Identifying Euler equations estimated by non-linear IV/GMM.

John Hunter Christos Ioannidis*

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

September 13, 2000

Abstract

In this article, the identi¯cation of instrumental variables and generalised method of moment (GMM) estimators is discussed. It is common that representations of such models are derived from the solution to linear quadratic optimisation problems. Here, it is shown that even though the rank condition on the Jacobian and the instrument set is valid, that the transversality condition may not be satis¯ed by the estimated model. Further, acceptance of the transversality condition does occur when identi¯cation fails or the forward model vanishes. As a result the parameters of such models irrespective of any correction for serial correlation may not be identi¯ed in a fundamental sense. This suggests that either forward looking models should be estimated directly or more complex non-linear restrictions should be imposed.

Keywords: Identi¯cation, Linear Operator Models, Order Condition, Rank Condition, Rational Expectations, Reduced Form, Structural Form

1 Introduction

There has been a lot of recent interest in the question of identification of parameters in structural or quasi-structural relationships (Arellano and Hansen (1999), Forchini and Hillier (1999)). Particular attention has focused on the GMM estimator either in the panel or the time series context as this has become the dominant approach to the estimation of Euler equations from forward looking models (West (1995), Arellano and Bond (1991), Bond and Meghir,(1994) Stock and Wright (2000)). By and large the use of such estimators is based on the "assumption" that the structural parameters are identified and that only consistent estimators are required.

A number of issues arise in this context which will cast doubt on the wisdom of such an approach. In particular, the articles by Blinder and Pesaran (1995) and Dufour (1997) have criticised this method. Dufour's principle objection relates to the use of the Wald Statistic applied to test rational expectations restrictions based on unrestricted dynamic models. An other practice,that has crept into Financial and Economic modelling is the use of differenced data and the choice of somewhat arbitrary instrument sets. In time series econometrics, the differenced VAR approach suggested by Sims (1980) has been superceded by the advent of cointegration or the partial over-differencing of sets of persistent series. With the exception of Blundell and Bond (1998) who accept the informational limitations imposed by using the differenced series, it is still common practice in Panel studies to difference the data and define the dynamic Euler equation in terms of such transformations. In the field of Financial Econometrics differencing is also prevalent, given the inter-relation between returns series that are uncorrelated and the notion of market efficiency.

The Euler equation is often estimated by Generalised Method of Moments (GMM). The application of this estimator requires the imposition of non-linear restrictions and corrections for auto-regressive or moving average error behaviour. In the simplest case GMM collapses to Instrumental Variables (IV), and IV can be readily modified to incorporate some of the above characteristics (Sargan (1983), Pesaran and Pesaran (1998)).

The question of identification in the IV and GMM cases revolves around the testing of the over-identifying restrictions. Instrument selection is key to the validity of these estimators Sargan(1964). In the rational expectations context a necessary condition for identification is the acceptance of order and rank conditions as discussed by Pesaran (1987).

A problem that arises in the context of the commonly used panel and financial data series is the order of the dynamic on the exogenous processes forcing the evolution of the dependent variable. Often, the test fails to reject provided that enough agent specific information is incorporated, without regard to whether the underlying rank condition is likely to be satisfied.

In the panel context, information is rich at the level of the agent, but poor in terms of dynamic structure. Hence, the ability to select instruments to explain the dynamic properties of the problem is limited especially when the agent specific information is discrete. Whereas for problems defined in a time series framework, there is only readily available information on the persistence of variables. In general, for identification it is imperative that the order of the dynamic process of the forcing variables should be longer than that on the endogenous variable. This criterion is often difficult to meet in the context of financial data as the dynamics are often short and the information set has a limited number of elements especially when high frequency data are used.

The above problems apply to any model estimated by limited information methods. Additionally, a more significant issue as far as identification of forward looking behaviour is concerned, is that both the IV and GMM estimators do not bind the solution based on the minimum of the optimisation problem to the restrictions associated with the terminal condition (Nickell (1985)).

Tests of over-identifying restrictions do not impose burdensome conditions on the estimator. The satisfaction of the necessary conditions follows without difficulty with the exception of highly non-persistent processes. This article, shows that although the necessary and sufficient conditions for identification are likely to be accepted in finite samples, this may not be true in the limit. As a result, the estimator may fail asymptotically. This leads to the proposition, that limited information estimators of forward behaviour fail to identify structural parameters. There are two substantive issues: (a) the key information about the solution resides in the residual which is not formally modelled. (b) Such models are observationally equivalent to backward looking processes associated with error correction behaviour. This is intimately related to the notion of Super Exogeneity which may negate the practical use value of the Lucas critique (Lucas (1973), Hendry (1988) and Hendry and Favero (1992)).

In such circumstances, efficient estimation requires a strategy that involves: (a) the estimation of the processes driving the exogenous variables, to determine lag order, significance of lags and derive innovations for the forcing variables. (b) The estimation of the quasi-structural model by a restricted instrumental variables estimator with MA errors or direct estimation of the forward looking model.

This article is as follows. Firstly, the literature on linear quadratic adjustment costs models(LQAC) is summarised in section 2. In section 3, some necessary conditions and instrument selection are discussed. The conditions for local identification of non-linear IV/AR systems presented in Sargan (1983) are adapted to handle the rational expectations (RE) problem in section 4. Then the question of the transversality condition and identification is discussed in section 5 and finally conclusions are drawn in section 6.

