind the parameters to the solution, while (2) contains the true expectations. These two RE model will only be comparable when the separation associated with (i) and (ii) holds or empirically:

$$\begin{aligned} (ia)\ rank(I - P_1 C) &= m \\ (iia)\ rank\{(J_j^2)' \otimes P_1 - J_j' \otimes I_m + I_m \otimes P_2\} &= m - 1 \end{aligned}$$

However, (ii) cannot be tested from estimates of (10) alone, because C will then be computed by stacking the stable roots of the z-transform of the polynomial $P^{IV}(z) = P_1 z^2 - z + P_2 = 0$, implying $P^{IV}(C) = P_1 C^2 - C + P_2 = 0$ or (iia) holds by construction. Otherwise, BP95 provide an approach that might be used to check the validity of the RE form and calculate an RE solution. Set $\bar{X}_t = \bar{x}_t - C x_{t-1}$ and rewrite (10):

$$(I - BC)\overline{X}_t - P_1 \overline{X}_{t+1} - (P_1 C^2 - C + P_2)\overline{x}_{t-1} = P_3 \overline{z}_t + \varsigma_{t+1}^*$$

When (i) or (ia) holds, then:

$$\begin{aligned} \overline{X}_t &= (I - BC)^{-1} P_1 \overline{X}_{t+1} + (I - BC)^{-1}(P_1 C^2 - C + P_2)\overline{x}_{t-1} + \\ &\quad (I - BC)^{-1} P_3 \overline{z}_t + (I - BC)^{-1} \varsigma_{t+1}^* \end{aligned} \tag{19}$$

If C has a canonical form $S \Lambda S^{-1}$ and $P(C) = P_1 C^2 - C + P_2 = 0$, then (19) can be used to compute both empirically and analytically a forward looking solution. When $P(C) = 0$, then (19) gives rise to:

$$\overline{X}_t = (I - BC)^{-1} P_1 \overline{X}_{t+1} + (I - BC)^{-1} P_3 \overline{z}_t + (I - BC)^{-1} \varsigma_{t+1}^* \tag{20}$$

Taking expectations at period $t + i$, where $E(\varsigma_{t+i}^* | I_t) = 0$ for $i = 1, ...N$ and substituting recursively for $E(\overline{X}_{t+i} | I_t)$ using $(I - BC)^{-1} P_1 E(\overline{X}_{t+1+i} | I_t) + (I - BC)^{-1} P_3 E(\overline{z}_{t+i} | I_t)$ yields the finite horizon solution

$$\begin{aligned} \overline{X}_t &= \overline{x}_t - C\overline{x}_{t-1} = (I - BC)^{-1} P_1 E(\overline{X}_{t+1+N} | I_t) + \\ &\quad \sum_{s=0}^{N} F^s (I_n - BC)^{-1} P_3 E(\bar{z}_{t+s} | I_t) \end{aligned}$$

where $(I - BC)^{-1}P_1E(\overline{X}_{t+1+N}|I_t) = \sum_{s=N}^{\infty} F^s(I_n - BC)^{-1}P_3E(\bar{z}_{t+s}|I_t)$. It follows that the finite horizon solution is the same as (6) when the truncation or terminal condition $\sum_{s=N}^{\infty} F^s(I_n - BC)^{-1}P_3E(\bar{z}_{t+s}|I_t) = 0$. From Proposition 1 in BP97, $\sum_{s=N}^{\infty} F^s(I_n - BC)^{-1}P_3E(\bar{z}_{t+s}|I_t) = 0$ when condition (ii) is accepted. While the isomorphism required for (2) and (10) to be observationally equivalent also depends on satisfaction of condition (ii), which implies that the transversality condition has been satisfied. A sufficient condition for proposition (ii) to hold is $\overline{X}_t \sim I(0)$. However, (ii) will fail when either $\overline{X}_t \nsim I(0)$ or C is ill-defined and the latter occurs, when C does not have an appropriate canonical form. For the solution to be partitioned into separate forward looking and backward looking solutions (i) and (ii) above are required to hold. If the parameters estimated from (9) are consistent with the existence of a forward looking solution, S^{-1} exists $P(C) = 0$ and $A = C - BC^2$.

Alternatively, Sargan (1983) states that consistency is generally sufficient for higher order identification. Thus (10) and (2) are isomorphic, when the matrix pair $[P_1, P_2]$ lies in a neighbourhood of some consistent estimates of $[\hat{B}, \hat{A}]$. To test this proposition requires $E(\overline{x}_{t+i}|I_t)$ in (2) to be replaced by a measure that is consistent with the transversality condition being satisfied.

