11 Panel Data

11.1 Wang (2003) estimator

$$
\begin{aligned}
0 &= \sum_{i=1}^{n}\sum_{t=1}^{T_i} K_h(\mathbf{x}_{it},\mathbf{x})\begin{pmatrix}1\\ \frac{\mathbf{x}_{it}-\mathbf{x}}{h}\end{pmatrix}\left\{\begin{array}{c}\sigma^{tt}\left[y_{it}-\widehat{m}_{[l]}(\mathbf{x})-\left(\frac{\mathbf{x}_{it}-\mathbf{x}}{h}\right)\widehat{\beta}_{[l]}(\mathbf{x})\right]\\ +\sum_{\substack{s=1\\ s\neq t}}^{T_i}\sigma^{st}\left[y_{is}-\widehat{m}_{[l-1]}(\mathbf{x}_{is})\right]\end{array}\right\} \\
&= \sum_{i=1}^{n}\sum_{t=1}^{T_i} K_h(\mathbf{x}_{it},\mathbf{x})\begin{pmatrix}1\\ \frac{\mathbf{x}_{it}-\mathbf{x}}{h}\end{pmatrix}\left\{\sigma^{tt}y_{it}+\sum_{\substack{s=1\\ s\neq t}}^{T_i}\sigma^{st}y_{is}\right\} \\
&\quad -\sum_{i=1}^{n}\sum_{t=1}^{T_i} K_h(\mathbf{x}_{it},\mathbf{x})\begin{pmatrix}1\\ \frac{\mathbf{x}_{it}-\mathbf{x}}{h}\end{pmatrix}\left\{\begin{array}{c}\sigma^{tt}\left[\widehat{m}_{[l]}(\mathbf{x})+\left(\frac{\mathbf{x}_{it}-\mathbf{x}}{h}\right)\widehat{\beta}_{[l]}(\mathbf{x})\right]\\ \sum_{\substack{s=1\\ s\neq t}}^{T_i}\sigma^{st}\widehat{m}_{[l-1]}(\mathbf{x}_{is})\end{array}\right\}
\end{aligned}
$$

Combining like terms and moving the second term to the left-hand-side of zero gives

$$
\begin{aligned}
&\left[\sum_{i=1}^{n}\sum_{t=1}^{T_i} K_h(\mathbf{x}_{it},\mathbf{x})\,\sigma^{tt}\begin{pmatrix}1\\ \frac{\mathbf{x}_{it}-\mathbf{x}}{h}\end{pmatrix}\begin{pmatrix}1 & \frac{\mathbf{x}_{it}-\mathbf{x}}{h}\end{pmatrix}\right]\begin{pmatrix}\widehat{m}_{[l]}(\mathbf{x})\\ \widehat{\beta}_{[l]}(\mathbf{x})\end{pmatrix} \\
=\ &\left[\sum_{i=1}^{n}\sum_{t=1}^{T_i} K_h(\mathbf{x}_{it},\mathbf{x})\begin{pmatrix}1\\ \frac{\mathbf{x}_{it}-\mathbf{x}}{h}\end{pmatrix}\right]\left[\sigma^{tt}y_{it}+\sum_{\substack{s=1\\ s\neq t}}^{T_i}\sigma^{st}\left(y_{is}-\widehat{m}_{[l-1]}(\mathbf{x}_{is})\right)\right]
\end{aligned}
$$

and hence

$$
\begin{aligned}
\begin{pmatrix}\widehat{m}_{[l]}(\mathbf{x})\\ \widehat{\beta}_{[l]}(\mathbf{x})\end{pmatrix} &= \left[\sum_{i=1}^{n}\sum_{t=1}^{T_i} K_h(\mathbf{x}_{it},\mathbf{x})\,\sigma^{tt}\begin{pmatrix}1\\ \frac{\mathbf{x}_{it}-\mathbf{x}}{h}\end{pmatrix}\begin{pmatrix}1 & \frac{\mathbf{x}_{it}-\mathbf{x}}{h}\end{pmatrix}\right]^{-1} \\
&\quad\times\left[\sum_{i=1}^{n}\sum_{t=1}^{T_i} K_h(\mathbf{x}_{it},\mathbf{x})\begin{pmatrix}1\\ \frac{\mathbf{x}_{it}-\mathbf{x}}{h}\end{pmatrix}\right]\left[\sigma^{tt}y_{it}+\sum_{\substack{s=1\\ s\neq t}}^{T_i}\sigma^{st}\left(y_{is}-\widehat{m}_{[l-1]}(\mathbf{x}_{is})\right)\right].
\end{aligned}
$$