2 LQAC Models

The Identification of RE models depends on the structure of the RE problem, the conditions for a solution and the relationship between different representations of the same system. The following non-separable, non-symmetric objective function is used here (Kollintzas (1985)).

$$E(\Im_t|\Omega_t) = \sum_{t=o} E\{\beta^t(\Delta y_t' K \Delta y_t + (y_t - z_t)' H (y_t - z_t))|\Omega_t\} \qquad (1)$$

Let (1) define a control problem (Chow (1978)), y_t is a g vector of endogenous variables and z_t a g vector of unobserved targets, that can be defined as a linear function of k exogenous variables, x_t, $z_t = Ax_t + w_t$ where A is a matrix of long-run multipliers and $w_t = z_t - E(z_t|\Omega_t)$ is a g vector of white noise innovations and β is the discount rate. With fixed initial conditions $y_o = \bar{y}$, then from Kollintzas (1985) the Lagrange-Euler first order condition after substituting out for z_t is:

$$E(\beta^t Q_0 y_t - \beta^{t+1} Q_1 y_{t+1} - \beta^t Q_1' y_{t-1} - \beta^t HAx_t|\Omega_t) = 0; \qquad (2)$$

where $Q_o = (1+\beta)K + H$ and $Q_1 = K$.

Consider the process when it approaches its terminal value (at T):

$$E(\beta^T Q_0 y_T - \beta^{T+1} Q_1 y_{T+1} - \beta^T Q_1' y_{T-1} - \beta^T HAx_T|\Omega_t) = 0. \qquad (3)$$

Stationarity is one pre-condition traditionally accepted for the transversality condition to be satisfied (Pesaran(1981)), but when the structure includes a discount factor this assumption is too strong. In general all that is required is for (3) to be bounded as $T \to \infty$.

The standard solution to quadratic problems is symmetric, to reveal such a solution (3) is scaled by $\beta^{-\frac{1}{2}(T+1)}$:

$$\begin{aligned} &E(\beta^{-\frac{1}{2}(T+1)}\beta^T Q_0 y_T - \beta^{-\frac{1}{2}(T+1)}\beta^{T+1} Q_1 y_{T+1} - \\ &\beta^{-\frac{1}{2}(T+1)}\beta^T Q_1' y_{T-1} - \beta^{-\frac{1}{2}(T+1)}\beta^T HAx_T|\Omega_t) = 0. \end{aligned} \qquad (4)$$

Simplifying (4):

$$\begin{aligned} &E(\beta^{-\frac{1}{2}} Q_0 \beta^{\frac{1}{2}(T)} y_T - \beta^{\frac{1}{2}(T+1)} Q_1 y_{T+1} - \beta^{\frac{1}{2}(T-1)} Q_1' y_{T-1} \\ &- \beta^{-\frac{1}{2}} HA\beta^{\frac{1}{2}T} x_T|\Omega_t) = 0. \end{aligned} \qquad (5)$$

Re-defining (5) in terms of $y_T^* = \beta^{\frac{1}{2}(T)} y_T$ and $x_T^* = \beta^{\frac{1}{2}(T)} x_T$ gives rise to the symmetric form:

$$E(\beta^{-\frac{1}{2}} Q_0 y_T^* - Q_1 y_{T+1}^* - Q_1' y_{T-1}^* - \beta^{-\frac{1}{2}} HAx_T^*|\Omega_t) = 0. \qquad (6)$$

In the limit (6) is bounded when the roots of the processes driving x_t and y_t are of mean order less than $\beta^{-\frac{1}{2}}$ as:

$$\lim_{T \to \infty} E(y_{T+1}^*|\Omega_t) \to 0 \text{ and } \lim_{T \to \infty} E(x_{T+1}^*|\Omega_t) \to 0$$

Notice, that (6) is bounded even when y and x have univariate time series representations that are non-stationary. Now consider the cointegration case. Dividing (2) by β^t and transforming, yields an error correction representation:

$$E(-\beta K \Delta y_{t+1} + K \Delta y_t + H(y_t - Ax_t)|\Omega_t) = 0. \qquad (7)$$

It follows that (7) is bounded in the limit when:

$$\lim_{T\to\infty} \{-\beta K \Delta y_{T+1} + K \Delta y_T + H(y_T - Ax_T)\} \to 0. \qquad (8)$$

Decomposing (8):

$$\lim_{T\to\infty} \{-\beta K \Delta y_{T+1} + K \Delta y_T + H(y_T - Ax_T)\}$$
$$= \lim_{T\to\infty} \{-\beta K \Delta y_{T+1} + K \Delta y_T\} + \lim_{T\to\infty} \{H(y_T - Ax_T)\} \to 0 \qquad (9)$$

The conditions for cointegration (Engle and Granger (1987)) are sufficient for (9) to be satisfied. That is $y_t \sim I(1)$ and $(y_t - Ax_t) \sim I(0)$ or:

$$\lim_{T\to\infty} \Delta y_{T+1} \to 0 \text{ and } \lim_{T\to\infty} H(y_T - Ax_T) \to 0,$$

y_t and x_t cointegrate.

