Proposal: Panel, Selection, and Control

by Eric Penner

Introduction

Given a partially linear fixed effects panel data model, suppose that for all cross-sections and time periods a known subset of both the linear in parameters regressors, and non parametrically defined regressors are correlated with the error term. This endogeneity problem will be resolved using the control function approach of Newey, Powell, and Vella (1999), with the additional complication of having the set of instrumental variables which are appropriate/available for at least one cross-section not identical to those available/appropriate for a non empty subset of other cross-sections. That is, suppose that for each time period $t \in \{1, 2, \ldots, T\}$, and cross-section $j \in \{1, 2, \ldots, q\}$ where $\{q, T\} \in \mathbb{N}$,

$$Y_{jt} = \beta_0 + [\ Z'_{1jt} \ \ Z'_{2jt}\ ]\beta_1 + h(X_{jt}) + e_j + \varepsilon_{jt} \tag{1}$$

$$Z_{1jt} = m_{1j}(Z_{1jt-1}, \ldots, Z_{1jt-c_1}, Z_{2jt}, W_{jt}) + V_{1jt} \tag{2}$$

$$X_{jt} = m_{2j}(X_{jt-1}, \ldots, X_{jt-c_2}, Z_{2jt}, W_{jt}) + V_{2jt} \tag{3}$$

$$E(\varepsilon_{jt}|Z_{1jt}, Z_{2jt}, X_{jt}) = E(\varepsilon_{jt}|Z_{1jt}, X_{jt}) \neq 0 \tag{4}$$

$$E(V_{1jt}|Z_{1jt-1}, \ldots, Z_{1jt-c_1}, Z_{2jt}, W_{jt}) = 0 \tag{5}$$

$$E(V_{2jt}|X_{jt-1}, \ldots, X_{jt-c_2}, Z_{2jt}, W_{jt}) = 0 \tag{6}$$

Where for $\{p_1, p_2, D_x, w, w_1, w_2, \ldots, w_q\} \subset \mathbb{N}$, Y_{jt}, and ε_{jt} are scalar random variables, Z_{2jt} is a p_2 dimesion vector of exogenous random variables, Z_{1jt} and X_{jt} are vectors of endogenous random variables having dimension p_1 and D_x respectively. e_j is a scalar fixed effect, ε_{jt} is a scalar error term, V_{1jt} and V_{2jt} are error vectors of dimension p_1 and p_2 respectively. $W_t = [\ W_{1t} \ \ W_{2t} \ \ \cdots \ \ W_{wt}\ ]'$ is vector of instrumental variables of dimension w, $W_{jt} = [\ W_{j1t} \ \ W_{j2t} \ \ \cdots \ \ W_{jw_jt}\ ]'$ is a vector of instrumental variables of dimension w_j where $\{W_{jit}\}_{i=1}^{w_j} \subset \{W_{it}\}_{i=1}^{w}$ and there exists at least one pair $j, j' \in \{1, 2, \ldots, p\}$, where $j \neq j'$, such that $\{W_{jit}\}_{i=1}^{w_j} \neq \{W_{j'it}\}_{i=1}^{w_{j'}}$. Lastly $h(\cdot)$ is an unknown function common to all cross-sections, $m_{1j}(\cdot)$ and $m_{2j}(\cdot)$ are vector of unknown functions unique to cross section j. β_0 is a scalar, β_1 is a $p_1 + p_2 = p$ dimensional vector of real numbers.

Remarks :

a.) We have in equation (1) the partially linear function of primary interest having endogenous random variable entering both parametrically, Z_{2jt} and nonparametrically X_{jt}. The focus of this proposal is the identification and estimation of finite dimensional parameter β_1.

b.) Given the time series nature of the model, I have defined equation (2) as a nonparametric autoregressive model of order c_1, or $AR(c_1)$, and equation (3) is a nonparametric $AR(c_2)$ model. Note I have not yet assumed anything particularly restrictive regarding the characteristics of the error terms V_{1jt} and V_{2jt}, this is intentional so as to leave room for homoskedastic errors, or conditional heteroskedastic errors so that equations (2) and (3) are nonparametric ARCH models.

c.) A fully parametric fixed effects panel data model is nested within the current model by taking the dimension of X_{jt} equal to zero, i.e. $D_x = 0$.

