Econometric Theory, **30**, 2014, 150–175.
doi:10.1017/S0266466613000157

MEASUREMENT ERRORS IN DYNAMIC MODELS

IVANA KOMUNJER
University of California, San Diego

SERENA NG
Columbia University

Static models that are not identifiable in the presence of white noise measurement errors are known to be potentially identifiable when the model has dynamics. However, few results are available for the plausible case of serially correlated measurement errors. This paper provides order and rank conditions for "limited information" identification of parameters in dynamic models with measurement errors where some aspects of the probability model are not fully specified or utilized. The key is to consider a model for the contaminated data that has richer dynamics than the model for the correctly observed data. Simply counting the total number of unknown parameters in the true model relative to the estimable model will not yield an informative order condition for identification. Implications for single-equation, vector autoregressive, and panel data models are studied.

1. INTRODUCTION

Empirical analysis often involves using an observable that is an inexact measure of the latent variable of interest. To distinguish the marginal effect of a covariate from measurement error, assumptions on the measurement error process are required. While the white noise assumption is convenient, it is rather restrictive in time series and panel data settings when the data are correlated over time. This paper provides general conditions for identification of dynamic models subject to serially correlated measurement errors. It is shown that the error-ridden data necessitate consideration of a model with higher-order dynamics than if the data were correctly observed. Furthermore, an informative order condition for identification cannot be obtained by simply comparing the total number of unknowns in the true model with that of the estimable model. Specific conditions are provided for vector autoregressive models with predetermined variables (VARX) and dynamic panel models with fixed effects.

This paper was presented at the 2010 SETA conference in Singapore. We thank the editor, three anonymous referees, and the seminar participants for helpful comments. Financial support from the National Science Foundation SES-0962473 and SES-0962431 is gratefully acknowledged. Address correspondence to Serena Ng, Columbia University, 420 W. 118 St. MC 3308, New York, NY 10027; e-mail: serena.ng@columbia.edu.

 © Cambridge University Press 2013

As is well documented, measurement errors are prevalent in survey data. Erickson and Whited (2000) argue that investment cash-flow regressions can be uninformative when Tobin's Q is measured with errors because average Q is no longer a good proxy for marginal Q. Devereux (2001) finds significant errors in survey reports of earnings and work hours. Quality variables (such as those of hospitals and the environment) are especially susceptible to measurement errors. Aggregate data are less susceptible but are not immune to measurement error problems. Since we do not observe variables such as the level of global economic activity or the state of the economy, filtered variables are often used as proxies. Except by coincidence, the deviations between the latent process and the filtered series are likely to be serially correlated. Furthermore, errors in data collection and reporting are inevitable. Wilcox (1992) discusses the issues in consumption measurements, especially at the monthly level. If the theoretical construct of consumption differs from its empirical counterpart, identification of the behavioral parameters and testing of consumption from observed data will be misleading. Ermini (1993) shows that allowing for serially uncorrelated measurement errors changes the measure of persistence in consumption growth. Falk and Lee (1990) suggest that measurement errors can explain rejections of the permanent income hypothesis. Sargent (1989) considers a model in which the data collecting agency observes data with serially correlated errors but reports a filtered version of the data. In the context of an accelerator model, he shows that when the data are error ridden, more steps are needed to estimate the parameters.

Given that economic time series are known to be persistent, it is somewhat surprising that measurement errors are usually assumed to be serially uncorrelated. Perhaps one explanation is that estimation with white noise measurement errors is no more difficult than when measurement errors are absent, provided valid instruments can be found. However, when measurement errors are serially correlated, short lags of the data will still be correlated with the composite error, while lags long enough to satisfy instrument exogeneity will likely be weak. Eberly, Rebelo, and Vincent (2009) estimate an AR(1) model for the measurement noise in Tobin's Q and note that the measurement error problem "cannot be corrected in investment regressions using instrumental regressions." Dealing with correlated measurement errors might well require state space modeling of all variables, which may not always be possible.

Indeed, while much has been written about errors in variables (EIV) in a static setting, the role of measurement errors in dynamic models is much less studied. Grether and Maddala (1973) examine the effects of measurement errors on the probability limits of the estimated coefficients in distributed lag models. They show that measurement errors in the exogenous variables may lead a researcher to detect long lags in adjustments. Ashley and Vaughan (1986) obtain an upper bound on the amount of measurement error in time series using a spectral decomposition method. Chanda (1996) studies the asymptotic properties of estimators based on Yule-Walker equations in autoregressive models with EIV. Tanaka (2002) develops tests to detect the existence of measurement errors in

long- and short-memory data. Staudenmayer and Buonaccorsi (2005) study the estimation of parameters in autoregressive models when the measurement errors are uncorrelated but possibly heteroskedastic. All of the above papers focus on single-equation analysis. Patriota, Sato, and Blas (2009) show that not accounting for measurement errors distorts Granger causality tests and consider estimation of VARs with white noise measurement errors.

We consider identification of single- and multiple-equation dynamic models with dynamic measurement errors under limited information. This can mean using the conditional model alone for identification, leaving the marginal distribution of the exogenous variables unspecified, as in autoregressive distributed lag models. Or it can mean that some but not all aspects of the probability distribution of the data are used for identification. To proceed, Section 2 sets up the VARX framework and reviews related work in the literature. Section 3 first presents general order conditions for identification and then specializes to ARX and VAR models. Rank conditions are derived in Section 4, and the results are illustrated via an example. Dynamic panel data models with measurement errors are considered in Section 5. Proofs are relegated to an Appendix. As a matter of notation, if $P(L)$ is an autoregressive polynomial matrix of order p, then $P(L) = I - \sum_{j=1}^{p} P_j L^j$, and if $Q(L)$ is a moving average polynomial matrix of order q, then $Q(L) = I + \sum_{j=1}^{q} Q_j L^j$. Finally, if $R(L)$ is a polynomial matrix of order r associated with the covariates, then $R(L) = \sum_{j=0}^{r} R_j L^j$.

2. VARX MODELS WITH MEASUREMENT ERRORS

2.1. Preliminaries

Consider a system of dynamic equations relating n_y endogenous variables $y_t = (y_{1t}, \ldots, y_{n_y t})'$ to n_x exogenous variables $x_t = (x_{1t}, \ldots, x_{n_x t})'$ and n_y disturbances $u_t = (u_{1t}, \ldots, u_{n_y t})'$. For all $t = 1, \ldots T$, the data generating process (DGP) is assumed to be

$$y_t = \sum_{j=1}^{p_y} A_j y_{t-j} + \sum_{j=0}^{r_x} B_j x_{t-j} + u_t. \qquad \textbf{(Model-y)}$$

We will henceforth refer to the DGP as Model-y. Then $\vec{A} \equiv (A_1, \ldots, A_{p_y})$, $\vec{B} \equiv (B_0, \ldots, B_{r_x})$, and Σ_u are the parameters of Model-y.

Suppose the observed data are contaminated by errors ϵ_t^y and ϵ_t^x so that for all t,

$$Y_t = y_t + \epsilon_t^y$$
$$X_t = x_t + \epsilon_t^x.$$

A correctly measured variable has a corresponding measurement error component of zero. Let M^y and M^x be matrices consisting of zeros and ones (assumed known to the econometrician) that select the n_ϵ^y and n_ϵ^x nonzero components of

ϵ_t^y and ϵ_t^x, respectively. When the jth component of y_t is measured without error, the corresponding component of ϵ_t^y is zero, as is the jth column of the $n_\epsilon^y \times n_y$ matrix M^y. When every component of y_t is measured with error, M^y is an identity matrix of dimension n_y.

Assumption A. (i) $u_t \sim WN(0, \Sigma_u)$ with Σ_u full rank; (ii) x_t is weakly stationary; (iii) for all t and $\tau \geqslant 0$, $\mathbb{E}[u_t x'_{t-\tau}] = 0$; and (iv) for all (t, s), $\mathbb{E}[\epsilon_t^y u'_s] = 0$, $\mathbb{E}[\epsilon_t^y x'_s] = 0$, $\mathbb{E}[\epsilon_t^x x'_s] = 0$, and $\mathbb{E}[\epsilon_t^x u'_s] = 0$.

Assumption B. The matrix polynomial $A(L)$ is stable; p_y and r_x are known.

Assumption C. Let $v_t^y \sim WN(0, \Sigma_v^y)$ and $v_t^x \sim WN(0, \Sigma_v^x)$, where Σ_v^y and Σ_v^x are full-rank diagonal matrices with $\mathbb{E}[v_t^y v_s^{x\prime}] = 0$ for all (t, s). The measurement errors are

$$\Phi^y(L)M^y\epsilon_t^y = \Theta^y(L)v_t^y,$$
$$\Phi^x(L)M^x\epsilon_t^x = \Theta^x(L)v_t^x,$$

where $\Phi^y(L)$, $\Theta^y(L)$, $\Phi^x(L)$, and $\Theta^x(L)$ are diagonal polynomial matrices of known orders p_ϵ^y, q_ϵ^y, p_ϵ^x, and q_ϵ^x, respectively. Furthermore, $\Phi^y(L)$ and $\Phi^x(L)$ are stable; $\Phi^y(L)$ and $\Theta^y(L)$ as well as $\Phi^x(L)$ and $\Theta^x(L)$ are left co-prime.

Under Assumption A, the vector white noise process u_t is orthogonal to the measurement errors ϵ_t^y and ϵ_t^x, and the latter are "classical" in the sense that ϵ_t^y is uncorrelated with y_t, and ϵ_t^x is uncorrelated with x_t. Moreover, u_t is uncorrelated with the current and past values of x_t, which is the usual exogeneity condition. Under Assumption B, all roots of $\det A(z)$ $(z \in \mathbb{C})$ are outside the closed unit disk. Under Assumption C, each element in $M^y\epsilon_t^y$ is a scalar ARMA$(p_\epsilon^y, q_\epsilon^y)$ process, while each element of $M^x\epsilon_t^x$ is a scalar ARMA$(p_\epsilon^x, q_\epsilon^x)$ process. These are mutually uncorrelated stationary processes. Multivariate processes complicate the analysis and are less likely to be empirically relevant. When Assumptions A, B, and C hold, $(Y'_t, X'_t)'$ is weakly stationary (see, e.g., Hannan, 1970).

Define the parameters in Model-y and the measurement error processes as:

$$\theta_{AR} \equiv \mathrm{vec}(A_1, \ldots, A_{p_y}, B_0, \ldots, B_{r_x}, \Phi_1^y, \ldots, \Phi_{p_\epsilon^y}^y, \Phi_1^x, \ldots, \Phi_{p_\epsilon^x}^x),$$
$$\theta_{MA} \equiv \mathrm{vec}(\Sigma_u, \Theta_1^y, \ldots, \Theta_{q_\epsilon^y}^y, \Theta_1^x, \ldots, \Theta_{q_\epsilon^x}^x, \Sigma_v^y, \Sigma_v^x).$$

The objective of the exercise is to identify θ_{AR} and θ_{MA}.

Consider first identification of the parameters of Model-y in the absence of measurement error. Let $\vec{x}_t = (x'_t, \ldots, x'_{t-r_x})'$. If lagged values of y_t are absent from Model-y, then $\vec{B}$ is identified if $\mathbb{E}[u_t\vec{x}'_t] = 0$ and $\mathbb{E}[\vec{x}_t\vec{x}'_t]$ is full rank, no matter what the dynamics of the unobservables u_t are. With lagged y_t's in the model, $\vec{A}$ and $\vec{B}$ are identified if for all t, $\mathbb{E}[u_t x'_{t-\tau}] = 0$ for all $\tau \geqslant 0$ and $\mathbb{E}[u_t u'_s] = 0$ for all $s < t$.[1] Since u_t is covariance stationary under Assumption A, time invariance of the second moments permits identification of Σ_u from the covariance structure of the residuals $A(L)y_t - B(L)x_t$.