It follows from the above discussion, that a regular solution to (6) exists, if and only if: (a) Q_o is symmetric; (b) K is non-singular; and (c) $\lambda < \beta^{-\frac{1}{2}}$. Dividing (2) by β^{t} yields the following difference equation:

$$E(Q_o y_t - \beta Q_1 y_{t+1} - Q_1 y_{t-1} - HAx_t | \Omega_t) = 0 \qquad (10)$$

Redefining (10) using the forward (L^{-1}) and backward (L) lag operators:

$$Q(L)E(y_t|\Omega_t) = HAE(x_t|\Omega_t) \qquad (11)$$

Now $Q(L) = (Q_o I - Q_1 L^{-1} - Q_1' L)$ has the following factorisation:

$$Q_1 Q(L) = (I - G_1 L^{-1})(I - FL)$$

where $G_1 = \beta F$, $F = D\Lambda D$ and Λ is a matrix whose diagonal elements are the stable eigen roots of the system. Substituting out for $Q_1 Q(L)$ yields:

$$(I - G_1 L^{-1})(I - FL)E(y_t|\Omega_t) = K^{-1} HAE(x_t|\Omega_t). \qquad (12)$$

It follows that the solution of the system can be written as:

$$y_t - Fy_{t-1} = \sum_{s=0}^{\infty} (G_1)^s F E(R_o A x_{t+s}|\Omega_t) + (G_1)^{-t} M_t + u_t \qquad (13)$$

(Sargent(1978)). Where $R_o = (\beta(F - I) + F^{-1} - I)$ and M_t satisfies the martingale property $E(M_{t+1}|\Omega_t) = (G_1)M_t$ (Pesaran(1981)).

Assuming, that their are no bubbles and a forcing process $\Gamma(L)x_t = w_t$ (w_t is white noise), then the backward solution comes from substituting $E(x_{t+s}|\Omega_t)$ using the Wiener-Kolmogorov prediction formula, that gives rise to the reduced form:

$$y_t - Fy_{t-1} = \Pi(L)x_t + u_t, \qquad (14)$$

where $\Xi(L) = (\Xi_0 + \Xi_1 L + \ldots \Xi_{s-1} L^{s-1})$ is a function of $\beta; H; K; A$ and $\Theta(L) = (I + \Theta_1 L + \ldots \Theta_s L^s)$.

When the saddle path property holds and $M_t = 0$, then estimators based on (13) and (14) use all the information associated with the full solution, but such estimators require the imposition of a number of non-linear restrictions involving both the structural parameters $(\beta; H; K; A)$ and the matrix polynomial $\Theta(L)$ for the exogenous variables. The implementation of this procedure is practicable in a time-series framework when g is small (Hunter (1989)). Whereas, in the Panel data context, thus far, the time dimension has not been sufficiently long to make the approach practicable.

Otherwise, estimators based on (13) use the Errors-in-Variables method due to Muth (1961) to replace $E(y_{t-1}|\Omega_t)$ by $y_{t-1}; E(y_{t+i}|\Omega_t)$ by $y_{t+i} - u_{t+i} - iHADw_{t+i}$ for $i = 0; 1$, and $E(x_t|\Omega_t)$ by x_t:

$$Q_o y_t - \beta Q_1 y_{t+1} - Q_1 y_{t-1} - HAx_t = \varepsilon_{t+1} \quad (15)$$

where $\varepsilon_{t+1} = Q_o u_t - \beta Q_1 u_{t+1} + HADw_t$ is an MA(1) error with an innovation in x. Equation (15) defines a Quasi-Structural Form (Q-SF) and the following linearisations can be estimated consistently by Instrumental Variables when an optimal instrument set exists (Sargan (1983)):

$$Q_o y_t - \beta Q_1 y_{t+1} - Q_1 y_{t-1} - A_1 x_t = \varepsilon_{t+1} \quad (16)$$

where Qo and Q_1 are defined above, $A_1 = HA$, and an unrestricted Quasi-Reduced Form (Q-RF) of (15) and (16) is:

$$y_{t+1} = P_1 y_t + P_2 y_{t-1} + P_3 x_t + v_{t+1}: \quad (17)$$

Unstable solutions cannot be ruled out by definition, but estimates of the Q-SF parameters based on (15), (16) and (17) are not affected by the non-uniqueness as the transform $(I - G_1 L)$ annihilates all bubble solutions. Identification of the Q-SF parameters $[\beta; K; H; A]$ comes via comparison with those in the Q-RF $[P_1; P_2; P_3]$.

3 Necessary Conditions and Instrument Selection

Consider the following structural representation of the forward looking model (13), given that K is non-singular[1]:

$$Ky_t - KFy_{t-1} = \sum_{s=0} H(G_1)^s F E(Ax_{t+s}|\Omega_t) + K(G_1)^{-t} M_t + Ku_t \quad (18)$$

[1] Should K be singular, then the dimension of the problem is reduced via an adding up constraint or the endogenous variables cointegrate.

where $KR_o = K(\beta(F - I) + F^{-1} - I) = H$ as $R_o = K^{-1}H$ commutes. In the light of this, K and H can be identified independently subject to the usual type of restriction. Essentially, identification of K follows from the additional restrictions, while identification of H follows from F, given knowledge of K and any additional restrictions on the system. Moving on to the Q-RF parameters, their identification is a pre-requisite for the identification of the Q-SF parameters. As this determines, the instrument selection criterion. From Rothenberg (1971) the following Theorem is sufficient to globally identify the Q-RF parameters.