11.2 Henderson, Carroll and Li (2008) estimator

The algorithm is linear in the y_{it}'s. By defining

$$H_{i,[\ell-1]} = \begin{pmatrix} y_{i2} - \widehat{m}_{[\ell-1]}(\mathbf{x}_{i2}) \\ \vdots \\ y_{iT_i} - \widehat{m}_{[\ell-1]}(\mathbf{x}_{iT_i}) \end{pmatrix} - \{y_{i1} - \widehat{m}_{[\ell-1]}(\mathbf{x}_{i1})\}\mathbf{i}_{T_i-1},$$

we get

$$\begin{aligned} 0 &= \sum_{i=1}^{n} K_h(\mathbf{x}_{i1}, \mathbf{x}) \begin{pmatrix} 1 \\ \frac{\mathbf{x}_{i1}-\mathbf{x}}{h} \end{pmatrix} \left[\begin{array}{c} -\mathbf{i}'_{T_i-1} V^{-1} H_{i,[\ell-1]} \\ +\mathbf{i}'_{T_i-1} V^{-1} \mathbf{i}_{T_i-1} \left\{ \widehat{m}_{[\ell-1]}(\mathbf{x}_{i1}) - \begin{pmatrix} 1 & \frac{\mathbf{x}_{i1}-\mathbf{x}}{h} \end{pmatrix} \binom{\alpha_0}{\alpha_1} \right\} \end{array} \right] \\ &+ \sum_{i=1}^{n} \sum_{t=2}^{T_i} K_h(\mathbf{x}_{it}, \mathbf{x}) \begin{pmatrix} 1 \\ \frac{\mathbf{x}_{it}-\mathbf{x}}{h} \end{pmatrix} \left[\begin{array}{c} c'_{t-1} V^{-1} H_{i,[\ell-1]} \\ +c'_{t-1} V^{-1} c_{t-1} \left\{ \widehat{m}_{[\ell-1]}(\mathbf{x}_{it}) - \begin{pmatrix} 1 & \frac{\mathbf{x}_{it}-\mathbf{x}}{h} \end{pmatrix} \binom{\alpha_0}{\alpha_1} \right\} \end{array} \right]. \end{aligned}$$

Solving for α_0 and α_1 leads to $\{\widehat{\alpha}_0(\mathbf{x}), \widehat{\alpha}_1(\mathbf{x})\}' = D_1^{-1}(D_2 + D_3)$

$$\begin{aligned} D_1 &= n^{-1} \sum_{i=1}^{n} \left\{ \begin{array}{c} \mathbf{i}'_{T_i-1} V^{-1} \mathbf{i}_{T_i-1} K_h(\mathbf{x}_{i1}, \mathbf{x}) \left(\begin{smallmatrix} 1 \\ \frac{\mathbf{x}_{i1}-\mathbf{x}}{h} \end{smallmatrix}\right) \begin{pmatrix} 1 & \frac{\mathbf{x}_{i1}-\mathbf{x}}{h} \end{pmatrix} \\ + \sum_{t=2}^{T_i} c'_{t-1} V^{-1} c_{t-1} K_h(\mathbf{x}_{it}, \mathbf{x}) \left(\begin{smallmatrix} 1 \\ \frac{\mathbf{x}_{it}-\mathbf{x}}{h} \end{smallmatrix}\right) \begin{pmatrix} 1 & \frac{\mathbf{x}_{it}-\mathbf{x}}{h} \end{pmatrix} \end{array} \right\}; \\ D_2 &= n^{-1} \sum_{i=1}^{n} \mathbf{i}'_{T_i-1} V^{-1} \mathbf{i}_{T_i-1} K_h(\mathbf{x}_{i1}, \mathbf{x}) \begin{pmatrix} 1 \\ \frac{\mathbf{x}_{i1}-\mathbf{x}}{h} \end{pmatrix} \widehat{m}_{[\ell-1]}(\mathbf{x}_{i1}) \\ &\quad + \sum_{t=2}^{T_i} c'_{t-1} V^{-1} c_{t-1} K_h(\mathbf{x}_{it}, \mathbf{x}) \begin{pmatrix} 1 \\ \frac{\mathbf{x}_{it}-\mathbf{x}}{h} \end{pmatrix} \widehat{m}_{[\ell-1]}(\mathbf{x}_{it}); \\ D_3 &= n^{-1} \sum_{i=1}^{n} \left\{ \begin{array}{c} \sum_{t=2}^{T_i} K_h(\mathbf{x}_{it}, \mathbf{x}) \left(\begin{smallmatrix} 1 \\ \frac{\mathbf{x}_{it}-\mathbf{x}}{h} \end{smallmatrix}\right) c'_{t-1} V^{-1} H_{i,[\ell-1]} \\ -K_h(\mathbf{x}_{i1}, \mathbf{x}) \left(\begin{smallmatrix} 1 \\ \frac{\mathbf{x}_{i1}-\mathbf{x}}{h} \end{smallmatrix}\right) \mathbf{i}'_{T_i-1} V^{-1} H_{i,[\ell-1]} \end{array} \right\}, \end{aligned}$$

The next step estimate of $m(\mathbf{x})$ is given by $\widehat{m}_{[\ell]}(\mathbf{x}) = \widehat{\alpha}_0(\mathbf{x})$.

