Appendix

Appendix 1: Introduction to Estimation of Panel Data

A common formulation of fixed effect regression assumes that differences across units can be captured in differences in the constant term. Thus, in Eq. (7), each α_i is an unknown parameter to be estimated. Let $\mathbf{y}_i$ and $\mathbf{X}_i$ be the T observations for the ith unit, and let $\boldsymbol{\varepsilon}_i$ be associated $T \times 1$ vector of disturbances. Then we can write Eq. (7) as

$$\mathbf{y}_i = \mathbf{i}\alpha_i + \mathbf{X}_i\beta + \boldsymbol{\varepsilon}_i . \tag{A1}$$

Collecting these terms gives

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_n \end{bmatrix} = \begin{bmatrix} \mathbf{i} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{i} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{i} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} + \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \vdots \\ \mathbf{X}_n \end{bmatrix} \beta + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_n \end{bmatrix} \tag{A2}$$

or

$$\mathbf{y} = \begin{bmatrix} \mathbf{d}_1 & \mathbf{d}_2 & \cdots & \mathbf{d}_n & \mathbf{X} \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \boldsymbol{\varepsilon} , \tag{A3}$$

where $\mathbf{d}_i$ is a dummy variable indicating the ith unit. Let the $nT \times n$ matrix $\mathbf{D} = \begin{bmatrix} \mathbf{d}_1 & \mathbf{d}_2 & \cdots & \mathbf{d}_n \end{bmatrix}$. Then, assembling all nT rows gives

$$\mathbf{y} = \mathbf{D}\,\alpha + \mathbf{X}\,\beta + \boldsymbol{\varepsilon} . \tag{A4}$$

It is a classical regression model and usually referred to as the least squares dummy variable (LSDV) model.

In other settings, random effect regression might be more appropriate because the individual specific constant terms are viewed as randomly distributed across cross-sectional units. Consider, then, a reformulation of the model

$$\mathbf{y}_{i,t} = \alpha + \beta' \mathbf{x}_{i,t} + u_{i,t} + \boldsymbol{\varepsilon}_{i,t} , \tag{A5}$$

in which there are K regressors in addition to the constant term. The component u_i is the random disturbance characterizing the *i*th observation and is constant through time. Hence, we assume further that

$$
\begin{aligned}
E[\varepsilon_{i,t}] &= E[u_{i,t}] = 0, \\
E[\varepsilon_{i,t}^2] &= \sigma_\varepsilon^2, \\
E[u_i^2] &= \sigma_u^2, \\
E[\varepsilon_{i,t} u_j] &= 0 \quad \text{for all } i, t, \text{and } j, \\
E[\varepsilon_{i,t} \varepsilon_{j,s}] &= 0 \quad \text{if } t \neq s \text{ or } i \neq j, \\
E[u_i u_j] &= 0 \quad \text{if } i \neq j.
\end{aligned}
\tag{A6}
$$

As well, it is useful to view the formulation of the model in blocks of T observations for observations i, $\mathbf{y}_i$, $\mathbf{X}_i$, $u_i\mathbf{i}$, and $\boldsymbol{\varepsilon}_i$. For these T observations, let $w_{i,t} = \varepsilon_{i,t} + u_i$ and $\mathbf{w}_i = [w_{i1}, w_{i2}, \cdots, w_{iT}]'$. In view of this form of $w_{i,t}$, we have what is often called an "error components model." Then, for this model,

$$
\begin{aligned}
E[w_{i,t}^2] &= \sigma_\varepsilon^2 + \sigma_u^2, \\
E[w_{i,t} w_{i,s}] &= \sigma_u^2, \ t \neq s.
\end{aligned}
\tag{A7}
$$

For the T observations for unit i, let $\mathbf{\Omega} = E[\mathbf{w}_i \mathbf{w}_j']$. Then

$$
\mathbf{\Omega} = \begin{bmatrix} \sigma_\varepsilon^2 + \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ \sigma_u^2 & \sigma_\varepsilon^2 + \sigma_u^2 & \sigma_u^2 & \cdots & \sigma_u^2 \\ & \vdots & & & \\ \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & & \sigma_\varepsilon^2 + \sigma_u^2 \end{bmatrix} = \sigma_\varepsilon^2 \mathbf{I}_T + \sigma_u^2 \mathbf{i}_T \mathbf{i}_T', \tag{A8}
$$

where $\mathbf{i}$ is a $T \times 1$ column vector of 1s. Since observations i and j are independent, the disturbance covariance matrix for the full nT observations is

$$
\mathbf{V} = \begin{bmatrix} \mathbf{\Omega} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{\Omega} & & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{\Omega} \end{bmatrix} = \mathbf{\Omega} \otimes \mathbf{I}_n. \tag{A9}
$$