Theorem 1 *rank$[\hat{y}_t : y_{t-1} : x_t] = 2g + k$ is sufficient for the identification of the $Q-RF$ parameters $\rho = vec[P_1, P_2, P_3]$.*

Proof. : [Rothenberg (1971)] ■

When Theorem 1 holds ρ is uniquely defined and each equation in the system can be estimated by Instrumental Variables. However, the local conditions for identification in Pesaran (1987)[2] are also of interest. Equation (15) is analogous to the representation used by Pesaran(1987):

$$Q_o y_t - \beta K y_{t+1} - K y_{t-1} - A_1 x_t = \varepsilon_{t+1} \tag{19}$$

When $Q_o = (1+\beta)K + H$ is non-singular, then (14) is also the RF and the optimal predictor is $y_t = F y_{t-1} + \Xi(L)x_t + u_t$. Hence, $E(y_t|\Omega_t)$ depends on y_{t-1} and x_{t-k}, for $k = 0, 1, 2..s-1$ and by projecting (14) forward one period and taking expectations:

$$E(y_{t+1}|\Omega_t) = FE(y_t|\Omega_t) + \Xi(L)E(x_{t+1}|\Omega_t).$$

Substituting for $E(y_t|\Omega_t)$ using (14) and $E(x_{t+1}|\Omega_t)$ using $(I + \Theta_1 L + \ldots \Theta_s L^s)x_t$, it follows, that:

$$E(y_{t+1}|\Omega_t) = F^2 y_{t-1} + \Upsilon(L)x_t + u_t. \tag{20}$$

Where $\Upsilon(L) = F\Xi(L)$. The following recursive relations can be derived from (14), (19) and (20):

$$\begin{aligned} Q_0 F - K - \beta K F^2 &= 0 \\ Q_0 \Upsilon_1 + A_1 + \beta K \Xi_0 &= 0 \\ Q_0 \Upsilon_i + \beta K \Xi_i &= 0 \text{ for } i = 2, 3, \ldots s-1 \end{aligned}$$

If $B = [Q_0, -K, A_1, -\beta K]$ and b_i is the vector of parameters for the i^{th} equation, then the parameters of a Q-RF equation are identified when matrix Q below has rank $3g + k - 1$.[3]

$$Q = \begin{bmatrix} \Phi_i & F & \Pi_0 & \Pi_1 & \ldots & \Pi_{s-1} \\ O & I_g & O & O & \ldots & O \\ O & O & I_k & O & \ldots & O \\ O & F^2 & \Xi_0 & \Xi_1 & \ldots & \Xi_{s-1} \end{bmatrix}$$

[2] Pesaran (1987) makes the correction to the results presented in his1981 article, suggested by Wegge and Feldman (1981).

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

If the rank$(Q) = 3g + k - 1$, then a necessary condition for identification is s; dimensioned so that the following order condition is satisfied: $3g + k - 1 < r + g + ks$ or $2g - 1 < r + k(s - 1)$.

Following Pesaran (1987) consider the matrix T

$$T = \begin{bmatrix} Q_0 & -K & A_1 & \beta K \\ O & I_g & O & O \\ O & O & I_k & O \\ O & O & O & I_g \end{bmatrix}$$

pre-multiplying Q by T:

$$TQ = \begin{bmatrix} \Phi B & O & O & \ldots & O \\ O & I_g & O & \ldots & O \\ O & O & I_k & \ldots & O \\ O & F^2 & \Xi_0 & \ldots & \Xi_{s-1} \end{bmatrix}$$

where $\Phi = [\Phi_i : O]$ and rank$(TQ) = g + k +$ rank(TQ^*):

$$TQ^* = \begin{bmatrix} \Phi B & O & \ldots & O \\ O & \Xi_1 & \ldots & \Xi_{s-1} \end{bmatrix}$$

Now rank$(TQ^*) = 2g - 1$, which is valid only when there are sufficient restrictions r or sufficient lags in the VAR generating x_t. As long as there is a valid set of instruments and appropriate restrictions then the parameters are identified. In the usual type of system, identification depends on the existence of $g - 1$ restrictions per equation. When the parameters of interest are associated with the forward looking model $(\beta; F; A])$, that is equivalent to assuming that ΦB contains $g - 1$ restrictions. This case revolves around the rank of the matrix associated with the instruments. Hence:

$$\text{rank}\begin{bmatrix} \Xi_1 & \ldots & \Xi_{s-1} \end{bmatrix} = g - 1;$$

which follows when there is sufficient information on the exogenous variables and enough lags. As a result, $s > 1$ and $g \leq k$. An analogous condition associated with models not dependent on the structure of an optimised objective function are considered in Pesaran (1981,1987).

