A Semiparametric Panel Model for unbalanced data with Application to Climate Change in the United Kingdom*

Alev Atak[†]
Queen Mary, University of London

Oliver Linton[‡]
London School of Economics

Zhijie Xiao[§]
Boston College

November 7, 2014

Abstract

This paper is concerned with developing a semiparametric panel model to explain the trend in UK temperatures and other weather outcomes over the last century. We work with the monthly averaged maximum and minimum temperatures observed at the twenty six Meteorological Office stations. The data is an unbalanced panel. We allow the trend to evolve in a nonparametric way so that we obtain a fuller picture of the evolution of common temperature in the medium timescale. Profile likelihood estimators (PLE) are proposed and their statistical properties are studied. The proposed PLE has improved asymptotic property comparing the the sequential two-step estimators. Finally, forecasting based on the proposed model is studied.

Keywords: Global warming; Kernel estimation; Semiparametric; Trend analysis

JEL classification: C13; C14; C21; D24

*We thank a guest-editor and two referees for helpful comments on an early version and Jaanaki Momaya for helpful research assistance.

†Department of Economics, Queen Mary, University of London, Mile End Road, London E1 4NS. Email: a.atak@qmul.ac.uk

‡Department of Economics, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom. E-mail address: o.linton@lse.ac.uk.

Thanks to the ESRC, the ERC, and the Leverhulme foundations for financial support. This paper was partly written while I was a Universidad Carlos III de Madrid-Banco Santander Chair of Excellence, and I thank them for financial support.

§Department of Economics, Boston College, Chestnut Hill, MA 02467, USA. Tel: 617-552-1709. Email: xiaoz@bc.edu. School of Economics and Management, Tsinghua University, Beijing, China.

Financial support from Boston College is gratefully acknowledged.

1 Introduction

The partially linear regression model was introduced in Engle, Granger, Rice and Weiss (1986),

$$y = \beta^T X + \theta(Z) + \varepsilon \tag{1}$$

where $\theta(.)$ is an unknown scalar function and ε is a zero mean error orthogonal to both X and $\theta(.)$. This model embodies a compromise between employing a general nonparametric specification $g(X, Z)$, which, if the conditioning variables are high dimensional, would lead to serious loss of precision, and a fully parametric specification which may result in badly biased estimators and inconsistent hypothesis tests. The implicit asymmetry between the effects of X and Z may be attractive when X consists of dummy or categorical variables, as in Stock (1989, 1991). This specification arises in various sample selection models, see Ahn and Powell (1993), Newey, Powell, and Walker (1990), and Lee, Rosenzweig and Pitt (1992). It is also the basis of a general specification test for functional form introduced in Delgado and Stengos (1994). The model has been used in a number of applications. We will use a panel data version of this model to model climate change.

The issue of global warming has received a great deal of attention recently. This paper is concerned with developing a semiparametric model to describe the trend in UK regional temperatures and other weather outcomes over the last century. The data we work with conditions the analysis we propose. We work with the monthly averaged maximum and minimum temperatures observed at the twenty six Meteorological Office stations. The data is an unbalanced panel. We propose a semiparametric partial linear panel model in which there is a common trend component that is allowed to evolve in a nonparametric way. This permits the most general possible pattern for the evolution of a common secular change in temperature. We also allow for a deterministic seasonal component in temperature, since we are working with monthly data. Gao and Hawthorne (2006) used a univariate partially linear model to explain annual global temperature in terms of a nonparametric time trend and a covariate the southern oscillation index (SOI). They applied existing theory to deduce the properties of their estimators and developed a new adaptive test of the shape of the trend function. See Campbell and Diebold (2005) for some alternative analysis of multivariate climate time series data. Peteiro-Lopez and Gonzalez-Manteiga (2006) worked with a multivariate model with cross-sectionally correlated errors and different trends for each series. They establish distribution theory for the parametric components and derive the bias and variance of the nonparametric components. Their setting is similar to ours except that we impose a common trend structure. Furthermore, the covariates in our parametric part are also common and deterministic, as they represent seasonality. Most importantly we allow for unbalanced dataset, which is important in applications. This difference has important implications for efficient estimation. The asymptotic framework we work with allows a non-trivial fraction of the data to be missing. We propose to use a profile likelihood method, which in the unbalanced case is different from the sequential two-step squares method proposed by Robinson (1998) in the univariate case and employed by Peteiro-Lopez and Gonzalez-Manteiga (2006) in the multivariate case. This method is fully efficient in the Gaussian case as established in Severini and Wong (1992). Finally, we allow for heteroskedasticity and serial correlation in the error terms.

We apply our methods to the UK dataset. We show the nonparametric trend in comparison with a more standard parametric approach. In both cases there is an upward trend over the last twenty years that is statistically significant. We compare our results with those obtained by Gao and Hawthorne (2006). We also use our model to forecast future temperature.

2 Model and Data

The subject that we are interested are monthly temperatures $\{y_{it}\}$, where i signifies different stations and t is the corresponding time when the temperature is recorded, $t = 1, \dots, T$ and $i = 1, \dots, n$. In practice, there may be missing data in the sense that some stations began keeping records before other stations. In our application, Oxford started in 1857, while Cardiff Bute Park only began in 1977. So we suppose that station i starts at time t_i, $i = 1, \dots, n$, thus records for station i are only available from time t_i to T. Order the stations by their starting point so that $t_1 \leq t_2 \leq \cdots \leq t_n < T$. The complete record occurs after t_n. At any point in time there are n_t stations available with n_t varying from one to n. The most general model we consider is of the following form

$$y_{it} = \alpha_i + \beta_i^\top D_t + \gamma_i^\top X_{it} + g_i(t/T) + \varepsilon_{it},$$

where $i = 1, \dots, n$ and $t = t_i, \dots, T$. Here, $D_t \in \mathbb{R}^d$ is a vector of seasonal dummy variables, X_{it} are a vector of observed covariates, and the error terms ε_{it} satisfy $E(\varepsilon_{it}|X_{it}) = 0$ a.s.. The functions $g_i(\cdot)$ are unknown but smooth. These represent the trend in temperatures at location i. We shall further assume that $g_i(\cdot) = g(\cdot)$, so that there is a single common trend, which imposes a standard way of thinking about climate change. For simplicity we also dispense with the additional covariates X (in our application we are concerned with documenting the temperature record rather than assigning changes to particular causes). The parameter vector $\theta = (\alpha_1, \dots, \alpha_n, \beta_1^\top, \dots \beta_n^\top)^\top$ is unknown and describes the seasonal and level effects for the different locations. The model is not identified as it stands, since one can add a constant to each α_i and subtract the same constant from $g(\cdot)$. For identification we suppose that $\sum_{i=1}^n \alpha_i = 0$, in which case the function $g(.)$ represents the common level of average temperature relative to average seasonal variation. According to Wikipedia (2009): "Climate change is any long-term significant change in the “average weather" of a region or the earth as a whole. Average weather may include average temperature, precipitation and wind patterns." Our model directly permits the measuring of this average weather trend through the function $g(\cdot)$.

In doing the asymptotics we suppose that $T \to \infty$ but n is fixed (in fact $n = 26$ in our application).

In conclusion the model we adopt for the application is as follows

$$y_{it} = \alpha_i + \beta_i^\top D_t + g(t/T) + \varepsilon_{it}, \tag{2}$$

where the error term may be heteroskedastic across i and serially correlated over time. Let $\beta_i^\top = (\beta_{i1}, \ldots, \beta_{id})$. We can write the model as

$$y = A\alpha + \sum_{j=1}^{d} C_j \beta_j + Bg + \varepsilon, \tag{3}$$

where y, ε is the $nT \times 1$ data,error vector with zeros in place of missing observations, while $\alpha \in \mathbb{R}^n$, $g = (g(1/T), \ldots, g(1))^\top \in \mathbb{R}^T$, and $\beta_j = (\beta_{1j}, \ldots, \beta_{nj}) \in \mathbb{R}^n$. In this case, A, B are matrices of conformable dimensions of zeros and ones that reflect the commonality and missingness as well, see below. The matrices C_j contains the dummy variable D_j. This representation is different from equation (2) of Peteiro-Lopez and Gonzalez-Manteiga (2006); it allows for the "missingness" of data in some observation units and preserves a simple algebraic structure that is useful in the sequel.

Suppose $n = 2$ and $T = 3$ and for simplicity that $d = 0$, i.e., no seasonal effect. Then

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ 0 \\ y_{22} \\ y_{23} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ 0 \\ \varepsilon_{22} \\ \varepsilon_{23} \end{bmatrix}.$$

3 Profile Likelihood Estimation

Our model may be estimated using different nonparametric methods. We consider in this paper the widely used kernel estimators. Specifically, we consider the Gaussian profile likelihood procedure for the general unbalanced case - see additional discussions in Remarks 2 - 3 for advantages of using profile likelihood estimation. This in general leads to semiparametrically efficient estimators, Severini and Wong (1992).

3.1 The Estimator of g

We first define the local profile likelihood in the local parameter $\eta \in \mathbb{R}$:

$$\begin{aligned}\mathcal{L}(\eta; t/T) &= \sum_{s=1}^{T}\sum_{i=1}^{n_s}\left(y_{is}-\alpha_i-\beta_i^\top D_s-\eta\right)^2 K_h((t-s)/T)\\ &= \sum_{i=1}^{n}\sum_{s=t_i}^{T}\left(y_{is}-\alpha_i-\beta_i^\top D_s-\eta\right)^2 K_h((t-s)/T),\end{aligned}$$

where I_s denotes the set of stations available at time s, which is of cardinality n_s and we assumed the ordering of the stations is consistently chosen. Here, K is a kernel function and h is a bandwidth so that $K_h(.) = K(./h)/h$. The first derivative with respect to η is given by

$$\frac{\partial \mathcal{L}(\eta; t/T)}{\partial \eta} = -2\sum_{s=1}^{T}\sum_{i\in I_s}\left(y_{is}-\alpha_i-\beta_i^\top D_s-\eta\right) K_h((t-s)/T),$$

so that

$$\begin{aligned}\widehat{\eta} &= \widehat{g}_\theta(t/T) = \frac{T^{-1}\sum_{i=1}^{n}\sum_{s=t_i}^{T}\left(y_{is}-\alpha_i-\beta_i^\top D_s\right)K_h((t-s)/T)}{T^{-1}\sum_{i=1}^{n}\sum_{s=t_i}^{T}K_h((t-s)/T)}\\ &= \frac{T^{-1}\sum_{s=1}^{T}K_h((t-s)/T)\sum_{i=1}^{n_s}\left(y_{is}-\alpha_i-\beta_i^\top D_s\right)}{T^{-1}\sum_{s=1}^{T}K_h((t-s)/T)n_s}\end{aligned}$$

Notice that if we standardize the kernel so that $T^{-1}\sum_{s=1}^{T}K_h(u-s/T)=1$, then, when T is large, $m_t = m$, where $m_t = T^{-1}\sum_{i=1}^{n}\sum_{s=t_i}^{T}K_h((s-t)/T)$, for all t with $t_m/T < t/T < t_{m+1}/T$.

3.2 The Estimator of θ

The global profile likelihood in the parameter vector θ is given by

$$\mathcal{L}(\theta; \widehat{g}_\theta) = \sum_{j=1}^{n}\sum_{t=t_i}^{T}\left(y_{jt}-\alpha_j-\beta_j^\top D_t-\widehat{g}_\theta(t/T)\right)^2.$$

We maximize this subject to the constraint that $\sum_{i=1}^{n}\alpha_i = 0$, equivalently finding the first order condition of the Lagrangian $L(\theta,\lambda) = \mathcal{L}(\theta;\widehat{g}_\theta) + \lambda\sum_{i=1}^{n}\alpha_i$.

The first derivatives of $\mathcal{L}$ with respect to θ are:

$$\frac{\partial \mathcal{L}(\theta;\widehat{g}_\theta)}{\partial \alpha_i} = 2\sum_{j=1}^{n}\sum_{t=t_j}^{T}\widehat{\varepsilon}_{jt}(\theta)\frac{\partial \widehat{\varepsilon}_{jt}(\theta)}{\partial \alpha_i} \quad ; \quad \frac{\partial \mathcal{L}(\theta;\widehat{g}_\theta)}{\partial \beta_i} = 2\sum_{j=1}^{n}\sum_{t=t_j}^{T}\widehat{\varepsilon}_{jt}(\theta)\frac{\partial \widehat{\varepsilon}_{jt}(\theta)}{\partial \beta_i},$$

where $\widehat{\varepsilon}_{jt}(\theta) = y_{jt} - \alpha_j - \beta_j^\top D_t - \widehat{g}_\theta(t/T)$ and

$$\frac{\partial \widehat{\varepsilon}_{jt}(\theta)}{\partial \alpha_i} = \begin{cases} -1 - \frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i} & if\ j = i \\ -\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i} & else \end{cases}$$

$$\frac{\partial \widehat{\varepsilon}_{jt}(\theta)}{\partial \beta_i} = \begin{cases} -D_t - \frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i} & if\ j = i \\ -\frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i} & else \end{cases}$$

for $i = 1, \ldots, n$, where

$$\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i} = -\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T} K_h((t-s)/T) \to \begin{cases} -\frac{1}{m_t}, & i \leq m_t \\ \mathbf{0}, & i > m_t \end{cases}, \text{ as } T \to \infty.$$

$$\frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i} = -\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T} K_h((t-s)/T) D_s \to \begin{cases} -\frac{1}{12 m_t}\mathbf{i}_{11}, & i \leq m_t \\ \mathbf{0}_{11}, & i > m_t \end{cases}, \text{ as } T \to \infty$$

do not depend on the unknown parameters. The profile likelihood equations are linear in θ and can be solved explicitly to give the constrained estimators $\widehat{\theta}$. We then define the nonparametric estimator $\widehat{g}(u) = \widehat{g}_{\widehat{\theta}}(u)$.

3.3 In Matrix Notation

We may re-write the vector of $\widehat{\eta}$ as

$$\widehat{g}_\theta = (\widehat{g}_\theta(1/T), \ldots, \widehat{g}_\theta(1))^\top = (i_n^\top \otimes \mathbf{K})\left(y - A\alpha - \sum_{j=1}^{d} C_j \beta_j\right), \tag{4}$$

where $\mathbf{K}$ is the $T \times T$ smoother matrix with typical element $K_{ts} = K_h((t-s)/T)/m_t T$, and $m_t = T^{-1}\sum_{i=1}^{n}\sum_{s=t_i}^{T} K_h((s-t)/T)$.