In the limit an estimator, which satisfies (14) will converge to zero and the associated model estimates will satisfy the Euler condition as $p\lim \frac{E'Z^+}{N}$ will not be statistically different from zero. Hence, any estimator satisfying a test of over-identifying restrictions will tend to parameter values that satisfy the transversality condition. However, the notion of convergence that is appropriate for the test of over-identification to be consistent with satisfaction of the transversality condition implies that the forecast horizon $T - N$ must collapses to zero as N becomes large. It was shown in section 2 that the separation into a forward and backward looking component occurs when $P(C) = 0$. But such a test requires one to check the consistency of any limited information estimates of parameters by comparison with equivalent parameters from a model that contains estimates of expectations that embody the proposition that agents are forward looking.

5 Conclusions

Estimation of the structural parameters of optimizing models has become enormously popular. It is legitimate to question the validity of the two approaches to the problem. The full information procedure is computationally burdensome, requires additional models for the exogenous processes and as with all likelihood based procedures is often viewed as being non-robust. However, it permits the restrictions associated with both the forward looking solution and RE to be imposed. Limited information procedures (IV/GMM) as typically applied, yield Q-SF, but they do not ordinarily permit testing of all the RE restrictions or the imposition of the transversality condition.

In this article, it is shown that the parameters of the structural model can be identified under the assumption that $P(C) = 0$. This proposition appears not to be testable in the context of models that estimate the first order condition via GMM or IV. When compared with testing for over-identifying restrictions (Arellano et al (2000), Hansen and Sargent (1982) and Sargan (1964)) and checking rank conditions (Rothenberg (1971) and Flôres and Szafarz (1994)), the approach due to Sargan (1983a) considers the impact of both the moments and the Jacobian Matrix. Pesaran (1987) derived conditions for identification, but they relate to linear models and in that case they boil down to the existence of sufficient instruments, which in practice are often found to be weak (Stock and Wright (2000)). Conventional tests of over-identifying restrictions or moment conditions may obviate some of the problems associated with weak instruments, but even these tests may be swamped by selection of enough lagged information. And according to Stock, Wright and Yogo (2002), the impact of weak instruments is worse when GMM is used.

Flôres and Szafarz (1994) have addressed the question of instrument interaction in the sense of the information set available, but their conditions pre-dominantly emphasize the role of the non-linearity in identification via the Jacobian matrix. In the Sargan approach the effect of the moments and the Jacobian matrix interact, which yields conditions reliant on non-linear structure and data dependence. Both the moment conditions and the Jacobian conditions are necessary, but not sufficient for identification. Especially when cointegration is important then the distinction between the impact of the Jacobian and the moment conditions is apparent The moment conditions will fail with cointegration amongst endogenous, exogenous and between endogenous and exogenous variables. This is a question of long-run structure as compared with the poor informational properties of the data.

A further issue not addressed in the conventional literature, which further distinguishes this work, relates to the existence of the forward looking representations, this follows from the Generalized Bézout Theorem. Partly, this relates to (i) which is necessary for the existence of RF and Q-RF equations and all the results considered in section 3, but it also relates to cointegration amongst the endogenous variable processes. Furthermore it links to the key condition for separation of the solution into backward and forward components that requires a C matrix satisfying the condition $P(C) = 0$.

In line with the concluding remarks made by Stock and Wright (2000), this article has extended the results presented in Sargan (1983a) for the Identification of non-linear IV estimators to the Euler equation case and augmented these conditions by the well known proposition that alternative estimates either of expectations or of model parameters must be consistent for identification.

6 Appendix A

Applying the forward Koych transformation $(I - FL^{-1})$ to (6) assuming that there exist measures for all lagged expectations in $\overline{x}_{t-1}$:

$$\begin{aligned}(I - FL^{-1})(\overline{x}_t - C\overline{x}_{t-1}) &= \sum_{s=0}^{\infty} F^s (I_n - BC)^{-1} D_0^{-1} E(\bar{\vartheta}_{t+s}|I_t) - \\ & F\sum_{s=0}^{\infty} F^s (I_n - BC)^{-1} D_0^{-1} E(\bar{\vartheta}_{t+s+1}|I_{t+1}) \end{aligned} \quad (21)$$

where $L^{-i}E(\overline{x}_{t+s}|I_t) = E(\overline{x}_{t+s+1}|I_{t+1})$. Re-ordering the indices on the summation signs and re-ordering terms gives rise to:

$$\begin{aligned}(I - FL^{-1})(\overline{x}_t - C\overline{x}_{t-1}) &= (I_n - BC)^{-1} D_0^{-1} \bar{\vartheta}_t + \\ & F\sum_{s=0}^{\infty} F^s (I_n - BC)^{-1} D_0^{-1} (E(\bar{\vartheta}_{t+s}|I_t) - E(\bar{\vartheta}_{t+s+1}|I_{t+1}))\end{aligned}$$