As (16) is linear the local conditions associated with Pesaran (1987) globally identify the Q-RF parameters. From the rank condition and the dimension of TQ^*, follows an order condition, that is necessary for local identification:

$$g + (s - 1)k > 2g \text{ or } s > \frac{g}{k} + 1:$$

This order condition implies, that for a given number of equations in the system (g), the lag length of the forcing variables depends on their number (k). The greater the number of such variables the shorter the lag length required to a lower limit of two.

When direct maximum likelihood is applied to (13) and the exogenous variable process is sufficiently long, then the rank condition ($\text{rank}(TQ^*) = g + k$) is likely to hold. Notice, that under the relatively weak criterion on the roots discussed in section 2, the transversality condition is also satisfied by this estimator. However, IV or GMM estimators applied to (19), even when the non-linear parameter restrictions are imposed are not likely to satisfy the transversality condition. Failure to satisfy this condition will result in loss of identification even though the rank condition is satisfied.

The following section develops the argument for the use of non-linear IV estimators instead of the more conventional GMM and linear IV estimators.

4 Local Conditions for Identification

A necessary and sufficient condition for local identification can be derived using the Jacobian matrix and the moment matrix of the data. Rothenberg(1971) deals with conditions for local identification of quite general models and we can extended those conditions when there are enough appropriate instruments for the endogenous variables and the future exogenous variables. In this section conditions for identification are derived for the IV estimator and it is shown that such conditions fail when certain moment conditions hold.

Consider the usual form of the first order condition associated with the symmetric quadratic loss function discussed in the previous section:

$$E(Q_0 y_t - \beta Q_1 y_{t+1} + 1 - Q_1' y_{t-1} - HAx_t | \Omega_t) = 0; \tag{21}$$

where $Q_0 = (1 + \beta)K + H$ and $Q_1 = K$. Substituting out for expectations and by actual values:

$$Q_0 y_t - \beta Q_1 y_{t+1} - Q_1 y_{t-1} - HAx_t = \varepsilon_{t+1} \tag{22}$$

where $\varepsilon_{t+1} = Q_0 u_t - \beta Q_1 u_{t+1} + HADw_t$ is an MA(1) error with a surprise in the x's. Equation (22) can be estimated consistently by minimising the following criterion when optimal instruments exist (Sargan (1983)) or the rank condition discussed above is satisfied.

$$V(\theta)\frac{\text{plim}(X^{*\prime}Z)}{N} = 0 \tag{23}$$

where $X^* = [Y\ X\ Y_{+1}\ Y_{-1}]$; $Z = \left[\hat{Y}\ X\ \hat{Y}_{+1} Y_{-1}\right]$ and $V(\theta) = [(1 + \beta)K + H : -HA : -\beta K : -K]$:

The criterion is made operational by replacing $\hat{Y}$, $\hat{Y}_{+1}$ and $\hat{X}_{+1}$ in Z and then using $Z^+ = [Y_{-1}; X; X_{-1}; X_{-2}; \ldots X_{-s}]$ Vectorising (23) and letting $\text{plim}(\frac{X^{*\prime}Z}{N}) = M'$ a matrix of constants:

$$\text{vec}(V(\theta)\text{plim}(\frac{X^{*\prime}Z}{N})) = (M \otimes I_g)\text{vec}(V(\theta)) = 0: \tag{24}$$

Sargan(1983) shows for a generalised instrumental variables system that a necessary and sufficient condition for local identification is given by looking at the first derivative of the probability limit specified above. Differentiating (24):

$$\frac{d\,vec(V(\theta)\,plim(\frac{X^{*\prime}Z}{N}))}{d\theta} = (M \otimes I_g)\frac{d\,vec(V(\theta))}{d\theta}; \tag{25}$$

gives rise to the rank condition:

$$rank\{(M \otimes I_g)\frac{d\,vec(V(\theta))}{d\theta}\} = m = gk + 2g^2: \tag{26}$$

The moment matrix of the data can be written as:

$$M = plim \left[\frac{Z^{*\prime}Y}{N} : \frac{Z^{*\prime}X}{N} : \frac{Z^{*\prime}Y_{+1}}{N} : \frac{Z^{*\prime}Y_{-1}}{N} \right]$$

$$= [M_0 : M_1 : M_2 : M_3]$$

If $\theta = [\beta : vec(A)' : vec(H)' : vec(K)']$, then:

$$\frac{d\,vec(V(\theta))}{d\theta} = \begin{bmatrix} vec(K) & 0 & I_{g^2} & (1+\beta)I_{g^2} \\ 0 & -(I_g \otimes H) & -(A' \otimes I_g) & 0 \\ -vec(K) & 0 & 0 & -\beta I_{g^2} \\ 0 & 0 & 0 & -I_{g^2} \end{bmatrix} :^4 \tag{27}$$

That the Jacobian matrix above has full rank is a necessary condition for global identification in its own right (Sargan(1983)). However, this condition is not sufficient. If one transforms the Jacobian by elementary row manipulation, then the following equivalent matrix is obtained:

$$\frac{d\,vec(V(\theta))^*}{d\theta} = \begin{bmatrix} vec(K) & 0 & I_{g^2} & (1+\beta)I_{g^2} \\ 0 & -(I_g \otimes H) & -(A' \otimes I_g) & 0 \\ 0 & 0 & I_{g^2} & I_{g^2} \\ 0 & 0 & 0 & -I_{g^2} \end{bmatrix}: \tag{28}$$