In matrix notation the profile likelihood estimator solves

$$\min_{\theta:\alpha^\top i_n = 1} \left(y - A\alpha - \sum_{j=1}^{d} C_j\beta_j - B\widehat{g}_\theta\right)^\top \left(y - A\alpha - \sum_{j=1}^{d} C_j\beta_j - B\widehat{g}_\theta\right)$$

or equivalently, since $\widehat{g}_\theta$ is linear in y,

$$\min_{\theta:\alpha^\top i_n = 1} \left(\widetilde{y} - \widetilde{X}\theta\right)^\top \left(\widetilde{y} - \widetilde{X}\theta\right),$$

where $\theta = (\alpha^\top, \beta_1^\top, \ldots, \beta_d^\top)^\top \in \mathbb{R}^{n(d+1)}$ and $\widetilde{X} = (\widetilde{A}, \widetilde{C}_1, \ldots, \widetilde{C}_d)$ is nT by $n(d+1)$, while: $\widetilde{y} = My$, $\widetilde{A} = MA$, and $\widetilde{C}_j = MC_j$ with $M = I_{nT} - B(i_n^\top \otimes \mathbf{K})$. Ignoring the restriction we can write the above first order conditions in the following matrix form $\widetilde{X}^\top\widetilde{X}\widehat{\theta} = \widetilde{X}^\top\widetilde{y}$, except that $\widetilde{X}^\top\widetilde{X}$ is singular. Define $q^\top = (1, \ldots, 1, 0, \ldots, 0)$, then the linear restriction is represented as $q^\top\theta = 0$. Then define the matrix R, which is a $k \times (k-1)$ matrix, where $k = n(d+1)$, such that (q, R) is non singular and $R^\top q = 0$, Amemiya (1985, §1.4). In this case, we can take

$$R = \begin{bmatrix} R_1 & R_2 \\ R_3 & R_4 \end{bmatrix}, \quad R_1 = \begin{bmatrix} I_{n-1} \\ -i_{n-1} \end{bmatrix}_{n \times n-1}; \quad R_4 = I_{nd \times nd},$$

where i_{n-1} is the $n-1 \times 1$ vector of ones, and R_2, R_3 are matrices of zeros of conformable dimensions. It follows that for the profile likelihood estimator subject to the linear restriction $q^\top \theta = 0$, we have

$$\widehat{\theta} = R\left(R^\top \widetilde{X}^\top \widetilde{X} R\right)^{-1} R^\top \widetilde{X}^\top \widetilde{y},$$

where $R^\top \widetilde{X}^\top \widetilde{X} R$ is non-singular.[1] Then,

$$\widehat{g} = (i_n^\top \otimes \mathbf{K})\left(y - A\widehat{\alpha} - \sum_{j=1}^{d} C_j \widehat{\beta}_j\right).$$

In computing the least squares estimators in our application we make some additional steps because T is very large, 1858 in fact. We partition $A = (A_1^\top, \ldots, A_n^\top)^\top$ and $B = (B_1^\top, \ldots, B_n^\top)^\top$, where A_j and B_j are $T \times n$ matrices and $T \times T$ matrices respectively. Then, for example, $MA = A - ((B_1 \mathbf{K} \sum_{j=1}^n A_j)^\top, \ldots, (B_n \mathbf{K} \sum_{j=1}^n A_j)^\top)^\top$, where $B_{j'} \mathbf{K} A_j$ is a $T \times n$ matrix. In this way one can avoid matrices of dimensions $nT \times nT$ or even $nT \times T$, which are too large to fit into memory.

4 Asymptotic Properties

In this section we present the asymptotic properties of the estimators defined above. The following conditions are quite standard in kernel estimation. For the convenience of asymptotic analysis, we introduce β-mixing (absolutely regular), which is defined as follows. A stationary process $\{(\xi_t, \mathcal{F}_t), -\infty < t < \infty\}$ is said to be β-mixing (or, absolutely regular) if the mixing coefficient $\beta(n)$ defined by

$$\beta(n) = E\Big\{ \sup_{A \in \mathcal{F}_{t+n}^\infty} |P(A|\,\mathcal{F}_{-\infty}^t) - P(A)| \Big\}$$

converges to zero as $n \to \infty$. β-mixing includes many linear and nonlinear time series models as special cases; see Doukhan (1994) for more discussion on mixing.

Assumptions A.

[1] Note that $R_1^\top \alpha = (\alpha_1, \ldots, \alpha_{n-1})^\top$. We can interpret the above as a reparameterizion to $\theta = (\alpha_1, \ldots, \alpha_{n-1}, \beta_1^\top, \ldots \beta_n^\top)^\top$ with $\alpha_n = -\sum_{i=1}^{n-1} \alpha_i$ and then changing $A \mapsto A^*$ in (3) to reflect the different structure. For example, in the special case given above, $A^* = (1, 1, 1, 0, -1, -1)^\top$. Then compute $\widehat{\theta}$ by an unconstrained regression.

1. *For each i, ε_{it} is a stationary β-mixing with mixing decay rate β_{it} with* $\limsup_t b^t \max_{1\leq i\leq n} \beta_{it} < \infty$ *for some* $b > 1$, $\sum_{h=-\infty}^{\infty} E\left(\varepsilon_{it}\varepsilon_{it+h}\right) = \omega_i^2$ *and* $s_i^2 = \sum_{k=-\infty}^{\infty} E\left(\varepsilon_{it}\varepsilon_{i,t+12k}\right)$ *with* $0 < \underline{\omega} \leq \min_{1\leq i\leq n} \omega_i \leq \max_{1\leq i\leq n} \omega_i \leq \overline{\omega} < \infty$.

2. *The function $g : [0,1] \to \mathbb{R}$, is continuously differentiable up to the order $\tau \geq p$.*

3. *The kernel K has support $[-1,1]$ and is symmetric about zero and satisfies $\int K(u)du = 1$. In addition, $\int u^j K(u)du = 0$, $j = 1,\ldots,p-1$, and $\int u^p K(u)du \neq 0$. Define $\mu_p(K) = \int u^p K(u)du$ and $||K||_2^2 = \int K^2(z)dz$.*

4. *The bandwidth satisfies:*

   (a) *As $T \to \infty$, $h \to 0$, and $Th \to \infty$, $Th^{2p} \to 0$*

   (b) $h = c_T T^{-1/2p+1}$ *with* $0 < \liminf_{T\to\infty} c_T \leq \limsup_{T\to\infty} c_T < \infty$.

Assumptions A1 is a typical assumption in the time series literature and ensures that ε_{it} is stationary with weak dependence and that appropriate limiting theory can be applied. This condition is useful in our technical development and, no doubt could be replaced by a range of similar assumptions. Assumption A2 concerns about the smoothness of the trend function and ensures a Taylor expansion to appropriate order. Assumption A3 for the kernel function and Assumption A4 for the bandwidth expansion are quite standard in nonparametric estimation: in part a, the bandwidth is chosen to ensure root-T asymptotics for parametric quantities; in part b, the bandwidth is chosen to be optimal for estimation of the nonparametric component.

The asymptotics depends on our assumptions about $t_1 \leq t_2 \leq \cdots \leq t_n$. In the simplest case when $t_1 \leq t_2 \leq \cdots \leq t_n$ are finite numbers, the asymptotic results are the same as those with complete data - the differences in the starting dates are asymptotically ignorable, thus the asymptotic distributions are unaffected by the difference of starting dates. We shall assume that $t_i \to \infty$ in such a way that

$$t_i = \lfloor r_i T \rfloor \text{, where } r_i \in (0,1), \tag{5}$$

for $i = 1,\ldots,n$, (and $r_{n+1} = 1$) in which case the starting time affects the estimators asymptotically.

To present the main result we need some notation. Let $a_{kj} = \sum_{s=j}^{n} \left(r_{s+1} - r_s\right)/s^k$, $k = 1,2,3,4$, $\delta_i = (1 - r_i - 2a_{1i} + a_{2i})$, $f_i = (n+2)a_{2,i} - 2a_{1,i} - na_{3,i}$, and $\lambda_i = (n^2 a_{4,i} - 4na_{3,i} + 4a_{2,i})$, and let

$$\begin{aligned}
\Omega_n &= \operatorname{diag}\left[\begin{array}{ccccc} \delta_1\omega_1^2, & \ldots, & \delta_i\omega_i^2, & \ldots, & \delta_n\omega_n^2 \end{array}\right], \\
S_n &= \operatorname{diag}\left[\begin{array}{ccccc} \delta_1 s_1^2, & \ldots, & \delta_i s_i^2, & \ldots, & \delta_n s_n^2 \end{array}\right] \\
\Delta_n &= \operatorname{diag}\left\{1,\ldots,1-r_i,\ldots,1-r_n\right\}.
\end{aligned}$$

In addition, let A_n be the $n\times n$ symmetric matrix whose (i,j)-th element is

$$[A_n]_{i,j}=\begin{cases}\lambda_i\sum_{j=1}^{i-1}\omega_j^2+\sum_{j=i+1}^{n}\lambda_j\omega_j^2, & i=j\\ f_i\left(\omega_j^2+\omega_i^2\right)+\lambda_i\sum_{l\neq j,l<i}\omega_l^2+\sum_{l>i}\lambda_l\omega_l^2, & j<i\end{cases},$$

$$G_n=\begin{bmatrix}\left(a_{21}-2a_{11}+\sum_{l=2}^{n}a_{2l}\right) & \dots & \left(ia_{2i}-2a_{1i}+\sum_{l=i+1}^{n}a_{2l}\right) & \dots & \left(na_{2,n}-2a_{1,n}\right)\\ \left(ia_{2i}-2a_{1i}+\sum_{l=i+1}^{n}a_{2l}\right) & & \left(ia_{2i}-2a_{1i}+\sum_{l=i+1}^{n}a_{2l}\right) & & \left(na_{2,n}-2a_{1,n}\right)\\ \left(na_{2,n}-2a_{1,n}\right) & & \left(na_{2,n}-2a_{1,n}\right) & & \left(na_{2n}-2a_{1n}\right)\end{bmatrix}.$$

Then define the matrices:

$$Q=\begin{bmatrix}\Delta_n+G_n & \left(\Delta_n+G_n\right)\otimes\frac{1}{12}\mathbf{i}_{11}\\ \left(\Delta_n+G_n\right)\otimes\frac{1}{12}\mathbf{i}_{11}^{\top} & \Delta_n\otimes\frac{1}{12}I_{11}+G_n\otimes\frac{1}{12^2}J_{11}\end{bmatrix}, \tag{6}$$

$$\Omega=\begin{bmatrix}\Omega_n+A_n & [\Omega_n+A_n]\otimes\frac{1}{12}\mathbf{i}_{11}^{\top}\\ [\Omega_n+A_n]\otimes\frac{1}{12}\mathbf{i}_{11}^{\top} & S_n\otimes\frac{1}{12}I_{11}+A_n\otimes\frac{1}{12^2}J_{11}\end{bmatrix}, \tag{7}$$

where i_{11} is a 11×1 vector of ones, and $J_{11}=i_{11}i_{11}^{\top}$ is a 11×11 matrix of ones, and

$$\mathbf{g}^*=\begin{bmatrix}b\\ b\otimes\frac{1}{12}\mathbf{i}_{11}\end{bmatrix}, b=\begin{bmatrix}b_1, & \dots, & b_i, & \dots, & b_n\end{bmatrix}^{\top},$$

$$b_i=\frac{1}{p!}\mu_p(K)\left[\sum_{l=1}^{n}\left(\int_{r_l}^{1}\delta(s)g^{(p)}(s)\,ds\right)-\left(\int_{r_i}^{1}g^{(p)}(s)\,ds\right)\right]$$

and $\delta(s)$ is a weighting function on $[0,1]$, $\delta(s)=1/j$, if $r_j<s<r_{j+1}$, $j=1,2,\dots,n$. We summarize the limiting distributions as follows.

Theorem 1. *Suppose that Assumptions A1 - A4 hold, and assume that the initial observation condition are given by (5). Then, as $T\to\infty$,*

$$\sqrt{T}\left(R^{\top}\widehat{\theta}-R^{\top}\theta+h^p\left(R^{\top}QR\right)^{-1}R^{\top}\mathbf{g}^*\right)\Rightarrow N\left(0,\left(R^{\top}QR\right)^{-1}R^{\top}\Omega R\left(R^{\top}QR\right)^{-1}\right).$$

Remark 1. The asymptotic distribution of the profile likelihood estimator is complicated largely due to the unbalanced data structure, which affects the limiting distributions under our assumptions.

Remark 2. The partial linear model that we study in this paper may be estimated by other methods - see an early version of this paper ALX(2008) for studies of other methods. Comparing the profile likelihood estimator with the other estimators, the profile likelihood estimator is a joint estimation for the nonparametric and parametric parts, while the other estimators such as the traditional methods used in the literature of partial linear regressions are sequential two-step estimators. It's easy to see that the profile likelihood estimator has a smaller bias term than the two step estimator.

REMARK 3. Heteroskedasticity across i, weak correlation over t, and seasonality all affect the limiting results. These effects are reflected through ω_i^2 and s_i^2 in the limits.

If we consider the special case with complete data, all observations start at $t = 1$, then $r_i = 0$, $i = 1, \ldots, n$, $r_{n+1} = 1$, and we have $\delta(s) = 1/n$, for $0 < s < 1$, $j = 1, 2, \ldots, n$. Consequently

$$b_i = \frac{1}{p!}\mu_p(K)\left[\sum_{l=1}^{n}\left(\int_0^1 \delta(s)g^{(p)}(s)\,ds\right) - \left(\int_0^1 g^{(p)}(s)\,ds\right)\right] = 0.$$

This cancellation occurs because of the recentering due to the parametric part of the model.

Thus we have the following simplified asymptotic results for the profile likelihood estimator with complete data. Let

$$\underline{Q} = \Sigma_X - \frac{1}{n}\Sigma_X^*, \tag{8}$$

$$\Sigma_X = \begin{bmatrix} I_n & \frac{1}{12} I_n \otimes \mathbf{i}_{11}^\top \\ \frac{1}{12} I_n \otimes \mathbf{i}_{11} & \frac{1}{12} I_{11n} \end{bmatrix}, \ \Sigma_X^* = \begin{bmatrix} J_n & \frac{1}{12} J_n \otimes \mathbf{i}_{11} \\ \frac{1}{12} J_n \otimes \mathbf{i}_{11}^\top & \frac{1}{12^2} J_{11n} \end{bmatrix},$$

and $\underline{\Omega}$ is defined by the same formula (7) with

$$\begin{aligned} \Omega_n &= \operatorname{diag}\left[\ \left(1-\tfrac{1}{n}\right)^2\omega_1^2, \quad \ldots, \quad \left(1-\tfrac{1}{n}\right)^2\omega_i^2, \quad \ldots, \quad \left(1-\tfrac{1}{n}\right)^2\omega_n^2\ \right], \\ S_n &= \operatorname{diag}\left[\ \left(1-\tfrac{1}{n}\right)^2 s_1^2, \quad \ldots, \quad \left(1-\tfrac{1}{n}\right)^2 s_i^2, \quad \ldots, \quad \left(1-\tfrac{1}{n}\right)^2 s_n^2\ \right], \end{aligned}$$

and the (i, j)-th element of A_n is given by

$$[A_n]_{i,j} = \begin{cases} \frac{1}{n^2}\sum_{j\neq i}\omega_j^2, & i = j \\ -\frac{1}{n}\left(1-\frac{1}{n}\right)\left(\omega_j^2+\omega_i^2\right) + \frac{1}{n^2}\sum_{l\neq j,i}\omega_l^2, & j < i \end{cases}.$$

COROLLARY 1. *Suppose that Assumptions A1 - A4 hold, in the case with complete data, the profile likelihood estimator has the following asymptotic distribution as* $T \to \infty$,

$$\sqrt{T}\left(R^\top\widehat{\theta} - R^\top\theta\right) \Rightarrow N\left(0, \left(R^\top\underline{Q}R\right)^{-1} R^\top\underline{\Omega}R\left(R^\top\underline{Q}R\right)^{-1}\right).$$

If we further assume that ε_{it} are iid distributed with mean zero and variance σ^2, $\Omega_n = S_n = \left(1-\frac{1}{n}\right)^2\sigma^2 I_n$where I_n is the n-dimensional identity matrix, and the (i, j)-th element of A_n is given by

$$[A_n]_{i,j} = \begin{cases} \frac{1}{n}\left(1-\frac{1}{n}\right)\sigma^2, & i = j \\ -\frac{1}{n}\sigma^2 & j \neq i \end{cases}.$$

We next analyze the estimator of the trend function. The asymptotic results of this estimator is summarized in Theorem 2 below whose proofs are again given in the Appendix.