With measurement errors, the model in terms of observables is

$$A(L)Y_t = B(L)X_t + u_t + A(L)\epsilon_t^y - B(L)\epsilon_t^x.$$

The three terms in $u_t + A(L)\epsilon_t^y - B(L)\epsilon_t^x$ are not separately observable. Identification of dynamic models in the presence of measurement errors is a nontrivial problem because the composite error term $u_t + A(L)\epsilon_t^y - B(L)\epsilon_t^x$ will in general fail to be a finite order moving average process, and common factor dynamics cannot be ruled out. This violates classical conditions for identification of vector autoregressive moving average (VARMA) processes derived in Hannan (1969), which require co-primeness even in the absence of exogenous regressors. Furthermore, the composite error is necessarily correlated with Y_t and/or X_t, violating the exogeneity conditions used, for example, in Hannan (1971) and Hannan, Dunsmuir, and Deistler (1980).

2.2. Related Literature

In a seminal paper, Reiersøl (1950) showed that identification in a single-regressor static model is possible only if (Y_t, X_t) are not jointly normally distributed. However, as discussed in Maravall and Aigner (1977) and Maravall (1979), a static model that is unidentified in the presence of serially uncorrelated measurement errors could be identifiable when the model has a dynamic structure. Intuitively, if ϵ_t^y and ϵ_t^x are white noise but the model has dynamics, $\vec{A}$ and $\vec{B}$ can, in principle, be identified by the suitable choice of instruments such as lags of Y_t and X_t. Aigner, Hsiao, Kapteyn, and Wansbeek (1984, p. 1324) pointed out that consistent parameter estimation is also possible with repeated observations on y_t, or strong a priori information.[2] Aigner et al. (p.1377) also noted that serially correlated measurement errors will complicate the identification problem, but the econometrics literature on this issue is quite sparse. Mcdonald and Darroch (1983) considered single-equation estimation of $\vec{A}$ and $\vec{B}$ with correlated measurement errors but did not consider separate identification of the equation error and the measurement noise.

There is, however, a reasonably large literature in control theory studying the identification of noise-corrupted "multiple input-multiple output" (MIMO) models; see Söderström (2007) for a survey. By defining a new, latent endogenous variable $y_t^* \equiv y_t - A(L)^{-1}u_t$, Model-y can be written as a MIMO model,

$$y_t^* = B^*(L)x_t,$$

where $B^*(L) \equiv A(L)^{-1}B(L)$. In a noise-corrupted analysis, x_t is referred to as the input vector and y_t is the output vector. The observed variables are $X_t = x_t + \epsilon_t^x$ and $Y_t = y_t^* + \epsilon_t^{y*}$, where the new measurement error $\epsilon_t^{y*} \equiv A(L)^{-1}u_t + \epsilon_t^y$ amalgamates the shocks u_t and the observational errors ϵ_t^y; see Solo (1986).

While a specification with no error term on the right-hand side is uncommon in econometric analysis, y_t has a MIMO representation and results from that

literature are relevant. The key result, shown for the case of a single input and single output (also known as SISO systems), is given in Anderson and Deistler (1984). They show that the transfer function $B^*(L)$ (or the Wold representation) of such systems is impossible to identify nonparametrically. This "essential nonidentifiability" holds even if the system is known to be causal. Identification is possible only if some structure is put on the measurement errors. One alternative is to treat ϵ_t^x as a dynamic factor, and $B(L)$ are the factor loadings. A second alternative is to parametrically specify the covariance structure of the residuals. Anderson and Deistler assume that the input and the measurement errors are autoregressive moving average (ARMA) processes and obtain conditions for identification from the second moments of the data without restricting the lag orders a priori. The multivariate extension to stationary and causal MIMO-EIV systems is considered in Nowak (1992, 1993).

The common feature of these results is that they all require a parametric model for x_t. For instance, Nowak (1992, 1993) assumes that x_t are VARMA processes. Parametric identification when the data are correctly observed dates back to Koopmans (1953) and Fisher (1963). Measurement errors impose the additional requirement that the model for x_t has to be specified. This is restrictive if x_t's are thought to be determined outside of the system defined by Model-y.

An alternative to MIMO models is state space systems that explicitly allow for measurement noise. Sargent (1989) shows how the Kalman filter can be used when the variables in a dynamic model are subject to autoregressive measurement errors. But, much like the MIMO models, the process for every variable in the model needs to be completely specified. There are in fact relatively few identification results for which the exogenous variables are not explicitly modeled. Solo (1986) considers an Autoregressive-Moving Average Models with Exogenous Variables (ARMAX) model without imposing assumptions on the exogenous variable and shows that the identification conditions agree with those derived by Maravall (1979) with parametric assumptions on the exogenous process. Our approach is similar in spirit to Solo (1986), with several considerations in mind.

First, if the primitive shocks of a structural VAR were to be subsequently recovered from the reduced-form errors u_t, then u_t needs to be separated from the measurement noise ϵ_t^y. Considering the composite error ϵ_t^{y*} as in MIMO systems is not enough. Second, we restrict attention to VARX models because VARMAX models are rarely used in economic applications. Requiring u_t to be white noise rules out observational equivalence that may arise from common factors between $A(L)$ and u_t. Third, we do not specify the distribution of the measurement errors or of the exogenous variables. The only restriction is that the measurement errors are scalar ARMA processes. Parametric assumptions on the measurement errors are necessary in view of the result of Anderson and Deistler (1984). One can think of this *limited information* approach as a semiparametric setup in which the distribution of the measurement errors and the exogenous variables is the nonparametric part of the model.

3. ORDER CONDITION FOR MISMEASURED VARX MODELS

3.1. General Results

Our point of departure is that when the measurement errors are serially correlated, the dynamic representation for Y_t will, in general, be different from that for y_t because Y_t is the sum of two latent dynamic processes. In a univariate setting, Granger and Morris (1976) showed that if y_t is ARMA(p_y, q_y) and ϵ_t^y is ARMA$(p_\epsilon^y, q_\epsilon^y)$, then $Y_t = y_t + \epsilon_t^y$ is ARMA(p, q) where $p \leqslant p_y + p_\epsilon^y$ and $q \leqslant \max\{p_y + q_\epsilon^y, p_\epsilon^y + q_y\}$.[3] The following lemma provides a general result for vector processes.

LEMMA 1. *Suppose Assumptions A, B, and C hold. Then the observed data can be represented as*

$$\mathcal{A}(L)Y_t = \mathcal{B}(L)X_t + U_t, \qquad \textbf{(Model-Y)}$$

where $\mathcal{A}(L)$ and $\mathcal{B}(L)$ are polynomials of orders p_Y and r_X whose coefficients only depend on θ_{AR}, U_t is a $VMA(q_U)$ error process that depends on both θ_{AR} and θ_{MA}, and

$$p_Y = p_y + \bar{p}$$
$$q_U \leqslant \bar{p} + \max\{0, [p_y + q_\epsilon^y - p_\epsilon^y] \cdot 1(n_\epsilon^y > 0), [r_x + q_\epsilon^x - p_\epsilon^x] \cdot 1(n_\epsilon^x > 0)\}$$
$$r_X = r_x + \bar{p},$$

for some $\bar{p} \leqslant n_\epsilon^y p_\epsilon^y + n_\epsilon^x p_\epsilon^x$. Moreover, $\mathbb{E}(U_t) = 0$, $\mathbb{E}(U_t Y_{t-\tau}') = 0$ for all $\tau > \tau_Y$, and $\mathbb{E}(U_t X_{t-\tau}') = 0$ for all $\tau > \tau_X$, where

$$\tau_Y \leqslant \bar{p} + \max\{0, [p_y + q_\epsilon^y - p_\epsilon^y] \cdot 1(n_\epsilon^y > 0)\},$$
$$\tau_X \leqslant \bar{p} + \max\{0, [r_x + q_\epsilon^x - p_\epsilon^x] \cdot 1(n_\epsilon^x > 0)\}.$$

Lemma 1 shows that Model-Y has richer dynamics than Model-y. The autoregressive part of Model-Y is of order $p_Y \geqslant p_y$. Furthermore, Y_t has a moving average (MA) component even if none is present in y_t and $\Gamma_U(j) \equiv \mathbb{E}(U_t U_{t-j}') \neq 0$ for all $j = 1, \ldots, q_U$. In consequence, the usual orthogonality conditions $\mathbb{E}(U_t Y_{t-\tau}') = 0$ and $\mathbb{E}(U_t X_{t-\tau}) = 0$ no longer hold for all $\tau > 0$. Instead, $\mathbb{E}(U_t Y_{t-\tau}') = 0$ only for $\tau > \tau_Y$, and $\mathbb{E}(U_t X_{t-\tau}') = 0$ only for $\tau > \tau_X$. The composite error U_t is a moving average process of order

$$q_U = \max\left[\tau_Y, \tau_X\right].$$

In general, $\tau_Y \neq \tau_X$ because p_y and r_x may differ, and so can the dynamics of the measurement errors in Y_t and X_t. Even though under Assumption C the measurement errors are mutually uncorrelated ARMA processes, U_t will follow a vector MA process whenever one of the polynomials $A(L)$ and $B(L)$ is not diagonal.

Although p_Y, r_X, and q_U can be identified from the data, point identification of $p_y, r_x, p_\epsilon^y, q_\epsilon^y, p_\epsilon^x, q_\epsilon^x$ necessitates a priori restrictions. For this reason, Assumptions B and C treat these parameters as fixed. Interestingly, Lemma 1 implies that

the true lags p_y and r_x are always set identified in the sense that $0 \leqslant p_y \leqslant p_Y$ and $0 \leqslant r_x \leqslant r_X$.

From the data Y_t and X_t and the assumed lag orders, the most we can uncover are the parameters in Model-Y:

$$\pi_{AR} \equiv \text{vec}(\mathcal{A}_1, \ldots, \mathcal{A}_{p_Y}, \mathcal{B}_0, \ldots, \mathcal{B}_{r_X}),$$
$$\pi_{MA} \equiv \text{vec}(\Gamma_U(0), \ldots, \Gamma_U(q_U)).$$

The question is whether knowing the parameters $\pi' \equiv (\pi'_{AR}, \pi'_{MA})$ enables identification of the parameters $\theta' \equiv (\theta'_{AR}, \theta'_{MA})$ without fully specifying the DGP for x_t. A key to our order condition for identification is to exploit the specific structure implied by Lemma 1:

$$\pi_{AR} = \pi_{AR}(\theta_{AR}),$$
$$\pi_{MA} = \pi_{MA}(\theta_{AR}, \theta_{MA}).$$

This observation leads to the following order conditions.

