Estimation of a Partially Linear Seemingly Unrelated Regressions Model

Application to a Translog Cost System

Xin Geng* & Kai Sun†

*Nankai University †Shanghai University

Motivation

- translog cost system with one input (simplified)

$$\ln C = \ln\Theta(Z) + \beta_1 \ln W + \frac{1}{2}\beta_2(\ln W)^2 + u_1$$
$$S = \beta_1 + \beta_2 \ln W + u_2$$

  - C: total cost; W: input price
  - $\Theta(\cdot)$: efficiency parameter of environmental factors Z
  - S: input share obtained by Shephard's lemma
- more efficient estimating the system as a whole

Abstract

- motivated by estimation of a translog cost system
- propose more efficient estimators for a partially linear SUR model
- combine profile least-square (Robinson, 1988) and SUR (Zellner, 1962)
- establish asymptotic normality and efficiency for both the linear and nonparametric estimators
- covariance decomposition method matters in terms of nonparametric efficiency, i.e., Cholesky prevails Spectral

A partially linear SUR model

Consider a system of m equations

$$y_{si} = \theta_s(z_{si}) + x_{si}'\beta_s + u_{si}, \tag{1}$$

for $i = 1, \cdots, n$ and $s = 1, \cdots, m$.

- θ_s is unknown and x_{si} enters linearly
- errors across equations are correlated s.t. $\mathbb{E}(u_{si}u_{li}) = \sigma_{sl}$ but not across time $\mathbb{E}(u_{si}u_{lt}) = 0$

Moment conditions

1. as $\mathbb{E}(u_{si} \mid z_{si}, x_{si}) = 0$, letting $g_{sw}(z) \equiv \mathbb{E}(w_{si} \mid z_{si} = z)$ for $w = x, y$,

$$g_{sy}(z_{si}) = \theta_s(z_{si}) + g_{sx}(z_{si})'\beta_s, \tag{2}$$
$$y_{si}^* = x_{si}^{*\prime}\beta_s + u_{si}, \tag{3}$$
$$\beta_s = \mathbb{E}\left(x_{si}^* x_{si}^{*\prime}\right)^{-1}\mathbb{E}\left(x_{si}^* y_{si}^*\right), \tag{4}$$

where $y_{si}^* = y_{si} - \mathbb{E}(y_{si} \mid z_{si})$ and $x_{si}^* = x_{si} - \mathbb{E}(x_{si} \mid z_{si})$.

2. GLS moment condition is

$$\beta = \mathbb{E}\left(x_i^* \Sigma_m^{-1} x_i^{*\prime}\right)^{-1}\mathbb{E}\left(x_i^* \Sigma_m^{-1} y_i^*\right), \tag{5}$$

where $\beta = (\beta_1', \cdots, \beta_m')'$, $y_i^* = (y_{1i}^*, \cdots, y_{mi}^*)'$, $x_i^* = \begin{pmatrix} x_{1i}^* & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & x_{mi}^* \end{pmatrix}$, $u_i = (u_{1i}, \cdots, u_{mi})'$, and $\Sigma_m = \mathrm{Var}(u_i) = \{\sigma_{sl}\}_{s,l=1}^{m,m}$.

Estimation

- by Robinson (1988), single-equation estimator for β_s is

$$\widehat{\beta}_s = \left(\sum_{i=1}^n \widehat{x}_{si}^* \widehat{x}_{si}^{*\prime}\right)^{-1} \sum_{i=1}^n \widehat{x}_{si}^* \widehat{y}_{si}^*. \tag{6}$$

- by Zellner (1962), our SUR estimator for β is

$$\widehat{\beta}_{\text{sur}} = \left(\sum_{i=1}^n \widehat{x}_i^* \widehat{\Sigma}_m^{-1} \widehat{x}_i^{*\prime}\right)^{-1}\left(\sum_{i=1}^n \widehat{x}_i^* \widehat{\Sigma}_m^{-1} \widehat{y}_i^*\right). \tag{7}$$

where $\widehat{x}_{si}^*$ and $\widehat{y}_{si}^*$ are residuals from single-equation nonparametric regression for $\mathbb{E}(x_{si} \mid z_{si})$ and $\mathbb{E}(y_{si} \mid z_{si})$, and $\widehat{\Sigma}_m = \{\widehat{\sigma}_{sl}\}_{s,l=1}^{m,m}$ with $\widehat{\sigma}_{sl} = \frac{1}{n}\sum_{i=1}^n \widehat{u}_{si}\widehat{u}_{li}$ and $\widehat{u}_{si} = \widehat{y}_{si}^* - \widehat{x}_{si}^{*\prime}\widehat{\beta}_s$.