THEOREM 2. *Suppose that Assumptions A1 - A4 hold, and assume that the initial observation condition are given by (5). Then, as $T \to \infty$,*

$$\sqrt{Th}\left[\widehat{g}(u) - g(u) - h^p b(u)\right] \;\Rightarrow\; N\left(0, \frac{1}{m}\overline{\omega}_m^2 ||K||_2^2\right), \text{ for } u \in [r_m, r_{m+1}),\ m = 1, \ldots, n-1,$$

$$\sqrt{Th}\left[\widehat{g}(u) - g(u) - h^p b(u)\right] \;\Rightarrow\; N\left(0, \frac{1}{n}\omega^2 ||K||_2^2\right), \text{ for } u > r_n.$$

where $b(u) = \frac{1}{p!} g^{(p)}(u)\mu_p(K)$, *while* $\overline{\omega}_m^2 = m^{-1}\sum_{i=1}^m \omega_i^2$, $\omega^2 = n^{-1}\sum_{i=1}^n \omega_i^2$.

In the special case with complete data, we have the following special result.

COROLLARY 2. *Suppose that Assumptions A1 - A4 hold and all observations start at $t = 1$. Then, as $T \to \infty$,*

$$\sqrt{Th}\left[\widehat{g}(u) - g(u) - h^p b(u)\right] \Rightarrow N\left(0, \frac{1}{n}\omega^2 ||K||_2^2\right). \tag{9}$$

REMARK 4. It is possible to extend the above results to allow for cross-sectional dependence as well, since the CLT is coming from the weak dependence in the large time series dimension. Suppose instead that $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})^\top = \Xi(t/T)^{1/2}\eta_t$, where the vector $\eta_t = (\eta_{1t}, \ldots, \eta_{nt})^\top$ is stationary β-mixing with the same decay rate as in assumption A1, while $\Xi(u)$ is a symmetric positive definite matrix of smooth functions. Let $\Psi(s) = E\eta_t\eta_{t+s}^\top$ and $\Psi_\infty = \sum_{s=-\infty}^{\infty}\Psi(s)$. Then the asymptotic variance in (9) becomes $||K||_2^2 i^\top \Xi(u)^{1/2}\Psi_\infty \Xi(u)^{1/2} i/n$, where $i = (1, 1, \ldots, 1)^\top$. However, the results for $\widehat{\theta}$ are much more complicated in this case.

REMARK 5. One can also expect that Theorem 2 continues to hold in the case where $n \to \infty$. In this case, the rate of convergence of $\widehat{g}(u)$ is of order $1/\sqrt{Tmh}$, and if $u > r_n$ this rate is $1/\sqrt{Tnh}$. The precise rates attainable depend on the distribution of the sequence $r_1, r_2, \ldots$ throughout $[0, 1]$. However, the asymptotic distribution is the same regardless of whether n is large or not. The corresponding results for $\widehat{\theta}$ have to be rethought in this case because the dimensions of this parameter vector increases.

5 Forecasting

In this section we consider forecasting based on the semiparametric model (2). In particular, we consider q-step forecasting, i.e. forecasting of $y_{i,T+q}$ based on information upto time T. Our primary interest is to forecast $y_{i,T+q}$ with finite q, although our analysis allows for forecasts with $q \to \infty$ under appropriate expansion rate of q. The common structure in our model allows us to exploit the forecasting gains entailed by these restrictions (reduction in forecasting variance), which amount to homogeneity restrictions in a panel-data environment. These restrictions were found to be helpful in the empirical application of Hoogstrate, Palm, and Pfann (2000) for GDP forecasts. In a recent paper, Issler and Lima (2009) have a theoretical explanation of why these restrictions might work in practice.

Notice that

$$y_{i,T+q} = \alpha_i + \beta_i^\top D_{T+q} + g(1+q/T) + \varepsilon_{i,T+q}.$$

Therefore, a simple forecast for $y_{i,T+q}$, that ignores the error dynamics, can be obtained based on estimators for α_i, β_i and a predictor of $g(1+q/T)$ based on observations $i = 1, \ldots, n$ and $t \leq T$. Since estimators for α_i, β_i are studied in the previous sections, we study forecasting of $g(1+q/T)$ in this section and construct a predictor of $y_{i,T+q}$ using the predicted $g(1+q/T)$. We are also interested in forecasting the average temperature, $\overline{y}_{T+q} = \sum_{i=1}^n y_{i,T+q}/n$, given by

$$\overline{y}_{T+q} = \overline{\beta}^\top D_{T+q} + g(1+q/T) + \overline{\varepsilon}_t, \tag{10}$$

where $\overline{\beta}^\top = \sum_{i=1}^n \beta_i/n$, and $\overline{\varepsilon}_{T+q} = \sum_{i=1}^n \varepsilon_{i,T+q}/n$.

We first consider the simple case when $\{\varepsilon_{it}\}_t$ are martingale difference sequences. Since forecasting of $g(1+q/T)$ is the key issue, we note that

$$E_T y_{i,T+q} = \alpha_i + \beta_i^\top D_{T+q} + g(1+q/T),$$

where E_T denotes conditional expectation given the data.

We make the following assumptions to facilitate forecasting the common trend.

A1' *For each i, ε_{it} is a martingale difference sequence, $E(\varepsilon_{it}^2) = \sigma_i^2$, and $0 < \underline{\sigma} \leq \min_{1\leq i\leq n} \sigma_i \leq \max_{1\leq i\leq n} \sigma_i \leq \overline{\sigma} < \infty$.*

A2' *The function $g : [0, 1+\epsilon] \to \mathbb{R}$, some $\epsilon > 0$, is continuously differentiable up to the order $\tau \geq p$.*

A5 *$\mathcal{K}$ is a one-sided kernel satisfying* (a) *$\mathcal{K}$ and $\mathcal{K}'$ are continuous on $[-1, 0]$;* (b) *$\mu_0^*(\mathcal{K}) > 0$ and $\mu_0^*(\mathcal{K})\mu_2^*(\mathcal{K}) - \mu_1^*(\mathcal{K})^2 > 0$, where $\mu_j^*(\mathcal{K}) = \int_{-1}^0 u^j \mathcal{K}(u)du$.*

A6 *The bandwidth h satisfies A4(a) and the bandwidth h_1 satisfies $h/h_1 \to 0$ as $T \to \infty$.*

We construct a local polynomial predictor for $g(1+q/T)$. Notice that $g(\cdot)$ is a smooth function under Assumption A2'; therefore, when $T \to \infty$, $q/T \to 0$, by a Taylor expansion of $g(\cdot)$ around $u = 1$ to the τ-th order ($\tau = p - 1$),

$$g(1+q/T) = \sum_{k=0}^{\tau} \frac{1}{k!} g^{(k)}(1) \left(\frac{q}{T}\right)^k + o\left(\left(\frac{q}{T}\right)^\tau\right) = \sum_{k=0}^{\tau} \gamma_k \cdot \left(\frac{q}{T}\right)^k + o\left(\left(\frac{q}{T}\right)^\tau\right).$$

As will be more clear later in this section, forecasting at time T is largely affected by data information close to time T. We let

$$\underline{y}_t = n^{-1}\sum_{i=1}^{n}(y_{it} - \widehat{\alpha}_i - \widehat{\beta}_i^\top D_t) = \overline{y}_t - \overline{\widehat{\beta}}^\top D_t,$$

for $t_n \le t \le T$. Let $\mathcal{K}(\cdot)$ be a one-sided kernel whose properties are defined in Assumption A5 above, we consider the following local polynomial estimation at the end point T:

$$\sum_{t=1}^{T}\mathcal{K}\left(\frac{T-t}{Th_1}\right)\left(\underline{y}_t - \sum_{k=0}^{\tau}\gamma_k\cdot\left(\frac{t-T}{T}\right)^k\right)^2. \tag{11}$$

where h_1 is a bandwidth parameter satisfying Assumption A6.

We summarize the asymptotic behavior of the local polynomial estimator (11) in the following Theorem. Let

$$B(\mathcal{K}) = \frac{1}{(\tau+1)!}g^{(\tau+1)}(1)\begin{pmatrix}\mu^*_{\tau+1}(\mathcal{K})\\ \mu^*_{\tau+2}(\mathcal{K})\\ \cdots\\ \mu^*_{2\tau+1}(\mathcal{K})\end{pmatrix}$$

$$M(\mathcal{K}) = \begin{bmatrix}\mu^*_0(\mathcal{K}) & \mu^*_1(\mathcal{K}) & \cdots & \mu^*_\tau(\mathcal{K})\\ \mu^*_1(\mathcal{K}) & \mu^*_2(\mathcal{K}) & & \mu^*_{\tau+1}(\mathcal{K})\\ \cdots & \cdots & & \cdots\\ \mu^*_\tau(\mathcal{K}) & \mu^*_{\tau+1}(\mathcal{K}) & \cdots & \mu^*_{2\tau}(\mathcal{K})\end{bmatrix}, V(\mathcal{K}) = \begin{bmatrix}\nu^*_0(\mathcal{K}) & \nu^*_1(\mathcal{K}) & \cdots & \nu^*_\tau(\mathcal{K})\\ \nu^*_1(\mathcal{K}) & \nu^*_2(\mathcal{K}) & & \nu^*_{\tau+1}(\mathcal{K})\\ \cdots & \cdots & & \cdots.\\ \nu^*_\tau(\mathcal{K}) & \cdots & & \nu^*_{2\tau}(\mathcal{K})\end{bmatrix},$$

and $\mu^*_k(\mathcal{K}) = \int_{-1}^{0}\mathcal{K}(u)\,u^k du$, $\nu^*_j(\mathcal{K}) = \int_{-1}^{0}u^j\mathcal{K}^2(u)du$. Let also $D_h = \mathrm{diag}\left(1, h, \ldots, h^\tau\right)$.

THEOREM 3. *Suppose that Assumptions A1, A2′, A3, A4, A5, and A6 hold, as* $T\to\infty$,

$$\sqrt{Th}D_h\left(\widehat{\gamma} - \gamma - h_1^{\tau+1}M(\mathcal{K})^{-1}B(\mathcal{K})\right) \Rightarrow N\left(0, \frac{1}{n}\sigma^2 M(\mathcal{K})^{-1}V(\mathcal{K})M(\mathcal{K})^{-1}\right),$$

where $\sigma^2 = n^{-1}\sum_{i=1}^{n}\sigma_i^2$.

The above result indicates that the leading bias effect of local polynomial estimation of $(\gamma_0, \gamma_1, \ldots, \gamma_\tau)$ is given by $h^{\tau+1}D_h M(\mathcal{K})^{-1}B(\mathcal{K})$, and the leading variance effect is given by $\omega^2 D_h^{-1}M(\mathcal{K})^{-1}V(\mathcal{K})M(\mathcal{K})^{-1}D_h^{-1}/nTh$. The local polynomial predictor for $g(1+q/T)$ is then given by

$$\widehat{g}(1+q/T) = \sum_{k=0}^{\tau}\widehat{\gamma}_k\cdot\left(\frac{q}{T}\right)^k,$$

and our predictor for $y_{i,T+q}$ is given by

$$\widehat{y}_{i,T+q} = \widehat{\alpha}_i + \widehat{\beta}_i^\top D_{T+q} + \widehat{g}(1+q/T). \tag{12}$$

The forecast for average temperature is just the average forecast, so

$$\widehat{\overline{y}}_{T+q} = \widehat{\overline{\beta}}^{\top} D_{T+q} + \widehat{g}(1 + q/T), \tag{13}$$

where $\widehat{\overline{\beta}} = n^{-1} \sum_{i=1}^{n} \widehat{\beta}_i^{\top}$.

The forecasting error is given in the following theorem. Let $\mathcal{P}_\tau = (1, (q/Th), \ldots, (q/Th)^\tau)^\top$. Let E_T^* denotes asymptotic conditional expectation given the data.

THEOREM 4. *Suppose that Assumptions A1, A2′, A3, A4, and A5 hold, as $T \to \infty$, the forecasting bias in $\widehat{y}_{i,T+q}$ is given by*

$$E_T^*[\widehat{y}_{i,T+q} - y_{i,T+q}] = b_g = h^{\tau+1} \left[\mathcal{P}_\tau^\top M(\mathcal{K})^{-1} B(\mathcal{K}) + o(1)\right],$$

and the forecasting error variance in $\widehat{y}_{i,T+q}$ is given by

$$E_T^* \left[(\widehat{y}_{i,T+q} - E_T \widehat{y}_{i,T+q})^2\right] = \sigma_i^2 + \left(\frac{1}{Tnh}\left[\mathcal{P}_\tau^\top M(\mathcal{K})^{-1} V(\mathcal{K}) M(\mathcal{K})^{-1} \mathcal{P}_\tau + o(1)\right]\right)\sigma^2,$$

where, σ^2 is defined in Theorem 3. For the forecast of average temperature, $\widehat{\overline{y}}_{T+q}$, the forecasting bias is the same as that of $\widehat{y}_{i,T+q}$ given by the above formula, and the forecasting error variance in $\widehat{\overline{y}}_{T+q}$ is given by

$$E_T^* \left[\left(\widehat{\overline{y}}_{T+q} - E_T^* \widehat{\overline{y}}_{T+q}\right)^2\right] = \frac{1}{n}\left(1 + \frac{1}{Th}\left[\mathcal{P}_\tau^\top M(\mathcal{K})^{-1} V(\mathcal{K}) M(\mathcal{K})^{-1} \mathcal{P}_\tau + o(1)\right]\right)\sigma^2.$$

The results of Theorems 3 and 4 indicate that the forecasting error of $\widehat{y}_{i,T+q}$ is dominated by that of the local polynomial forecaster of $\widehat{g}(1+q/T)$. In particular, for the leading case of forecasting with finite q, the bias term is dominated by the first term in b_g : $h^{\tau+1} B_0$, where B_0 is the first element in the $(\tau+1)$-vector $M(\mathcal{K})^{-1} B(\mathcal{K})$. The forecasting error variance is dominated by $\sigma_i^2 + V_0 \sigma^2 / Tnh$, where V_0 is the (1,1)-element of matrix $M(\mathcal{K})^{-1} V(\mathcal{K}) M(\mathcal{K})^{-1}$. Similar result can be obtained for the average temperature forecaster $\widehat{\overline{y}}_{T+q}$. These results also hold for more general cases as long as $q/Th \to 0$.

If we allow that $q \to \infty$, the order of magnitude of the forecasting error is determined jointly by the bandwidth h and the forecasting distance q/T. In the case of $\widehat{y}_{i,T+q}$, if $q/Th \to 0$, the bias term is dominated by the first term in b_g : $h^{\tau+1} B_0$, and the forecasting error variance is dominated by $\sigma_i^2 + V_0 \sigma^2 / Tnh$, where B_0 and V_0 are defined in the same way as above. If $q/Th \to \delta \in (0, \infty)$, the leading bias term is affected by all terms in b_g : $h^{\tau+1} \Delta_\tau^\top M(\mathcal{K})^{-1} B(\mathcal{K})$, where $\Delta_\tau = (1, \delta, \ldots, \delta^\tau)^\top$. The leading variance terms is giving by: $\sigma_i^2 + \Delta_\tau^\top M(\mathcal{K})^{-1} V(\mathcal{K}) M(\mathcal{K})^{-1} \Delta_\tau \sigma^2 / Tnh$. If $q/Th \to \infty$, our theory is not applicable.