PROPOSITION 1. *Suppose Assumptions A, B, and C hold. Assume that $\pi(\theta)$ is continuously differentiable and that the rank of $\frac{\partial \pi}{\partial \theta}$ remains constant in a neighborhood of θ. The necessary order conditions for local identification of θ_{AR} and θ_{MA} jointly are*

$$AR: \quad \dim(\theta_{AR}) \leqslant \dim(\pi_{AR}) + \dim(\pi_{MA}) - \dim(\theta_{MA}),$$
$$MA: \quad \dim(\theta_{MA}) \leqslant \dim(\pi_{MA}).$$

The results exploit the triangular structure of the problem and are new to the literature. Since θ_{MA} appears in π_{MA} and not in π_{AR}, θ_{MA} can only be identified from π_{MA}, and the MA condition follows. If θ_{MA} is overidentified, the additional information can be used to identify θ_{AR}, leading to the AR condition. These order restrictions are more stringent than just requiring that $\dim(\theta_{AR}) + \dim(\theta_{MA}) \leqslant \dim(\pi_{AR}) + \dim(\pi_{MA})$. Proposition 1 can be made more explicit upon substituting in the dimensions of the parameters:

$$\dim(\theta_{AR}) = n_y^2 \cdot p_y + n_y \cdot n_x \cdot (r_x + 1) + n_\epsilon^y \cdot p_\epsilon^y + n_\epsilon^x \cdot p_\epsilon^x,$$
$$\dim(\theta_{MA}) = \frac{n_y(n_y+1)}{2} + n_\epsilon^y + n_\epsilon^y \cdot q_\epsilon^y + n_\epsilon^x + n_\epsilon^x \cdot q_\epsilon^x,$$
$$\dim(\pi_{AR}) = n_y^2 \cdot p_Y + n_y \cdot n_x \cdot (r_X + 1),$$
$$\dim(\pi_{MA}) = \frac{n_y(n_y+1)}{2} + n_y^2 \cdot q_U.$$

The AR order condition holds whenever the following two conditions are satisfied:

$$n_y^2(p_Y - p_y) \geqslant n_\epsilon^y p_\epsilon^y \quad \text{and} \quad n_y n_x (r_X - r_x) \geqslant n_x^\epsilon p_\epsilon^x.$$

Although simple instrumental variables arguments can be used to establish that $\mathcal{A}_j$ and $\mathcal{B}_j$ are identified, $\mathcal{A}_j \neq A_j$ and $\mathcal{B}_k \neq B_k$ in the presence of measurement errors. The above conditions ensure that the dynamics of Model-Y are sufficiently

richer than those of Model-y to disentangle the effects of lagged y_t and x_t from the measurement noise. The required difference in dynamics depends on the number of variables measured with errors. The conditions are, however, stronger than necessary.

The MA order condition requires that

$$n_\epsilon^y(q_\epsilon^y+1)+n_\epsilon^x(q_\epsilon^x+1)$$
$$\leqslant n_y^2\left[n_\epsilon^y p_\epsilon^y+n_\epsilon^x p_\epsilon^x+\max\{0,[p_y+q_\epsilon^y-p_\epsilon^y]\cdot 1(n_\epsilon^y>0),[r_x+q_\epsilon^x-p_\epsilon^x]\cdot 1(n_\epsilon^x>0)\}\right],$$

which evidently depends on the lag structure of the model (as given by p_y and r_x) and the dynamics of the measurement errors (as given by $p_\epsilon^y, q_\epsilon^y, p_\epsilon^x, q_\epsilon^x$). If the condition fails, then θ_{MA} cannot be recovered from the variance and the q_U autocovariances of U_t collected in π_{MA}. As our subsequent examples will show, the MA order condition is more likely to be violated than the AR condition, even in simple models. It is therefore possible that $\dim(\theta_{AR})+\dim(\theta_{MA})\leqslant \dim(\pi_{AR})+\dim(\pi_{MA})$ and yet $\dim(\theta_{MA})>\dim(\pi_{MA})$.

The results of Proposition 1 have implications for identification of the structural parameters in the simultaneous equations system,

$$\mathbb{A}_0 y_t=\mathbb{A}_1 y_{t-1}+\cdots+\mathbb{A}_{p_y} y_{t-p_y}+\mathbb{B}_0 x_t+\cdots+\mathbb{B}_{r_x} x_{t-r_x}+\mathbb{U}_t,$$

with $\mathbb{A}_0\neq I$, and whose reduced form is Model-y. Hsiao (1979) develops exclusion restrictions for the case of white noise measurement errors. Because identification of Model-y is necessary for identification of $(\mathbb{A}_1,\ldots,\mathbb{A}_{p_y})$, $(\mathbb{B}_0,\ldots,\mathbb{B}_{r_x})$, and $\Sigma_{\mathbb{U}}$, the order conditions presented above are also necessary for identification of the structural parameters.

PROPOSITION 2. *Suppose Assumptions A, B, and C hold. Assume that $\pi(\theta)$ is continuously differentiable, and that the ranks of $\frac{\partial\pi}{\partial\theta}$, $\frac{\partial\pi}{\partial\theta_{AR}}$, and $\frac{\partial\pi}{\partial\theta_{MA}}$ remain constant in a neighborhood of θ. The necessary order conditions for local identification of θ_{AR} alone and of θ_{MA} alone are, respectively,*

$$AR: \quad \dim(\theta_{AR})\leqslant\dim(\pi_{AR})+\dim(\pi_{MA}),$$
$$MA: \quad \dim(\theta_{MA})\leqslant\dim(\pi_{MA}).$$

If one is interested in θ_{MA} alone, the necessary order condition for identification is simply $\dim(\theta_{MA})\leqslant\dim(\pi_{MA})$. This is hardly surprising, since θ_{MA} can only be recovered from π_{MA}. However, θ_{AR} can potentially be recovered from all information available. This leads to the weaker order condition: $\dim(\theta_{AR})\leqslant \dim(\pi_{AR})+\dim(\pi_{MA})$.

3.2. Two Special Cases: ARX Models and VAR

This subsection studies the implications of Proposition 1 for important special cases: scalar ARX models and pure VARs. We focus on the joint identification of θ_{AR} and θ_{MA}.

Case 1 (**ARX**(p_y, r_x) **with** $n_x = 1, n_\epsilon^y = 0, n_\epsilon^x = 1$). When x_t is scalar, Model-y is

$$y_t = A_1 y_{t-1} + \cdots + A_{p_y} y_{t-p_y} + B_0 x_t + \cdots + B_{r_x} x_{t-r_x} + u_t, \qquad u_t \sim WN(0, \Sigma_u). \tag{1}$$

If y_t is correctly measured but x_t is measured with errors, Model-Y is $ARMA(p_Y, q_U, r_X)$ with $p_Y = p_y + p_\epsilon^x$, $q_U = p_\epsilon^x + \max\{0, r_x + q_\epsilon^x - p_\epsilon^x\}$, and $r_X = r_x + p_\epsilon^x$.[4] The model is quite commonly used, as latent variables such as expected inflation, permanent income, wealth, and marginal Q play important roles in economic analysis. For example, the Phillips curve has x_t being the output gap and y_t being inflation. As the output gap is latent, its proxy can be thought of as a mismeasured variant of x_t. If this proxy is constructed by filtering an observed variable, it is likely that ϵ_t^x will be serially correlated.

Applying Lemma 1, we have $\dim(\theta_{AR}) = p_y + (r_x + 1) + p_\epsilon^x$, $\dim(\theta_{MA}) = 2 + q_\epsilon^x$, $\dim(\pi_{AR}) = p_y + (r_x + 1) + 2p_\epsilon^x$, and $\dim(\pi_{MA}) = 1 + p_\epsilon^x + \max\{0, r_x + q_\epsilon^x - p_\epsilon^x\}$. Consider the cases in Table 1.

When there are no lags of the dependent variable (hence $p_y = 0$) and $r_x = 0$, Model-y is static. The autoregressive dynamics in Model-Y alone are not sufficient for identification, since the residual autocovariances need to provide information for recovering the components of θ_{MA}. The MA order condition fails except when ϵ_t^x is AR(1). The condition also fails if $p_y = 1$ or $p_y = 2$, with $r_x = 0$ because there are fewer nonzero lagged autocovariances of U_t than are needed to recover all MA parameters. Identification is more likely to obtain when Model-y is dynamic. As indicated from Table 1, the MA order condition holds for any value of p_y whenever $r_x \geqslant 1$. The identification conditions can be formalized as follows.

TABLE 1. *ARX(p_y, r_x) with $n_\epsilon^y = 0$.*

ϵ_t^x	p_Y	q_U	r_X	AR	MA	p_Y	q_U	r_X	AR	MA
	$p_y = 0, r_x = 0$			Order Condition		$p_y = 1, r_x = 0$			Order Condition	
WN	0	0	0	pass	fail	1	0	0	pass	fail
MA(1)	0	1	0	pass	fail	1	1	0	pass	fail
AR(1)	1	1	1	pass	pass	2	1	1	pass	pass
ARMA(1,1)	1	1	1	pass	fail	2	1	1	pass	fail
	$p_y = 2, r_x = 0$			Order Condition		$p_y = 1, r_x = 1$			Order Condition	
WN	2	0	0	pass	fail	1	1	1	pass	pass
MA(1)	2	0	0	pass	fail	1	2	1	pass	pass
AR(1)	3	1	1	pass	pass	2	1	2	pass	pass
ARMA(1,1)	3	1	1	pass	fail	2	2	2	pass	pass

COROLLARY 1. *Consider an ARX(p_y, r_x) model in which the scalar variable x_t is observed with error, and let all the assumptions of Proposition 1 hold. The AR order condition for joint local identification of θ_{AR} and θ_{MA} always holds; the MA order condition for joint local identification of θ_{AR} and θ_{MA} requires* $\max\{p_\epsilon^x - q_\epsilon^x, r_x\} \geqslant 1$.

The corollary makes clear that the AR order condition always holds, but the necessary condition for identifying θ_{MA} can fail. More precisely, $\dim(\pi_{AR}) + \dim(\pi_{MA})$ can exceed $\dim(\theta_{AR}) + \dim(\theta_{MA})$, and yet the order condition for θ_{MA} fails. In the case of an ARX(1,0) model with ARMA(1,1) measurement error, for example, $\dim(\pi) = \dim(\theta) = 6$, but $\dim(\pi_{MA}) - \dim(\theta_{MA}) = -1$. Although the total number of parameters in Model-Y and Model-y are equal, this total number is uninformative and misleading about the identifiability of Model-y. The issue is that π_{AR} tends to overidentify θ_{AR}, but the overidentifying conditions are of no help in identifying θ_{MA}. For this reason, the MA and AR order conditions need to be considered jointly.

Case 2 (**ARX**(p_y, r_x) **with** $n_x = n_\epsilon^y = n_\epsilon^x = 1$). When x_t and y_t are both mismeasured, the dynamic structures of y_t and ϵ_t^y will both play a role in identification of θ_{MA}. This is illustrated for the case when $\epsilon_t^y \sim AR(1)$ in Table 2.

COROLLARY 2. *Consider an ARX(p_y, r_x) model in which the scalar variables y_t and x_t are observed with errors and ϵ_t^y is an AR(1) process, and let all the assumptions of Proposition 1 hold. The AR order condition for joint local identification of θ_{AR} and θ_{MA} always holds; the MA order condition for joint local identification of θ_{AR} and θ_{MA} requires* $\max\{p_\epsilon^x - q_\epsilon^x, p_y - 1 + p_\epsilon^x - q_\epsilon^x, r_x\} \geqslant 1$.

The order condition for identifying θ_{MA} holds with white noise but not with MA(1) errors in x_t. This is because q_U is the same in both cases, but the lags of U_t provide no additional information to identify the MA(1) parameter in the measurement error process. In such a case, the order condition requires either that r_x be greater than one or that the measurement error be an ARMA(1,1).

Case 3 (**VAR**(p_y)). Vector regressions are widely used in macroeconomic analysis. In the presence of serially correlated measurement errors, a VAR(p_y) model becomes a VARMA(p_Y, q_U).

