1 Nonparametric Estimation of Panel Data Models

Aim: to develop the nonparametric and semiparametric estimation of both the fixed and random effects panel data models which are robust to the misspecification in the functional form.

- **Pooled Kernel Estimation**

Consider the nonparametric panel data model

$$y_{it} = m(x_{it}) + u_{it}, \quad i = 1, 2, \cdots, n; \quad t = 1, 2, \cdots, T. \tag{1}$$

where y_{it} is the dependent variable, $x_{it} \in R^q$ is the row vector of explanatory variables and u_{it} is the error term, i.i.d. and satisfying $E[u_{it}|x_{it}] = 0$, $V[u_{it}^2|x_{it}] = \sigma^2(x_{it})$. The data (y_{it}, x_{it}) are assumed to be i.i.d. The sample n is large and T is small. The function $m(\cdot)$ is left unspecified, which is to be estimated. $E[y_{it}|x_{it} = x] = m(x)$.

1) **Local Constant Nonparametric Estimator** By a Taylor expansion,

$$y_{it} = m(x) + u_{it} + O(|x_{it} - x|) = m(x) + u_{it} + O(|h|) \quad \text{with} \lim_{n\to\infty} h = 0,$$

which still is denoted as

$$y_{it} = m(x) + u_{it},$$

that is, the term $O(|h|)$ is added with u_{it} so that for large n it is still the case that the conditional expectation of the combined error is zero. In a matrix form,

$$Y = l_{nT}m(x) + u,$$

where

$Y = (y_{11,} \cdots, y_{1T}, y_{21}, \cdots, y_{2T}, \cdots, y_{n1}, \cdots, y_{nT})' = (y_1', y_2', \cdots, y_n')'$;
$u = (u_{11,} \cdots, u_{1T}, u_{21}, \cdots, u_{2T}, \cdots, u_{n1}, \cdots, u_{nT})' = (u_1', u_2', \cdots, u_n')'$;
l_{nT} is an $nT \times 1$ vector of unit elements.
Minimize

$$\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} - a)^2 k\left(\frac{x_{it} - x}{h}\right) = (Y - al_{nT})'K(x)(Y - al_{nT})$$

and obtain **the local constant nonparametric estimator** of $m(x)$:

$$\hat{m}(x) = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T} y_{it}k\left(\frac{x_{it}-x}{h}\right)}{\sum_{i=1}^{n}\sum_{t=1}^{T} k\left(\frac{x_{it}-x}{h}\right)} = (l_{nT}'K(x)l_{nT})^{-1}l_{nT}'K(x)Y,$$

where $K(x)$ is the $nT \times nT$ diagonal matrix with the diagonal elements $K_{it} = k\left((x_{it} - x)/h\right)$, which is the kernel or weight function taking low values for x_{it} far away from x but high values for x_{it} close to x. Therefore, the nonparametric estimate $\hat{m}(x)$ is the smoothed average of y values which correspond to the x_{it} values in a small interval of x such that $(|x_{it} - x|) = O(|h|)$.

Give nonparametric estimation of **the derivatives of** $m(x)$ and **the average derivatives**: ...... for details see Pagan and Ullah (1999: P164-173), "Nonparametric Econometrics", Cambridge University Press.

The pointwise estimator of the partial derivative of $m(x)$ with respect to the j-th regressor x_j:

(i) numerical derivative:

$$\tilde{\beta}_j(x) = [\hat{m}(x + e_j h) - \hat{m}(x - e_j h)]/(2h),$$

where $e_j = (0, \cdots, 0, 1, 0, \cdots, 0)$ with the j-th element being 1;

(ii) analytical derivative:

$$\hat{\beta}_j(x) = \partial \hat{m}(x)/\partial x_j = [\partial \hat{g}(x)/\partial x_j - \hat{m}(x)\partial \hat{f}(x)/\partial x_j]/\hat{f}(x),$$

where

$$\hat{f}(x) = \frac{1}{nTh^q} \sum_{i=1}^{n} \sum_{t=1}^{T} k\left(\frac{x_{it} - x}{h}\right)$$

$$\partial \hat{f}(x)/\partial x_j = -\frac{1}{nTh^{q+1}} \sum_{i=1}^{n} \sum_{t=1}^{T} k_j^{(1)}\left(\frac{x_{it} - x}{h}\right)$$

$$\partial \hat{g}(x)/\partial x_j = -\frac{1}{nTh^{q+1}} \sum_{i=1}^{n} \sum_{t=1}^{T} y_{it} k_j^{(1)}\left(\frac{x_{it} - x}{h}\right)$$

and $k_j^{(1)}(v) = \partial k(v)/\partial v_j$ with $v = (v_1, v_2, ..., v_q)$. The two estimators are approximately the same, i.e. $\tilde{\beta}_j(x) \approx \hat{\beta}_j(x)$, since

$$\hat{m}(x \pm e_j h) = \hat{m}(x) \pm h\frac{\partial \hat{m}(x)}{\partial x_j} + O(h^2) = \hat{m}(x) \pm h\hat{\beta}_j(x) + O(h^2).$$