REMARK 4. In the general case when $\{\varepsilon_{it}\}_t$ are weakly dependent,

$$E_T y_{i,T+q} = \alpha_i + \beta_i^\top D_{T+q} + g(1 + q/T) + E_T \varepsilon_{i,T+q},$$

where E_T denotes conditional expectation given the data. Under our condition A1, $E_T\varepsilon_{i,T+q} \neq 0$ (although $E_T\varepsilon_{i,T+q} \to 0$ as $q \to \infty$). To forecast $E_T\varepsilon_{i,T+q}$, we should fit a time series model (say, an ARMA model as Box and Jenkins) to the error term, and using the existing forecasting method to construct a predictor. In this case, we may detrend and remove the seasonal components from $y_{i,t}$ using our estimates $\widehat{\alpha}_i$, $\widehat{\beta}_i$, and $\widehat{g}(t/T)$, i.e.

$$\widehat{\varepsilon}_{i,t} = y_{i,t} - \widehat{\alpha}_i - \widehat{\beta}_i^\top D_t - \widehat{g}(t/T)$$

and then fit the estimated stochastic component $\widehat{\varepsilon}_{i,t}$ by an appropriate ARMA model to obtain forecast of $\varepsilon_{i,T+q}$, say, $\widehat{E}_T\varepsilon_{i,T+q}$. A predictor for $y_{i,T+q}$ can then be constructed by $\widehat{g}(1+q/T)$ that we obtained earlier in this section together with other components, i.e.

$$\widetilde{y}_{i,T+q} = \widehat{\alpha}_i + \widehat{\beta}_i^\top D_{T+q} + \widehat{g}(1+q/T) + \widehat{E}_T\varepsilon_{i,T+q}.$$

In the AR(1) special case $\varepsilon_{i,t} = \rho\varepsilon_{i,t-1} + \eta_{it}$, where η_{it} is iid, we have $E_T\varepsilon_{i,T+q} = \rho^q\varepsilon_{i,T}$. More generally, for ARMA process errors one could use the standard linear forecasting techniques associated with Box and Jenkins. Alternatively, we may ignore the error dynamics and simple construct forecasts for $y_{i,T+q}$ and $\overline{y}_{T+q}$ by (12) and (13). Such predictors are asymptotically equivalent to predictors that takes into account the weak correlation in $\varepsilon_{i,t}$ for long-run forecasting (the case $q \to \infty$), but are less efficient for short-run forecasting than predictors that utilize the correlation property.

6 Application

Our dataset contains the average maximum temperature within a month ($TMAX$), the average minimum temperature within a month ($TMIN$), the difference between the average maximum and minimum temperatures within a month b($TRANGE$), all measured in degrees Celsius and also the number of hours of sunshine and the number of millimeters of rainfall. The primary data source is the met office web site for each of the twenty six stations.[2] The first observations were taken in 1853 at Armagh and Oxford so that we have a total of 1858 time series records.

In the working paper version of this paper we provide the full results of a univariate parametric analysis based on a quadratic trend. This shows evidence of seasonality and an upward trend for all stations. There is also some evidence of serial correlation in the residuals but little evidence of GARCH effects. The error correlation does not affect the estimation of the regression coefficients and changes only slightly the standard errors. Similar results were obtained for both maximum and minimum temperature. We also report results for the range. These are somewhat different.

[2] The data are available at http://www.metoffice.gov.uk/climate/uk/stationdata/

Specifically, the trend coefficients are significant in only nine cases, with seven of those cases having a similar upward trend, whereas the other two actually have a negative trend in range. Range has also a significant seasonal effect and a significant autocorrelation coefficient in most cases. The results for sunshine hours are not so consistent as for temperature. There are seven stations with significant trends, six of them with increasing trend. Overall though many other stations have negative, albeit insignificant, trends. With rainfall, the trend is not significant in any station.

One critique of such a parametric analysis is that the implied trend is a little unrealistic and poorly estimated. Extrapolating beyond the sample implies an outrageously high temperature twenty years from now, which is just not credible. This is why we have advocated a semiparametric approach.

We next present the results of the semiparametric analysis. In Tables 1 and 2 we give the estimated values of θ and the associated standard errors for TMAX and TMIN. The parameter values are strongly significant and show evidence of geographic variability in the level of temperature and seasonality. These results are broadly consistent with the individual purely parametric results we gave in the working paper version.

We present in Figures 2 and 3 the implied trend from the parametric analysis. The jagged nature of the graph is caused by the introduction of new stations. Also note that the implied trend at the end of the period is quite extreme. Our results are somewhat different from those obtained in Gao and Hawthorne (2006) for example, since we find evidence of trend starting much later. In Figures 4 and 5 we give the estimated nonparametric trend over the same period. The trend is much more moderate especially at the end of the period. In Figures 6 and 7 we give the trend just for the recent period by only considering the balanced subset of the data. Even though the nonparametric trend indicates some variation i.e., some downward movements, but generally it climbs upward, this being more pronounced after 1995. In both cases, balanced and unbalanced, we can easily claim that there is an upward trend for the TMAX and TMIN values. These were implemented using a Gaussian kernel and Silverman's rule of thumb bandwidth (which in this case yield $h \simeq 0.05$). As we remarked in the text, the estimation of the common trend is purely local and unaffected by earlier data. The standard errors for the nonparametric estimators of TMAX and TMIN over the shown period are 0.476709, 0.48602 respectively, indicating the level of significance of the estimated curves.

We next present the result of an out of sample analysis. We compute the estimated forecast based on local linear smoothing. We report the absolute error for the p-step forecast, where $p = 1, 2, \ldots, 12$, so forecasting out to one year ahead. The forecast errors given in Figures 8 and 9 appear reasonable and are better than the corresponding parametric results, which substantially overpredict the temperature in this period.

****Figures and Tables Here***

7 Conclusion

In conclusion, we have developed a semiparametric model we think is appropriate for modelling the changes in temperatures observed at a cross section of locations. The model and methods are defined for the important practical case of unbalanced data. The methods we develop give similar results to a parametric analysis and help to confirm the main finding of a gradual upward trend in temperature in the UK, although with somewhat less trend obtained by the nonparametric method than the parametric one.

8 Appendix

8.1 Proof of Theorems

Proof of Theorem 1. The first order condition (FOC) for θ is

$$\begin{aligned}\frac{\partial \mathcal{L}(\theta)}{\partial \alpha_i} &= -\sum_{j\neq i}\sum_{t=t_j}^{T}\left(y_{jt}-\widehat{\alpha}_j-\widehat{\beta}_j^\top D_t-\widehat{g}_\theta(t/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\\ &\quad -\sum_{t=t_i}^{T}\left(y_{it}-\widehat{\alpha}_i-\widehat{\beta}_i^\top D_t-\widehat{g}_\theta(t/T)\right)\left(1+\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right)=0\end{aligned}$$

$$\begin{aligned}\frac{\partial \mathcal{L}(\theta)}{\partial \beta_i} &= -\sum_{j\neq i}\sum_{t=t_j}^{T}\left(y_{jt}-\widehat{\alpha}_j-\widehat{\beta}_j^\top D_t-\widehat{g}_\theta(t/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i}\\ &\quad -\sum_{t=t_i}^{T}\left(y_{it}-\widehat{\alpha}_i-\widehat{\beta}_i^\top D_t-\widehat{g}_\theta(t/T)\right)\left(D_t+\frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i}\right)=0,\end{aligned}$$

where:

$$\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i} = -\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)\rightarrow\begin{cases}-\frac{1}{m_t}, & i\leq m_t\\ \mathbf{0}, & i>m_t\end{cases}$$

$$\frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i} = -\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)D_s\rightarrow\begin{cases}-\frac{1}{12m_t}\mathbf{i}_{11}, & i\leq m_t\\ \mathbf{0}_{11}, & i>m_t\end{cases}.$$

Thus, for $i=1,\ldots,n$,

$$\begin{aligned}&\sum_{l\neq i}\sum_{t=t_l}^{T}\left(y_{lt}-\widehat{\alpha}_l-\widehat{\beta}_l^\top D_t-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}\left(y_{js}-\widehat{\alpha}_j-\widehat{\beta}_j^\top D_s\right)K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}+\\ &\sum_{t=t_i}^{T}\left(y_{it}-\widehat{\alpha}_i-\widehat{\beta}_i^\top D_t-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}\left(y_{js}-\widehat{\alpha}_j-\widehat{\beta}_j^\top D_s\right)K_h((t-s)/T)\right)\left(1+\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right)=0,\end{aligned}$$

$$\sum_{l\neq i}\sum_{t=t_l}^{T}\left(y_{lt}-\widehat{\alpha}_l-\widehat{\beta}_l^{\top}D_t-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}\left(y_{js}-\widehat{\alpha}_j-\widehat{\beta}_j^{\top}D_s\right)K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}$$

$$\sum_{t=t_i}^{T}\left(y_{it}-\widehat{\alpha}_i-\widehat{\beta}_i^{\top}D_t-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}\left(y_{js}-\widehat{\alpha}_j-\widehat{\beta}_j^{\top}D_s\right)K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)=0,$$

Substitute the true model $y_{it}=\alpha_i+\beta_i^{\top}D_t+g(t/T)+\varepsilon_{it}$ into the above FOC, notice that

$$y_{it}-\widehat{\alpha}_i-\widehat{\beta}_i^{\top}D_t=\varepsilon_{it}+g(t/T)-(\widehat{\alpha}_i-\alpha_i)-\left(\widehat{\beta}_i^{\top}-\beta_i^{\top}\right)D_t,$$

thus we have, for $i=1,\ldots,n$, the corresponding FOC w.r.t. α_i is given by

$$
\begin{aligned}
& \sum_{l\neq i}\left[\sum_{t=t_l}^{T}\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right](\widehat{\alpha}_l-\alpha_l)+\sum_{l\neq i}\left[\sum_{t=t_l}^{T}\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}D_t\right]\left(\widehat{\beta}_l^{\top}-\beta_l^{\top}\right) \\
& -\sum_{j\neq i}\left[\frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_j}^{T}K_h((t-s)/T)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right)\right](\widehat{\alpha}_j-\alpha_j) \\
& -\left[\frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right](\widehat{\alpha}_i-\alpha_i) \\
& -\sum_{j\neq i}\left(\widehat{\beta}_j^{\top}-\beta_j^{\top}\right)\left[\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T}D_sK_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right] \\
& -\left(\widehat{\beta}_i^{\top}-\beta_i^{\top}\right)\left[\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T}D_sK_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right] \\
& +(\widehat{\alpha}_i-\alpha_i)\sum_{t=t_i}^{T}\left(1-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right) \\
& +\left(\widehat{\beta}_i^{\top}-\beta_i^{\top}\right)\sum_{t=t_i}^{T}\left(D_t-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}D_sK_h((t-s)/T)\right)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right) \\
& -\sum_{j\neq i,j=1}^{n}(\widehat{\alpha}_j-\alpha_j)\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right) \\
& -\sum_{j\neq i,j=1}^{n}\left(\widehat{\beta}_j^{\top}-\beta_j^{\top}\right)\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T}D_sK_h((t-s)/T)\right)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right) \\
= & \sum_{l\neq i}\sum_{t=t_l}^{T}\left(\varepsilon_{lt}-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}\varepsilon_{js}K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i} \\
& +\sum_{l\neq i}\sum_{t=t_l}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i} \\
& +\sum_{t=t_i}^{T}\left(\varepsilon_{it}-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}\varepsilon_{is}K_h((t-s)/T)\right)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right) \\
& +\sum_{t=t_i}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right) \\
& -\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{j\neq i,j=1}^{n}\sum_{s=t_j}^{T}\varepsilon_{js}K_h((t-s)/T)\right)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right)
\end{aligned}
$$

and the corresponding FOC w.r.t. β_i is

$$
\begin{aligned}
& \sum_{l\neq i}\left[\sum_{t=t_l}^{T}\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right](\widehat{\alpha}_l-\alpha_l)+\sum_{l\neq i}\left[\sum_{t=t_l}^{T}\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}D_t\right]\left(\widehat{\beta}_l^\top-\beta_l^\top\right)\\
& -\sum_{j\neq i}\left[\frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_j}^{T}K_h((t-s)/T)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\right](\widehat{\alpha}_j-\alpha_j)\\
& -\left[\frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right](\widehat{\alpha}_i-\alpha_i)\\
& -\sum_{j\neq i}\left(\widehat{\beta}_j^\top-\beta_j^\top\right)\left[\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T}D_sK_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right]\\
& -\left(\widehat{\beta}_i^\top-\beta_i^\top\right)\left[\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T}D_sK_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right]\\
& +(\widehat{\alpha}_i-\alpha_i)\sum_{t=t_i}^{T}\left(1-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\\
& +\left(\widehat{\beta}_i^\top-\beta_i^\top\right)\sum_{t=t_i}^{T}\left(D_t-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}D_sK_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\\
& -\sum_{j\neq i,j=1}^{n}(\widehat{\alpha}_j-\alpha_j)\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\\
& -\sum_{j\neq i,j=1}^{n}\left(\widehat{\beta}_j^\top-\beta_j^\top\right)\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T}D_sK_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\\
= \quad & \sum_{l\neq i}\sum_{t=t_l}^{T}\left(\varepsilon_{lt}-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}\varepsilon_{js}K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\\
& +\sum_{l\neq i}\sum_{t=t_l}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\\
& +\sum_{t=t_i}^{T}\left(\varepsilon_{it}-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}\varepsilon_{is}K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\\
& +\sum_{t=t_i}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\\
& -\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{j\neq i,j=1}^{n}\sum_{s=t_j}^{T}\varepsilon_{js}K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right).
\end{aligned}
$$

If we denote:

$$C_{T,a} = \begin{bmatrix} C_{a,11} & \dots & C_{a,1n} \\ \dots & & \dots \\ C_{a,n1} & \dots & C_{a,nn} \end{bmatrix}, C_{T,b} = \begin{bmatrix} C_{b,11} & \dots & C_{b,1n} \\ \dots & & \dots \\ C_{b,n1} & \dots & C_{b,nn} \end{bmatrix}$$

$$C_{T,A} = \begin{bmatrix} C_{A,11} & \dots & C_{A,1n} \\ \dots & & \dots \\ C_{A,n1} & \dots & C_{A,nn} \end{bmatrix}, C_{T,B} = \begin{bmatrix} C_{B,11} & \dots & C_{B,1n} \\ \dots & & \dots \\ C_{B,n1} & \dots & C_{B,nn} \end{bmatrix}$$

$$d_a = \begin{bmatrix} d_{a,1} \\ \dots \\ d_{a,n} \end{bmatrix}, d_A = \begin{bmatrix} d_{A,1} \\ \dots \\ d_{A,n} \end{bmatrix}, \; e_a = \begin{bmatrix} e_{a,1} \\ \dots \\ e_{a,n} \end{bmatrix}, e_A \begin{bmatrix} e_{A,1} \\ \dots \\ e_{A,n} \end{bmatrix},$$

$$\begin{aligned}
C_{a,ii} &= \frac{1}{T} \left\{ \begin{array}{c} \sum_{t=t_i}^{T} \left(1 - \frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T} K_h((t-s)/T)\right)\left(1 + \frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right) \\ -\left[\frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_i}^{T} K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right] \end{array} \right\} \\
C_{a,ij} &= \frac{1}{T} \left[\begin{array}{c} \sum_{t=t_j}^{T} \frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i} - \frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_j}^{T} K_h((t-s)/T)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right) \\ -\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T} K_h((t-s)/T)\right)\left(1 + \frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right) \end{array} \right] \\
C_{b,ii} &= \frac{1}{T} \left\{ \begin{array}{c} \sum_{t=t_i}^{T} \left(D_t^\top - \frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T} D_s^\top K_h((t-s)/T)\right)\left(1 + \frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right) \\ -\left[\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T} D_s^\top K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right] \end{array} \right\} \\
C_{b,ij} &= \frac{1}{T} \left[\begin{array}{c} \sum_{t=t_j}^{T} D_t^\top \frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i} - \frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_j}^{T} D_s^\top K_h((t-s)/T)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right) \\ -\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T} D_s K_h((t-s)/T)\right)\left(1 + \frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right) \end{array} \right]
\end{aligned}$$

$$\begin{aligned}
d_{a,i} &= \frac{1}{\sqrt{T}} \sum_{l\neq i}\sum_{t=t_j}^{T}\left(g(t/T) - \frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T} g(s/T)K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i} \\
&\quad + \sum_{t=t_i}^{T}\left(g(t/T) - \frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T} g(s/T)K_h((t-s)/T)\right)\left(1 + \frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right)
\end{aligned}$$