For notational convenience let $D_v = D_x + p_1$ and define the following equivalences,

$$Z_{jt} = [\ Z'_{1jt}\ \ Z'_{2jt}\ ]' \quad \text{and} \quad V_{jt} = [\ V'_{1jt}\ \ V'_{2jt}\ ]'$$

As mentioned above $E(\varepsilon_{jt}|Z_{1jt}, X_{jt}) \neq 0$ means that the error term ε_{jt} is not orthogonal to the space of functions of Z_{1jt} and, X_{jt}. Following the control function approach of Newey, Powell, and Vella (1999), I assume that there exists a random variable u_{jt} and unknown function $f_j(\cdot)$ such that;

$$\varepsilon_{jt} = f_j(V_{jt}) + u_{jt} \qquad \text{where} \quad E(u_{jt}|\{Z_{j_k}, W_{jk}, X_{jk}, V_{jk}\}_{k=1}^{t}) = 0$$

Furthermore I assume that all unknown functions heretofore define are additive, that is,

$$h(X_{jt}) = \sum_{d=1}^{D_x} h_d(X_{jdt}) \qquad\qquad f_j(V_{jt}) = \sum_{l=1}^{D_v} f_{jl}(V_{jlt})$$

Where $h_d(\cdot)$ and $f_{jl}(\cdot)$ are vectors of unknown real valued univariate functions. Also assume,

$$m_{1j}(Z_{1jt-1}, \ldots, Z_{1jt-c_1}, Z_{2jt}, W_{jt}) = \sum_{k=1}^{c_1} m_{11jk}(Z_{1jt-k}) + \sum_{k=1}^{p_2} m_{12jk}(Z_{2jkt}) + \sum_{k=1}^{w_j} m_{13jk}(W_{jkt})$$
$$m_{2j}(X_{1jt-1}, \ldots, X_{1jt-c_2}, Z_{2jt}, W_{jt}) = \sum_{k=1}^{c_2} m_{21jk}(X_{1jt-k}) + \sum_{k=1}^{p_2} m_{22jk}(Z_{2jkt}) + \sum_{k=1}^{w_j} m_{23jk}(W_{jkt})$$

Where, for all $a \in \{1, 2\}$ and $b \in \{1, 2, 3\}$, $m_{abjk}(\cdot)$ is a vectors of unknown real valued univariate functions. Now using variation of the panel data differencing technique employed by (citation later), defining $c^* = \max\{c_1, c_2\}$ where $t > c^*$, we have the following,

$$Y_{jt} - Y_{jc^*} = [Z'_{jt} - Z'_{jc^*}]\beta_1 + \sum_{d=1}^{D_x}[h_d(X_{jdt}) - h_d(X_{jdc^*})] + \sum_{l=1}^{D_v}[f_{jl}(V_{jlt}) - f_{jl}(V_{jlc^*})] + u_{jt} - u_{jc^*}$$

Adopting the notation that for any finite dimensional random vector A_{jt} define $\Delta A_{jt} \equiv A_{jt} - A_{jc^*}$ we can rewrite the preceding as,

$$\Delta Y_{jt} = \Delta Z'_{jt}\beta_1 + \sum_{d=1}^{D_x} \Delta h_j(X_{jt}) + \sum_{l=1}^{D_v} \Delta f_{jl}(V_{jlt}) + \Delta u_{jt}$$

As a result of the foregoing assumptions and definitions, we can rewrite the model as the following,

$$\Delta Y_{jt} = \Delta Z'_{jt}\beta_1 + \sum_{d=1}^{D_x} \Delta h_d(X_{jt}) + \sum_{l=1}^{D_v} \Delta f_{jl}(V_{jlt}) + \Delta u_{jt} \tag{7}$$

$$Z_{1jt} = \sum_{k=1}^{c_1} m_{11jk}(Z_{1jt-k}) + \sum_{k=1}^{p_2} m_{12jk}(Z_{2jpt}) + \sum_{k=1}^{w} m_{13jk}(W_{jkt}) + V_{1jt} \tag{8}$$

$$X_{jt} = \sum_{k=1}^{c_2} m_{21jk}(X_{1jt-k}) + \sum_{k=1}^{p_2} m_{22jk}(Z_{2jpt}) + \sum_{k=1}^{w} m_{23jk}(W_{jkt}) + V_{2jt} \tag{9}$$