The estimator of the average derivative $\beta = E\beta(x)$:

(i) A direct estimator for β: $\hat{\beta} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\tilde{\beta}(x_{it})$ or $\frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\beta}(x_{it})$;

(ii) A average-based estimator by $\beta = E\beta(x) = -E[yf'(x)/f(x)]$:

$$\hat{\beta} = -\frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} y_{it} \frac{\hat{f}'(x_{it})}{\hat{f}(x_{it})}.$$

The two average derivative estimators have the following asymptotic property:

$$\sqrt{n}(\hat{\beta} - \beta) \to N(0, \Sigma/T),$$

where

$$\Sigma = E\left[\sigma_u^2(x)\left(\frac{f'(x)}{f(x)}\right)^2\right] + V(\beta(x))$$

which can be estimated by

$$\hat{\Sigma} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[\hat{\sigma}_u^2(x_{it})\left(\frac{\hat{f}'(x_{it})}{\hat{f}(x_{it})}\right)^2 + \left(\hat{\beta}(x_{it}) - \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}\hat{\beta}(x_{it})\right)^2\right].$$

Optimal bandwidth for estimating $m(x)$ and $\beta(x)$: Note that

$$\begin{aligned}
E(\hat{m}(x)|x_{it}) &= (l'_{nT}K(x)l_{nT})^{-1}l'_{nT}K(x)m^*, \\
V(\hat{m}(x)|x_{it}) &= (l'_{nT}K(x)l_{nT})^{-1}l'_{nT}\Omega(x)l_{nT}(l'_{nT}K(x)l_{nT})^{-1},
\end{aligned}$$

where $m^* = (m(x_{11}), \cdots, m(x_{nT})))'$, $\Omega(x) = K(x)\Sigma_1 K(x)$ and Σ_1 is a diagonal matrix with the diagonal elements $\sigma_u^2(x_{it}) = E[u_{it}^2|x_{it}]$. A similar argument to that in Chapter 2 (on the nonparametric estimation of regression function in cross-section case) can deduce that

$$E(\hat{m}(x)) - m(x) = \frac{h^2}{2}\kappa_2\left[m''(x) + 2m'(x)\frac{f'(x)}{f(x)}\right] + o(h^2)$$

$$V(\hat{m}(x)) = \frac{1}{nh^q}\kappa\frac{\sigma_u^2(x)}{Tf(x)} + o\left(\frac{1}{nh^q}\right)$$

and

$$E(\hat{\beta}(x)) - \beta(x) = \frac{h^2}{2}\kappa_2\left[m^{(4)}(x) - 2\beta(x)\left(\frac{f'^2(x)}{f^2(x)} - \frac{f^{(2)}(x)}{f(x)}\right) + 2\frac{m^{(3)}(x)f'(x)}{f(x)}\right] + o(h^2)$$

$$V(\hat{\beta}(x)|x_{it}) = \frac{1}{nh^{q+2}}\frac{\sigma_u^2(x)}{Tf(x)}\int k'^2(v)dv + o\left(\frac{1}{nh^{q+2}}\right)$$

for any given $x \in Supp(X)$. Therefore, the optimal bandwidth that minimize the integrated MSE of $\hat{m}(x)$ and $\hat{\beta}(x)$, respectively, are

$$h_0 = O(n^{-1/(q+4)}) \qquad and \qquad h_1 = O(n^{-1/(q+6)}).$$

For example, for $q = 1$, $h_0 = O(n^{-1/5})$ *and* $h_1 = O(n^{-1/7})$. The asymptotic properties of the local constant nonparametric estimator $\hat{m}(x)$ requires that $x \in (Supp(X))^o$, which behaves in the same way as in the cross-section case in Chapter 2. That is, the estimator has a boundary effect. Hence necessity of the local linear estimator below.