However, it follows that this transformation implies a re-ordering of elements in $M \otimes I$. Therefore

$$\frac{d\,vec(V(\theta)^*plim(\frac{X^{*\prime}Z}{N})^*)}{d\theta} = (M^* \otimes I_g)\frac{d\,vec(V(\theta))^*}{d\theta}$$

where $M^* = [M_0 - M_2 : M_1 : M_2 : M_3]$. Notice, that the transformed Jacobian is dimensioned, $gk + 3g^2$ by $1 + gk + 2g^2$: The maximum rank is therefore $1 + gk + 2g^2$. Given the quasi diagonal structure of the new matrix, then

[4] It is useful to remember that:

$$-(A' \otimes I)(vec(H)) = -(I \otimes H)vec(A)$$

independence of the last $2g^2$ columns follows from the reordering. Any rank deficiency of the upper two blocks is simply due to the nullity of the first two blocks which either require K to be nilpotent or H to be rank deficient.

If as is common in the literature the data has been differenced, then inherent stationarity of y negates the possibility that H is rank deficient. For identification of β (the discount rate), all that is required is a single column of K to be non null.

The sufficient condition for identification depends on the rank of M (the matrix of second moments). One way of testing this is via the well known Bassman-Sargan tests of over-identifying restrictions. Again, it is common practice to simply extend the matrix of instruments to the point at which the test fails to reject the null.[5] As a consequence, this condition may always be satisfied. However, the necessary and sufficient condition for local identification require:

$$\operatorname{rank}\{(M^* \otimes I_g)\frac{d\operatorname{vec}(V(\theta))^*}{d\theta}\} \leq \min[\operatorname{rank}((M^* \otimes I_g)); \operatorname{rank}\left(\frac{d\operatorname{vec}(V(\theta))^*}{d\theta}\right)] \tag{29}$$

Satisfaction of either the Jacobian or the rank condition on the moments is only necessary for identification as it is possible for either of them to hold and yet the model fails the rank condition. Consider:

$$(M^* \otimes I)\frac{d\operatorname{vec}(V(\theta))^*}{d\theta} = [((M_o - M_2) \otimes I_g)\operatorname{vec}(K) : -(M_1 \otimes H) :$$

$$((M_o - M_2) \otimes I_g) - (A'M_1 \otimes I_g) + (M_2 \otimes I_g)) : (1+\beta)((M_o - M_2) \otimes I_g) + (M_2 \otimes I_g) - (M_3 \otimes I_g)]$$

$$= [((M_o - M_2) \otimes I_g)\operatorname{vec}(K) : -(M_1 \otimes H) : ((M_o - A'M_2) \otimes I_g) :$$

$$\beta((M_o - M_2) \otimes I_g) + ((M_o - M_3) \otimes I_g)]. \tag{30}$$

The rank condition is violated when any of the following moment conditions hold:

I) $\operatorname{rank}[((M_o - M_2) \otimes I_g)\operatorname{vec}(K) : (M_1 \otimes H)] \leq g^2$
II) $M_o - A'M_1 = 0$
III) $\beta(M_o - M_2) + (M_o - M_3) = 0$

At this juncture, it is assumed, that the instruments are adequate.[6] Having assumed that the simple reason for failure of identification is rejected, which is

[5] In particular, the simulation evidence presented by Hall et al (1996) suggests that selecting instrument sets may exacerbate the poor small sample properties of the test. While Stock and Wright (2000) impute for monthly data, that the asymptotic properties of test statistics are only likely to hold when a Century of data is amassed.

[6] Hence, the Bassman Sargan criterion is satisfied for the linearised case or equivalently the rank condition in the previous section is accepted. Should this not be the case, then the moment matrices will be null and the estimator will fail as a result.

often the case when sufficient lagged information is included, then conditions (I)-(III) may be viewed in two ways.

Firstly, (I)-(III) provide a set of distinct rank and moment conditions, which should they hold will lead to the model parameters being unidentified. Notice, that these conditions are different from the conditions developed by Pesaran (1987) and the more usual over-identifying restrictions. As a result, (I) and the conditions, which lie at the heart of (III) hold when $M_o = M_2$ and $M_o = M_3$ as contemporaneous and future correlations tend in the limit to the same value. In such circumstances, it will not be possible to identify β and K[7]. It follows that (II) holds when $M_o = A'M_1$ and identification fails when the data are cointegrated or a super-consistency result holds. As a result H is not identified. Identification of A fails when rank(M_1 - H) $< g^2$ and either elements of M_1 are dependent, H is not identified or (II) holds.

In this section it has been shown that local identification requires the imposition of additional restrictions on the behaviour of the moments of the data to satisfy the necessary and sufficient conditions. It should be noticed that some of these distinctions may become fine when the data is highly persistent and has a short order dynamic as their may be little to distinguish between M_o, M_2 and M_3. Hence, identification may be lost when long time series of financial data are considered (Stock and Wright (2000)).

In the next section, a further issue arrises, that relates to the imposition of the transversality condition. It is shown, that when some of the conditions discussed above hold, this will lead to a loss of asymptotic identification.

5 Local identification and the transversality condition

Conditions (I)-(III) are related to the transversality condition, which also needs to hold for the estimator to be distinguished as a forward looking representation of the data as compared with being backward looking.