The particular form of our problem means that many of the terms have simple expressions.

In particular,

$$\begin{aligned}
\mathcal{L}_{i,11m} &= -e_{t-1}'V^{-1}e_{t-1} = -(T_i-1)/(T_i\sigma_v^2); \\
\mathcal{L}_{i,ttm} &= -c_{t-1}'V^{-1}c_{t-1} = -(T_i-1)/(T_i\sigma_v^2) \qquad \text{for } t\geq 2; \\
\mathcal{L}_{i,1tm} &= -c_{t-1}'V^{-1}e_{m-1} = -\frac{1}{T_i\sigma_v^2} \qquad \text{for } t\geq 2; \\
\mathcal{L}_{i,tsm} &= -c_{t-1}'V^{-1}c_{s-1} = \frac{1}{T_i\sigma_v^2} \qquad \text{for } t,s\geq 2 \text{ and } t\neq s;
\end{aligned}$$

If we incorporate these values, we can simplify our estimator as

$$\begin{aligned}
D_1 &= \frac{1}{n\sigma_v^2}\sum_{i=1}^{n}\left\{\begin{array}{c} K_h(\mathbf{x}_{i1},\mathbf{x})\binom{1}{\frac{\mathbf{x}_{i1}-\mathbf{x}}{h}}\left(1\quad \frac{\mathbf{x}_{i1}-\mathbf{x}}{h}\right) \\ +\sum_{t=2}^{T_i}\frac{(T_i-1)}{T_i}K_h(\mathbf{x}_{it},\mathbf{x})\binom{1}{\frac{\mathbf{x}_{it}-\mathbf{x}}{h}}\left(1\quad \frac{\mathbf{x}_{it}-\mathbf{x}}{h}\right)\end{array}\right\}; \\
D_2 &= \frac{1}{n\sigma_v^2}\sum_{i=1}^{n}\left\{\begin{array}{c} K_h(\mathbf{x}_{i1},\mathbf{x})\binom{1}{\frac{\mathbf{x}_{i1}-\mathbf{x}}{h}}\widehat{m}_{[\ell-1]}(\mathbf{x}_{i1}) \\ +\sum_{t=2}^{T_i}\frac{(T_i-1)}{T_i}K_h(\mathbf{x}_{it},\mathbf{x})\binom{1}{\frac{\mathbf{x}_{it}-\mathbf{x}}{h}}\widehat{m}_{[\ell-1]}(\mathbf{x}_{it})\end{array}\right\}; \\
D_3 &= \frac{1}{n}\sum_{i=1}^{n}\left\{\begin{array}{c} \sum_{t=2}^{T_i}K_h(\mathbf{x}_{it},\mathbf{x})\binom{1}{\frac{\mathbf{x}_{it}-\mathbf{x}}{h}}c_{t-1}'V^{-1}H_{i,[\ell-1]} \\ -K_h(\mathbf{x}_{i1},\mathbf{x})\binom{1}{\frac{\mathbf{x}_{i1}-\mathbf{x}}{h}}\mathbf{i}_{T_i-1}'V^{-1}H_{i,[\ell-1]}\end{array}\right\},
\end{aligned}$$

and this simplifies to

$$\begin{aligned}
D_1 &= \frac{1}{\sigma_v^2}\sum_{i=1}^{n}\sum_{t=1}^{T_i}\frac{(T_i-1)}{T_i}K_h(\mathbf{x}_{it},\mathbf{x})\binom{1}{\frac{\mathbf{x}_{it}-\mathbf{x}}{h}}\left(1\quad \frac{\mathbf{x}_{it}-\mathbf{x}}{h}\right) \\
D_2 &= \frac{1}{\sigma_v^2}\sum_{i=1}^{n}\sum_{t=1}^{T_i}\frac{(T_i-1)}{T_i}K_h(\mathbf{x}_{it},\mathbf{x})\binom{1}{\frac{\mathbf{x}_{it}-\mathbf{x}}{h}}\widehat{m}_{[\ell-1]}(\mathbf{x}_{it}); \\
D_3 &= \sum_{i=1}^{n}\left\{\begin{array}{c} \sum_{t=2}^{T_i}K_h(\mathbf{x}_{it},\mathbf{x})\binom{1}{\frac{\mathbf{x}_{it}-\mathbf{x}}{h}}c_{t-1}'V_i^{-1}H_{i,[\ell-1]} \\ -K_h(\mathbf{x}_{i1},\mathbf{x})\binom{1}{\frac{\mathbf{x}_{i1}-\mathbf{x}}{h}}\mathbf{i}_{T_i-1}'V_i^{-1}H_{i,[\ell-1]}\end{array}\right\}.
\end{aligned}$$