2) **Local Linear Nonparametric Estimator** By a Taylor expansion,

$$\begin{aligned} y_{it} &= m(x) + (x_{it} - x)\beta(x) + u_{it} + O\left((x_{it} - x)^2\right) \\ &= m(x) + (x_{it} - x)\beta(x) + u_{it} + O(h^2) \quad \text{with } \lim_{n\to\infty} h = 0, \end{aligned}$$

which still is denoted as

$$y_{it} = m(x) + (x_{it} - x)\beta(x) + u_{it},$$

that is, the term $O(h^2)$ is added with u_{it} so that for large n it is still the case that the conditional expectation of the combined error is zero. In a matrix form,

$$Y = Z(x)\delta(x) + u,$$

where

$Z(x)$ is an $nT \times (q+1)$ matrix with it-th element $z_{it} = (1, x_{it} - x)$;

$\delta(x) = (m(x), \beta(x)')'$ is a $(q+1) \times 1$ parametric vector.

Minimize

$$\begin{aligned} &\sum_{i=1}^{n}\sum_{t=1}^{T}(y_{it} - a - (x_{it} - x)b)^2\, k\left(\frac{x_{it} - x}{h}\right) \\ = &\left[Y - a l_{nT} - (X - l_{nT} \otimes x)b\right]' K(x) \left[Y - a l_{nT} - (X - l_{nT} \otimes x)b\right] \\ = &\left(Y - Z(x)\begin{pmatrix} a \\ b \end{pmatrix}\right)' K(x) \left(Y - Z(x)\begin{pmatrix} a \\ b \end{pmatrix}\right) \end{aligned}$$

with respect to a and b, and obtain **the local linear nonparametric estimator** of $\delta(x) = (m(x), \beta(x)')'$:

$$\begin{aligned} \tilde{\delta}(x) &= \left[\sum_{i=1}^{n}\sum_{t=1}^{T} z_{it}' z_{it} k\left(\frac{x_{it} - x}{h}\right)\right]^{-1} \sum_{i=1}^{n}\sum_{t=1}^{T} z_{it}' y_{it} k\left(\frac{x_{it} - x}{h}\right) \\ &= (Z'(x)K(x)Z(x))^{-1} Z'(x)K(x)Y, \end{aligned}$$

where $K(x)$ is defined as above. This amounts to doing LS of $\sqrt{K_{it}}y_{it}$ on $\sqrt{K_{it}}$ and $\sqrt{K_{it}}(x_{it} - x)$. Then

$$\tilde{m}(x) = (1, \mathbf{0}_q)\tilde{\delta}(x), \quad \tilde{\beta}(x) = (0, \mathbf{1}_q)\tilde{\delta}(x).$$

For the poolablity of the panel data, see Badi H. Baltagi, Javier Hidalgo and Qi Li (1996), "A nonparametric test for poolability using panel data"■Journal of Econometrics 75: 345-367.

- **A Nonparametric Fixed Effects Estimator**

The nonparametric panel data model is

$$y_{it} = \alpha_i + m(x_{it}) + u_{it}, \tag{2}$$

where y_{it} is the dependent variable, $x_{it} \in R^q$ is the row vector of explanatory variables, α_i is the individual fixed effects, and u_{it} is the error term, i.i.d. and satisfying $E[u|x] = 0$, $V[u^2|x] = \sigma^2$. $i = 1, 2, \cdots, n$; $t = 1, 2, \cdots, T$. The function $m(\cdot)$ is left unspecified, which is to be estimated.

Denote $\beta(x) = \nabla m(x)$.By a Taylor expansion in (2),

$$y_{it} = \alpha_i + m(x) + (x_{it} - x)\beta(x) + u_{it}. \tag{3}$$

Denote $\bar{y}_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} y_{it}$, $\bar{x}_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} x_{it}$ and $\bar{u}_{i\cdot} = \frac{1}{T}\sum_{t=1}^{T} u_{it}$. From (3),

$$\bar{y}_{i\cdot} = \alpha_i + m(x) + (\bar{x}_{i\cdot} - x)\beta(x) + \bar{u}_{i\cdot} \tag{4}$$

which gives

$$y_{it} - \bar{y}_{i\cdot} = (x_{it} - \bar{x}_{i\cdot})\beta(x) + u_{it} - \bar{u}_{i\cdot} \tag{5}$$