Consider the first order condition (22), stacked across the sample:

$$((1+\beta)K + H)Y' - \beta K Y'_{+1} - K Y'_{-1} - HAX' = E' \qquad (31)$$

Post multiplying by the instrument matrix (Z^+); gives rise to:

$$((1+\beta)K + H)Y'Z^+ - \beta K Y'_{+1}Z^+ - K Y'_{-1}Z^+ - HAX'Z^+ = E'Z^+. \qquad (32)$$

[7] Sargan (1992) and Gregory et al (1993) address the question of identification of the discount rate. If g = 1, then β and the single coefficient in $K^{-1}H$ are indistinguishable without further restriction. When there is more than one equation, then this is simply the requirement to normalise everything with respect to one parameter. Recall, that identification of K usually requires additional restrictions, when systems of equations of the form described by (18) are considered. It follows from the global condition discussed in Hunter (1989, 1992), that failure of identification of β follows when K is nilpotent or vec(K) = 0. Furthermore, nilpotency is less a condition of the data than the parameters of the objective function.

It is well known that the IV estimator minimises the distance between the error vector and instruments, and for linear models this condition is exactly zero when the number of instruments matches the number of endogenous variables. Otherwise, the test in Sargan (1964) determines whether the instrument set is valid or whether $E'Z^+$ is significantly different from zero:

Dividing (32) by N and taking limits:

$$\begin{aligned} &((1+\beta)K + H)\operatorname{plim}\frac{Y'Z^+}{N} - \beta K\operatorname{plim}\frac{Y'_{+1}Z^+}{N} \\ &- K\operatorname{plim}\frac{Y'_{-1}Z^+}{N} - HA\operatorname{plim}\frac{X'Z^+}{N} = \operatorname{plim}\frac{E'Z^+}{N} \end{aligned} \tag{33}$$

Re-ordering the elements of (33) above and replacing the probability limits by the appropriate elements of M:

$$\beta K(M_o' - M_2') + K(M_o' - M_3') + H(M_o' - AM_1') = \operatorname{plim}\frac{E'Z^+}{N} \tag{34}$$

In the limit an estimator, which satisfies a Sargan type criterion will satisfy the condition associated with the Euler condition as $\operatorname{plim}\frac{Z^{+\prime}E}{N}$ will not be statistically different from zero. Across the sample, any estimator satisfying the test of over-identifying restrictions will be bound to satisfy the transversality condition. However, this presents a number of difficulties, (i) satisfying the transversality condition is technically consistent with a failure of identification, (ii) it is only in the limit when the sample is large that acceptance of the over-identifying restrictions test is consistent with acceptance of the transversality condition.

Consider (34) and substitute out conditions (II) and (III). Then the left hand side of (34) is always zero irrespective of the satisfaction of the Sargan condition. The same result obtains when instead of (II), either H is nilpotent or instead of (III), K is nilpotent. Otherwise, both H and K might be nilpotent and by implication, from the structure of the objective function the forward looking explanation cannot underlie (22) as K^{-1} does not exist. A similar conclusion can be drawn when the discount rate is not well defined.

Secondly, in finite samples, satisfying the test of over-identifying restrictions is necessary, but not sufficient for accepting the transversality condition. It is not sufficient for the instruments to minimise the error variance within sample as the transversality condition must hold beyond the estimation sample.[8]

If the right hand side of (34) is set to zero, then subject to (I)-(III) being accepted:

$$K[\beta(M_o - M_2)' + (M_o - M_3)'] = -H(M_o - M_1A')' \tag{35}$$

[8] In practice, the prediction tests or parameter stability tests discussed by Hendry (1980) and Hendry and Richard (1982) are required for the satisfaction of the transversality condition. In line with the argument suggested by Ericsson et al (1999), this would appear to yield a contradiction. The parameter estimates need to be stable when the process forcing the exogenous variables may well not be.

If the discount rate is ignored or $\beta \approx 1$, then the following condition is relevant for the existence of well defined parameters in the objective function:

$$K^{-1}H = -(M_o - M_1A')'[(M_o - M_2)' + (M_o - M_3)']^{-1} \qquad (36)$$

From (36) it follows that $(M_o - M_2)' + (M_o - M_3)'$ cannot converge at a faster rate than $(M_o - M_1A')$.

Recall that $K^{-1}H = (\beta(F - I) + F^{-1} - I)$ directly relates to the roots of the process driving y_t. In the case where there are unit roots, it will not be possible to identify all the elements of H as $(\beta(F - I) + F^{-1} - I)$ has some zero roots. In this case there is a loss of identification. However, reduced form estimation is possible provided that parameter estimation fully accounts for any non-stationarity amongst the y variables. Subject to estimation under this restriction the solution still exists when there is a discount factor.

6 Conclusions

Estimation of the structural parameters of optimising 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 rational expectations to be imposed. Limited information procedures (IV/GMM) as typically applied, yield Quasi-Structural Forms, but they do not ordinarily permit testing of the rational expectations restrictions or the imposition of the transversality condition.

The practice, that persistence in the error term can be filtered confounds the issue further as the power of the tests of over-identifying restrictions is then called into question. Any filter is implicitly re-using the instruments. Limited information estimators have no procedure for imposing the restrictions associated with the transversality condition, hence they are not identified.