$$
\begin{aligned}
e_{a,i} &= \frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}\left(\varepsilon_{it}-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}\varepsilon_{is}K_h((t-s)/T)\right)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right)\\
&\quad -\frac{1}{\sqrt{T}}\sum_{s=t_i}^{T}\left(\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\frac{1}{T}K_h((t-s)/T)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right)\varepsilon_{is}\\
&\quad +\frac{1}{\sqrt{T}}\sum_{j\neq i}\left(\sum_{t=t_j}^{T}\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right)\varepsilon_{jt}-\sum_{j\neq i,j=1}^{n}\frac{1}{T}\sum_{s=t_j}^{T}\left(\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}K_h((t-s)/T)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right)\varepsilon_{js}\\
&\quad -\frac{1}{\sqrt{T}}\sum_{j\neq i,j=1}^{n}\frac{1}{T}\sum_{s=t_j}^{T}\left(\sum_{t=t_i}^{T}\frac{1}{m_t}K_h((t-s)/T)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right)\right)\varepsilon_{js}
\end{aligned}
$$

$$
\begin{aligned}
C_{A,ii} &= \frac{1}{T}\left\{\begin{array}{c}\sum_{t=t_i}^{T}\left(1-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\\ -\left[\frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right]\end{array}\right\}\\
C_{A,ij} &= \frac{1}{T}\left[\begin{array}{c}\sum_{t=t_j}^{T}\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}-\frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_j}^{T}K_h((t-s)/T)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\\ -\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\end{array}\right]\\
C_{B,ii} &= \frac{1}{T}\left\{\begin{array}{c}\sum_{t=t_i}^{T}\left(D_t^\top-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}D_s^\top K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\\ -\left[\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T}D_s^\top K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right]\end{array}\right\}\\
C_{B,ij} &= \frac{1}{T}\left[\sum_{t=t_j}^{T}D_t^\top\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}-\frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_j}^{T}D_s^\top K_h((t-s)/T)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\right]\\
&\quad -\frac{1}{T}\left[\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T}D_sK_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
d_{A,i} &= \frac{1}{\sqrt{T}}\sum_{l\neq i}\sum_{t=t_j}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\\
&\quad +\sum_{t=t_i}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)
\end{aligned}
$$

$$
\begin{aligned}
e_{A,i} &= \frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}\left(\varepsilon_{it}-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}\varepsilon_{is}K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right) \\
&\quad -\frac{1}{\sqrt{T}}\sum_{s=t_i}^{T}\left(\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\frac{1}{T}K_h((t-s)/T)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right)\varepsilon_{is} \\
&\quad +\frac{1}{\sqrt{T}}\sum_{j\neq i}\left(\sum_{t=t_j}^{T}\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\varepsilon_{jt}-\sum_{j\neq i,j=1}^{n}\frac{1}{T}\sum_{s=t_j}^{T}\left(\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}K_h((t-s)/T)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\varepsilon_{js} \\
&\quad -\frac{1}{\sqrt{T}}\sum_{j\neq i,j=1}^{n}\frac{1}{T}\sum_{s=t_j}^{T}\left(\sum_{t=t_i}^{T}\frac{1}{m_t}K_h((t-s)/T)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\right)\varepsilon_{js},
\end{aligned}
$$

then we have

$$
\begin{bmatrix} C_{T,a} & C_{T,b} \\ C_{T,A} & C_{T,B} \end{bmatrix}\begin{bmatrix} \sqrt{T}\left(\widehat{\alpha}-\alpha\right) \\ \sqrt{T}\left(\widehat{\beta}-\beta\right) \end{bmatrix}=\begin{bmatrix} d_a \\ d_A \end{bmatrix}+\begin{bmatrix} e_a \\ e_A \end{bmatrix}.
$$

Let

$$
C_T=\begin{bmatrix} C_{T,a} & C_{T,b} \\ C_{T,A} & C_{T,B} \end{bmatrix},\ d_T=\begin{bmatrix} d_a \\ d_A \end{bmatrix},\ e_T=\begin{bmatrix} e_a \\ e_A \end{bmatrix},
$$

the FOC can be written as:

$$
C_T\sqrt{T}\left(\widehat{\theta}-\theta\right)=d_T+e_T. \tag{14}
$$

Thus the profile likelihood estimator subject to the linear restriction $q^\top\theta=0$ satisfies

$$
\sqrt{T}\left(\widehat{\theta}-\theta\right)=R\left(R^\top C_T R\right)^{-1}R^\top d_T+R\left(R^\top C_T R\right)^{-1}R^\top e_T,
$$

where R is the $K\times(K-1)$ normalized orthogonal complements of q.

By results of Lemmas 1 and 2, as $T\to\infty$:

$$
C_{T,a}\Rightarrow\begin{bmatrix} c_{11} & \dots & c_{1i} & \dots & c_{1n} \\ & & & & \\ c_{i1} & & c_{ii} & & c_{in} \\ & & & & \\ c_{n1} & & c_{ni} & & c_{nn} \end{bmatrix}=\Delta_n+G_n=C_n,
$$

$$\begin{aligned}
C_{T,b} &\rightarrow C_n \otimes \left(\frac{1}{12}\mathbf{i}_{11}^{\top}\right) \\
&= \begin{bmatrix} c_{11} & \dots & c_{1i} & \dots & c_{1n} \\ \\ c_{i1} & & c_{ii} & & c_{in} \\ \\ c_{n1} & & c_{ni} & & c_{nn} \end{bmatrix} \otimes \left(\frac{1}{12}\mathbf{i}_{11}^{\top}\right) = (\Delta_n + G_n) \otimes \left(\frac{1}{12}\mathbf{i}_{11}^{\top}\right),
\end{aligned}$$

$$\begin{aligned}
C_{T,A} &\rightarrow C_n \otimes \left(\frac{1}{12}\mathbf{i}_{11}\right) \\
&= \begin{bmatrix} c_{11} & c_{1i} & c_{1n} \\ \\ c_{i1} & c_{ii} & c_{in} \\ \\ c_{n1} & c_{ni} & c_{nn} \end{bmatrix} \otimes \left(\frac{1}{12}\mathbf{i}_{11}\right) = (\Delta_n + G_n) \otimes \left(\frac{1}{12}\mathbf{i}_{11}\right),
\end{aligned}$$

and

$$C_{T,B} \rightarrow \Delta_n \otimes \frac{1}{12} I_{11} + G_n \otimes \frac{1}{12^2}\mathbf{i}_{11}^{\top}\mathbf{i}_{11}.$$

Thus

$$C_T \rightarrow Q = \begin{bmatrix} \Delta_n + G_n & (\Delta_n + G_n) \otimes \frac{1}{12}\mathbf{i}_{11} \\ (\Delta_n + G_n) \otimes \frac{1}{12}\mathbf{i}_{11}^{\top} & \Delta_n \otimes \frac{1}{12} I_{11} + G_n \otimes \frac{1}{12^2}\mathbf{i}_{11}^{\top}\mathbf{i}_{11} \end{bmatrix}.$$

By Lemma 3, the bias terms are

$$\begin{bmatrix} d_a \\ d_A \end{bmatrix} = -\sqrt{T}h^p \begin{bmatrix} b \\ b \otimes \frac{1}{12}\mathbf{i}_{11} \end{bmatrix} + o(\sqrt{T}h^p),$$

where:

$$b = \begin{bmatrix} b_1, & \dots, & b_i, & \dots, & b_n \end{bmatrix}^{\top}$$

$$b_i = \frac{1}{p!}\mu_p(K) \left[\sum_{l \neq i} \left(\int_{r_l}^{1} \delta(s) g^{(p)}(s)\, ds \right) - \left(\int_{r_i}^{1} w(s) g^{(p)}(s)\, ds \right) \right],$$

where $w(s)$ and $\delta(s)$ are weighting functions on [0,1] :

$$\begin{aligned}
\delta(s) &= \frac{1}{j}, \text{ if } r_j < s < r_{j+1},\ j = 1, 2, \dots, n. \\
w(s) &= 1 - \delta(s) = 1 - \frac{1}{j}, \text{ if } r_j < s < r_{j+1},\ j = 1, 2, \dots, n.
\end{aligned}$$

By Lemma 4, the stochastic term e_T converge in distribution to a multivariate normal with covariance matrix

$$\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} = \begin{bmatrix} \Omega_n + A_n & [\Omega_n + A_n] \otimes \frac{1}{12}\mathbf{i}_{11}^{\top} \\ [\Omega_n + A_n] \otimes \frac{1}{12}\mathbf{i}_{11}^{\top} & S_n \otimes \frac{1}{12} I_{11} + A_n \otimes \frac{1}{12^2} J_{11} \end{bmatrix}.$$

■

Proof of Theorem 2. Consider

$$\widehat{g}_P(u) = \frac{T^{-1}\sum_{i=1}^{n}\sum_{s=t_i}^{T}\left(y_{is} - \widehat{\alpha}_i - \widehat{\beta}_i^{\top} D_s\right) K_h(u - s/T)}{T^{-1}\sum_{i=1}^{n}\sum_{s=t_i}^{T} K_h(u - s/T)}.$$

If $\frac{t_m}{T} < u < \frac{t_{m+1}}{T}$, $\sum_{i=1}^{n}\sum_{t=t_i}^{T} K_h(u - t/T)/T = \sum_{i=1}^{m}\sum_{t=t_i}^{T} K([u - t/T]/h)/Th = m$. Therefore,

$$\begin{aligned}
\widehat{g}_P(u) &= \frac{1}{Tm}\sum_{i=1}^{n}\sum_{t=t_i}^{T} K_h(u - t/T)\left(y_{it} - \widehat{\alpha}_i - \widehat{\beta}_i^{\top} D_t\right) \\
&= \frac{1}{Tmh}\sum_{i=1}^{m}\sum_{t=t_i}^{T} K(u - t/T)\left(y_{it} - \widehat{\alpha}_i - \widehat{\beta}_i^{\top} D_t\right) \\
&= \frac{1}{Tmh}\sum_{i=1}^{m}\sum_{t=t_i}^{T} K([u - t/T]/h)\left(y_{it} - \alpha_i - \beta_i^{\top} D_t - (\widehat{\alpha}_i - \alpha_i) - \left(\widehat{\beta}_i^{\top} - \beta_i^{\top}\right) D_t\right) \\
&= \frac{1}{Tmh}\sum_{i=1}^{m}\sum_{t=t_i}^{T} K([u - t/T]/h)\left(g(t/T) + \varepsilon_{it} - (\widehat{\alpha}_i - \alpha_i) - \left(\widehat{\beta}_i^{\top} - \beta_i^{\top}\right) D_t\right) \\
&= \frac{1}{Tmh}\sum_{i=1}^{m}\sum_{t=t_i}^{T} K([u - t/T]/h) g(t/T) + \frac{1}{Tmh}\sum_{i=1}^{m}\sum_{t=t_i}^{T} K([u - t/T]/h)\varepsilon_{it} \\
&\quad - \frac{1}{Tmh}\sum_{i=1}^{m}\sum_{t=t_i}^{T} K([u - t/T]/h)(\widehat{\alpha}_i - \alpha_i) \\
&\quad - \frac{1}{Tmh}\sum_{i=1}^{m}\sum_{t=t_i}^{T} K([u - t/T]/h)\left(\widehat{\beta}_i^{\top} - \beta_i^{\top}\right) D_t.
\end{aligned}$$

For the first stochastic term,

$$\frac{1}{Tmh}\sum_{i=1}^{m}\sum_{t=t_i}^{T} K([u - t/T]/h)\varepsilon_{it} = \frac{1}{m}\sum_{i=1}^{m}\left\{\frac{1}{Th}\sum_{t=t_i}^{T} K([u - t/T]/h)\varepsilon_{it}\right\}$$

Again, for each i, $\sum_t K([u - t/T]/h)\varepsilon_{it}$ is a weighted sum of weakly correlated random variables and a CLT applies,

$$\frac{1}{\sqrt{Th}}\sum_{t=t_i}^{T} K([u - t/T]/h)\varepsilon_{it} \Rightarrow \omega_i ||K||_2^{21/2}\xi_i.$$

The second term is simply a kernel smoothed estimator of $g(u)$,

$$\begin{aligned}
&\frac{1}{Tmh}\sum_{i=1}^{m}\sum_{t=t_i}^{T}K([u-t/T]/h)g(t/T)\\
=\;&\frac{1}{m}\sum_{i=1}^{m}\frac{1}{Th}\sum_{t=t_i}^{T}K([u-t/T]/h)g(t/T)\\
=\;&\frac{1}{m}\sum_{i=1}^{m}\frac{1}{Th}\sum_{t=t_i}^{T}K([u-t/T]/h)\left\{g(u)+\sum_{j=1}^{p}\frac{1}{j!}h^j\left(\frac{u-t/T}{h}\right)^j g^{(j)}(u)\right\}+o(h^p)\\
=\;&\frac{1}{m}\sum_{i=1}^{m}\left(g(u)+\frac{1}{p!}h^p g^{(p)}(u)\int_0^1 z^p K(z)dz+o(h^p).\right)\\
=\;&g(u)+\frac{1}{p!}h^p g^{(p)}(u)\int_0^1 z^p K(z)dz+o(h^p).
\end{aligned}$$

For the third and fourth terms,

$$\frac{1}{Tmh}\sum_{i=1}^{m}\sum_{t=t_i}^{T}K([u-t/T]/h)\left(\widehat{\alpha}_i-\alpha_i\right)=o_p\left(\frac{1}{\sqrt{Th}}\right),$$
$$\frac{1}{Tmh}\sum_{i=1}^{m}\sum_{t=t_i}^{T}K([u-t/T]/h)\left(\widehat{\beta}_i^\top-\beta_i^\top\right)D_t=o_p\left(\frac{1}{\sqrt{Th}}\right),$$

the preliminary estimation of θ does not affect the first order asymptotics for this estimator.