The local fixed effects estimator of $\beta(x)$ can then be obtained by minimizing

$$\sum_{i=1}^{n}\sum_{t=1}^{T} \left(y_{it} - \bar{y}_{i\cdot} - (x_{it} - \bar{x}_{i\cdot})\beta(x)\right)^2 k\left(\frac{x_{it} - x}{h}\right)$$

or in a matrix form

$$(M_D Y - M_D X\beta(x))' K(x)(M_D Y - M_D X\beta(x)).$$

When $q = 1$, the estimator is

$$\hat{\beta}_{FE}(x) = \frac{\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it} - \bar{y}_{i\cdot}\right)\left(x_{it} - \bar{x}_{i\cdot}\right)k\left(\frac{x_{it}-x}{h}\right)}{\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \bar{x}_{i\cdot})^2 k\left(\frac{x_{it}-x}{h}\right)}.$$

When $q \geq 1$,

$$\begin{aligned}
&\hat{\beta}_{FE}(x) \\
= &\left[\sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \bar{x}_{i\cdot})'(x_{it} - \bar{x}_{i\cdot})k\left(\frac{x_{it} - x}{h}\right)\right]^{-1} \sum_{i=1}^{n}\sum_{t=1}^{T}(x_{it} - \bar{x}_{i\cdot})'\left(y_{it} - \bar{y}_{i\cdot}\right)k\left(\frac{x_{it} - x}{h}\right)
\end{aligned}$$

or in a matrix form

$$\hat{\beta}_{FE}(x) = \left(X' M_D K(x) M_D X\right)^{-1} X' M_D K(x) M_D Y,$$

where

X is an $nT \times q$ matrix;

$D = I_n \otimes l_T$ is an $nT \times n$ matrix, l_T is a $T \times 1$ vector of unit elements;

$M_D = I - DD'/T$, I is an $nT \times nT$ identity matrix;

$K(x)$ is the $nT \times nT$ diagonal matrix with the diagonal elements $K_{it} = k\left((x_{it} - x)/h\right)$, which is the kernel or weight function taking low values for x_{it} far away from x but high values for x_{it} close to x.

The estimator $\hat{\beta}_{FE}(x)$ is essentially the LS of $\sqrt{K_{it}}\,(y_{it} - \bar{y}_{i\cdot})$ on $\sqrt{K_{it}}(x_{it} - \bar{x}_{i\cdot})$. If $h = \infty$, then $K_{it} = K(0)$ and the estimator becomes the parametric fixed effect estimator. Conditional on x_{it}, for any given $x \in Supp(X)$,

$$E(\hat{\beta}_{FE}(x)) = (X'M_D K(x) M_D X)^{-1} X'M_D K(x) M_D m^*$$

$$V(\hat{\beta}_{FE}(x)) = \sigma_u^2 \left(X'M_D K(x) M_D X\right)^{-1} \left(X'M_D K^2(x) M_D X\right) \left(X'M_D K(x) M_D X\right)^{-1}$$

where $m^* = (m(x_{11}), \cdots, m(x_{nT})))'$. A feasible estimator of $V(\hat{\beta}_{FE}(x))$ is obtained by replacing σ_u^2 with $s_u^2 = \sum_{i=1}^n \sum_{t=1}^T \hat{u}_{it}^{*2}/(nT)$, where $\hat{u}_{it}^*$ is the residual from the regression in (5). By a Taylor expansion of m^* with large n, we can get the asymptotic bias and variance of $\hat{\beta}_{FE}(x)$. However, this, along with the asymptotic normality, still is an open problem.

Note: Assume that the fixed effects satisfy $\sum_{i=1}^n \alpha_i = 0$. Once $\hat{\beta}(x)$ is estimated, from (4), the function $m(x)$ can be estimated by $\hat{m}(x) = \bar{y} - (\bar{x} - x)\hat{\beta}(x)$, where $\bar{y}$ and $\bar{x}$ are the pooled sample averages of $\{y_{it}\}$ and $\{x_{it}\}$,respectively. And then the fixed effect of i is obtained: $\hat{\alpha}_i = \bar{y}_{i\cdot} - \hat{m}(x) + (\bar{x}_{i\cdot} - x)\hat{\beta}(x)$, which may change in x, implying that the fixed effects may be different in different levels of the regressors.