Should long time series be available for the investigation, then the above results suggest direct estimation of the forward model in combination with the exogenous variable processes. This is subject to the restriction implied by the transversality condition and a sufficiently long lag order for the driving processes. In the case where panel data is used and time series observations are short, then further moment conditions should be imposed, even though the transversality condition might still not be satisfied by the estimated model parameters. Notice, that such practice takes us beyond standard linear estimation and requires the imposition of additional restrictions. However, the imposition of such restrictions will guarantee that the necessary conditions are fully satisfied. Even so, there is no guarantee that such models provide reliable estimates of the Q-SF parameters. It was shown in section 4 that the necessary conditions are technically violated when either certain moments are zero or matrices H and K from the objective function are null. In the latter case, failure of identi¯cation stems from the failure of the forward looking optimising model to provide a legitimate explanation of the estimated Q-SF. Such problems will surface when it is possible to show that the transversality condition has been violated. One sign of such a violation is parameter instability. Hence, holding back a sub-sample of time series observations would permit tests of parameter stability, while the observation that estimated parameters were stable would provide additional support for the proposition that the Q-SF parameters are identi¯ed.

In line with the concluding remarks made by Stock and Wright (2000), this article has extended the results presented in Sargan (1983) for the Identi¯cation of non-linear IV estimators to the Euler equation case.

7 References

Arellano, M. and Bond, S.R., (1991), Some Tests of Speci¯cation 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. (1999) Underidentification? paper presented at the Econometrics Study Group Conference, Bristol July 1999.

Blinder, M. and Pesaran, M.H. (1995) Multivariate Rational Expectations Models and Macroeconomic Modelling: A Review and Some New Results. Handbook of Applied Econometrics, Blackwell, pp139-187.

Blundell, R. and Bond, S.R. (1998), Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, Journal of Econometrics, 87, 115-43.

Bond, S.R. and Meghir, C. (1994) Dynamic Investment Models and the Firm's Financial Policy. Review of Economic Studies, 61, 197-222.

Chow, G.C.(1978), Analysis and Control of Dynamic Economic Systems. John Wiley.

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

Engle, R. F., and Granger, W. J. (1987) Cointegration and Error-Correction: Representation, Estimation and Testing. Econometrica, 55, 251-276.

Ericsson, N.R. and Hendry, D.F. (1999) Encompassing and Rational Expectations: How Sequential Corroboration Can Imply Refutation. Empirical Economics. 24 (1). 1-21.

Favero, C. and Hendry, D.F. (1992) Testing the Lucas Critique: A Review. Econometric Reviews, 11 (3), 265-306.

Forchini, G. and Hillier, G. (1999) Conditional Inference for Possibly Unidentified Structural Equations. paper presented at the Econometrics Study Group Conference, Bristol July 1999.

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. p 220-39.

Hall, A.R., Rudebusch, G.D. and Wilcox, D.W. (1996) Judging Instrument Relevance in Instrumental Variables Estimation. International Economic Review. 37 283-98.

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

Hendry, D.F. (1980) Predictive failure and econometric modelling. In P. Ormerod (ed), Macroeconomics; the transactions demand for money, in Economic Modelling; p 217-42, Heinemann, London.

Hendry, D.F. and Richard, J.F. (1982) On the formulation of empirical models in dynamic econometrics. Journal of Econometrics; 20, 3-33.

Hendry, D.F. (1988) The Encompassing Implications of Feedback versus Feedforward Mechanisms in Econometrics. Oxford Economic Papers; No 40, 132-149.

Hunter, J.(1989) 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 also presented at ESEM93, August 1993 in Uppsala.

Lucas, R.E. (1976) Econometric Policy Evaluation: A Critique. In Brunner K. and Meltzer A.H. (eds), The Philips Curve and Labor Markets; Carnegie-Rochester Conference Series on Public Policy, 1. North-Holland, Amsterdam.

Kollintzas, T. (1985) The Symmetric Linear Rational Expectations Model\, Econometrica, 53, 963-76.

Muth, J.F.,(1961) Estimation of Economic Relationships Containing Latent Expectations Variables. In Lucas, R.E. and Sargent, T.J.(1981), Rational Expectations and Econometric Practice, George Allen and Unwin, London, p 321-328

Nickell, S.J.(1985) Error-correction, partial adjustment and all that: an expository note. Oxford Bulletin of Economics and Statistics, 47, 119-29.

Pesaran, M.H.,(1981) Identi¯cation of Rational Expectations Models. Journal of Econometrics, 16, 375-98.

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

Pesaran, B. and Pesaran, M.H. (1998), Microfit 4: Oxford Electronic Publishing, Oxford.

Rothenberg, T.J. (1971) Identi¯cation 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 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, pp.169-205.

Sargent, T.J.(1978) Estimation of Dynamic Labour Demand Schedules Under Rational Expectations. Journal of Political Economy, 86, 1009-44.

Sims, C., (1980), Macroeconomics and Reality, Econometrica 48, 11-48.

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

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

Wegge, L.L. and Feldman, M.(1983) Comment to the Editor. Journal of Econometrics, 21, 255-6.