Asymptotic normality

Theorem 1. Under Assumptions A1–A4, we have

$$\sqrt{n}\left(\widehat{\beta}_{\text{sur}} - \beta\right) \xrightarrow{d} \mathcal{N}(0, V), \tag{8}$$

where $V = \left(\mathbb{E}\left(x_i^* \Sigma_m^{-1} x_i^{*\prime}\right)\right)^{-1}$.

Nonparametric estimator

$$\widehat{\theta}_s(z) = \widehat{g}_{sy}(z) - \widehat{g}_{sx}(z)'\widehat{\beta}_s, \qquad \widetilde{\theta}_s(z) = \widehat{g}_{sy}(z) - \widehat{g}_{sx}(z)'\widehat{\beta}_{s,\text{sur}}.$$

Theorem 2. Under Assumption A1–A4 and assuming that $\mathbb{E}\left(|u_{si}|^{2+\delta} \mid z_{si}, x_{si}\right) \leq C$ for some $\delta > 0$, we have

$$\sqrt{nh_s^{p_s}}\left(\widetilde{\theta}_s(z) - \theta_s(z) - b_{s,1}(z)\right) \xrightarrow{d} \mathcal{N}\left(0, \mathcal{V}_{s,1}(z)\right), \tag{9}$$

where $b_{s,1}(z) = O_p(h_s^{p_s})$ and $\mathcal{V}_{s,1}(z) \equiv f_{sz}^{-1}(z)\sigma_{ss}\int K_s^2(\gamma)\,\mathrm{d}\gamma$.

Efficiency discussion

- by Zellner (1962), $\widehat{\beta}_{\text{sur}}$ is efficient relative to $\widehat{\beta}_s$ as $\mathrm{AVar}(\widehat{\beta}_{\text{sur}}) \leq \mathrm{AVar}(\widehat{\beta}_s)$.
- $\widehat{\theta}_s(z)$ and $\widetilde{\theta}_s(z)$ are asymptotically equivalent; cross-equation correlation is not effectively explored.

More efficient estimation of $\theta_s(\cdot)$

- by Martins-Filho and Yao (2009) and Su et al. (2013), pre-whitening (rendering errors spherical) for nonparametric estimation also matters
- let $Y_s = (y_{s1}, ..., y_{sn})'$, $X_s = (x_{s1}, ..., x_{sn})'$, $\Theta_s(Z_s) = (\theta_s(z_{s1}), ..., \theta_s(z_{sn}))'$,

$$Y_s - X_s\beta_s = \Theta_s(Z_s) + U_s,$$
$$Y - X\beta = \Theta(Z) + U, \tag{10}$$

where $Y = \left(Y_1', ..., Y_m'\right)'$, $X = \begin{pmatrix} X_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & X_m \end{pmatrix}$, $\Theta(Z) = \left(\Theta_1(Z_1)', ..., \Theta_m(Z_m)'\right)'$.

- let $\Sigma \equiv \mathbb{E}(UU') = PP'$, $V \equiv P^{-1}$, $\mathcal{E} \equiv VU$ with $\mathbb{E}(\mathcal{E}\mathcal{E}') = I_{mn}$.
- as $\Sigma = \Sigma_m \otimes I_n$, we have $P = P_m \otimes I_n$ and $V = V_m \otimes I_n$
- let $Y^* \equiv H\Theta(Z) + \mathcal{E}$ where $H \equiv \mathrm{diag}(V)$, for each equation

$$Y_s^* \equiv v_{ss}\Theta_s(Z_s) + \mathcal{E}_s. \tag{11}$$

- local linear estimation of an estimated Y_{si}^*/v_{ss} on z_{si} would yield our SUR nonparametric estimator $\widetilde{\theta}_{s,\text{sur}}(z)$.
- Y_{si}^*/v_{ss} can be estimated by $\widehat{Y}_{si}^*/\widehat{v}_{ss} \equiv \widehat{\theta}_s(z_{si}) + \widehat{\mathcal{E}}_{si}/\widehat{v}_{ss}$ and $\widehat{\mathcal{E}}_{si} \equiv \sum_{l=1}^m \widehat{v}_{sl}\widehat{U}_{li}$.