- **A Nonparametric Random Effects Estimator**

The nonparametric panel data model with random effects is

$$y_{it} = m(x_{it}) + v_i + u_{it}, \tag{6}$$

where y_{it} is the dependent variable, $x_{it} \in R^q$ is the row vector of explanatory variables, v_i is (i.i.d.) the individual random effects with mean zero and variance σ_v^2, and u_{it} is the error term, i.i.d. and satisfying $E[u|x] = 0$, $V[u^2|x] = \sigma_u^2$. v_i and u_{jt} are not correlated, $\forall i, j$ and t. And $E[v_i + u_{it}|x_{it}] = 0$.

As before, expand m around $x_{it} = x$ and use the combined error still denoted as u_{it}:

$$\begin{aligned} y_{it} &= m(x) + (x_{it} - x)\beta(x) + v_i + u_{it} \\ &= (1, (x_{it} - x))(m(x), \beta'(x))' + v_i + u_{it} \\ &\equiv z_{it}\delta(x) + \varepsilon_{it}, \end{aligned} \tag{7}$$

where

$$\begin{aligned}
\varepsilon_{it} &= v_i + u_{it} \\
z_{it} &= (1, (x_{it} - x)) \\
\delta(x) &= (m(x), \beta'(x))'.
\end{aligned}$$

Denote

$$y_i = (y_{i1}, y_{i2}, \cdots, y_{iT})', z_i = (z'_{i1}, z'_{i2}, \cdots, z'_{iT})', \quad \varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, \cdots, \varepsilon_{iT})'$$

and

$$Y = (y'_1, y'_2, \cdots, y'_n)', \quad Z(x) = (z'_1, z'_2, \cdots, z'_n)', \quad \varepsilon = (\varepsilon'_1, \varepsilon'_2, \cdots, \varepsilon'_n)'.$$

Then the matrix form of the model (7) is

$$Y = Z(x)\delta(x) + \varepsilon. \tag{8}$$

Since

$$V \equiv E\varepsilon_i\varepsilon'_i = \begin{pmatrix} \sigma_v^2 + \sigma_u^2 & \sigma_v^2 & \cdots & \sigma_v^2 \\ \sigma_v^2 & \sigma_v^2 + \sigma_u^2 & \cdots & \sigma_v^2 \\ \vdots & \vdots & & \vdots \\ \sigma_v^2 & \sigma_v^2 & \cdots & \sigma_v^2 + \sigma_u^2 \end{pmatrix} = \sigma_v^2 J_T + \sigma_u^2 I_T$$

and

$$E\varepsilon_i\varepsilon'_j = 0_{T\times T} \quad (i \neq j)$$

the variance-covariance matrix of ε is

$$\begin{aligned}
\Omega \equiv E\varepsilon\varepsilon' &= \begin{pmatrix} \sigma_v^2 J_T + \sigma_u^2 I_T & 0 & \cdots & 0 \\ 0 & \sigma_v^2 J_T + \sigma_u^2 I_T & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_v^2 J_T + \sigma_u^2 I_T \end{pmatrix} \\
&= I_n \otimes V.
\end{aligned}$$

where $J_T = l_T l'_T$. Then $\Omega^{-1} = I_n \otimes V^{-1}$, where $V^{-1} = \frac{1}{\sigma_u^2}\left[I_T - (1-\lambda)l_T l'_T/T\right]$ and $\lambda = \sigma_u^2/(\sigma_u^2 + T\sigma_v^2)$. Since $V^{-1/2} = I_T - (1-\lambda^{1/2})l_T l'_T/T$, we have

$$\begin{aligned}
\Omega^{-1/2} &= I_n \otimes V^{-1/2} = I_n \otimes \left(I_T - (1-\lambda^{1/2})l_T l'_T/T\right) \\
&= I_{nT} - (1-\lambda^{1/2})\left[I_n \otimes (l_T l'_T)\right]/T \\
&= I_{nT} - (1-\lambda^{1/2})(I_n \otimes l_T)(I_n \otimes l'_T)/T \\
&\equiv I_{nT} - (1-\lambda^{1/2})DD'/T,
\end{aligned}$$

where $D = I_n \otimes l_T$ and $\Omega^{-1/2}\Omega^{-1/2} = \Omega^{-1}$. Then from (8),

$$\Omega^{-1/2}Y = \Omega^{-1/2}Z(x)\delta(x) + \Omega^{-1/2}\varepsilon$$

denoted as

$$Y^* = Z^*(x)\delta(x) + \varepsilon^* \tag{9}$$

satisfying $E(\varepsilon^*\varepsilon^{*\prime}) = E\left[\Omega^{-1/2}\varepsilon\left(\Omega^{-1/2}\varepsilon\right)'\right] = \Omega^{-1/2}E\left[\varepsilon\varepsilon'\right]\Omega^{-1/2} = I_{nT}$. The matrix form (9) can be also written as