$$E(\Delta u_{jt}|\{Z_{j_k}, W_{jk}, X_{jk}, V_{jk}\}_{k=1}^{t}) = 0 \tag{10}$$

$$E(V_{1jt}|Z_{1jt-1}, \ldots, Z_{1jt-c_1}, Z_{2jt}, W_{jt}) = 0 \tag{5}$$

$$E(V_{2jt}|X_{1jt-1}, \ldots, X_{1jt-c_2}, Z_{2jt}, W_{jt}) = 0 \tag{6}$$

Remarks :

a.) In equation (7), by differencing equation (1), including the control function, and imposing additivity, we have eliminated fixed effect e_j , but in return we now have a collection of equations $\{f_{j1}(\cdot), f_{j2}(\cdot), \ldots, f_{jD_v}(\cdot)\}$ unique to each cross-section j, a fate seemingly even worse than the presence of fixed effects.

b.) In the following I will show that using a variation on the identification technique of Manzan and Zerom (2005) SaPL, it is possible under mild conditions to identify and estimate finite dimensional parameter β_1 and hopefully achieve $\sqrt{n}$ asymptotic normality without having to jointly estimate either $h(X_{jt})$ or $f_j(V_{jt})$,

c.) Note the imposition of additivity in (8) and (9) is intended to avoid the curse of dimensionality, and to take advantage of the increases in efficiency that typically are associated with imposing an additive structure.

Next in preparation for the following lemma regarding the identification of parameter vector β_1 I define the following density ratios,

$$\phi_{jt} = \frac{\prod_{d=1}^{D_x} p(X_{jdt}, X_{jdc^*}) \prod_{l=1}^{D_v} p(V_{jlt}, V_{jlc^*})}{p(X_{jt}, X_{jc^*}, V_{jt}, V_{jc^*})}$$

where for any random vector A_j, $p(A_j)$ is the joint density of that vector. Define following conditional expectations,

$$H^1_{jd}(\Delta Z_{jt}) = E[\phi_{jt}\Delta Z_{jt}|X_{jdt}, X_{jdc^*}] \qquad H^2_{jl}(\Delta Z_{jt}) = E[\phi_{jt}\Delta Z_{jt}|V_{jlt}, V_{jlc^*}]$$

$$H^1_{jd}(\Delta Y_{jt}) = E[\phi_{jt}\Delta Y_{jt}|X_{jdt}, X_{jdc^*}] \qquad H^2_{jl}(\Delta Y_{jt}) = E[\phi_{jt}\Delta Y_{jt}|V_{jlt}, V_{jlc^*}]$$

$$H_j(\Delta Z_{jt}) = \sum_{d=1}^{D_x} H^1_{jd}(\Delta Z'_{jt}) + \sum_{l=1}^{D_v} H^2_{jl}(\Delta Z_{jt}) \qquad H_j(\Delta Y_{jt}) = \sum_{d=1}^{D_x} H^1_{jd}(\Delta Y_{jt}) + \sum_{l=1}^{D_v} H^2_{jl}(\Delta Y_{jt})$$

Lastly, define the following collection of vectors,

$$H(\Delta Y_t) = [\ H_1(\Delta Y_{1t})\ \ H_2(\Delta Y_{2t})\ \ \cdots\ \ H_q(\Delta Y_{qt})\ ]' \qquad H(\Delta Z_t) = [\ H_1(\Delta Z_{1t})\ \ H_2(\Delta Z_{2t})\ \ \cdots\ \ H_q(\Delta Z_{qt})\ ]'$$

$$\Delta Y_t = [\ \Delta Y_{1t}\ \ \Delta Y_{2t}\ \ \cdots\ \ \Delta Y_{qt}\ ]' \qquad \Delta Z_t = [\ \Delta Z_{1t}\ \ \Delta Z_{2t}\ \ \cdots\ \ \Delta Z_{qt}\ ]' \qquad \Delta u_t = [\ \Delta u_{1t}\ \ \Delta u_{2t}\ \ \cdots\ \ \Delta u_{qt}\ ]'$$