Thus for $t_m/T<u<t_{m+1}/T$, $m=1,\ldots,n-1$,

$$\sqrt{Th}\left[\widetilde{g}^*(u)-g(u)-h^p b(u)\right]\Rightarrow N\left(0,\frac{1}{m}\left(\frac{1}{m}\sum_{i=1}^{m}\omega_i^2\right)||K||_2^2\right).$$

For $u>t_n/T$,

$$\sqrt{Th}\left[\widetilde{g}^*(u)-g(u)-h^p b(u)\right]\Rightarrow N\left(0,\frac{1}{n}\left(\frac{1}{n}\sum_{i=1}^{n}\omega_i^2\right)||K||_2^2\right).$$

$\blacksquare$

Proof of Theorems 3 and 4. Notice that when $q/T\to 0$, as $T\to\infty$, under Assumption A2′, by a Taylor expansion,

$$g(1+q/T)=\sum_{k=0}^{\tau}\frac{1}{k!}g^{(k)}(1)\left(\frac{q}{T}\right)^k+o\left(\left(\frac{q}{T}\right)^\tau\right)=\sum_{k=0}^{\tau}\gamma_k\cdot\left(\frac{q}{T}\right)^k+o\left(\left(\frac{q}{T}\right)^\tau\right).$$

The local polynomial estimation at the *end point* T is given as follows:

$$\sum_{t=1}^{T}\mathcal{K}\left(\frac{T-t}{Th}\right)\left(\underline{y}_t-\gamma^\top x_t\right)^2,\quad\text{where }\gamma=\begin{pmatrix}\gamma_0\\ \vdots\\ \gamma_\tau\end{pmatrix},\; x_t=\begin{pmatrix}1\\ \vdots\\ \left(\frac{t-T}{T}\right)^\tau\end{pmatrix}.$$

The local polynomial estimator can be written as

$$\begin{aligned}
\widehat{\gamma} &= \gamma + \left[\sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) x_t x_t^\top\right]^{-1} \sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) x_t \overline{\varepsilon}_t \\
&\quad - \left[\sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) x_t x_t^\top\right]^{-1} \sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) x_t \left(\left(\widehat{\overline{\beta}}^\top - \overline{\beta}^\top\right) D_t\right) \\
&\quad + \left[\sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) x_t x_t^\top\right]^{-1} \frac{1}{(\tau+1)!} g^{(\tau+1)}(1) \sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) \left(x_t \left(\frac{t-T}{T}\right)^{\tau+1}\right).
\end{aligned}$$

By result of ALX(2008),

$$\begin{aligned}
\widehat{\gamma} &= \gamma + \left[\sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) x_t x_t^\top\right]^{-1} \sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) x_t \overline{\varepsilon}_t \\
&\quad + \left[\sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) x_t x_t^\top\right]^{-1} \frac{h^{\tau+1}}{(\tau+1)!} g^{(\tau+1)}(1) \sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) \left(x_t \left(\frac{t-T}{Th}\right)^{\tau+1}\right) + o_p((Th)^{-1/2} + h^{\tau+1})
\end{aligned}$$

Notice that, under Assumption 5,

$$\frac{1}{Th} \sum_{t=1}^{T} \mathcal{K}\left(\frac{t-T}{Th}\right) \left(\frac{t-T}{Th}\right)^{k} \to \int_{-1}^{0} \mathcal{K}(u)\, u^k du = \mu_k^*(\mathcal{K}),$$

and thus

$$\frac{1}{Th} \sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) x_t x_t^\top \to \left[\begin{array}{cccc} \mu_0^*(\mathcal{K}) & \mu_1^*(\mathcal{K}) & \dots & \mu_\tau^*(\mathcal{K}) \\ \mu_1^*(\mathcal{K}) & \mu_2^*(\mathcal{K}) & & \mu_{\tau+1}^*(\mathcal{K}) \\ \dots & \dots & & \dots \\ \mu_\tau^*(\mathcal{K}) & \mu_{\tau+1}^*(\mathcal{K}) & \dots & \mu_{2\tau}^*(\mathcal{K}) \end{array}\right] = M(\mathcal{K}).$$

Notice that, although with incomplete data, when we consider the end point T and neighbourhood around T, observations from all i are available,

$$\frac{1}{\sqrt{Th}} \sum_{t=1}^{T} \mathcal{K}\left(\frac{T-t}{Th}\right) \left(\frac{t-T}{Th}\right)^{k} \left(\frac{1}{n}\sum_{i=1}^{n} \varepsilon_{it}\right) \Rightarrow \frac{1}{n}\sum_{i=1}^{n} N\left(0, \omega_i^2 \nu_{2k}^*(\mathcal{K})\right) = N\left(0, \frac{\nu_{2k}^*(\mathcal{K})}{n^2} \sum_{i=1}^{n} \omega_i^2\right),$$

and

$$\frac{1}{\sqrt{Th}} \sum_{t=1}^{T} \mathcal{K}\left(\frac{t-T}{Th}\right) x_t \overline{\varepsilon}_t \Rightarrow N\left(0, \frac{\sum_{i=1}^{n}\omega_i^2}{n^2} \left[\begin{array}{cccc} \nu_0^*(\mathcal{K}) & \nu_1^*(\mathcal{K}) & \dots & \nu_\tau^*(\mathcal{K}) \\ \nu_1^*(\mathcal{K}) & \nu_2^*(\mathcal{K}) & & \nu_{\tau+1}^*(\mathcal{K}) \\ \dots & \dots & & \dots \\ \nu_\tau^*(\mathcal{K}) & \dots & & \nu_{2\tau}^*(\mathcal{K}) \end{array}\right]\right) = N\left(0, \frac{1}{n}\omega^2 V(\mathcal{K})\right),$$

where $\omega^2 = \sum_{i=1}^n \omega_i^2/n$, since

$$\begin{aligned} & \mathrm{E}\left(\frac{1}{\sqrt{Th}}\sum_{t=1}^{T}\mathcal{K}\left(\frac{T-t}{Th}\right)\left(\frac{t-T}{Th}\right)^k\varepsilon_{it}\right)\left(\frac{1}{\sqrt{Th}}\sum_{s=1}^{T}\mathcal{K}\left(\frac{T-s}{Th}\right)\left(\frac{s-T}{Th}\right)^l\varepsilon_{is}\right) \\ \to & \left(\sum_{j=-\infty}^{\infty}\gamma_{\varepsilon_i}(j)\right)\int\mathcal{K}(u)^2 u^{l+k}du = \omega_i^2\nu_{l+k}^*(\mathcal{K}). \end{aligned}$$

The variance term of the local polynomial estimator is

$$\begin{aligned} \left[\frac{1}{Th}\sum_{t=1}^{T}\mathcal{K}\left(\frac{T-t}{Th}\right)x_t x_t^\top\right]^{-1}\frac{1}{\sqrt{Th}}\sum_{t=1}^{T}\mathcal{K}\left(\frac{T-t}{Th}\right)x_t\bar{\varepsilon}_t & \Rightarrow M(\mathcal{K})^{-1}N\left(0,\frac{1}{n}\omega^2 V(\mathcal{K})\right) \\ & = N\left(0,\frac{1}{n}\omega^2 M(\mathcal{K})^{-1}V(\mathcal{K})M(\mathcal{K})^{-1}\right). \end{aligned}$$

And the bias term

$$\frac{1}{(\tau+1)!}g^{(\tau+1)}(1)\frac{1}{Th}\sum_{t=1}^{T}\mathcal{K}\left(\frac{T-t}{Th}\right)\left(x_t\left(\frac{t-T}{Th}\right)^{\tau+1}\right) \to \frac{1}{(\tau+1)!}g^{(\tau+1)}(1)\begin{pmatrix}\mu_{\tau+1}^*(\mathcal{K}) \\ \mu_{\tau+2}^*(\mathcal{K}) \\ \cdots \\ \mu_{2\tau+1}^*(\mathcal{K})\end{pmatrix} = B(\mathcal{K}).$$

Thus

$$\sqrt{Th}\left(\widehat{\gamma}-\gamma-h^{\tau+1}M(\mathcal{K})^{-1}B(\mathcal{K})\right) \Rightarrow N\left(0,\frac{1}{n}\omega^2 M(\mathcal{K})^{-1}V(\mathcal{K})M(\mathcal{K})^{-1}\right).$$

Notice that

$$\widehat{\gamma}_k-\gamma_k = h^{\tau-k+1}B_k + \frac{1}{\sqrt{T}h^{k+1/2}}U_k,$$

and our forecaster for $g(1+q/T)$ is given by

$$\widehat{g}(1+q/T) = \sum_{k=0}^{\tau}\widehat{\gamma}_k\cdot\left(\frac{q}{T}\right)^k.$$

Thus, the forecasting error is

$$\begin{aligned} & \widehat{g}(1+q/T)-g(1+q/T) \\ = & \sum_{k=0}^{\tau}\left(h^{\tau-k+1}\left(\frac{q}{T}\right)^k B_k\right)+o\left(\left(\frac{q}{T}\right)^\tau\right)+\sum_{k=0}^{\tau}\left(\frac{1}{\sqrt{T}h^{k+1/2}}\left(\frac{q}{T}\right)^k U_k\right)+o\left(\left(\frac{q}{T}\right)^\tau\right). \end{aligned}$$

The bias and variance terms are given by

$$\begin{aligned} b_g & = \sum_{k=0}^{\tau}\left(h^{\tau-k+1}\left(\frac{q}{T}\right)^k B_k\right) = h^{\tau+1}\sum_{k=0}^{\tau}\left(\left(\frac{q}{Th}\right)^k B_k\right), \\ v_g & = \sum_{k=0}^{\tau}\frac{1}{\sqrt{T}h^{k+1/2}}\left(\frac{q}{T}\right)^k U_k = \frac{1}{\sqrt{Th}}\sum_{k=0}^{\tau}\left(\frac{q}{Th}\right)^k U_k. \end{aligned}$$

whose order of magnitude are jointly determined by the bandwidth h and the forecasting distance q/T. In particular, the prediction error is given by

$$y_{i,T+q} - \widehat{y}_{i,T+q} = \varepsilon_{i,T+q} - (\widehat{\alpha}_i - \alpha_i) - \left(\widehat{\beta}_i^\top - \beta_i^\top\right) D_{T+q} - \left[\widehat{g}(1+q/T) - g(1+q/T)\right],$$

Since the parameter estimates are of smaller error, for any fixed q,

$$y_{i,T+q} - \widehat{y}_{i,T+q} = \varepsilon_{i,T+q} - h^{\tau+1}B_0 - \frac{1}{\sqrt{Th}}U_0 + o_p\left(h^{\tau+1} + \frac{1}{\sqrt{Th}}\right).$$

Thus, the forecasting bias is of order $O(h^{\tau+1})$, with leading term $h^{\tau+1}B_0$, and the leading term of forecasting variance is

$$\omega_i^2 + \frac{1}{Th}V_0,$$

where V_0 is the (1,1)-element in the matrix $\frac{1}{n}\omega^2 M(\mathcal{K})^{-1}V(\mathcal{K})M(\mathcal{K})^{-1}$. ■

8.2 Lemmas

Lemma 1. For each i, as $T \to \infty$:

$$\begin{aligned}
C_{a,ii} &\to c_{ii} = 1 - r_i - 2a_{1i} + ia_{2i} + \sum_{l=i+1}^{n} a_{2l}, \\
C_{b,ii} &\to c_{ii}\left(\frac{1}{12}\mathbf{i}_{11}^\top\right) = \left[(1-r_i) - 2a_{1i} + ia_{2i} + \sum_{l=i+1}^{n} a_{2l}\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right), \\
C_{A,ii} &\to c_{ii}\left(\frac{1}{12}\mathbf{i}_{11}\right) = \left(1 - r_i - 2a_{1i} + ia_{2i} + \sum_{l=i+1}^{n} a_{2l}\right)\left(\frac{1}{12}\mathbf{i}_{11}\right), \\
C_{B,ii} &\to \widetilde{C}_{ii} = (1-r_i)\frac{1}{12}I_{11} - 2a_{1i}\frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11} + ia_{2i}\frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11} + \sum_{l=i+1}^{n} a_{2l}\left(\frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11}\right) \\
&= (1-r_i)\frac{1}{12}I_{11} + \left(ia_{2i} - 2a_{1i} + \sum_{l=i+1}^{n} a_{2l}\right)\left(\frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11}\right).
\end{aligned}$$

LEMMA 2. For $i \neq j$, as $T \to \infty$:

$$
\begin{aligned}
C_{a,ij} &\to c_{ij} = (\max(i,j)-1)a_{2,\max(i,j)} + \sum_{l=\max(i,j)}^{n} a_{2l} - 2a_{1,\max(i,j)}, \\
C_{b,ij} &\to c_{ij}\left(\frac{1}{12}\mathbf{i}_{11}^{\top}\right) = \left[(\max(i,j)-1)a_{2,\max(i,j)} + \sum_{l=\max(i,j)}^{n} a_{2l} - 2a_{1,\max(i,j)}\right]\left(\frac{1}{12}\mathbf{i}_{11}^{\top}\right), \\
C_{A,ij} &\to c_{ij}\left(\frac{1}{12}\mathbf{i}_{11}\right) = \left[(\max(i,j)-1)a_{2,\max(i,j)} - 2a_{1,\max(i,j)} + \sum_{l=\max(i,j)}^{n} a_{2l}\right]\left(\frac{1}{12}\mathbf{i}_{11}\right), \\
C_{B,ij} &= c_{ij}\left(\frac{1}{12^2}\mathbf{i}_{11}^{\top}\mathbf{i}_{11}\right) = \left[(\max(i,j)-1)a_{2,\max(i,j)} - 2a_{1,\max(i,j)} + \sum_{l=\max(i,j)}^{n} a_{2l}\right]\left(\frac{1}{12^2}\mathbf{i}_{11}^{\top}\mathbf{i}_{11}\right).
\end{aligned}
$$

LEMMA 3. For each i, as $T \to \infty$:

$$
\begin{aligned}
d_{a,i} &= -\sqrt{T}h^p b_i + o(\sqrt{T}h^p) \\
&= -\sqrt{T}h^p \frac{1}{p!}\mu_p(K)\left[\sum_{l\neq i}\left(\int_{r_l}^{1}\delta(s)g^{(p)}(s)\,ds\right) - \left(\int_{r_i}^{1} w(s)g^{(p)}(s)\,ds\right)\right] + o(\sqrt{T}h^p),
\end{aligned}
$$

$$
\begin{aligned}
d_{A,i} &= -\sqrt{T}h^p b_i\left(\frac{1}{12}\mathbf{i}_{11}\right) + o(\sqrt{T}h^p) \\
&= -\sqrt{T}h^p \frac{1}{p!}\mu_p(K)\left[\sum_{l\neq i}\left(\int_{r_l}^{1}\delta(s)g^{(p)}(s)\,ds\right) - \left(\int_{r_i}^{1} w(s)g^{(p)}(s)\,ds\right)\right]\left(\frac{1}{12}\mathbf{i}_{11}\right) + o(\sqrt{T}h^p),
\end{aligned}
$$

where $w(s)$ and $\delta(s)$ are weighting functions on [0,1] :

$$
\begin{aligned}
\delta(s) &= \frac{1}{j}, \text{ if } r_j < s < r_{j+1},\ j = 1,2,\ldots,n. \\
w(s) &= 1-\delta(s) = 1-\frac{1}{j}, \text{ if } r_j < s < r_{j+1},\ j = 1,2,\ldots,n.
\end{aligned}
$$

LEMMA 4. For each i, as $T \to \infty$,

$$
e_{a,i} \Rightarrow N\left(0,\sigma_a^2\right),\ e_{A,i} \Rightarrow N\left(0, \frac{1}{12}\sigma_{A1}^2 I_{11} + \frac{1}{12^2}\sigma_{A2}^2 J_{11}\right),
$$

where:

$$
\begin{aligned}
\sigma_a^2 &= (1-r_i-2a_{1i}+a_{2i})\,\omega_i^2 + (na_{4,i}-4na_{3,i}+4a_{2,i})\sum_{j=1}^{i-1}\omega_j^2 + \sum_{j=i+1}^{n}(na_{4,j}-4na_{3,j}+4a_{2,j})\,\omega_j^2 \\
\sigma_{A1}^2 &= s_j^2\,(1-r_i-2a_{1i}+a_{2i})\,,\ \sigma_{A2}^2 = (na_{4i}-4na_{3i}+4a_{2i})\sum_{j<i}\omega_j^2 + \sum_{j>i}(na_{4j}-4na_{3j}+4a_{2j})\,\omega_j^2.
\end{aligned}
$$