$$y_{it}^* = z_{it}^*\delta(x) + \varepsilon_{it}^*.$$

The local nonparametric random effector estimator of $m(x)$ and $\beta(x)$ in (9) is defined by minimizing

$$\begin{aligned}
&\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it}^* - z_{it}^*\delta(x)\right)^2 k\left(\frac{x_{it}-x}{h}\right) \\
=\ & \left(Y^* - Z^*(x)\delta(x)\right)' K(x)\left(Y^* - Z^*(x)\delta(x)\right) \\
=\ & \left(Y - Z(x)\delta(x)\right)'\Omega^{-1/2}K(x)\Omega^{-1/2}\left(Y - Z(x)\delta(x)\right)
\end{aligned}$$

where

$$\begin{aligned}
Y^* &= \Omega^{-1/2}Y, \quad Z^*(x) = \Omega^{-1/2}Z(x), \\
y_{it}^* &= y_{it} - (1-\lambda^{1/2})\bar{y}_{i\cdot}, \quad z_{it}^* = z_{it} - (1-\lambda^{1/2})\bar{z}_{i\cdot}
\end{aligned}$$

and

$$\lambda = \frac{\sigma_u^2}{\sigma_u^2 + T\sigma_v^2}.$$

Therefore,

$$\begin{aligned}
\hat{\delta}(x) &= \left[\sum_{i=1}^{n}\sum_{t=1}^{T} z_{it}^{*\prime}z_{it}^* k\left(\frac{x_{it}-x}{h}\right)\right]^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T} z_{it}^{*\prime}y_{it}^* k\left(\frac{x_{it}-x}{h}\right) && (10)\\
&= \left(Z^{*\prime}(x)K(x)Z^*(x)\right)^{-1}Z^{*\prime}(x)K(x)Y^*. \\
&= \left[Z'(x)\Omega^{-1/2}K(x)\Omega^{-1/2}Z(x)\right]^{-1}\left[Z'(x)\Omega^{-1/2}K(x)\Omega^{-1/2}Y\right] && (11)
\end{aligned}$$

To obtain a feasible estimator of $\delta(x)$, we replace λ with its estimator

$$\hat{\lambda} = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + T\hat{\sigma}_v^2},$$

where $\hat{\sigma}_u^2$ is estimated by

$$\hat{\sigma}_u^2 = \frac{1}{n(T-1)}\sum_{i=1}^{n}\sum_{t=1}^{T}\left[y_{it} - \bar{y}_{i\cdot} - (x_{it} - \bar{x}_{i\cdot})\hat{\beta}_{FE}(x)\right]^2$$

from the regression model

$$y_{it} - \bar{y}_{i\cdot} = (x_{it} - \bar{x}_{i\cdot})\beta(x) + u_{it} - \bar{u}_{i\cdot}$$

while $\hat{\sigma}_v^2$ is defined as $\hat{\sigma}_v^2 \equiv \hat{\sigma}_\eta^2 - \hat{\sigma}_u^2/T$, where $\hat{\sigma}_\eta^2 = \sum_{i=1}^n (\bar{y}_{i\cdot} - \tilde{m}(\bar{x}_{i\cdot}))^2/n$ and $\tilde{m}(x)$ is obtained by performing local LS estimation on the model

$$\bar{y}_{i\cdot} = m(x) + (\bar{x}_{i\cdot} - x)\beta(x) + v_i + \bar{u}_{i\cdot}.$$

Conditional on x_{it}, for any given $x \in Supp(x)$,

$$V(\hat{\delta}(x)) = \left(Z^{*\prime}(x)K(x)Z^*(x)\right)^{-1} Z^{*\prime}(x)K^2(x)Z^*(x)\left(Z^{*\prime}(x)K(x)Z^*(x)\right)^{-1}.$$

Then the nonparametric random effects estimators of $m(x)$ and $\beta(x)$, respectively, are

$$\begin{aligned} \hat{m}_{RE}(x) &= (1, 0_q')\hat{\delta}(x), \\ \hat{\beta}_{RE}(x) &= (0, 1_q')\hat{\delta}(x). \end{aligned}$$

- **Semiparametric FE and RE Estimators**

The semiparametric panel data model is

$$y_{it} = x_{it}\beta + m(z_{it}) + v_i + u_{it}, \tag{12}$$

where y_{it} is the dependent variable, $x_{it} \in R^q$ and $z_{it} \in R^p$ are two row vectors of regressors. $E[v_i + u_{it}|x_{it}, z_{it}] = 0$.