TABLE 2. *$ARX(p_y, r_x)$ with $(p_\epsilon^y, q_\epsilon^y) = (1, 0)$.*

	$p_y = 2, r_x = 0$			Order Condition		$p_y = 0, r_x = 2$			Order Condition	
ϵ_t^x	p_Y	q_U	r_X	AR	MA	p_Y	q_U	r_X	AR	MA
WN	3	2	1	pass	pass	1	3	3	pass	pass
MA(1)	3	2	1	pass	fail	1	4	3	pass	pass
AR(1)	4	3	2	pass	pass	2	4	4	pass	pass
ARMA(1,1)	4	3	2	pass	pass	2	4	4	pass	pass

COROLLARY 3. *Consider a VAR(p_y) process in which some (possibly all) components of the n_y vector y_t are observed with error, and let all the assumptions of Proposition 1 hold. The AR order condition for joint local identification of θ_{AR} and θ_{MA} is always satisfied; the MA order condition for joint local identification of θ_{AR} and θ_{MA} requires that*

$$\max\left\{\frac{n_y^2}{n_\epsilon^y}p_y+\left(\frac{n_y^2}{n_\epsilon^y}-1\right)q_\epsilon^y+\frac{n_y^2}{n_\epsilon^y}(n_\epsilon^y-1)p_\epsilon^y\ ,\ n_y^2p_\epsilon^y-q_\epsilon^y\right\}\geqslant 1.$$

Note that since $n_\epsilon^y \leqslant n_y$, the first term inside the max operator is always greater than or equal to p_y. Thus, if $p_y \geqslant 1$, the order condition for identification of θ_{MA} always holds. This comes from the fact that our measurement errors are univariate ARMA and not VARMA processes. Thus, while $\dim(\pi_{MA}) = O(n_y^2)$, the number of measurement error parameters is $O(n_\epsilon^y)$. The order conditions for VAR(p_y) models with $p_y \geqslant 1$ are thus satisfied for all n_y. Interestingly, the result implies that seemingly unrelated systems might make it possible to identify parameters that are otherwise nonidentifiable in a single-equation setup. In the case of an $AR(p_y)$ model (i.e. $n_y = 1$), the order condition for identifying θ_{MA} simplifies to $\max\{p_y, p_y^\epsilon - q_\epsilon^y\} \geqslant 1$.

4. RANK CONDITION FOR MISMEASURED VARX MODELS

4.1. General Result

The preceding analysis has focused on conditions that are necessary for identifying Model-y. The order conditions are not sufficient because Model-Y may not be identified. Showing that Model-Y is identifiable is not trivial for two reasons. First, the covariates X_t are no longer strongly exogenous. Second, Model-Y has a VARMAX structure that is susceptible to the presence of common factors that render the parameters in Model-Y nonidentifiable. However, deriving a canonical form for Model-Y that is free of common factors is often not a trivial task (see, e.g., Hannan, 1971).

The following lemma gives sufficient conditions for identification of Model-Y that do not require canonical forms. Let $\Gamma_{YY}(j) \equiv \mathbb{E}(Y_tY'_{t-j})$, $\Gamma_{XX}(j) \equiv \mathbb{E}(X_tX'_{t-j})$, $\Gamma_{XY}(j) \equiv \mathbb{E}(X_tY'_{t-j})$, and $\Gamma_{YX}(j) \equiv \mathbb{E}(Y_tX'_{t-j})$ be the autocovariances and cross-covariances of Y_t and X_t. Define

$$H_{YY} \equiv \Big(\Gamma_{YY}(j-i+\tau_Y)\Big)_{1\leqslant i,j\leqslant p_Y},$$

$$H_{YX} \equiv \Big(\Gamma_{YX}(j-i+\tau_X)\Big)_{1\leqslant i\leqslant p_Y,1\leqslant j\leqslant r_X+1},$$

$$H_{XY} \equiv \Big(\Gamma_{XY}(j-i+\tau_Y+1)\Big)_{1\leqslant i\leqslant r_X+1,1\leqslant j\leqslant p_Y},$$

$$H_{XX} \equiv \Big(\Gamma_{XX}(j-i+\tau_X+1)\Big)_{1\leqslant i,j\leqslant r_X+1}.$$

LEMMA 2. *Suppose Assumptions A, B, and C hold. Then the parameters π_{AR} and π_{MA} of Model-Y are jointly globally identifiable if and only if*

$$rank \begin{pmatrix} H_{YY} & H_{YX} \\ H_{XY} & H_{XX} \end{pmatrix} = n_y p_Y + n_x(r_X + 1).$$

The key insight of Lemma 2 is that the orthogonality restrictions $\mathbb{E}(U_t Y'_{t-\tau}) = 0$ for $\tau > \tau_Y$ and $\mathbb{E}(U_t X'_{t-\tau}) = 0$ for $\tau > \tau_X$ put enough structure to guarantee identification of Model-Y. The result requires, however, that there be enough information in the autocovariances and cross-covariances of Y_t and X_t to satisfy the rank condition. Since Model-Y is linear in π_{AR}, the rank condition is both necessary and sufficient, and the identification result is global.

As an example, consider an ARX(1,0) model in which both y_t and x_t are subject to white noise measurement errors. In that case, $Y_t = A_1 Y_{t-1} + B_0 X_t + u_t + v_t^y - A_1 v_{t-1}^y - B_0 v_t^x$ is an ARMAX(1,1,0) with $\tau_Y = 1$ and $\tau_X = 0$. Now if the latent covariates x_t are themselves a white noise process, the matrix in Lemma 2 reduces to

$$\begin{pmatrix} \Gamma_{YY}(1) & \Gamma_{YX}(0) \\ \Gamma_{XY}(2) & \Gamma_{XX}(1) \end{pmatrix} = \begin{pmatrix} \Gamma_{yy}(1) & \Gamma_{yx}(0) \\ 0 & 0 \end{pmatrix},$$

which clearly fails to satisfy the rank condition. Thus, we recover the well-known result that identification of EIV systems fails when the covariates are white noise. Lemma 2 is an extension of the results in Solo (1986) and Hsiao (1979) to multivariate systems with independent ARMA measurement errors.

Once the parameters of Model-Y are identified, we can proceed with the identification of Model-y. Sufficient rank conditions for θ_{AR} and θ_{MA} to be jointly locally identified from π_{AR} and π_{MA} are summarized in the following proposition.

PROPOSITION 3. *Suppose Assumptions A, B, and C hold. Assume in addition that $\pi(\theta)$ is continuously differentiable and that the rank condition of Lemma 2 holds. Then the sufficient rank conditions for local identification of θ_{AR} and θ_{MA} jointly are*

$$AR: \quad rank \left(\frac{\partial \pi_{AR}(\theta_{AR})}{\partial \theta_{AR}} \right) = \dim(\theta_{AR}),$$

$$MA: \quad rank \left(\frac{\partial \pi_{MA}(\theta_{AR}, \theta_{MA})}{\partial \theta_{MA}} \right) = \dim(\theta_{MA}).$$

The result of Proposition 3 exploits the triangular structure of the system. The practical implication of Proposition 3 is that identification of the autoregressive parameters precedes identification of the variances and moving average parameters of the measurement error process. Even if the AR rank condition is satisfied, the MA rank for condition may fail.

The following result provides separate rank conditions for partial identification of either θ_{AR} or θ_{MA} alone.

PROPOSITION 4. *Suppose Assumptions A, B, and C hold. Assume in addition that $\pi(\theta)$ is continuously differentiable and that the rank condition of Lemma 2 holds. Then the sufficient rank conditions for local identification of θ_{AR} alone and of θ_{MA} alone are, respectively,*

$$AR: \quad rank\left(\frac{\partial \pi_{AR}(\theta_{AR})}{\partial \theta_{AR}}\right) = \dim(\theta_{AR})$$

$$MA: \quad rank\left(\frac{\partial \pi(\theta_{AR},\theta_{MA})}{\partial \theta}\right) = rank\left(\frac{\partial \pi(\theta_{AR},\theta_{MA})}{\partial \theta_{AR}}\right) + \dim(\theta_{MA}).$$

Since π_{AR} is a function of θ_{AR} alone, the same rank condition as in Proposition 3 obtains. The condition is sufficient to identify θ_{AR} irrespective of θ_{MA}. Thus, it is possible that θ_{AR} is identified and yet θ_{MA} fails to be so. The rank condition for θ_{MA} now takes into account that θ_{AR} may no longer be identified.

4.2. Example

There are several ways in which the rank conditions may be checked. In simple models in which the mapping from θ to π is easy to compute analytically, evaluating the rank conditions is straightforward, though formal testing is not. In general, however, numerical approaches may be easier to implement. To illustrate how the rank condition presented in Proposition 3 is to be used, consider the model $y_t = A_1 y_{t-1} + B_0 x_t + B_1 x_{t-1} + u_t$ with $u_t \sim WN(0, \sigma_u^2)$. While y_t is correctly measured, we only observe $X_t = x_t + \epsilon_t$ where ϵ_t is ARMA(1,1), $\epsilon_t = \Phi\epsilon_{t-1} + v_t + \Theta v_{t-1}$, $v_t \sim WN(0, \sigma_v^2)$. The parameters $\theta' = (\theta'_{AR}, \theta'_{MA})$ of Model-y are

$$\theta_{AR} \equiv (A_1, B_0, B_1, \Phi)' \quad \text{and} \quad \theta_{MA} \equiv (\sigma_u^2, \Theta, \sigma_v^2)'.$$

The observed variables follow Model-Y, which is an ARMAX(2,2,2) given by

$$\begin{aligned} y_t &= (\Phi + A_1)y_{t-1} - \Phi A_1 y_{t-2} + B_0 X_t + (B_1 - \Phi B_0)X_{t-1} - \Phi B_1 X_{t-2} + U_t \\ &= \mathcal{A}_1 y_{t-1} + \mathcal{A}_2 y_{t-2} + \mathcal{B}_0 X_t + \mathcal{B}_1 X_{t-1} + \mathcal{B}_2 X_{t-2} + U_t. \end{aligned}$$

The parameters $\pi' = (\pi'_{AR}, \pi'_{MA})$ in Model-Y are

$$\pi_{AR} \equiv (\mathcal{A}_1, \mathcal{A}_2, \mathcal{B}_0, \mathcal{B}_1, \mathcal{B}_2)',$$
$$\pi_{MA} \equiv (\Gamma_U(0), \Gamma_U(1), \Gamma_U(2))',$$

where $\Gamma_U(0) = (1+\Phi^2)\sigma_u^2 + [B_0^2 + (B_1+\Theta)^2 + B_1^2\Theta^2]\sigma_v^2$, $\Gamma_U(1) = -\Phi\sigma_u^2 + (B_0 + B_1\Theta)(B_1+\Theta)\sigma_v^2$, and $\Gamma_U(2) = B_0 B_1 \Theta \sigma_v^2$.

First consider the identifiability of π. For this, let $W_t \equiv (y_{t-1}, y_{t-2}, X_t, X_{t-1}, X_{t-2})'$, $Z_t \equiv (y_{t-1}, y_{t-2}, X_{t-3}, X_{t-4}, X_{t-5})'$. Then Model-Y can be written as $y_t = W_t'\pi_{AR} + U_t$ with $\mathbb{E}[Z_t U_t] = 0$. Clearly, π_{AR} can be identified provided rank $\mathbb{E}[Z_t W_t'] = 5$. This is precisely the rank condition in Lemma 2. Once π_{AR} is identified, π_{MA} is determined from the autocovariances of the residuals $y_t - W_t'\pi_{AR}$.

Now consider the identification of $\theta' = (\theta'_{AR}, \theta'_{MA})$ from π, and note first that since $\mathcal{B}_0 = B_0$, B_0 is identified whether or not any of the other components of θ are identifiable. However, we need to identify the nuisance parameter Φ to identify (A_1, B_1). If σ_u^2 is of interest, then we also need to identify the nuisance parameters (Θ, σ_v^2).

The sufficient AR condition in Proposition 3 depends on

$$\frac{\partial \pi_{AR}}{\partial \theta_{AR}} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ -\Phi & 0 & 0 & -A_1 \\ 0 & 1 & 0 & 0 \\ 0 & -\Phi & 1 & -B_0 \\ 0 & 0 & -\Phi & -B_1 \end{pmatrix}.$$

Observe that rank $\left(\frac{\partial \pi_{AR}}{\partial \theta_{AR}}\right) = 4$ except when $B_0 = 1$ and $A_1 = B_1 = \Phi$, in which case the rank is 3 and identification of θ_{AR} from π_{AR} fails. When measurement errors are present, identification of the components of $A(L)$ and $B(L)$ requires that those polynomials be co-prime. This is unlike the case in which all the variables are correctly observed.