Efficiency of $\widetilde{\theta}_{s,\text{sur}}(z)$

Theorem 3. Under Assumptions A1–A3, and if for any $s, l = 1, \cdots, m$, $h_{1s}/h_{2l} \to 0$, $nh_{1s}^{p_s} \to \infty$ and $nh_{2s}^{p_s+2r_s} \to C \in [0, \infty]$ as $n \to \infty$, we have

$$\sqrt{nh_{2s}^{p_s}}\left(\widetilde{\theta}_{s,\text{sur}}(z) - \theta_s(z) - b_{s,2}(z)\right) \xrightarrow{d} \mathcal{N}\left(0, \mathcal{V}_{s,2}(z)\right), \tag{12}$$

where $b_{s,2}(z) = O_p(h_{2s}^{r_s})$ and $\mathcal{V}_{s,2}(z) \equiv f_{sz}^{-1}(z)v_{ss}^{-2}\int K_s^2(\gamma)\,\mathrm{d}\gamma$.

- $\widetilde{\theta}_{s,\text{sur}}(z)$ is more efficient relative to $\widehat{\theta}_s(z)$ as $\mathcal{V}_{s,2} \leq \mathcal{V}_{s,1}$ given that $\Sigma_m = P_m P_m'$ and $\sigma_{ss} = \sum_{l=1}^m p_{sl}^2 \geq p_{ss}^2 = v_{ss}^{-2}$.

Efficiency dependency on square root choice

- efficiency of $\widetilde{\theta}_{s,\text{sur}}(z)$ depends on $\mathcal{V}_{s,2}$ via v_{ss}, which varies with the square root choice for Σ_m
- denote terms induced using the Spectral decomp. by adding a superscript S, and those without for the Cholesky decomp. e.g, $\Sigma_m = P_m^S P_m^S$ with $P_m^S = P_m^{S\prime}$.
- Theorem 3 remains true for both $\widetilde{\theta}_{s,\text{ sur}}(z)$ and $\widetilde{\theta}_{s,\text{ sur}}^S(z)$.
- we can show algebraically that
  1. moving the position of the s^{th} equation to a later spot in the system reduces the corresponding $(v_{ss})^{-2}$.
  2. $(v_{mm})^{-2} \leq \left(v_{mm}^S\right)^{-2}$.
- this result suggests that it is optimal to estimate the nonparametric part using the Cholesky decomposition and always place the equation of interest at the end of the system.

Simulations

Consider the following DGPs

$$y_{1i} = \theta_1(z_{1i}) + \beta_1 x_{1i} + u_{1i},$$
$$y_{2i} = \theta_2(z_{2i}) + \beta_2 x_{2i} + u_{2i},$$

where z_{1i} and z_{2i} are i.i.d. $\mathcal{U}[0, 2]$, $\theta_1(z_{1i}) = \sin(z_{1i})$, $\theta_2(z_{2i}) = \cos(z_{2i})$, $\beta_1 = 1$, $\beta_2 = 2$, $x_{si} = \varrho z_{si} + e_{si}$, $\varrho = 0.6$, $e_{si} \sim$ i.i.d. $\mathcal{N}(1, 0.5^2)$ $\forall s = 1, 2$, and $(u_{1i}, u_{2i})' \sim$ i.i.d. multivariate normal $\mathcal{N}(0, \Omega)$ with $\Omega = \{\sigma_{sl}\}_{s,l=1}^{2,2}$, $\sigma_{11} = \sigma_{22} = 1$, and $\sigma_{12} = \sigma_{21} = 0.6$.