(i) For $v_i = 0$, we have

$$y_{it} - E[y_{it}|z_{it}] = (x_{it} - E[x_{it}|z_{it}])\,\beta + u_{it}$$

and β is estimated (with pooled data) by

$$\hat{\beta} = \left[\sum_{i=1}^n \sum_{t=1}^T (x_{it} - \hat{x}_{it})' (x_{it} - \hat{x}_{it})\right]^{-1} \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \hat{x}_{it})' (y_{it} - \hat{y}_{it})\, 1\{\hat{f}(z_{it}) \geq b\}.$$

Then $m(z)$ can be estimated by analyzing the model $y_{it}^* = m(z_{it}) + u_{it}$ as in the pooled model case, where $y_{it}^* = y_{it} - x_{it}\hat{\beta}$.

(ii) For $v_i \neq 0$ and fixed, the estimator $\hat{\beta}$ above is not affected. Once $\hat{\beta}$ is estimated the estimation of $\nabla m(z)$ can be carried out by analyzing the fixed effects model y_{it}^* $= m(z_{it}) + v_i + u_{it}$, where $y_{it}^* = y_{it} - x_{it}\hat{\beta}$.

When $p = 1$, the estimator of $\nabla m(z)$ is

$$\hat{\gamma}_{SFE}(z) = \frac{\sum_{i=1}^n \sum_{t=1}^T (y_{it}^* - \bar{y}_{i\cdot}^*)\,(z_{it} - \bar{z}_{i\cdot})k\left((z_{it} - z)/h\right)}{\sum_{i=1}^n \sum_{t=1}^T (z_{it} - \bar{z}_{i\cdot})^2 k\left((z_{it} - z)/h\right)}.$$

When $p \geq 1$,

$$\begin{aligned}
&\hat{\gamma}_{SFE}(z) \\
= &\left[\sum_{i=1}^{n}\sum_{t=1}^{T}(z_{it}-\bar{z}_{i\cdot})'(z_{it}-\bar{z}_{i\cdot})k\left(\frac{x_{it}-x}{h}\right)\right]^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}(z_{it}-\bar{z}_{i\cdot})'\left(y_{it}^{*}-\bar{y}_{i\cdot}^{*}\right)k\left(\frac{x_{it}-x}{h}\right)
\end{aligned}$$

or in a matrix form

$$\hat{\gamma}_{SFE}(z) = \left(Z'M_D K(z) M_D Z\right)^{-1} Z'M_D K(z) M_D Y^{*}.$$

(iii) For $v_i \neq 0$ and random, by

$$y_{it} - E[y_{it}|z_{it}] = \left(x_{it} - E[x_{it}|z_{it}]\right)\beta + v_i + u_{it}$$

or

$$R_{it}^{yz} = R_{it}^{xz}\beta + v_i + u_{it}, \tag{13}$$

we have

$$\hat{\beta}_{SRE} = \left[\sum_{i=1}^{n}\sum_{t=1}^{T} R_{it}^{xz*\prime} R_{it}^{xz*}\right]^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T} R_{it}^{xz*\prime} R_{it}^{yz*}$$

where $R_{it}^{xz*} = R_{it}^{xz} - (1-\lambda^{1/2})\bar{R}_{i\cdot}^{xz}$, $R_{it}^{yz*} = R_{it}^{yz} - (1-\lambda^{1/2})\bar{R}_{i\cdot}^{yz}$ and $\lambda = \sigma_u^2/(\sigma_u^2 + T\sigma_v^2)$, and $E[y_{it}|z_{it}]$ and $E[x_{it}|z_{it}]$ are replaced by their kernel estimators $\hat{E}[y_{it}|z_{it}]$ and $\hat{E}[x_{it}|z_{it}]$. Denote $y_{it}^{*} = y_{it} - x_{it}\hat{\beta}_{SRE}$. Then use the following random effects model

$$y_{it}^{*} = m(z_{it}) + v_i + u_{it}$$

to estimate $\delta(z) = (m(z), \gamma'(z))'$, the same way as in (10), where $\gamma(z) = \nabla m(z)$.