Now that all necessary definitions have been made, the following lemma gives the sufficient conditions for the identification of parameter β_1,

Lemma 1. *Letting $\phi_t = diag\big(\{\phi_{jt}\ \}_{j=1}^q\big)$, if*

i.) For all $d \in \{1, 2, \ldots, D_x\}$, $l \in \{1, 2, \ldots, D_v\}$, $j \in \{1, 2, \cdots, q\}$, and $t \in \{c^* + 1, c^* + 2, \ldots, T\}$

$$E\big[h_d(X_{jdt})\big] = E\big[h_d(X_{jdc^*})\big] \quad \textit{and} \quad E\big[f_{jl}(V_{jlt})\big] = E\big[f_{jl}(V_{jlc^*})\big]$$

ii.) The following matrix is positive semi definite,

$$E\Big([\Delta Z_t - H(\Delta Z_t)]'[\Delta Z_t - H(\Delta Z_t)]\Big)$$

Then β_1 is identified, in particular,

$$\beta_1 = E\Big([\Delta Z_t - H(\Delta Z_t)]'\phi_t[\Delta Z_t - H(\Delta Z_t)]\Big)^{-1} E\Big([\Delta Z_t - H(\Delta Z_t)]'\phi_t[\Delta Y_t - H(\Delta Y_t)]\Big)$$

Proof. I begin with a series of preliminary results, which combine to show the main result. First for $d \neq k$ consider,

$$E\big[\phi_{jt}\Delta h_k(X_{jkt})|X_{jdt}, X_{jdc^*}\big] = \int \phi_{jt}\Delta h_k(X_{jkt})p(X_{jt}, X_{jc^*}, V_{jt}, V_{jc^*})p(X_{jdt}, X_{jdc^*})^{-1}dX_{j-dt}X_{j-dc^*}V_{jt}V_{jc^*}$$

$$= \int \Delta h_k(X_{jkt})p(X_{jkt}, X_{jkc^*})dX_{jkt}X_{jkc^*} \prod_{a\notin\{k,d\}} \int p(X_{jat}, X_{jac^*})dX_{jat}X_{jac^*} \prod_{l=1}^{D_v} \int p(V_{jlt}, V_{jlc^*})dV_{jlt}V_{jlc^*}$$

$$= E[\Delta h_k(X_{jkt})] = E[h_k(X_{jkt})] - E[h_k(X_{jkc^*})] = 0$$

Similarly by repeating arguments and letting $d \neq l$ consider,

$$E\big[\phi_{jt}\Delta f_{jl}(V_{jlt})|V_{jdt}, V_{jdc^*}\big] = E\big[\Delta f_{jl}(V_{jlt})\big] = E\big[f_{jl}(V_{jlt})\big] - E\big[f_{jl}(V_{jlc^*})\big] = 0$$

Next consider

$$\begin{aligned}
E\big[\phi_{jt}\Delta h_k(X_{jkt})|V_{jlt}, V_{jlc^*}\big] &= \int \phi_{jt}\Delta h_k(X_{jkt})p(X_{jt}, X_{jc^*}, V_{jt}, V_{jc^*})p(V_{jlt}, V_{jlc^*})^{-1}dX_{jt}X_{jc^*}V_{j-lt}V_{j-lc^*} \\
&= \int \Delta h_k(X_{jkt})p(X_{jkt}, X_{jkc^*})dX_{jkt}X_{jkc^*}\prod_{a\neq k}\int p(X_{jat}, X_{jac^*})dX_{jat}X_{jac^*}\prod_{b\neq l}\int p(V_{jbt}, V_{jbc^*})dV_{jbt}V_{jbc^*} \\
&= E[\Delta h_k(X_{jkt})] = E[h_k(X_{jkt})] - E[h_k(X_{jkc^*})] = 0
\end{aligned}$$

Similarly by repeating arguments consider,

$$E\big[\phi_{jt}\Delta f_{jl}(V_{jlt})|X_{jdt}, X_{jdc^*}\big] = E\big[\Delta f_{jl}(V_{jlt})\big] = E\big[f_{jl}(V_{jlt})\big] - E\big[f_{jl}(V_{jlc^*})\big] = 0$$