When the exogenous processes have a Wold representation, $z_t = \sum_{r=1}^{\infty} \theta_i L^i v_t$, where θ_i is a $G \times G$ matrix of fixed parameters, v_t is a G vector of white noise residuals and L is the lag operator then:

$$\begin{aligned}(E(\bar{\vartheta}_{t+s}|I_t) - E(\bar{\vartheta}_{t+s}|I_{t+1})) &= (E(\Psi\overline{z}_{t+s} + \overline{\varepsilon}_{t+s}|I_t) - E(\Psi\overline{z}_{t+s} + \overline{\varepsilon}_{t+s}|I_{t+1})) \\ &= (\Psi E(\overline{z}_{t+s}|I_t) + E(\overline{\varepsilon}_{t+s}|I_t) - \Psi E(\overline{z}_{t+s}|I_{t+1}) \\ &\quad - E(\overline{\varepsilon}_{t+s}|I_{t+1})) \\ &= (\Psi E(\overline{z}_{t+s}|I_t) - \Psi E(\overline{z}_{t+s}|I_{t+1})) + (E(\overline{\varepsilon}_{t+s}|I_t) \\ &\quad - E(\overline{\varepsilon}_{t+s}|I_{t+1})) \\ &= -\Psi R_s \bar{\zeta}_{t+1} - \overline{\varepsilon}_{t+1}\end{aligned}$$

where, $E(\overline{\varepsilon}'_{t+s}|I_{t+i}) = 0$ for $s > i$, $\bar{\zeta}_t = (\zeta_t^{'}, 0_{n\times 1}^{'}, ...0_{n\times 1}^{'})^{'}$, $\zeta_t = (v_t^{'}, 0_{G\times 1}^{'}, ...0_{G\times 1}^{'})^{'}$; $\bar{\zeta}_t$ is an $Kn \times 1$ matrix $0_{n\times 1}^{'}$ is an $n \times 1$ vector of zeros and ζ_t is of dimension $n \times 1$, and R_s is a square matrix defined as

$$R_s = \begin{bmatrix} \Theta_s & 0_n & \cdots & 0_n \\ 0_n & 0_n & \cdots & 0_n \\ & & \ddots & \\ 0_n & 0_n & \cdots & 0_n \end{bmatrix} \text{ and } \Theta_s = \begin{bmatrix} \theta_s & 0_G & \cdots & 0_G \\ 0_G & 0_G & \cdots & 0_G \\ & & \ddots & \\ 0_G & 0_G & \cdots & 0_G \end{bmatrix} \text{ for } i = 1, ...\infty$$

A forward solution follows by substituting out for updated expectations:

$$(I - FL^{-1})(\overline{x}_t - C\overline{x}_{t-1}) = (I_n - BC)^{-1}D_0^{-1}\bar{\vartheta}_t - F\sum_{s=0}^{\infty} F^s(I_n - BC)^{-1}D_0^{-1}(\Psi R_s\bar{\zeta}_{t+1} + \overline{\varepsilon}'_{t+1})$$

Multiplying through by the forward term on the Koych lead and replacing $\bar{\vartheta}_t$ by $\Psi\overline{z}_t + \overline{\varepsilon}_t$,reveals a Quasi-Structural Form (Q-SF):

$$(I + FC)\overline{x}_t - FE(\overline{x}_{t+1}|I_{t+1}) - C\overline{x}_{t-1} - (I_n - BC)^{-1}D_0^{-1}(\Psi\overline{z}_t + \overline{\varepsilon}_t) = F(\Psi R\bar{\zeta}_{t+1} + \overline{\varepsilon}'_{t+1})$$

where $F = (I_n - BC)^{-1}B$ and $R = \sum_{s=0}^{\infty} F^s(I_n - BC)^{-1}D_0^{-1}R_s$.

7 References

Arellano, M, (2002) Sargan's Instrumental Variables estimation and the Generalized Method of Moments, *Journal of Business and Economic Statistics*, 20, 450-459.