The sufficient MA condition depends on

$$\frac{\partial \pi_{MA}}{\theta_{MA}} = \begin{pmatrix} (1+\Phi^2) & 2[B_1 + \Theta + B_1^2\Theta]\sigma_v^2 & B_0^2 + (B_1+\Theta)^2 + B_1^2\Theta^2 \\ -\Phi & (B_0 + B_1^2 + 2B_1\Theta)\sigma_v^2 & (B_0 + B_1\Theta)(B_1+\Theta) \\ 0 & B_0 B_1 \sigma_v^2 & B_0 B_1 \Theta \end{pmatrix},$$

whose determinant

$$\det\left(\frac{\partial \pi_{MA}}{\theta_{MA}}\right) = -B_0 B_1 \sigma_v^2 \Big(B_0^2\Phi + B_0 B_1 \Phi^2 + B_0 B_1 - B_1^2 \Phi\Theta^2 + B_1^2\Phi \\ - B_1\Phi^2\Theta^2 - B_1\Theta^2 - \Phi\Theta^2\Big)$$

vanishes whenever $B_0 = 0$ or $B_1 = 0$. This occurs when y_t is truly an AR(1) instead of an assumed ARX(1,1). When $B_0 \neq 0$, $B_1 \neq 0$, and the measurement errors are AR(1) with $\Theta = 0$, the determinant vanishes whenever $B_0 B_1 + (B_0^2 + B_1^2)\Phi + B_0 B_1 \Phi^2 = 0$. The solutions to this quadratic equation in Φ are $\{-B_0/B_1, -B_1/B_0\}$. If $|B_0| \neq |B_1|$, there is one solution strictly inside the unit circle. If, on the other hand, the measurement errors are MA(1) with $\Phi = 0$, the determinant vanishes whenever $B_0 = \Theta^2$.

5. DYNAMIC PANEL DATA MODELS WITH MEASUREMENT ERRORS

This section extends the identification analysis to dynamic panel data models. In theory, identification is possible from repeated observations of the data over time if the time variations of the measurement errors differ from those of the variables

that they contaminate. The problem is complicated by the fact that unobserved heterogeneity is characteristic of panel data, and transformations that remove the individual effects tend to remove more signal than noise from the data. Griliches and Hausman (1986) propose to combine information of the maximum likelihood (also known as the within or the LSDV) estimator with the first difference (FD) estimator to consistently estimate a static panel model with one covariate. Wansbeek and Koning (1991) and Biorn (1992,1996) extend the idea in various ways under the maintained assumption that the measurement errors are white noise. While the assumption is convenient, Bell and Wilcox (1993) find that the sampling error of retail sales is highly serially correlated, while Bound and Krueger (1991) find that the measurement error in reported earnings is correlated over time.

While there is a huge literature on measurement errors for cross section or time series data, a comprehensive treatment of which is given in Wansbeek and Meijer (2000), there are few results for dynamic panel models. Holtz-Eakin, Newey, and Rosen (1988) consider panel vector autoregressions but limit their attention to white noise measurement errors. Wansbeek (2001) derives moment conditions by considering the covariances between the dependent variables and the regressors over time, projecting out the parameters of the possibly serially correlated measurement error process. Biorn (2008) also allows for moving-average-type measurement errors. However, both studies adopt a static setup.

We consider a general panel autoregressive distributed lag ARX(p_y, r_x) model with a dependent variable y_{it} and a scalar covariate x_{it}. Assume there are N individual units indexed by i that are observed over T time periods indexed by t. For all i and t, the data are generated by

$$y_{it} = \lambda_i + \sum_{j=1}^{p_y} A_j y_{i,t-j} + \sum_{j=0}^{r_x} B_j x_{i,t-j} + u_{it}, \qquad \textbf{(Model-py)}$$

with unobserved heterogeneity captured by λ_i and possibly correlated with the predetermined variables $y_{i,t-\tau-1}$ and $x_{i,t-\tau}$ $(\tau \geqslant 0)$.

The data are contaminated by classical measurement errors. For $i = 1, \ldots N$ and $t = 1, \ldots T$,

$$Y_{it} = y_{it} + \epsilon_{it}^{y},$$
$$X_{it} = x_{it} + \epsilon_{it}^{x}.$$

Let $n_\epsilon^y = 1$ (resp. $n_\epsilon^x = 1$) if y_{it} (resp. x_{it}) is measured with error and zero otherwise. We do not allow for measurement errors to occur in only a subset of the units in the panel. The following assumptions are used.

Assumption PA. (i) For all i, $u_{it} \sim WN(0, \sigma_u^2)$; (ii) for all i, x_{it} is weakly stationary; (iii) for all (i,t) and $\tau \geqslant 0$, $\mathbb{E}[u_{it} x_{i,t-\tau}] = 0$ and $\mathbb{E}(u_{it}\lambda_i) = 0$; and (iv) for all (i,t,s), $\mathbb{E}[\epsilon_{it}^y u_{is}] = 0$, $\mathbb{E}[\epsilon_{it}^y x_{is}] = 0$, $\mathbb{E}[\epsilon_{it}^x x_{is}] = 0$, and $\mathbb{E}[\epsilon_{it}^x u_{is}] = 0$.

Assumption PB. The polynomial $A(L)$ is stable, and p_y and r_x are known.

Assumption PC. For all i, let $v_{it}^y \sim WN(0, \sigma_v^{y2})$ and $v_{it}^x \sim WN(0, \sigma_v^{x2})$, with $\mathbb{E}[v_{it}^y v_{is}^x] = 0$ for all (t, s). The measurement errors follow:

$$\Phi^y(L)\epsilon_{it}^y = \Theta^y(L)v_{it}^y, \quad \text{if } n_\epsilon^y = 1,$$
$$\Phi^x(L)\epsilon_{it}^x = \Theta^x(L)v_{it}^x, \quad \text{if } n_\epsilon^x = 1,$$

where $\Phi^y(L)$, $\Theta^y(L)$, $\Phi^x(L)$, and $\Theta^x(L)$ are polynomials of known orders p_ϵ^y, q_ϵ^y, p_ϵ^x, and q_ϵ^x, respectively. Moreover, $\Phi_0^y = \Theta_0^y = 1$, $\Phi_0^x = \Theta_0^x = 1$, $\Phi^y(L)$ and $\Phi^x(L)$ are stable, and $\Phi^y(L)$ and $\Theta^y(L)$ as well as $\Phi^x(L)$ and $\Theta^x(L)$ are co-prime.

Under Assumption PA(i), the disturbances are uncorrelated across time with variances that are homogeneous across units. The uncorrelatedness of u_{it} together with Assumption PA(iii) ensure that lagged values of y_{it} as well as x_{it} and its lags qualify as instrumental variables in Model-py. While the latent fixed effect λ_i is assumed uncorrelated with the error u_{it}, $\mathbb{E}(x_{it}\lambda_i)$ is left unrestricted. As in Holtz-Eakin et al. (1988), we do not impose any dependence restrictions across units. Cross-sectional independence can lead to additional orthogonality conditions beyond those specified in Assumption PA(iii). Assumption PA(iv) states that the measurement errors are classical. Assumption PB imposes stability on the model. When combined with Assumptions PA and PC, this assumption guarantees that for every i, $(Y_{it}, X_{it})'$ are weakly stationary. Under Assumption PC, the measurement errors are uncorrelated ARMA processes. The coefficients and the lag order of $A(L)$, $B(L)$, $\Phi^x(L)$, $\Phi^y(L)$, $\Theta^x(L)$, $\Theta^y(L)$, as well as σ_v^x and σ_v^y are homogeneous across units. The main difference with a VAR is thus that the parameters in the panel data model are common across units.

The parameters of Model-py are $\theta' \equiv (\theta_{AR}', \theta_{MA}')$ with

$$\theta_{AR} \equiv \left(A_1, \ldots, A_{p_y}, B_0, \ldots, B_{r_x}, \Phi_1^y, \ldots, \Phi_{p_\epsilon^y}^y, \Phi_1^x, \ldots, \Phi_{p_\epsilon^x}^x\right)',$$
$$\theta_{MA} \equiv \left(\sigma_u^2, \Theta_1^y, \ldots, \Theta_{q_\epsilon^y}^y, \Theta_1^x, \ldots, \Theta_{q_\epsilon^x}^x, \sigma_v^{y2}, \sigma_v^{x2}\right)',$$

and the objective of the exercise is to identify θ_{AR} and θ_{MA}.

The model in terms of the error-ridden data is

$$A(L)Y_{it} = \lambda_i + B(L)X_{it} + u_{it} + A(L)\epsilon_{it}^y - B(L)\epsilon_{it}^x.$$

Under Assumptions PA, PB, and PC, the same reasoning as in Lemma 1 shows that the observed data can be represented as

$$\mathcal{A}(L)Y_{it} = \lambda_i + \mathcal{B}(L)X_{it} + U_{it}, \qquad \textbf{(Model-PY)}$$

where $\mathcal{A}(L)$ and $\mathcal{B}(L)$ are polynomials of orders p_Y and r_X, respectively, with $p_Y = p_y + \bar{p}$, $q_U \leqslant \bar{p} + \max\{0, [p_y + q_\epsilon^y - p_\epsilon^y]n_\epsilon^y, [r_x + q_\epsilon^x - p_\epsilon^x]n_\epsilon^x\}$, $r_X = r_x + \bar{p}$, and $\bar{p} \leqslant n_\epsilon^y p_\epsilon^y + n_\epsilon^x p_\epsilon^x$. The composite error $U_{it} = u_{it} + A(L)\epsilon_{it}^y - B(L)\epsilon_{it}^x$ is an

$MA(q_U)$ even though u_{it} is white noise. A triangular structure obtains when the coefficients of $\mathcal{A}(L)$ and $\mathcal{B}(L)$ depend on θ_{AR} alone, while U_t depends on both θ_{AR} and θ_{MA}.

In addition to the possible correlation between U_{it} and the predetermined variables in Model-PY, identification of θ is now complicated by the presence of the unobserved heterogeneity term λ_i. Model-PY must further be transformed to eliminate the individual effect. Let $\Delta Y_{it} \equiv Y_{it} - Y_{i,t-1}$. First differencing Model-PY gives

$$\mathcal{A}(L)\Delta Y_{it} = \mathcal{B}(L)\Delta X_{it} + \Delta U_{it}. \qquad \textbf{(Model-}\Delta\textbf{PY)}$$

Since U_{it} is MA(q_U), it follows that ΔU_{it} is $MA(q_U+1)$. Model-py can be identified from Model-ΔPY only if Model-ΔPY is itself identifiable.

Letting $\Gamma_{\Delta U}(j) \equiv \mathbb{E}(\Delta U_{it}\Delta U_{i,t-j})$,[5] the parameters of Model-ΔPY are $\pi' \equiv (\pi'_{AR}, \pi'_{MA})$, with

$$\pi_{AR} \equiv \left(\mathcal{A}_1, \ldots, \mathcal{A}_{p_Y}, \mathcal{B}_0, \ldots, \mathcal{B}_{r_X}\right)',$$
$$\pi_{MA} \equiv (\Gamma_{\Delta U}(0), \ldots, \Gamma_{\Delta U}(q_U+1)))'.$$

In order to identify the $p_Y + (r_X+1)$ components of π_{AR}, we use the orthogonality conditions that hold under Model-ΔPY:

$$\begin{aligned}
&\mathbb{E}(\Delta U_{it}) = 0, \\
&\mathbb{E}(\Delta U_{it}\Delta Y_{i,t-\tau}) = 0, \quad \text{for all } \tau > \tau_Y + 1, \quad \tau_Y \leqslant \bar{p} + \max\{0, [p_y + q_\epsilon^y - p_\epsilon^y]n_\epsilon^y\}, \\
&\mathbb{E}(\Delta U_{it}\Delta X_{i,t-\tau}) = 0, \quad \text{for all } \tau > \tau_X + 1, \quad \tau_X \leqslant \bar{p} + \max\{0, [r_x + q_\epsilon^x - p_\epsilon^x]n_\epsilon^x\}.
\end{aligned}$$

The second set of moments yields $T - \tau_Y - 2$ orthogonality conditions, while the third set yields $T - \tau_X - 2$ conditions. The number of conditions differs because the lag lengths p_y and r_x are not necessarily identical, and the measurement errors for y_{it} and x_{it} may have different dynamics.[6] Once π_{AR} is identified, π_{MA} is identifiable from the autocovariances of the residuals $\mathcal{A}(L)\Delta Y_{it} - \mathcal{B}(L)\Delta X_{it}$ provided Model-py has $T - \max\{p_Y, r_X\} \geqslant q_U + 2$ observations. This leads to the following result.