Specifically, write

$$\begin{aligned}
y_{it}^{*} &= m(z) + (z_{it} - z)\nabla m(z) + v_i + u_{it} \\
&= (1, z_{it} - z)\delta(z) + v_i + u_{it} \\
&\equiv r_{it}^{*}(z)\delta(z) + v_i + u_{it},
\end{aligned} \tag{14}$$

where $r_{it}^{*}(z) = (1, z_{it} - z)$ and $\delta(z) = (m(z), \gamma'(z))'$. This gives the estimators of $\delta(z)$:

$$\begin{aligned}
\hat{\delta}_{SRE}(z) &= \left[\sum_{i=1}^{n}\sum_{t=1}^{T} r_{it}^{**\prime} r_{it}^{**} k\left(\frac{z_{it}-z}{h}\right)\right]^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T} r_{it}^{**\prime} y_{it}^{**} k\left(\frac{z_{it}-z}{h}\right) \\
&= \left(R^{**\prime}(z) K(z) R^{**}(z)\right)^{-1} R^{**\prime}(z) K(z) Y^{**}
\end{aligned}$$

where

$$\begin{aligned} Y^{**} &= \Omega^{-1/2}Y^*, \quad R^{**}(z) = \Omega^{-1/2}R^*(z), \\ y_{it}^{**} &= y_{it}^* - (1-\lambda^{1/2})\bar{y}_{i\cdot}^*, \quad r_{it}^{**} = r_{it}^* - (1-\lambda^{1/2})\bar{r}_{i\cdot}^* \end{aligned}$$

and

$$\lambda = \frac{\sigma_u^2}{\sigma_u^2 + T\sigma_v^2}.$$

Then the semiparametric random effects estimators of $m(z)$ and $\nabla m(z)$, respectively, are

$$\begin{aligned} \hat{m}_{SRE}(z) &= (1, 0_p')\hat{\delta}_{SRE}(z), \\ \hat{\gamma}_{SRE(}x) &= (0, 1_p')\hat{\delta}_{SRE}(z). \end{aligned}$$

There are two methods to estimate λ :

a) Based on (13): Estimate $\tilde{\sigma}_u^2$ by

$$\tilde{\sigma}_u^2 = \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[R_{it}^{yz} - \bar{R}_{i\cdot}^{yz} - (R_{it}^{xz} - \bar{R}_{i\cdot}^{xz})\hat{\beta}_{FE} \right]^2$$

from the regression model

$$R_{it}^{yz} - \bar{R}_{i\cdot}^{yz} = (R_{it}^{xz} - \bar{R}_{i\cdot}^{xz})\beta + u_{it} - \bar{u}_{i\cdot}$$

and estimate $\tilde{\sigma}_v^2$ from $\tilde{\sigma}_v^2 \equiv \tilde{\sigma}_\eta^2 - \tilde{\sigma}_u^2/T$, where $\tilde{\sigma}_\eta^2 = \sum_{i=1}^n (\bar{R}_{i\cdot}^{yz} - \bar{R}_{i\cdot}^{xz}\hat{\beta}_B)^2/n$ and $\hat{\beta}_B$ is the between estimator obtained by performing the LS estimation on the model

$$\bar{R}_{i\cdot}^{yz} = \bar{R}_{i\cdot}^{xz}\beta + v_i + \bar{u}_{i\cdot}.$$

b) Based on (14): Estimate $\hat{\sigma}_u^2$ by

$$\hat{\sigma}_u^2 = \frac{1}{n(T-1)} \sum_{i=1}^{n} \sum_{t=1}^{T} \left[y_{it}^* - \bar{y}_{i\cdot}^* - (z_{it} - \bar{z}_{i\cdot})\hat{\gamma}_{SFE}(z) \right]^2$$

from the regression model

$$y_{it}^* - \bar{y}_{i\cdot}^* = (z_{it} - \bar{z}_{i\cdot})\gamma_{SFE}(z) + u_{it} - \bar{u}_{i\cdot}$$

and estimate $\hat{\sigma}_v^2$ from $\hat{\sigma}_v^2 \equiv \hat{\sigma}_\eta^2 - \hat{\sigma}_u^2/T$, where $\hat{\sigma}_\eta^2 = \sum_{i=1}^n (\bar{y}_{i\cdot}^* - \tilde{m}(\bar{z}_{i\cdot}))^2/n$ and $\tilde{m}(z)$ is obtained by performing local LS estimation on the model

$$\bar{y}_{i\cdot}^* = m(z) + (z_{it} - \bar{z}_{i\cdot})\gamma(z) + v_i + \bar{u}_{i\cdot}.$$