Arellano, M. and Bond, S.R., (1991), Some Tests of Specification for Panel Data: Monte Carlo Evidence and an application to Employment equations, *Review of Economic Studies*, 58, 277-297.

Arellano, M., Hansen, L.P., and Sentana, E. (2000) *Underidentification*? paper presented at EC^2 Conference, Trinity College, Dublin December 2000.

Bhargava, A and Sargan, J.D (1983) Estimation of Dynamic Random Effects Models from Panel Data Covering Short Time periods, *Econometrica*, 51, 1635-59

Blinder, M. and Pesaran, M.H. (1995) Multivariate Rational Expectations Models and Macroeconomic Modelling: A Review and Some New Results. *Handbook of Applied Econometrics*, Basil Blackwell, Oxford, 139-187.

Blinder, M. and Pesaran, M.H. (1997) Multivariate Rational Expectations Models and Macroeconomic Modelling: *Econometrics Theory*

Broze, L and Szafarz, A,(1991), The Econometric Analysis of Non-Uniqueness in Rational Expectations Models, Amsterdam, North Holland.

Broze, L, Gourieroux, C, and Szafarz, A, (1997), Solutions to Multivariate Rational Expectations Models, *Econometrics Theory*, 11, 229-257.

Dufour, J-M., (1997), Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models, *Econometrica*, 65 (6), 1365-87.

Flôres and Szafarz.(1994), Identification of Rational Expectations Models, Revista de Economia, 14, 71-87.

Gregory, A.W., Pagan, A.R., and Smith, G.W., (1993) Estimating Linear Quadratic Models with Integrated Processes. In Phillips, P.C.B. (ed), *Models, methods, and applications of econometrics : Essays in honor of A. R. Bergstrom*, Cambridge, Mass. and Oxford: Blackwell. 220-39.

Hansen, L.P. and Sargent, T.J., (1982) Instrumental variables procedures for estimating linear rational expectations models. *Journal of Monetary Economics*, 9, 263-96.

Gantmacher, F.R. (1960) Matrix Theory Vol. I. New York: Chelsea Publishing Company.

Hunter, J. (1989). The effect of Cointegration on solutions to Rational Expectations models, paper presented at ESEM89, September 1989 in Munich.

Hunter, J.(1989a) *Dynamic Modelling of Expectations : with Particular Reference to the Labour Market*. unpublished Ph.D. manuscript, London School of Economics.

Hunter, J.(1992) *Global Identification of Linear Rational Expectations Models*. CERF discussion paper 92-03, presented at ESEM93, August 1993 in Uppsala.

Motzkin, T.S. and Taussky, O.(1954) Pairs of Matrices with Property L, *Transactions of the American Mathematical Society*, 73, 108-114.

Pesaran, M. H.(1987) *The Limits to Rational Expectations*. Basil Blackwell, Oxford.

Phillips, P.C.B. (2003), Vision and Influence in Econometrics : John Denis Sargan, Econometric Theory, 19, 495-511

Rothenberg, T.J. (1971) Identification in Parametric Models. *Econometrica*, 39, 577-91.

Sargan, J.D. (1964) Wages and prices in the UK: A study in econometric methodology. In P.E. Hart, G Mills and J.K. Whitaker (eds), *Econometric analysis for National Economic Planning*, Butterworths.

Sargan, J.D. (1982) *Some problems with the Muellbauer method of specifying and estimating a rational expectations model*. Florida University Discussion Paper 68.

Sargan, J.D., (1983), Identification and lack of Identification, *Econometrica*, 51, 1605-33

Sargan, J.D. (1983a) Identification in Models with Autoregressive Errors. In S.Karlin, T. Amemiya and L.A. Goodman (eds), *Studies in Econometrics, Time Series and Multivariate Statistics in honour of T.W. Anderson*, Academic Press, New York, 169-205.

Sargan, J.D. (1988), *Lectures on Advanced Econometric Theory*, Basil Blackwell, Oxford.

Schneider, H. (1955), A Pair of Matrices with Property P, *The American Mathematical Monthly*, 62, 247-249

Stock, J.H. and Wright, J. (2000) GMM with Weak Identification. *Econometrica*, 68, 1055-1096.

Stock, J.H, Wright, J and M.Yogo (2002) A Survey of Weak Instruments, Weak Identification in GMM. *Journal of Business and Economic Statistics*, 20, 518-529.

West, K.D. (1995) Inventory Models: The Estimation of Euler Equations. *Handbook of Applied Econometrics*, Blackwell, 188-220.