For generalized least squares (GLS), we require $\mathbf{V}^{-1/2} = \mathbf{I} \otimes \mathbf{\Omega}^{-1/2}$. Therefore, we need only find $\mathbf{\Omega}^{-1/2}$, which is

$$
\mathbf{\Omega}^{-1/2} = (1/\sigma_\varepsilon)[\mathbf{I} - (\theta/T)\mathbf{ii}'], \tag{A10}
$$

where $\theta = 1 - \left(\sigma_\varepsilon \big/ \sqrt{\sigma_\varepsilon^2 + T\sigma_u^2}\right)$. Then, the transformation of $\mathbf{y}_i$ and $\mathbf{X}_i$ for GLS is

$$\mathbf{\Omega}^{-1/2}\mathbf{y}_i = \frac{1}{\sigma_\varepsilon}\begin{bmatrix} y_{i1} - \theta\,\bar{y}_{i\cdot} \\ \mathrm{y}_{i2} - \theta\,\bar{y}_{i\cdot} \\ \vdots \\ y_{iT} - \theta\,\bar{y}_{i\cdot} \end{bmatrix}, \tag{A11}$$

and likewise for the rows of $\mathbf{X}_i$. For the data set as a whole, GLS is computed by the regression of these partial deviations of $y_{i,t}$ on the same transformations of $\mathbf{x}_{i,t}$.

To specify which estimator is consistent and efficient, Hausman test is frequently used to judge. Under the hypothesis of no correlation, both OLS in the LSDV model and GLS are consistent, but OLS is inefficient; whereas, under the alternative, OLS is consistent, but GLS is not. Hausman statistic, based on Wald criterion, can be

$$W = \chi^2[K] = \left[\mathbf{b} - \hat{\beta}\right]'\hat{\mathbf{\Sigma}}^{-1}\left[\mathbf{b} - \hat{\beta}\right]. \tag{A12}$$

For $\hat{\mathbf{\Sigma}}$, we used the estimated covariance matrices of the slope estimator in the LSDV model and the estimated covariance matrix in the random effect model, excluding the constant term. W is asymptotically distributed as chi-squared with K degree of freedom under the null hypothesis.

Appendix 2: The Sample List of Provinces and Cities in the Study

Number	Province / City	Location	Number	Province / City	Location
1	Beijing City	Coastal	16	Henan	Inland
2	Tianjin City	Coastal	17	Hubei	Inland
3	Hebei	Coastal	18	Hunan	Inland
4	Shanxi	Inland	19	Guangdong	Coastal
5	Inner Mongolia	Inland	20	Guangxi	Coastal
6	Liaoning	Coastal	21	Hainan	Coastal
7	Jilin	Inland	22	Sichuan	Inland
8	Heilongjiang	Inland	23	Guizhou	Inland
9	Shanghai City	Coastal	24	Yunnan	Inland
10	Jiangsu	Coastal	25	Shaanxi	Inland
11	Zhejiang	Coastal	26	Gansu	Inland
12	Anhui	Inland	27	Qinghai	Inland
13	Fujian	Coastal	28	Ningxia	Inland
14	Jiangxi	Inland	29	Xinjiang	Inland
15	Shandong	Coastal			

Note: Chongqing City and Tibet are excluded in this table for lack of important data of variables.

Appendix 3: Goodness of Fit against Each Trial Inflation Threshold Rate

π^*	$\overline{R}^2$	π^*	$\overline{R}^2$	π^*	$\overline{R}^2$	π^*	$\overline{R}^2$	π^*	$\overline{R}^2$
0.00	0.470	5.00	0.480	10.00	0.473	15.00	0.471	20.00	0.470
0.25	0.474	5.25	0.480	10.25	0.473	15.25	0.471	20.25	0.470
0.50	0.476	5.50	0.480	10.50	0.473	15.50	0.471	20.50	0.470
0.75	0.477	5.75	0.480	10.75	0.473	15.75	0.471	20.75	0.470
1.00	0.479	6.00	0.480	11.00	0.473	16.00	0.471	21.00	0.470
1.25	0.479	6.25	0.479	11.25	0.473	16.25	0.471	21.25	0.470
1.50	0.480	6.50	0.479	11.50	0.473	16.50	0.471	21.50	0.470
1.75	0.480	6.75	0.478	11.75	0.472	16.75	0.471	21.75	0.470
2.00	0.48093	7.00	0.478	12.00	0.472	17.00	0.471	22.00	0.471
2.25	0.48138	7.25	0.477	12.25	0.472	17.25	0.471		
2.50	**0.48141**	7.50	0.477	12.50	0.472	17.50	0.471		
2.75	0.48130	7.75	0.476	12.75	0.472	17.75	0.471		
3.00	0.48124	8.00	0.476	13.00	0.472	18.00	0.471		
3.25	0.48101	8.25	0.476	13.25	0.472	18.25	0.471		
3.50	0.48078	8.50	0.475	13.50	0.472	18.50	0.471		
3.75	0.48065	8.75	0.475	13.75	0.472	18.75	0.471		
4.00	0.480	9.00	0.474	14.00	0.472	19.00	0.471		
4.25	0.480	9.25	0.474	14.25	0.471	19.25	0.470		
4.50	0.480	9.50	0.474	14.50	0.471	19.50	0.470		
4.75	0.480	9.75	0.474	14.75	0.471	19.75	0.470		

Note: Each adjusted R-square is from Eq. (5) with spline regression for the given inflation threshold. π^* is in terms of percentage. The inflation rate 2.5% strikes the highest adjusted R-square, that is, 0.48141. All R-square values are rounded to 0.001 except for some which approximates 0.481 and presented in more detailed number.