Table 1: Finite Sample Performance *with* Cross-Equation Correlation ($\boldsymbol{\sigma_{12} = 0.6}$)

	β_1			β_2			$\theta_1(\cdot)$			$\theta_2(\cdot)$		
n	Bias	Var	MSE	Bias	Var	MSE	Abias	Avar	AMSE	Abias	Avar	AMSE
Partially linear SUR — $\widehat{\beta}_{s,\text{sur}}, \widetilde{\theta}_s(\cdot)$												
100	0.0046	0.0298	0.0298	0.0097	0.0309	0.0310	-0.0033	0.0421	0.1292	-0.0093	0.0414	0.1306
200	0.0050	0.0128	0.0128	-0.0017	0.0140	0.0140	0.0007	0.0223	0.0590	0.0104	0.0234	0.0634
400	-0.0011	0.0063	0.0063	0.0032	0.0066	0.0066	0.0052	0.0118	0.0306	-0.0011	0.0121	0.0322
800	0.0022	0.0034	0.0034	0.0039	0.0033	0.0034	-0.0021	0.0072	0.0170	-0.0031	0.0071	0.0169
1600	0.0012	0.0016	0.0016	0.0020	0.0017	0.0017	-0.0008	0.0040	0.0088	0.0001	0.0040	0.0090
Robinson's single-equation — $\widehat{\beta}_s, \widehat{\theta}_s(\cdot)$												
100	0.0025	0.0469	0.0469	0.0087	0.0464	0.0465	0.0000	0.0438	0.1754	-0.0078	0.0435	0.1727
200	0.0063	0.0198	0.0198	-0.0045	0.0211	0.0211	-0.0014	0.0230	0.0785	0.0148	0.0242	0.0822
400	0.0015	0.0094	0.0094	0.0045	0.0100	0.0100	0.0010	0.0122	0.0393	-0.0034	0.0124	0.0413
800	0.0028	0.0051	0.0051	0.0052	0.0052	0.0052	-0.0031	0.0073	0.0214	-0.0052	0.0073	0.0220
1600	0.0010	0.0025	0.0025	0.0002	0.0024	0.0024	-0.0005	0.0041	0.0113	0.0029	0.0040	0.0111

Nonparametric SUR estimator — $\widetilde{\theta}_{s,\text{sur}}(\cdot)$

	First equation first						First equation second					
	$\theta_1(\cdot)$			$\theta_2(\cdot)$			$\theta_1(\cdot)$			$\theta_2(\cdot)$		
n	Abias	Avar	AMSE	Abias	Avar	AMSE	Abias	Avar	AMSE	Abias	Avar	AMSE
100	-0.0240	0.0376	0.1261	-0.0215	0.0241	0.1142	-0.0216	0.0281	0.1164	-0.0243	0.0329	0.1234
200	-0.0218	0.0207	0.0586	-0.0031	0.0130	0.0535	-0.0201	0.0141	0.0518	-0.0060	0.0180	0.0586
400	-0.0167	0.0110	0.0304	-0.0154	0.0068	0.0276	-0.0136	0.0073	0.0265	-0.0170	0.0096	0.0305
800	-0.0228	0.0057	0.0162	-0.0168	0.0036	0.0139	-0.0186	0.0039	0.0142	-0.0183	0.0055	0.0158
1600	-0.0171	0.0030	0.0081	-0.0117	0.0020	0.0073	-0.0137	0.0019	0.0069	-0.0142	0.0029	0.0082

Note: (1) This table compares the effects of sample size on the partially linear SUR and Robinson equation-by-equation models when there exists cross-equation correlation (i.e., $\sigma_{12} = 0.6$). (2) For the partially linear SUR model, the parametric components are estimated with the feasible SUR estimator, and the nonparametric components are estimated with the two-step nonparametric estimator.

References

Martins-Filho, C., Yao, F., 2009. Nonparametric regression estimation with general parametric error covariance. Journal of Multivariate Analysis 100 (3), 309–333.

Robinson, P., 1988. Root-n-consistent semiparametric regression. Econometrica 56 (4), 931–954.

Su, L., Ullah, A., Wang, Y., 2013. Nonparametric regression estimation with general parametric error covariance: A more efficient two-step estimator. Empirical Economics 45, 1009–1024.

Zellner, A., 1962. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. Journal of the American Statistical Association 57 (298), 348–368.

Xin Geng (x.geng@outlook.com)

sites.google.com/site/xingeng2015

Nankai University