LEMMA 3. *Suppose Assumptions PA, PB, and PC hold. The order conditions for identification of π_{AR} and π_{MA} jointly are*

$$\begin{aligned}
AR: &\quad T \geqslant \frac{1}{2}\left[\tau_Y + \tau_X + p_y + r_x + 1\right] + \bar{p} + 2, \\
MA: &\quad T \geqslant q_U + \bar{p} + \max\{p_y, r_x\} + 2.
\end{aligned}$$

A useful observation can be made if π_{AR} (and not π_{MA}) is of interest, as in Holtz-Eakin et al. (1988), who consider serially uncorrelated measurement errors in a panel VAR. In this case $\bar{p} = 0$, $\tau_Y = p_y$, $\tau_X = r_x - 1$, and our AR order condition reduces to $T \geqslant p_y + r_x + 2$. This is equivalent to their condition (p. 1377) of $T \geqslant 2m + 2$, since $p_y = r_x = m$. Our results are more general, as we allow the

measurement errors to be serially correlated. Once π_{AR} and π_{MA} are identified, one can proceed with the local identification of θ_{AR} and θ_{MA} from π_{AR} and π_{MA}. The order and rank conditions are the same as before.

6. CONCLUSION

Mismeasured variables arise either as a result of the data collection process or when the latent variables in the model are replaced by proxies. This paper shows that in dynamic models with measurement errors, comparing the total number of unknowns in Model-y with the number of parameters in Model-Y does not provide an informative order condition for identification. The results have implications for estimation of dynamic models in which measurement errors are often inevitable.

NOTES

1. The exogeneity condition $\mathbb{E}[u_t \vec{x}_t'] = 0$ is no longer sufficient as $\vec{y}_{t-1} = (y_{t-1}', \ldots, y_{t-p_y}')'$ is a function of all the lagged x_t's and u_t's.
2. Erickson and Whited (2002) used higher-order moments, while Willassen (1977) used the joint distribution of the observables for identification. Hsiao (1979) studies identification of dynamic simultaneous equations models from the second-moment properties of the data when the measurement errors are white noise.
3. For example, if y_t and ϵ_t^y are $AR(1)$ with different coefficients, then Y_t is $ARMA(2, 1)$.
4. When all the variables in Model-y are scalar, $n_\epsilon^y = 0$ and $B(L) \neq \Phi^x(L)$, then the results of Lemma 1 hold with equalities in place of (upper) inequalities.
5. Note that because of our homogeneity restrictions, $\Gamma_{\Delta U}(j)$ does not vary across individuals.
6. Biorn (2008) considers orthogonality conditions for the first difference of Model-py without the lagged dependent variable and suggests using lagged levels and/or the differenced data as instruments. Our orthogonality conditions are based on Model-ΔPY. Note that one observation is lost from first differencing.

REFERENCES

Aigner, D., C. Hsiao, A. Kapteyn, & T. Wansbeek (1984) Latent variable models in econometrics. In Grichiles, Z. and M.D. Intriligator (eds.), *Handbook of Econometrics*, vol. 1, no 2. North Holland.

Anderson, B. & M. Deistler (1984) Identifiability in dynamic errors-in-variables models. *Journal of Time Series Analysis* 5:1, 1–13.

Ashley, R. & D. Vaughan (1986) Measuring measuring error in time series. *Journal of Business and Economic Statistics* 4:1, 95–103.

Bell, W. R. & D. Wilcox (1993) The effect of sampling on the time series behavior of consumption data. *Journal of Econometrics*, 55, 235–265.

Biorn, E. (1992) The bias of some estimators for panel models with measurement errors. *Empirical Economics* 17, 51–66.

Biorn, E. (1996) Panel data with measurement errors. In Matyas, L., & P. Sevestre (eds.), *The Econometrics of Panel Data*, Dordrecht. Kluwer.

Biorn, E. (2008) Measurement Error in a Dynamic Panel Data Analysis: A Synthesis on Modeling and GMM Estimation. Mimeo, University of Oslo.

Bound, J. & A. Krueger (1991) The extent of measurement error in longitudinal earnings data: Do two wrongs make a right? *Journal of Labor Economics* 9, 1–24.

Chanda, K. (1996) Asymptotoic properties of estimators for autoregressive models with errors in variables. *Annals of Statistics* 24:1, 423–430.

Devereux, P. (2001) The cyclicality of real wages within employer-employee matches. *Industrial and Labor Relations Review* 54:4, 835–50.

Eberly, J., S. Rebelo, & N. Vincent (2009) Investment and Value: A Neoclassical Benchmark, CIRPEE Working paper 09-08.

Erickson, T. & T. Whited (2000) Measurement error and the relationship between investment and q. *Journal of Political Economy* 108, 1027–57.

Erickson, T. & T. Whited (2002) Two-Step GMM estimation of the errors-in-variables model using higher order moments. *Econometric Theory* 18, 776–799.

Ermini, L. (1993) Effects of transitory consumption and temporal aggregation on the permanent income hypothesis. *Review of Economics and Statistics* 75:5, 736–74.

Falk, B. & B. Lee (1990) Time series implications of friedman's permanent income hypothesis. *Journal of Monetary Economics* 26:2, 267–283.

Fisher, F. (1963) Uncorrelated disturbances and identifiability criteria. *International Economic Review* 4(2), 134–152.

Fisher, F. (1966) *The Identification Problem in Econometrics*. McGraw-Hill.

Granger, C. W. & M. J. Morris (1976) Time series modelling and interpretation. *Journal of the Royal Statistical Association Series A* 139, 246–257.

Grether, D. & G. Maddala (1973) Errors in variables and serially correlated disturbances in distributed lag models. *Econometrica* 41:2, 255–262.

Griliches, Z. & J. Hausman (1986) Errors in variables in panel data. *Journal of Econometrics* 32:3, 93–118.

Hannan, E. (1969) The identification of vector mixed autoregressive-moving average systems. *Biometrika* 56(1), 223–225.

Hannan, E. (1970) *Multiple Time Series*. Wiley.

Hannan, E. (1971) The identification problem for multiple equation systems with moving average errors. *Econometrica* 39:5, 751–765.

Hannan, E., W. Dunsmuir, & M. Deistler (1980) Estimation of vector ARMAX models. *Journal of Multivariate Analysis* 10, 275–295.

Holtz-Eakin, D., W. Newey, & H. Rosen (1988) Estimating vector autoregressions with panel data. *Econometrica* 56, 1371–1395.

Hsiao, C. (1979) Measurement error in a dynamic simultaneous equations model with stationary disturbances. *Econometrica* 47:2, 475–494.

Koopmans, T. (1953) Identification problems in economic model construction. In W. Hood & T. C. Koopmans (eds.), *Studies in Econometric Method*, vol. 14 of Cowles Commission Monograph, chap. II. Wiley.

Maravall, A. (1979) *Identification in Dynamic Shock-Error Models*. Springer-Verlag.

Maravall, A. & D. Aigner (1977) Identification of the dynamic shock-error model: The case of dynamic regression. In D. Aigner & A. Goldberger (eds.), *Latent Variables in Socio-Economic Models*, North Holland.

Mcdonald, J. & J. Darroch (1983) Consistent estimation of equations with composite moving average disturbance terms. *Journal of Econometrics* 23, 253–267.

Nowak, E. (1992) Identifiability in multivariate dynamic linear error-in-variables models. *Journal of the American Statistical Association* 87, 714–723.

Nowak, E. (1993) The identification of multivariate linear dynamic errors-in-variables models. *Journal of Econometrics* 59:3, 237.

Patriota, A., J. Sato, & B. Blas (2009) Vector Autoregressive Models with Measurement Errors for Testing Granger Causality. Working paper, Universidade de Sao Paulo.

Reiersøl, O. (1950) Identifiability of a linear relation between variables which are subject to error. *Econometrica* 23, 375–389.

Sargent, T. (1989) Two models of measurements and the investment accelerator. *Journal of Political Economy* 97:2.

Söderström, T. (2007) Errors-in-variables methods in system identification. *Automatica* 43, 939–958.

Solo, V. (1986) Identifiability of time series models with errors in variables. *Journal of Applied Probability* 23, 63–71.

Staudenmayer, J. & J. Buonaccorsi (2005) Measurement error in linear autoregressive models. *Journal of the American Economic Association* 100, 841–852.

Tanaka, K. (2002) A unified approach to the measurement error problem in time series models. *Econometric Theory* 18, 278–296.

Wansbeek, T. (2001) GMM estimation in panel data models with measurement error. *Journal of Econometrics* 104, 259–268.

Wansbeek, T. & R. Koning (1991) Measurement error and panel data. *Statistica Needlandica* 45, 85–92.

Wansbeek, T. & E. Meijer (2000) *Measurement Error and Latent Variables in Econometrics*. Elsevier.

Wilcox, D. (1992) The construction of U.S. consumption data: Some facts and their implications for empirical work. *American Economic Review* 82:4, 992–941.

Willassen, Y. (1977) On identifiability of stochastic difference equations with errors-in-variables in relation to identifiability of the classical errors-in-variables models. *Scandinavan Journal of Statistics* 4, 119–124.

APPENDIX

Proof of Lemma 1. Proof is done in two steps. We first discuss the case in which at least one component of y_t and one component of x_t are measured with errors.

CASE 1. $n_\epsilon^y \geqslant 1$ AND $n_\epsilon^x \geqslant 1$.