11.3 Poolability test statistic

Following the logic from Chapter 6, we can construct a consistent test for poolability based on $E[uE(u|\mathbf{x})] = E\left\{[E(u|\mathbf{x})]^2\right\} \geq 0$ and the equality will hold if and only if the null is true. Again, we choose to estimate a density weighted version of $E[uE(u|\mathbf{x})]$ to avoid the random denominator problem in nonparametric estimation. The density weighted version of

$E\left[uE\left(u|\mathbf{x}\right)\right]$ is given by

$$E\left\{uf\left(\mathbf{x}\right)E\left[uf\left(\mathbf{x}\right)|\mathbf{x}\right]f\left(\mathbf{x}\right)\right\}.$$

The sample analogue of $uf\left(\mathbf{x}\right)$ is

$$\widehat{u}_{it}\widehat{f}\left(\mathbf{x}_{it}\right)$$

and that of $E\left(uf\left(\mathbf{x}\right)|\mathbf{x}\right)f\left(\mathbf{x}\right)$ is

$$\frac{1}{\left(n-1\right)\left(h_1h_2\cdots h_q\right)}\sum_{\substack{j=1\\j\neq i}}^{n}\widehat{u}_{jt}\widehat{f}\left(\mathbf{x}_{jt}\right)K_h\left(\mathbf{x}_{it},\mathbf{x}_{jt}\right)$$

using n cross-sectional data for a fixed value of t. Hence, our test statistic is based on

$$\begin{aligned}\widehat{J}_n &= \frac{1}{\left(nT-n\right)\left(h_1h_2\cdots h_q\right)}\sum_{i=1}^{n}\sum_{t=1}^{T_i}\widehat{u}_{it}\widehat{f}\left(\mathbf{x}_{it}\right)\left\{\sum_{\substack{j=1\\j\neq i}}^{n}\widehat{u}_{jt}\widehat{f}\left(\mathbf{x}_{jt}\right)K_h\left(\mathbf{x}_{it},\mathbf{x}_{jt}\right)\right\}\\ &= \frac{1}{\left(nT-n\right)\left(h_1h_2\cdots h_q\right)}\sum_{i=1}^{n}\sum_{t=1}^{T_i}\sum_{\substack{j=1\\j\neq i}}^{n}\widehat{u}_{it}\widehat{f}\left(\mathbf{x}_{it}\right)\widehat{u}_{jt}\widehat{f}\left(\mathbf{x}_{jt}\right)K_h\left(\mathbf{x}_{it},\mathbf{x}_{jt}\right).\end{aligned}$$

Baltagi, Hidalgo and Li (1996) show that $n\left(h_1h_2\cdots h_q\right)^2\widehat{J}_n$ is distributed normal, mean zero, with variance $2\sigma_0^2$ under the null. When the data are i.i.d. in the subscript i

$$\sigma_0^2=\left[\int k^2\left(u\right)du\right]\left\{\sum_{t=1}^{T_i}\frac{1}{T_i^2}E\left[\sigma^4\left(\mathbf{x}_{1t}\right)f^5\left(\mathbf{x}_{1t}\right)\right]\right\},$$

and a consistent estimate of this unknown parameter can be obtained as

$$\widehat{\sigma}_0^2=\frac{1}{\left(nT-n\right)\left(h_1h_2\cdots h_q\right)}\sum_{i=1}^{n}\sum_{t=1}^{T_i}\sum_{\substack{j=1\\j\neq i}}^{n}\left[\widehat{u}_{it}\widehat{f}\left(\mathbf{x}_{it}\right)\right]^2\left[\widehat{u}_{jt}\widehat{f}\left(\mathbf{x}_{jt}\right)\right]^2K_h^2\left(\mathbf{x}_{it},\mathbf{x}_{jt}\right)$$

Thus, the normalization

$$\widehat{T}_n=\frac{n\left(h_1h_2\cdots h_q\right)^2\widehat{J}_n}{\sqrt{2\widehat{\sigma}_0^2}}$$

is distributed standard normal under the null.