8.3 Proof of Lemmas

Proof of Lemma 1. Notice that

$$\sum_{t=t_l}^{t_{i-1}} \frac{1}{m_t} \left(\sum_{s=t_i}^{T} K_h((t-s)/T) \right) \underbrace{\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}}_{\to 0} \to 0,$$

we have

$$\begin{aligned}
C_{a,ii} &= \frac{1}{T}\sum_{t=t_i}^{T}\left(1-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\left(1-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)\right) \\
&\quad -\frac{1}{T}\left[\frac{1}{T}\sum_{l<i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right] \\
&\quad -\frac{1}{T}\left[\frac{1}{T}\sum_{l>i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right] \\
&= \frac{1}{T}\sum_{t=t_i}^{T}\left[1-\frac{1}{m_t}\right]\left[1-\frac{1}{m_t}\right]-\frac{1}{T}\left[\sum_{l<i}\sum_{t=t_i}^{T}\frac{1}{m_t}\left(\underbrace{\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)}_{=1}\right)\underbrace{\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}}_{\to -1/m_t}\right] \\
&\quad -\frac{1}{T}\left[\sum_{l>i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right] \\
&= \frac{1}{T}\sum_{t=t_i}^{T}\left[1-\frac{1}{m_t}\right]\left[1-\frac{1}{m_t}\right]-\frac{1}{T}\left[-\sum_{l<i}\sum_{t=t_i}^{T}\frac{1}{m_t^2}\right]-\frac{1}{T}\left[-\sum_{l>i}\sum_{t=t_l}^{T}\frac{1}{m_t^2}\right] \\
&= \frac{1}{T}\sum_{t=t_i}^{T}\left[1-\frac{2}{m_t}+\frac{1}{m_t^2}\right]+\frac{1}{T}\left[(i-1)\sum_{t=t_i}^{T}\frac{1}{m_t^2}\right]+\frac{1}{T}\left[\sum_{l=i+1}^{n}\sum_{t=t_l}^{T}\frac{1}{m_t^2}\right] \\
&= \frac{1}{T}\sum_{t=t_i}^{T}\left[1-\frac{2}{m_t}\right]+\frac{1}{T}\left[i\sum_{t=t_i}^{T}\frac{1}{m_t^2}\right]+\frac{1}{T}\left[\sum_{l=i+1}^{n}\sum_{t=t_l}^{T}\frac{1}{m_t^2}\right] \\
&= 1-r_i-2\sum_{j=i}^{n}\frac{1}{j}\left(r_{j+1}-r_j\right)+i\sum_{j=i}^{n}\frac{1}{j^2}\left(r_{i+1}-r_i\right)+\sum_{l=i+1}^{n}\sum_{j=l}^{n}\frac{1}{j^2}\left(r_{i+1}-r_i\right) \\
&= 1-r_i-2a_{1i}+ia_{2i}+\sum_{l=i+1}^{n}a_{2l},
\end{aligned}$$

$$
\begin{aligned}
C_{b,ii} &= \frac{1}{T}\sum_{t=t_i}^{T}\left(D_t^\top - \frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T} D_s^\top K_h((t-s)/T)\right)\left(1+\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right) \\
&\quad -\frac{1}{T}\left[\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T} D_s^\top K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right] \\
&= \frac{1}{T}\sum_{t=t_i}^{T}\left(D_t^\top - \frac{1}{12}\mathbf{i}_{11}^\top\frac{1}{m_t}\right)\left(1-\frac{1}{m_t}\right) \\
&\quad -\frac{1}{T}\left[\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T} D_s^\top K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right] \\
&= \frac{1}{T}\sum_{t=t_i}^{T} D_t^\top\left(1-\frac{1}{m_t}\right) - \frac{1}{T}\sum_{t=t_i}^{T}\frac{1}{12}\mathbf{i}_{11}^\top\left(\frac{1}{m_t}\right)\left(1-\frac{1}{m_t}\right) \\
&\quad -\frac{1}{T}\left[\sum_{l< i}\sum_{t=t_i}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T} D_s^\top K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right] \\
&\quad -\frac{1}{T}\left[\sum_{l> i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T} D_s^\top K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right] \\
&= \frac{1}{T}\sum_{t=t_i}^{T} D_t^\top - \frac{1}{T}\sum_{t=t_i}^{T}\frac{1}{m_t}D_t^\top - \frac{1}{T}\sum_{t=t_i}^{T}\frac{1}{12}\mathbf{i}_{11}^\top\left(\frac{1}{m_t}\right) + \frac{1}{T}\sum_{t=t_i}^{T}\frac{1}{12}\mathbf{i}_{11}^\top\left(\frac{1}{m_t}\right)\frac{1}{m_t} \\
&\quad -\frac{1}{T}\left[(i-1)\sum_{t=t_i}^{T}\left(-\frac{1}{m_t^2}\right)\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\right] - \frac{1}{T}\left[\sum_{l>i}\sum_{t=t_l}^{T}\left(-\frac{1}{m_t^2}\right)\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\right] \\
&= \frac{1}{12}(1-r_i)\mathbf{i}_{11}^\top - \frac{2}{12}\left(\sum_{l=i}^{n}\frac{1}{l}\left(r_{l+1}-r_l\right)\right)\mathbf{i}_{11}^\top + \frac{1}{12}\mathbf{i}_{11}^\top\left(\sum_{l=i}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\right) \\
&\quad +\left[(i-1)\sum_{l=i}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right) + \left[\sum_{l>i}\sum_{k=l}^{n}\frac{1}{k^2}\left(r_{k+1}-r_k\right)\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right) \\
&= \left[(1-r_i) - 2a_{1i} + ia_{2i} + \sum_{l=i+1}^{n} a_{2l}\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)
\end{aligned}
$$

$$
\begin{aligned}
C_{A,ii} &= \frac{1}{T}\left\{\begin{array}{c} \sum_{t=t_i}^{T}\left(1-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right) \\ -\left[\frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right] \end{array}\right\} \\
&= \frac{1}{T}\sum_{t=t_i}^{T}\left(1-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\left(D_t-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)D_s\right) \\
&\quad -\frac{1}{T}\left[\frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\left(-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)D_s\right)\right] \\
&= \frac{1}{T}\sum_{t=t_i}^{T}\left(1-\frac{1}{m_t}\right)\left(D_t-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}\right) \\
&\quad -\frac{1}{T}\left[\frac{1}{T}\sum_{l<i}\sum_{t=t_i}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\left(-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)D_s\right)\right] \\
&\quad -\frac{1}{T}\left[\frac{1}{T}\sum_{l>i}\sum_{t=t_i}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)\right)\left(-\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T)D_s\right)\right] \\
&= \frac{1}{T}\sum_{t=t_i}^{T}D_t-\frac{1}{T}\sum_{t=t_i}^{T}\frac{1}{m_t}D_t-\frac{1}{T}\sum_{t=t_i}^{T}\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}+\frac{i}{T}\sum_{t=t_i}^{T}\frac{1}{m_t^2}\frac{1}{12}\mathbf{i}_{11}+\left[\frac{1}{T}\sum_{l=i+1}^{n}\sum_{t=t_i}^{T}\frac{1}{m_t^2}\left(\frac{1}{12}\mathbf{i}_{11}\right)\right] \\
&\to (1-r_i)\frac{1}{12}\mathbf{i}_{11}-2\sum_{j=i}^{n}\frac{1}{j}\left(r_{j+1}-r_j\right)\frac{1}{12}\mathbf{i}_{11} \\
&\quad +i\sum_{j=i}^{n}\frac{1}{j^2}\left(r_{i+1}-r_i\right)\frac{1}{12}\mathbf{i}_{11}+\sum_{l=i+1}^{n}\sum_{j=l}^{n}\frac{1}{j^2}\left(r_{i+1}-r_i\right)\left(\frac{1}{12}\mathbf{i}_{11}\right) \\
&= \left[1-r_i-2\sum_{j=i}^{n}\frac{1}{j}\left(r_{j+1}-r_j\right)+i\sum_{j=i}^{n}\frac{1}{j^2}\left(r_{i+1}-r_i\right)+\sum_{l=i+1}^{n}\sum_{j=l}^{n}\frac{1}{j^2}\left(r_{i+1}-r_i\right)\right]\left(\frac{1}{12}\mathbf{i}_{11}\right) \\
&= \left(1-r_i-2a_{1i}+ia_{2i}+\sum_{l=i+1}^{n}a_{2l}\right)\left(\frac{1}{12}\mathbf{i}_{11}\right).
\end{aligned}
$$

If $j>i$,

$$
\begin{aligned}
C_{A,ij} &= \frac{1}{T}\left[\sum_{t=t_j}^{T}\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i} - \frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_j}^{T}K_h((t-s)/T)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right)\right] \\
&\quad -\frac{1}{T}\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right) \\
&= \frac{1}{T}\sum_{t=t_j}^{T}\left(-\frac{1}{m_t}\left(\frac{1}{12}\mathbf{i}_{11}\right)\right) - \frac{1}{T}\left[\sum_{l<i}\sum_{t=t_i}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right] \\
&\quad -\frac{1}{T}\left[\sum_{l>i}\left[\sum_{t=t_l}^{t_{j-1}}+\sum_{t=t_j}^{T}\right]\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right] \\
&\quad -\frac{1}{T}\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right) \\
&\to -\sum_{l=j}^{n}\frac{1}{l}\left(r_{l+1}-r_l\right)\left(\frac{1}{12}\mathbf{i}_{11}\right) \\
&\quad +\left[(i-1)\sum_{l=j}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\left(\frac{1}{12}\mathbf{i}_{11}\right)\right]+(j-i-1)\left[\sum_{l=j}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\right]\left(\frac{1}{12}\mathbf{i}_{11}\right) \\
&\quad +\sum_{l=j}^{n}\sum_{k=l}^{n}\frac{1}{k^2}\left(r_{k+1}-r_k\right)\left(\frac{1}{12}\mathbf{i}_{11}\right)-\sum_{l=j}^{n}\frac{1}{l}\left(r_{l+1}-r_l\right)\left(\frac{1}{12}\mathbf{i}_{11}\right)+\sum_{l=j}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\left(\frac{1}{12}\mathbf{i}_{11}\right) \\
&= \left[(j-1)a_{2j}-2a_{1j}+\sum_{l=j}^{n}a_{2l}\right]\left(\frac{1}{12}\mathbf{i}_{11}\right)
\end{aligned}
$$

$$
\begin{aligned}
C_{B,ii} &= \frac{1}{T}\sum_{t=t_i}^{T}\left(D_t^\top - \frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T} D_s^\top K_h((t-s)/T)\right)\left(D_t + \frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i}\right) \\
&\quad -\frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T} D_s^\top K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i} \\
&= \frac{1}{T}\sum_{t=t_i}^{T}\left(D_t^\top - \frac{1}{12}\mathbf{i}_{11}^\top\frac{1}{m_t}\right)\left(D_t - \frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}\right) \\
&\quad -\frac{1}{T}\left[\sum_{l<i}\sum_{t=t_i}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T} D_s^\top K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i}\right] \\
&\quad -\frac{1}{T}\left[\sum_{l>i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_i}^{T} D_s^\top K_h((t-s)/T)\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i}\right] \\
&= \frac{1}{T}\sum_{t=t_i}^{T} D_t^\top D_t - \frac{1}{T}\sum_{t=t_i}^{T}\frac{1}{12}\frac{1}{m_t}\mathbf{i}_{11}^\top D_t - \frac{1}{T}\sum_{t=t_i}^{T}\frac{1}{12}\frac{1}{m_t}D_t^\top\mathbf{i}_{11} + \frac{1}{T}\sum_{t=t_i}^{T}\left(\frac{1}{12}\frac{1}{m_t}\mathbf{i}_{11}^\top\right)\left(\frac{1}{12}\frac{1}{m_t}\mathbf{i}_{11}^\top\right) \\
&\quad -\frac{1}{T}\left[\sum_{l<i}\sum_{t=t_i}^{T}\frac{1}{m_t}\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\left(-\frac{1}{12}\mathbf{i}_{11}^\top\frac{1}{m_t}\right)\right] - \frac{1}{T}\left[\sum_{l>i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\left(-\frac{1}{12}\mathbf{i}_{11}^\top\frac{1}{m_t}\right)\right] \\
&= \frac{1}{12}(1-r_i)I_{11} - \frac{2}{12^2}\left(\sum_{l=i}^{n}\frac{1}{l}\left(r_{l+1}-r_l\right)\right)\mathbf{i}_{11}^\top\mathbf{i}_{11} + \frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11}\left(\sum_{l=i}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\right) \\
&\quad + \left[(i-1)\sum_{l=i}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\right]\left(\frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11}\right) + \left[\sum_{l>i}\sum_{k=l}^{n}\frac{1}{k^2}\left(r_{k+1}-r_k\right)\right]\left(\frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11}\right) \\
&= \frac{1}{12}(1-r_i)I_{11} - \frac{2}{12^2}a_{1i}\mathbf{i}_{11}^\top\mathbf{i}_{11} + ia_{2i}\frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11} + \sum_{l=i+1}^{n} a_{2l}\left(\frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11}\right).
\end{aligned}
$$

If $j > i$,

$$
\begin{aligned}
C_{B,ij} &= \frac{1}{T}\sum_{t=t_j}^{T} D_t^\top \frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i} - \frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T} D_s^\top K_h((t-s)/T)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i}\right) \\
&\quad -\frac{1}{T}\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T} D_s K_h((t-s)/T)\right)\left(D_t + \frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i}\right) \\
&= \frac{1}{T}\sum_{t=t_j}^{T} D_t^\top\left(-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}\right) - \frac{1}{T}\left[\sum_{l<i}\sum_{t=t_i}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T} K_h((t-s)/T) D_s^\top\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i}\right] \\
&\quad -\frac{1}{T}\left[\sum_{l>i}\left[\sum_{t=t_l}^{t_{j-1}} + \sum_{t=t_j}^{T}\right]\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T} K_h((t-s)/T) D_s^\top\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \beta_i}\right] \\
&\quad -\frac{1}{T}\sum_{t=t_j}^{T}\left(\frac{1}{m_t}\left(\frac{1}{12}\mathbf{i}_{11}\right)\right)\left(D_t - \frac{1}{m_t}\left(\frac{1}{12}\mathbf{i}_{11}\right)\right) \\
&= -\sum_{l=j}^{n}\frac{1}{l}\left(r_{l+1}-r_l\right)\left(\frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11}\right) \\
&\quad -\frac{1}{T}\left[\sum_{l<i}\sum_{t=t_j}^{T}\frac{1}{m_t}\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\left(-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}\right)\right] - \frac{1}{T}\left[\sum_{l=i+1}^{j-1}\sum_{t=t_j}^{T}\frac{1}{m_t}\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\left(-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}\right)\right] \\
&\quad -\frac{1}{T}\left[\sum_{l=j}^{n}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\left(-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}\right)\right] - \frac{1}{T}\sum_{t=t_j}^{T}\left(\frac{1}{m_t}\left(\frac{1}{12}\mathbf{i}_{11}\right)\right)\left(D_t - \frac{1}{m_t}\left(\frac{1}{12}\mathbf{i}_{11}\right)\right) \\
&= -a_j\left(\frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11}\right) + \left[(i-1)\left[\sum_{l=j}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\left(\frac{1}{12}\mathbf{i}_{11}\right)\right] \\
&\quad + \left[(j-i-1)\left[\sum_{l=j}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\left(\frac{1}{12}\mathbf{i}_{11}\right)\right] \\
&\quad + \left[\sum_{l=j}^{n}\sum_{k=l}^{n}\frac{1}{k^2}\left(r_{k+1}-r_k\right)\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\left(\frac{1}{12}\mathbf{i}_{11}\right) \\
&\quad -\frac{1}{T}\sum_{t=t_j}^{T}\left(\left(\frac{1}{12}\mathbf{i}_{11}\right)\frac{1}{m_t}\right)D_t + \frac{1}{T}\sum_{t=t_j}^{T}\frac{1}{m_t}\frac{1}{m_t}\left(\frac{1}{12}\mathbf{i}_{11}\right)\left(\frac{1}{12}\mathbf{i}_{11}\right) \\
&= \left[(j-1)a_{2j} - 2a_{1j} + \sum_{l=j}^{n} a_{2l}\right]\left(\frac{1}{12^2}\mathbf{i}_{11}^\top\mathbf{i}_{11}\right).
\end{aligned}
$$