Under Assumption C, the components of $\{M^y \epsilon_t^y\}$ and $\{M^x \epsilon_t^x\}$ are independent ARMA$(p_\epsilon^y, q_\epsilon^y)$ and ARMA$(p_\epsilon^x, q_\epsilon^x)$ processes, respectively:

$$\Phi^y(L) M^y \epsilon_t^y = \Theta^y(L) v_t^y, \quad v_t^y \sim WN(0, \Sigma_v^y), \tag{A.1}$$

$$\Phi^x(L) M^x \epsilon_t^x = \Theta^x(L) v_t^x, \quad v_t^x \sim WN(0, \Sigma_v^x), \quad \mathbb{E}[v_t^y v_s^{x\prime}] = 0 \quad \text{for all } (t,s), \tag{A.2}$$

where $\Phi^y(z) = I - \left[\sum_{j=1}^{p_\epsilon^y} \Phi_j^y z^j\right]$, $\Theta^y(z) = I + \sum_{j=1}^{q_\epsilon^y} \Theta_j^y z^j$, $\Phi^x(z) = I - \left[\sum_{j=1}^{p_\epsilon^x} \Phi_j^x z^j\right]$, and $\Theta^x(z) = I + \sum_{j=1}^{q_\epsilon^x} \Theta_j^x z^j$ are diagonal polynomial matrices. The matrices Φ_j^y, Θ_j^y, and Σ_v^y are of dimensions $n_\epsilon^y \times n_\epsilon^y$. Similarly, Φ_j^x, Θ_j^x, and Σ_v^x are of dimensions $n_\epsilon^x \times n_\epsilon^x$. Given that $\Phi^y(z)$ and $\Phi^x(z)$ have determinants with zeros outside the unit circle, i.e. for any $z \in \mathbb{C}$, $\det\Phi^y(z) = 0$ only if $|z| > 1$, and $\det\Phi^x(z) = 0$ only if $|z| > 1$, we have

$$M^y \epsilon_t^y = \Phi^y(L)^{-1} \Theta^y(L) v_t^y \quad \text{and} \quad M^x \epsilon_t^x = \Phi^x(L)^{-1} \Theta^x(L) v_t^x.$$

Whenever measurement errors are present, i.e., $n_\epsilon^y \neq 0$, we can "reconstruct" the zeros in ϵ_t^y by pre-multiplying $M^y \epsilon_t^y$ by $M^{y\prime}$. This gives

$$\epsilon_t^y = M^{y\prime} \Phi^y(L)^{-1} \Theta^y(L) v_t^y \quad \text{and} \quad \epsilon_t^x = M^{x\prime} \Phi^x(L)^{-1} \Theta^x(L) v_t^x.$$

Now use the fact that for an arbitrary stable polynomial matrix $\Phi(L)$, $\Phi(z)^{-1} = \frac{1}{\det[\Phi(z)]}\mathrm{Co}[\Phi(z)]'$ where $\mathrm{Co}[\Phi(z)]$ denotes the matrix of co-factors of $\Phi(z)$. Then,

$$\begin{aligned} u_t + A(L)\epsilon_t^y - B(L)\epsilon_t^x &= u_t + \frac{1}{\det[\Phi^y(L)]} A(L)M^{y\prime}\mathrm{Co}[\Phi^y(L)]'\Theta^y(L)v_t^y \\ &\quad - \frac{1}{\det[\Phi^x(L)]} B(L)M^{x\prime}\mathrm{Co}[\Phi^x(L)]'\Theta^x(L)v_t^x \\ &\equiv u_t + \frac{1}{p^y(L)} P^y(L)v_t^y - \frac{1}{p^x(L)} P^x(L)v_t^x, \end{aligned}$$

where $p^y(L)$ and $P^y(L)$ (and similarly $p^x(L)$ and $P^x(L)$) are a scalar and a matrix polynomial, respectively, that are co-prime. Specifically, letting $\pi^y(L)$ be the scalar polynomial that is the greatest common divisor (GCD) of $\det[\Phi^y(L)]$ and $A(L)M^{y\prime}\mathrm{Co}[\Phi^y(L)]'\Theta^y(L)$, $p^y(L)$ and $P^y(L)$ are such that

$$\det[\Phi^y(L)] = \pi^y(L)p^y(L) \quad \text{and} \quad A(L)M^{y\prime}\mathrm{Co}[\Phi^y(L)]'\Theta^y(L) = \pi^y(L)P^y(L).$$

Hence,

$$\deg\{p^y(L)\} \leqslant n_\epsilon^y p_\epsilon^y \quad \text{and} \quad \deg\{P^y(L)\} \leqslant p_y + \left(n_\epsilon^x - 1\right)p_\epsilon^y + q_\epsilon^y,$$

with an analogous result holding for $p^x(L)$ and $P^x(L)$. In addition, note that

$$\begin{aligned} \deg\left\{P^y(L)\right\} - \deg\left\{p^y(L)\right\} &= \deg\left\{\frac{1}{p^y(L)}P^y(L)\right\} \\ &= \deg\{A(L)\} + \deg\{\Theta^y(L)\} - \deg\{\Phi^y(L)\} \\ &= p_y + q_\epsilon^y - p_\epsilon^y, \end{aligned}$$

with an analogous result holding for $p^x(L)$ and $P^x(L)$.

Let $\varphi^U(L)$ be the scalar polynomial that is the least common multiplier (LCM) of $p^y(L)$ and $p^x(L)$, $\varphi^U(L) = LCM\left(p^y(L), p^x(L)\right)$ and denote by $\psi^y(L)$ and $\psi^x(L)$ the polynomials such that

$$p^y(L)\psi^y(L) = \varphi^U(L) \quad \text{and} \quad p^x(L)\psi^x(L) = \varphi^U(L).$$

For example, if $p^y(L)$ and $p^x(L)$ are co-prime, then $\varphi^U(L) = p^y(L)p^x(L)$, $\psi^y(L) = p^x(L)$, and $\psi^x(L) = p^y(L)$. It follows that

$$\begin{aligned} \varphi^U(L)\left[u_t + A(L)\epsilon_t^y - B(L)\epsilon_t^x\right] &= \varphi^U(L)u_t + \psi^y(L)P^y(L)v_t^y - \psi^x(L)P^x(L)v_t^x \\ &\equiv \widetilde{u}_t + \widetilde{v}_t^y + \widetilde{v}_t^x. \end{aligned}$$

Define the $n_y \times n_y$ and the $n_y \times n_x$ matrix polynomials $\mathcal{A}(L)$ and $\mathcal{B}(L)$ by

$$\mathcal{A}(L) \equiv \varphi^U(L)A(L) \quad \text{and} \quad \mathcal{B}(L) \equiv \varphi^U(L)B(L).$$

Then $\mathcal{A}(L)Y_t = \mathcal{B}(L)X_t + U_t$, with $U_t = \widetilde{u}_t + \widetilde{v}_t^y + \widetilde{v}_t^x$. Note that the coefficients of $p^y(L)$ (resp. $p^x(L)$) are functions of the coefficients in $\Phi^y(L)$ (resp. $\Phi^x(L)$), so the coefficients of $\varphi^U(L)$ are functions of those in $\Phi^y(L)$ and $\Phi^x(L)$. Hence, the coefficients of $\mathcal{A}(L)$ and $\mathcal{B}(L)$ are functions of θ_{AR} alone. On the other hand, U_t depends on both θ_{AR} and θ_{MA}.

To determine the degrees of various polynomials above, let $\bar{p} \equiv \deg\{\varphi^U(L)\}$ and note that $\bar{p} \leqslant n_\epsilon^y p_\epsilon^y + n_\epsilon^x p_\epsilon^x$. Then,

$$\begin{aligned}
\deg\{\mathcal{A}(L)\} &= p_y + \bar{p}, \qquad \deg\{\mathcal{B}(L)\} = r_x + \bar{p},\\
\widetilde{u}_t &\sim VMA(q_{\widetilde{u}}) \quad \text{with} \quad q_{\widetilde{u}} = \bar{p},\\
\widetilde{v}_t^y &\sim VMA(q_{\widetilde{w}}^y) \quad \text{with} \quad q_{\widetilde{w}}^y = \bar{p} - \deg\{p^y(L)\} + \deg\{P^y(L)\}\\
&\qquad\qquad\qquad\qquad\qquad\quad = \bar{p} + p_y + q_\epsilon^y - p_\epsilon^y,\\
\widetilde{v}_t^x &\sim VMA(q_{\widetilde{w}}^x) \quad \text{with} \quad q_{\widetilde{w}}^x = \bar{p} - \deg\{p^x(L)\} + \deg\{P^x(L)\}\\
&\qquad\qquad\qquad\qquad\qquad\quad = \bar{p} + r_x + q_\epsilon^x - p_\epsilon^x.
\end{aligned}$$

Since u_t, v_t^y, and v_t^x are mutually uncorrelated white noise processes, $\mathbb{E}[u_t v_s^{y\prime}] = 0$, $\mathbb{E}[u_t v_s^{x\prime}] = 0$, $\mathbb{E}[v_t^y v_s^{x\prime}] = 0$ for all (t, s), and the sequences $\widetilde{u}_t$, $\widetilde{v}_t^y$, and $\widetilde{v}_t^x$ are mutually uncorrelated VMA processes. Hence, $U_t \sim VMA(q_U)$ of order $q_U \leqslant \max\{q_{\widetilde{u}}, q_{\widetilde{w}}^y, q_{\widetilde{w}}^x\}$.

In addition, $\mathbb{E}(U_t) = 0$ and $\mathbb{E}(U_t Y_{t-\tau}') = 0$ whenever $\tau > \tau_Y$ with $\tau_Y \leqslant \bar{p} + \max\{0, p_y + q_\epsilon^y - p_\epsilon^y\}$, and $\mathbb{E}(U_t X_{t-\tau}') = 0$ whenever $\tau > \tau_X$ with $\tau_X \leqslant \bar{p} + \max\{0, r_x + q_\epsilon^x - p_\epsilon^x\}$. This establishes the result of Lemma 1 when $n_\epsilon^y \geqslant 1$ and $n_\epsilon^x \geqslant 1$.

CASE 2. $n_\epsilon^y = 0$ or $n_\epsilon^x = 0$

If $n_\epsilon^y = 0$ and $n_\epsilon^x \geqslant 1$, then following the same reasoning as above, we have $\mathcal{A}(L)Y_t = \mathcal{B}(L)X_t + U_t$, with

$$U_t = \widetilde{u}_t + \widetilde{v}_t^x, \qquad \widetilde{u}_t \equiv p^x(L)u_t, \qquad \widetilde{v}_t^x \equiv -P^x(L)v_t^x,$$

where the $n_y \times n_y$ and the $n_y \times n_x$ matrix polynomials $\mathcal{A}(L)$ and $\mathcal{B}(L)$ are now defined by

$$\mathcal{A}(L) \equiv p^x(L)A(L) \quad \text{and} \quad \mathcal{B}(L) \equiv p^x(L)B(L).$$

The coefficients of $\mathcal{A}(L)$ and $\mathcal{B}(L)$ are functions of θ_{AR} alone. On the other hand, U_t depends on both θ_{AR} and θ_{MA}. In addition, $\bar{p} \leqslant n_\epsilon^x p_\epsilon^x$ and

$$\begin{aligned}
\deg\{\mathcal{A}(L)\} &= p_y + \bar{p}, \qquad \deg\{\mathcal{B}(L)\} = r_x + \bar{p},\\
q_U &\leqslant \bar{p} + \max\{0, r_x + q_\epsilon^x - p_\epsilon^x\}.
\end{aligned}$$

Further, $\mathbb{E}(U_t) = 0$, $\mathbb{E}(U_t Y_{t-\tau}') = 0$ for all $\tau > \tau_Y$ with $\tau_Y \leqslant \bar{p}$, and $\mathbb{E}(U_t X_{t-\tau}') = 0$ for all $\tau > \tau_X$ with $\tau_X \leqslant \bar{p} + \max\{0, r_x + q_\epsilon^x - p_\epsilon^x\}$.

If on the other hand, $n_\epsilon^y \geqslant 1$ and $n_\epsilon^x = 0$, then $\mathcal{A}(L)Y_t = \mathcal{B}(L)X_t + U_t$ with

$$U_t = \widetilde{u}_t + \widetilde{v}_t^y, \qquad \widetilde{u}_t \equiv p^y(L)u_t, \qquad \widetilde{v}_t^y \equiv P^y(L)v_t^y,$$

and the $n_y \times n_y$ and the $n_y \times n_x$ matrix polynomials $\mathcal{A}(L)$ and $\mathcal{B}(L)$ are given by

$$\mathcal{A}(L) \equiv p^y(L)A(L) \quad \text{and} \quad \mathcal{B}(L) \equiv p^y(L)B(L).$$

Then $\bar{p} \leqslant n_\epsilon^y p_\epsilon^y$ and

$$\begin{aligned}
\deg\{\mathcal{A}(L)\} &= p_y + \bar{p}, \qquad \deg\{\mathcal{B}(L)\} = r_x + \bar{p},\\
q_U &\leqslant \bar{p} + \max\{0, p_y + q_\epsilon^y - p_\epsilon^y\}.
\end{aligned}$$

Since $\mathbb{E}(U_t) = 0$, while $\mathbb{E}(U_t Y'_{t-\tau}) = 0$ for all $\tau > \tau_Y$ with $\tau_Y \leqslant \bar{p} + \max\{0, p_y + q_\epsilon^y - p_\epsilon^y\}$ and $\mathbb{E}(U_t X'_{t-\tau}) = 0$ for all $\tau > \tau_X$ with $\tau_X \leqslant \bar{p}$. Again, the coefficients of $\mathcal{A}(L)$ and $\mathcal{B}(L)$ are functions of θ_{AR} alone, while U_t depends on both θ_{AR} and θ_{MA}. This completes the proof of Lemma 1. ■

Proof of Propositions 1, 2, 3, and 4. The proof of Propositions 1, 2, 3, and 4 is based on the following lemma.