Furthermore,

$$\begin{aligned}
E\big[\phi_{jt}|X_{jdt}, X_{jdc^*}\big] &= \int \phi_{jt}p(X_{jt}, X_{jc^*}, V_{jt}, V_{jc^*})p(X_{jdt}, X_{jdc^*})^{-1}dX_{j-dt}X_{j-dc^*}V_{jt}V_{jc^*} \\
&= \prod_{a\neq d}\int p(X_{jat}, X_{jac^*})dX_{jat}X_{jac^*}\prod_{l=1}^{D_v}\int p(V_{jbt}, V_{jbc^*})dV_{jbt}V_{jbc^*} \\
&= 1
\end{aligned}$$

Consequently,

$$\begin{aligned}
H^1_{jd}(\Delta Y_{jt}) &= E\big[\phi_{jt}\Delta Y_{jt}|X_{jdt}, X_{jdc^*}\big] \\
&= E\big[\phi_{jt}\Delta Z'_{jt}|X_{jdt}, X_{jdc^*}\big]\beta_1 + \Delta h_d(X_{jdt})E\big[\phi_{jt}|X_{jdt}, X_{jdc^*}\big] + \sum_{a\neq d}^{D_x}E\big[\phi_{jt}\Delta h_a(X_{jat})|X_{jdt}, X_{jdc^*}\big] \\
&\qquad + \sum_{l=1}^{D_v}E\big[\phi_{jt}\Delta f_{jl}(V_{jlt})|X_{jdt}, X_{jdc^*}\big] + E\big[\phi_{jt}\Delta u_{jt}|X_{jdt}, X_{jdc^*}\big] \\
&= H^1_{jd}(\Delta Z'_{jt})\beta_1 + \Delta h_d(X_{jdt}) + E\big[\phi_{jt}E\big(\Delta u_{jt}|\{Z_{j_k}, W_{jk}, X_{jk}, V_{jk}\}^t_{k=1}\big)|X_{jdt}, X_{jdc^*}\big] \\
&= H^1_{jd}(\Delta Z'_{jt})\beta_1 + \Delta h_d(X_{jdt})
\end{aligned}$$

Similarly by repeating arguments,

$$H^2_{jl}(\Delta Y_{jt}) = H^2_{jd}(\Delta Z_{jt})\beta_1 + \Delta f_{jl}(V_{jlt})$$

Furthermore

$$
\begin{aligned}
H_j(\Delta Y_{jt}) &= \sum_{d=1}^{D_x} H^1_{jd}(\Delta Y_{jt}) + \sum_{l=1}^{D_v} H^2_{jl}(\Delta Y_{jt}) \\
&= \left[\sum_{d=1}^{D_x} H^1_{jd}(\Delta Z'_{jt}) + \sum_{l=1}^{D_v} H^2_{jl}(\Delta Z'_{jt})\right]\beta_1 + \sum_{d=1}^{D_x} \Delta h_d(X_{jt}) + \sum_{l=1}^{D_v} \Delta f_{jl}(V_{jlt}) \\
&= H_j(\Delta Z_{jt})\beta_1 + \sum_{d=1}^{D_x} \Delta h_d(X_{jt}) + \sum_{l=1}^{D_v} \Delta f_{jl}(V_{jlt})
\end{aligned}
$$

Consequently

$$
\sqrt{\phi_{jt}}[\Delta Y_{jt} - H_j(\Delta Y_{jt}) = \sqrt{\phi_{jt}}\big[\Delta Z_{jt} - H_j(\Delta Z_{jt})\big]\beta_1 + \sqrt{\phi_{jt}}\Delta u_{jt}
$$

and stacking these equations into a vector,

$$
\sqrt{\phi_t}[\Delta Y_t - H(\Delta Y_t) = \sqrt{\phi_t}\big[\Delta Z_t - H(\Delta Z_t)\big]\beta_1 + \sqrt{\phi_t}\Delta u_t
$$

The result follows since $E\big[\sqrt{\phi_t}E\big(\Delta u_t|\{Z_{j_k}, W_{jk}, X_{jk}, V_{jk}\}_{k=1}^{t}\big)\big] = 0$ □

Estimation

Step One:

- Case 1: (Selection), For all $j \in \{1, 2, \ldots, q\}$, W_{jt} is an unknown subset of W_t, consequently equations (2) and (3) are estimated with the incorporation of a variable selection procedure as a first/preliminary step.
- Case 2: (No selection), For all $j \in \{1, 2, \ldots, q\}$, W_{jt} is a known subset of W_t, consequently equations (2) and (3) are estimated without variable selection perhaps as a two step procedure, like the Spline Kernel Backfitting procedure of Wang and Yang (2007) AoS or one of its successors since their paper has mistakes.
- In either case the estimation of (2) and (3) will generate residual vectors;

$$
\hat{V}_t = [\ \hat{V}'_{1t}\ \ \hat{V}'_{2t}\ \ \cdots\ \ \hat{V}'_{qt}\ ]'
$$

Step Two:

- 2.1 : Obtain Rosenblatt Kernel Density Estimates of $p(X_{jdt}, X_{jdc^*})$, $p(V_{jlt}, V_{jlc^*})$, $p(X_{jt}, X_{jc^*}, V_{jt}, V_{jc^*})$ using $\{X_{jt}, X_{jc^*}, \hat{V}_{jt}, \hat{V}_{jc^*}\}_{t=c^*+1}^{T}$.

- 2.2 : Form estimated density ratio $\hat{\phi}_{jt}$ with densities estimated in previous step.

Step Three:

- 3.1: Obtain Nadaraya Watson estimates of $H^1_{jd}(\Delta Z_{jt})$, $H^2_{jl}(\Delta Z_{jt})$, $H^1_{jd}(\Delta Y_{jt})$, $H^2_{jl}(\Delta Y_{jt})$, and construct

$$\hat{H}_j(\Delta Z_{jt}) = \sum_{d=1}^{D_x} \hat{H}^1_{jd}(\Delta Z'_{jt}) + \sum_{l=1}^{D_v} \hat{H}^2_{jl}(\Delta Z_{jt}) \qquad \hat{H}_j(\Delta Y_{jt}) = \sum_{d=1}^{D_x} \hat{H}^1_{jd}(\Delta Y_{jt}) + \sum_{l=1}^{D_v} \hat{H}^2_{jl}(\Delta Y_{jt})$$

- 3.2: Let $T^* = T - c^* - 1$ and construct the following vectors using estimates from previous steps.

$$\begin{aligned}
\hat{H}(\Delta Y_t) &= [\ \hat{H}_1(\Delta Y_{1t})\ \ \hat{H}_2(\Delta Y_{2t})\ \ \cdots\ \ \hat{H}_q(\Delta Y_{qt})\ ]' \\
\hat{H}(\Delta Z_t) &= [\ \hat{H}_1(\Delta Z_{1t})\ \ \hat{H}_2(\Delta Z_{2t})\ \ \cdots\ \ \hat{H}_q(\Delta Z_{qt})\ ]' \\
\Delta Z &= [\ \Delta Z'_{c^*+1}\ \ \Delta Z'_{c^*+2}\ \ \cdots\ \ \Delta Z'_T\ ]' \\
\Delta Y &= [\ \Delta Y_{c^*+1}\ \ \Delta Y_{c^*+2}\ \ \cdots\ \ \Delta Y_T\ ]' \\
\hat{H}(\Delta Z) &= \big[\ \hat{H}(\Delta Z_{c^*+1})'\ \ \hat{H}(\Delta Z_{c^*+2})'\ \ \cdots\ \ \hat{H}(\Delta Z_T)'\ \big] \\
\hat{H}(\Delta Y) &= \big[\ \hat{H}(\Delta Y_{c^*+1})\ \ \hat{H}(\Delta Y_{c^*+2})\ \ \cdots\ \ \hat{H}(\Delta Y_T)\ \big]
\end{aligned}$$

Step Four:

- 4.1: Let $\hat{\phi} = diag\big(\{\hat{\phi}_t\}^T_{c^*+1}\big)$ Calculate $\hat{\beta}_1$ as follows,

$$\hat{\beta}_1 = \Big([\Delta Z - \hat{H}(\Delta Z)]'\hat{\phi}[\Delta Z - \hat{H}(\Delta Z)]\Big)^{-1}\Big([\Delta Z - \hat{H}(\Delta Z)]'\hat{\phi}[\Delta Y - \hat{H}(\Delta Y)]\Big)$$