LEMMA 4. *Let $g : D \to \mathbb{R}^{r_1}$ and $f : D \to \mathbb{R}^{r_2}$ (with $D \subseteq \mathbb{R}^{d_1+d_2}$, $d_1 \geqslant 1$, $d_2 \geqslant 1$) be two continuously differentiable mappings on D, and define $\varphi : D \to \mathbb{R}^{r_1+r_2}$ to be a mapping such that $\varphi(x_1, x_2) \equiv \big(g(x_1), f(x_1, x_2)\big)$ for each $x = (x'_1, x'_2)' \in D$ (with $\dim x_1 = d_1$ and $\dim x_2 = d_2$). Then,*

(i) rank $\left(\frac{\partial g(x_1)}{\partial x_1}\right) = d_1$ and rank $\left(\frac{\partial f(x_1,x_2)}{\partial x_2}\right) = d_2$ imply rank $\left(\frac{\partial \varphi(x_1,x_2)}{\partial x}\right) = d_1 + d_2$,
(ii) rank $\left(\frac{\partial \varphi(x_1,x_2)}{\partial x}\right) = d_1 + d_2$ implies rank $\left(\frac{\partial f(x_1,x_2)}{\partial x_2}\right) = d_2$.

Using Lemma 1, we can write the parameters of Model-Y as

$$(\pi_{AR}, \pi_{MA}) = (\pi_{AR}(\theta_{AR}), \pi_{MA}(\theta_{AR}, \theta_{MA})).$$

Hence, the mapping $\pi = \pi(\theta)$ has a triangular structure as in Lemma 4.

The order conditions in Proposition 1 follow by a well-known result (see, e.g., Fisher, 1966) stating: If $\pi(\theta)$ is continuously differentiable and the rank of $\partial\pi/\partial\theta$ remains constant in a neighborhood of θ, then a necessary condition for θ to be a locally identified is that rank $(\partial\pi/\partial\theta) = \dim(\theta)$. The latter implies $\dim(\theta) \leqslant \dim(\pi)$, which is the AR order condition in Proposition 1. Combining rank $(\partial\pi/\partial\theta) = \dim(\theta)$ with Lemma 4(ii) further gives

$$\text{rank}\left(\frac{\partial \pi_{MA}(\theta_{AR}, \theta_{MA})}{\partial \theta_{MA}}\right) = \dim(\theta_{MA}) \quad \text{so} \quad \dim(\theta_{MA}) \leqslant \dim(\pi_{MA}),$$

which is the MA order condition in Proposition 1.

To derive the order conditions in Proposition 2, we use the second well-known result (see, e.g., Fisher, 1966) stating: If $\pi(\theta)$ is continuously differentiable and the ranks of $\partial\pi/\partial\theta$ and $\partial\pi/\partial\theta_{MA}$ remain constant in a neighborhood of θ, then a necessary condition for θ_{AR} alone to be locally identified is that rank $(\partial\pi/\partial\theta) =$ rank $(\partial\pi/\partial\theta_{MA}) + \dim(\theta_{AR})$. The above implies $\dim(\theta_{AR}) \leqslant \dim(\pi)$, which is the AR condition in Proposition 2. For the MA condition, we use an analogous result that says: If $\pi(\theta)$ is continuously differentiable and the ranks of $\partial\pi/\partial\theta$ and $\partial\pi/\partial\theta_{AR}$ remain constant in a neighborhood of θ, then a necessary condition for θ_{MA} alone to be locally identified is that rank $(\partial\pi/\partial\theta) =$ rank $(\partial\pi/\partial\theta_{AR}) + \dim(\theta_{MA})$. Now rank $(\partial\pi/\partial\theta) \leqslant$ rank $(\partial\pi/\partial\theta_{AR}) +$ rank $(\partial\pi/\partial\theta_{MA}) =$ rank $(\partial\pi/\partial\theta_{AR}) +$ rank $(\partial\pi_{MA}/\partial\theta_{MA})$, where the second equality follows by the triangular structure of the problem. Hence, a necessary condition is that $\dim(\theta_{MA}) \leqslant$ rank $(\partial\pi_{MA}/\partial\theta_{MA}) \leqslant \dim(\pi_{MA})$, which is the MA condition in Proposition 2.

The rank conditions in Proposition 3 are an immediate application of the implicit function theorem, which says that the rank condition rank $(\partial\pi/\partial\theta) = \dim\theta$ is also sufficient

for θ to be locally identified. Using Lemma 4(i), sufficient conditions are

$$\text{rank}\left(\frac{\partial \pi_{AR}(\theta_{AR})}{\partial \theta_{AR}}\right) = \dim(\theta_{AR}) \quad \text{and} \quad \text{rank}\left(\frac{\partial \pi_{MA}(\theta_{AR}, \theta_{MA})}{\partial \theta_{MA}}\right) = \dim(\theta_{MA}),$$

which are the conditions in Proposition 3.

Finally, for Proposition 4, note that since π_{AR} is a function of θ_{AR} alone, the AR condition follows by the implicit function theorem. The MA condition uses the fact that the above rank condition rank $(\partial\pi/\partial\theta) = \text{rank}\ (\partial\pi/\partial\theta_{AR}) + \dim(\theta_{MA})$ is also sufficient for partial local identification of θ_{MA}. ■

Proof of Lemma 4. To show (i), note first that for every $x = (x_1', x_2')' \in D$ (with $\dim x_1 = d_1$ and $\dim x_2 = d_2$),

$$\frac{\partial \varphi(x_1, x_2)}{\partial x} = \begin{pmatrix} \frac{\partial g(x_1)}{\partial x_1} & 0_{r_1 \times d_2} \\ \frac{\partial f(x_1,x_2)}{\partial x_1} & \frac{\partial f(x_1,x_2)}{\partial x_2} \end{pmatrix}.$$

If the $r_1 \times d_1$ matrix $\partial g(x_1)/\partial x_1$ is of rank d_1, then $r_1 \geqslant d_1$ and there exists a square $d_1 \times d_1$ matrix $M_g(x_1)$ obtained by deleting $(r_1 - d_1)$ rows of $\partial g(x_1)/\partial x_1$ that is such that $\det M_g(x_1) \neq 0$. Similarly, if the $r_2 \times d_2$ matrix $\partial f(x_1, x_2)/\partial x_2$ is of rank d_2, then $r_2 \geqslant d_2$ and there exists a square $d_2 \times d_2$ matrix $M_f(x_1, x_2)$ obtained by deleting $(r_2 - d_2)$ rows of $\partial f(x_1, x_2)/\partial x_2$ such that $\det M_f(x_1, x_2) \neq 0$. Now let $N_f(x_1, x_2)$ be the $d_2 \times d_1$ submatrix of $\partial f(x_1, x_2)/\partial x_1$ obtained by deleting the same rows. Then, consider the following $(d_1 + d_2) \times (d_1 + d_2)$ submatrix of $\partial \varphi(x_1, x_2)/\partial x$,

$$M_\varphi(x_1, x_2) = \begin{pmatrix} M_g(x_1) & 0_{d_1 \times d_2} \\ N_f(x_1, x_2) & M_f(x_1, x_2) \end{pmatrix}.$$

Since $\det\big(M_\varphi(x_1, x_2)\big) = \det\big(M_g(x_1)\big)\det\big(M_f(x_1, x_2)\big) \neq 0$, $\text{rank}\,(\partial\varphi(x_1, x_2)/\partial x) = d_1 + d_2$.

We now show (ii). If $\text{rank}\,(\partial\varphi(x_1, x_2)/\partial x) = d_1 + d_2$, then $\partial\varphi(x_1, x_2)/\partial x$ is of full column rank, which implies that

$$\text{rank}\begin{pmatrix} \frac{\partial g(x_1)}{\partial x_1} \\ \frac{\partial f(x_1,x_2)}{\partial x_1} \end{pmatrix} = d_1 \quad \text{and} \quad \text{rank}\begin{pmatrix} 0_{r_1 \times d_2} \\ \frac{\partial f(x_1,x_2)}{\partial x_2} \end{pmatrix} = d_2.$$

The second equality can only hold if $\text{rank}\,(\partial f(x_1, x_2)/\partial x_2) = d_2$. ■

Proof of Lemma 2. From Lemma 1, Model-Y is

$$\mathcal{A}(L)Y_t = \mathcal{B}(L)X_t + U_t,$$

with disturbances U_t that satisfy $\mathbb{E}(U_t Y_{t-\tau}') = 0$ for $\tau > \tau_Y$, and $\mathbb{E}(U_t X_{t-\tau}') = 0$ for all $\tau > \tau_X$. Now post-multiplying the above equation by $Y_{t-\tau}'$ for $\tau_Y + 1 \leqslant \tau \leqslant \tau_Y + p_Y$ and taking expectations gives

$$\big(\mathcal{A}_1 \ldots \mathcal{A}_{p_Y}\ \mathcal{B}_0 \ldots \mathcal{B}_{r_X}\big)\begin{pmatrix} H_{YY} \\ H_{XY} \end{pmatrix} = \mathcal{C}_Y, \tag{A.3}$$

where $\mathcal{C}_Y \equiv \big(\Gamma_{YY}(\tau_Y+1)\ldots\Gamma_{YY}(\tau_Y+p_Y)\big)$. Similarly, post-multiplying by $X'_{t-\tau}$ for $\tau_X+1 \leqslant \tau \leqslant \tau_X+r_X+1$ and taking expectations gives

$$\big(\mathcal{A}_1 \ldots \mathcal{A}_{p_Y}\ \mathcal{B}_0 \ldots \mathcal{B}_{r_X}\big)\begin{pmatrix} H_{YX} \\ H_{XX}\end{pmatrix} = \mathcal{C}_X, \tag{A.4}$$

where $\mathcal{C}_X \equiv \big(\,\Gamma_{YX}(\tau_X+1)\ldots\Gamma_{YX}(\tau_X+r_X+1)\big)$. Combining (A.3) and (A.4) and applying the vec operator to both sides gives

$$\left[\begin{pmatrix} H_{YY} & H_{YX} \\ H_{XY} & H_{XX}\end{pmatrix}' \otimes I_{n_y}\right] \mathrm{vec}\big(\mathcal{A}_1 \ldots \mathcal{A}_{p_Y}\ \mathcal{B}_0 \ldots \mathcal{B}_{r_X}\big) = \mathrm{vec}\big(\mathcal{C}_Y\ \mathcal{C}_X\big).$$

The above system of equations is linear in $\pi_{AR} = \mathrm{vec}\big(\mathcal{A}_1 \ldots \mathcal{A}_{p_Y}\ \mathcal{B}_0 \ldots \mathcal{B}_{r_X}\big)$. Hence, it has a globally unique solution if and only if

$$\left[\begin{pmatrix} H_{YY} & H_{YX} \\ H_{XY} & H_{XX}\end{pmatrix}' \otimes I_{n_y}\right] \quad \text{is nonsingular, i.e., rank} \begin{pmatrix} H_{YY} & H_{YX} \\ H_{XY} & H_{XX}\end{pmatrix} = n_y p_Y + n_x(r_X+1).$$

Once all the parameters in $\mathcal{A}(L)$ and $\mathcal{B}(L)$ are identified, the disturbances U_t are identified, and so are their auto-covariances $(\Gamma_U(0),\ldots,\Gamma_U(q_U))$. Hence, the global identification of π_{MA} follows. ■