■

Proof of Lemma 2. We have

$$
\begin{aligned}
C_{a,ij} &= \frac{1}{T}\left[\sum_{t=t_j}^{T}\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i} - \frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_j}^{T}K_h((t-s)/T)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right)\right] \\
&\quad -\frac{1}{T}\left[\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right)\right] \\
&= \frac{1}{T}\sum_{t=t_j}^{T}\left(-\frac{1}{m_t}\right) - \frac{1}{T}\left[\sum_{l<i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right] \\
&\quad -\frac{1}{T}\left[\sum_{l>i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right] \\
&\quad -\frac{1}{T}\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right) \\
&= -\frac{1}{T}\sum_{t=t_j}^{T}\frac{1}{m_t} + \frac{1}{T}\left[\sum_{l<i}\sum_{t=t_j}^{T}\frac{1}{m_t^2}\right] - \frac{1}{T}\sum_{t=t_j}^{T}\frac{1}{m_t}\left(1-\frac{1}{m_t}\right) \\
&\quad -\frac{1}{T}\left[\left[\sum_{l=i+1}^{j-1}+\sum_{l=j}^{n}\right]\left[\sum_{t=t_l}^{t_{j-1}}+\sum_{t=t_j}^{T}\right]\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right] \\
&= -\sum_{l=j}^{n}\frac{1}{l}\left(r_{l+1}-r_l\right) + \left[(i-1)\sum_{l=j}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\right] + \left[(j-i-1)\left[\sum_{l=j}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\right]\right] \\
&\quad + \left[\sum_{l=j}^{n}\sum_{k=l}^{n}\frac{1}{k^2}\left(r_{k+1}-r_k\right)\right] - \sum_{l=j}^{n}\frac{1}{l}\left(r_{l+1}-r_l\right) + \sum_{l=j}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right) \\
&= (j-1)a_{2j} + \sum_{l=j}^{n}a_{2l} - 2a_{1j}.
\end{aligned}
$$

For $C_{b,ij}$, if $j > i$,

$$\begin{aligned}
C_{b,ij} &= \frac{1}{T}\left[\sum_{t=t_j}^{T} D_t^\top \frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i} - \frac{1}{T}\sum_{l\neq i}\sum_{t=t_l}^{T}\frac{1}{m_t}\left(\sum_{s=t_j}^{T} D_s^\top K_h((t-s)/T)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right)\right] \\
&\quad -\frac{1}{T}\left[\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T} D_s K_h((t-s)/T)\right)\left(1+\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right)\right] \\
&= -\frac{1}{T}\sum_{t=t_j}^{T}\frac{1}{m_t}D_t^\top - \frac{1}{T}\left[\sum_{l<i}\sum_{t=t_i}^{T}\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)D_s^\top\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right] \\
&\quad -\frac{1}{T}\left[\sum_{l>i}\left[\sum_{t=t_l}^{t_{j-1}}+\sum_{t=t_j}^{T}\right]\frac{1}{m_t}\left(\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)D_s^\top\right)\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right] \\
&\quad -\frac{1}{T}\sum_{t=t_i}^{T}\left(\frac{1}{m_t}\frac{1}{T}\sum_{s=t_j}^{T}K_h((t-s)/T)D_s^\top\right)\left(1+\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i}\right) \\
&= -\sum_{l=j}^{n}\frac{1}{l}\left(r_{l+1}-r_l\right)\left(\frac{1}{12}\mathbf{i}_{11}^\top\right) \\
&\quad +(i-1)\left[\sum_{l=j}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)+\left[(j-i-1)\left[\sum_{l=j}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\right] \\
&\quad +\left[\sum_{l=j}^{n}\sum_{k=l}^{n}\frac{1}{k^2}\left(r_{k+1}-r_k\right)\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)-\sum_{l=j}^{n}\frac{1}{l}\left(1-\frac{1}{l}\right)\left(r_{l+1}-r_l\right)\left(\frac{1}{12}\mathbf{i}_{11}^\top\right) \\
&= -a_{1j}\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)+(i-1)a_{2j}\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)+\left[(j-i-1)a_{2j}\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)\right] \\
&\quad +\left[\sum_{l=j}^{n}a_{2l}\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)-\sum_{l=j}^{n}\frac{1}{l}\left(r_{l+1}-r_l\right)\left(\frac{1}{12}\mathbf{i}_{11}^\top\right)+\sum_{l=j}^{n}\frac{1}{l^2}\left(r_{l+1}-r_l\right)\left(\frac{1}{12}\mathbf{i}_{11}^\top\right) \\
&= \left[(j-1)a_{2j}+\sum_{l=j}^{n}a_{2l}-2a_{1j}\right]\left(\frac{1}{12}\mathbf{i}_{11}^\top\right).
\end{aligned}$$

■

Proof of Lemma 3. Notice that

$$\frac{\partial \widehat{g}_\theta(t/T)}{\partial \alpha_i} = -\frac{1}{m_t}\frac{1}{T}\sum_{s=t_i}^{T}K_h((t-s)/T) \to \begin{cases} -\frac{1}{m_t}, & i \leq m_t \\ \mathbf{0}, & i > m_t \end{cases}$$

and

$$\begin{aligned}
& \frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right)\left(g\left(\frac{t}{T}\right)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{m_t}\sum_{s=t_j}^{T}K_h((t-s)/T)g\left(\frac{s}{T}\right)\right) \\
= \; & \sqrt{T}h^p\frac{1}{p!}\mu_p(K)\left(\int_{r_i}^{1}w(s)g^{(p)}(s)\,ds\right)+o(\sqrt{T}h^p),
\end{aligned}$$

we have

$$\begin{aligned}
d_{a,i} \; = \; & \frac{1}{\sqrt{T}}\sum_{l\neq i}\sum_{t=t_l}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i} \\
& +\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\left(1+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\alpha_i}\right) \\
= \; & -\frac{1}{p!}\mu_p(K)\sqrt{T}h^p\sum_{l\neq i}\left(\int_{r_l}^{1}\delta(s)g^{(p)}(s)\,ds\right)+\sqrt{T}h^p\frac{1}{p!}\mu_p(K)\left(\int_{r_i}^{1}w(s)g^{(p)}(s)\,ds\right)+o(\sqrt{T}h^p) \\
= \; & -\frac{1}{p!}\mu_p(K)\sqrt{T}h^p\left[\sum_{l\neq i}\left(\int_{r_l}^{1}\delta(s)g^{(p)}(s)\,ds\right)-\left(\int_{r_i}^{1}w(s)g^{(p)}(s)\,ds\right)\right]+o(\sqrt{T}h^p),
\end{aligned}$$

$$\begin{aligned}
d_{A,i} \; = \; & \frac{1}{\sqrt{T}}\sum_{l\neq i}\sum_{t=t_j}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i} \\
& +\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\left(D_t+\frac{\partial\widehat{g}_\theta(t/T)}{\partial\beta_i}\right) \\
= \; & \frac{1}{\sqrt{T}}\sum_{l\neq i}\sum_{t=t_j}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\left(-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}\right) \\
& +\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}\left(g(t/T)-\frac{1}{m_t}\frac{1}{T}\sum_{j=1}^{n}\sum_{s=t_j}^{T}g(s/T)K_h((t-s)/T)\right)\left(D_t-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}\right) \\
= \; & -\frac{1}{p!}\mu_p(K)\sqrt{T}h^p\sum_{l\neq i}\left(\int_{r_l}^{1}\delta(s)g^{(p)}(s)\,ds\right)\frac{1}{12}\mathbf{i}_{11} \\
& +\sqrt{T}h^p\frac{1}{p!}\mu_p(K)\left(\int_{r_i}^{1}w(s)g^{(p)}(s)\,ds\right)\left(\frac{1}{12}\mathbf{i}_{11}\right)+o(\sqrt{T}h^p).
\end{aligned}$$

$\blacksquare$

Proof of Lemma 4. Notice that

$$e_a = \begin{bmatrix} e_{a,1} \\ \dots \\ e_{a,n} \end{bmatrix}, \text{ and } e_A \begin{bmatrix} e_{A,1} \\ \dots \\ e_{A,n} \end{bmatrix},$$

$$\begin{aligned} e_{a,i} &= \frac{1}{\sqrt{T}} \sum_{t=t_i}^{t_{i+1}-1} \left(1 - \frac{2}{m_t} + \frac{i}{m_t^2}\right) \varepsilon_{it} + \frac{1}{\sqrt{T}} \sum_{t=t_{i+1}}^{t_{i+2}-1} \left(1 - \frac{2}{m_t} + \frac{i+1}{m_t^2}\right) \varepsilon_{it} \\ &\quad + \dots + \frac{1}{\sqrt{T}} \sum_{t=t_n}^{T} \left(1 - \frac{2}{m_t} + \frac{n}{m_t^2}\right) \varepsilon_{it} \\ &\quad + \frac{1}{\sqrt{T}} \sum_{j=1}^{i-1} \left(\sum_{t=t_i}^{T} \left(n \left(\frac{1}{m_t^2} \right) - \frac{2}{m_t} \right) \varepsilon_{jt} \right) + \frac{1}{\sqrt{T}} \sum_{j=i+1}^{n} \left(\sum_{t=t_j}^{T} \left(n \left(\frac{1}{m_t^2} \right) - \frac{2}{m_t} \right) \varepsilon_{jt} \right) \\ &= \frac{1}{\sqrt{T}} \sum_{t=t_i}^{T} c_{it} \varepsilon_{it} + \sum_{j \neq i} \left(\frac{1}{\sqrt{T}} \sum_{t=t_i}^{T} c_{ijt} \varepsilon_{jt} \right). \end{aligned}$$

First,

$$\frac{1}{\sqrt{T}} \sum_{t=t_i}^{T} c_{it} \varepsilon_{it} \Rightarrow N_i \left(0, \omega_i^2 \sum_{j=i}^{n} (r_{j+1} - r_j) \left(1 - \frac{1}{j}\right)^2 \right) = N_i \left(0, (1 - r_i - 2a_{1i} + a_{2i}) \omega_i^2\right),$$

since

$$\begin{aligned} &\frac{1}{\sqrt{T}} \sum_{t=t_i}^{T} c_{it} \varepsilon_{it} \\ &= \frac{1}{\sqrt{T}} \sum_{t=t_i}^{t_{i+1}-1} \left(1 - \frac{2}{m_t} + \frac{i}{m_t^2}\right) \varepsilon_{it} + \frac{1}{\sqrt{T}} \sum_{t=t_{i+1}}^{t_{i+2}-1} \left(1 - \frac{2}{m_t} + \frac{i+1}{m_t^2}\right) \varepsilon_{it} + \ldots + \frac{1}{\sqrt{T}} \sum_{t=t_n}^{T} \left(1 - \frac{2}{m_t} + \frac{n}{m_t^2}\right) \varepsilon_{it}. \end{aligned}$$

Under assumption A1,

$$\begin{aligned} &\lim \operatorname{Var} \left(\frac{1}{\sqrt{T}} \sum_{t=t_i}^{T} c_{it} \varepsilon_{it} \right) \\ &= \lim \operatorname{Var} \left[\frac{1}{\sqrt{T}} \sum_{t=t_i}^{t_{i+1}-1} \left(1 - \frac{2}{m_t} + \frac{i}{m_t^2}\right) \varepsilon_{it} \right] + \lim \operatorname{Var} \left[\frac{1}{\sqrt{T}} \sum_{t=t_{i+1}}^{t_{i+2}-1} \left(1 - \frac{2}{m_t} + \frac{i+1}{m_t^2}\right) \varepsilon_{it} \right] \\ &\quad + \dots + \lim \operatorname{Var} \left[\frac{1}{\sqrt{T}} \sum_{t=t_n}^{T} \left(1 - \frac{2}{m_t} + \frac{n}{m_t^2}\right) \varepsilon_{it} \right] + o(1) \\ &= \sum_{j=i}^{n} (r_{j+1} - r_j) \left(1 - \frac{1}{j}\right)^2 \omega_i^2. \end{aligned}$$

Next,
$$\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T} c_{ijt}\varepsilon_{jt} \Rightarrow N_j\left(0,\omega_j^2\left(na_{4,\max(i,j)}-4na_{3,\max(i,j)}+4a_{2,\max(i,j)}\right)\right),$$
in particular, for $j<i$,
$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T} c_{ijt}\varepsilon_{jt}\\
=\ &\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}\left(n\left(\frac{1}{m_t^2}\right)-\frac{2}{m_t}\right)\varepsilon_{jt}=\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T} c_{ijt}\varepsilon_{jt}\\
\Rightarrow\ &N_j\left(0,\omega_j^2\sum_{j=i}^{n}\left(r_{j+1}-r_j\right)\left(\frac{n}{j^2}-\frac{2}{j}\right)^2\right)=N_j\left(0,\omega_j^2\left(na_{4i}-4na_{3i}+4a_{2i}\right)\right),
\end{aligned}$$
since
$$\begin{aligned}
&\lim \mathrm{Var}\left[\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T} c_{ijt}\varepsilon_{jt}\right]\\
=\ &\frac{1}{T}\sum_{t=t_i}^{T}\sum_{s=t_i}^{T}\left(n\left(\frac{1}{m_t^2}\right)-\frac{2}{m_t}\right)\left(n\left(\frac{1}{m_s^2}\right)-\frac{2}{m_s}\right)\gamma_j(t-s)\\
\Rightarrow\ &\omega_j^2\sum_{s=i}^{n}\left(r_{s+1}-r_s\right)\left(\frac{n}{s^2}-\frac{2}{s}\right)^2,
\end{aligned}$$
and for $j>i$,
$$\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T} c_{ijt}\varepsilon_{jt}\\
=\ &\frac{1}{\sqrt{T}}\sum_{t=t_j}^{T}\left(n\left(\frac{1}{m_t^2}\right)-\frac{2}{m_t}\right)\varepsilon_{jt}=\frac{1}{\sqrt{T}}\sum_{t=t_j}^{T} c_{ijt}\varepsilon_{jt}\\
\Rightarrow\ &N_j\left(0,\omega_j^2\sum_{l=j}^{n}\left(r_{l+1}-r_l\right)\left(\frac{n}{l^2}-\frac{2}{l}\right)^2\right)=N_j\left(0,\omega_j^2\left(na_{4j}-4na_{3j}+4a_{2j}\right)\right).
\end{aligned}$$
Thus
$$e_{a,i}\Rightarrow N\left(0,\left(1-r_i-2a_{1i}+a_{2i}\right)\omega_i^2+\left(na_{4,i}-4na_{3,i}+4a_{2,i}\right)\sum_{j=1}^{i-1}\omega_j^2+\sum_{j=i+1}^{n}\left(na_{4,j}-4na_{3,j}+4a_{2,j}\right)\omega_j^2\right).$$

Next, we consider

$$
\begin{aligned}
e_{A,i} &= \frac{1}{\sqrt{T}}\sum_{t=t_i}^{t_{i+1}-1}\left(D_t-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}-\frac{1}{m_t}D_t+\frac{i}{m_t^2}\frac{1}{12}\mathbf{i}_{11}\right)\varepsilon_{it}\\
&+\frac{1}{\sqrt{T}}\sum_{t=t_{i+1}}^{t_{i+2}-1}\left(D_t-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}-\frac{1}{m_t}D_t+\frac{i+1}{m_t^2}\frac{1}{12}\mathbf{i}_{11}\right)\varepsilon_{it}\\
&+\frac{1}{\sqrt{T}}\sum_{t=t_{i+2}}^{t_{i+3}-1}\left(D_t-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}-\frac{1}{m_t}D_t+\frac{i+2}{m_t^2}\frac{1}{12}\mathbf{i}_{11}\right)\varepsilon_{it}+\ldots\\
&+\frac{1}{\sqrt{T}}\sum_{t=t_n}^{T}\left(D_t-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}-\frac{1}{m_t}D_t+\frac{n}{m_t^2}\frac{1}{12}\mathbf{i}_{11}\right)\\
&+\left(\frac{1}{12}\mathbf{i}_{11}\right)\frac{1}{\sqrt{T}}\sum_{j=1}^{i-1}\sum_{t=t_i}^{T}\left[\frac{n}{m_t^2}-\frac{2}{m_t}\right]\varepsilon_{jt}+\left(\frac{1}{12}\mathbf{i}_{11}\right)\frac{1}{\sqrt{T}}\sum_{j=i+1}^{n}\left(\sum_{t=t_j}^{T}\left[\frac{n}{m_t^2}-\frac{2}{m_t}\right]\right)\varepsilon_{jt}\\
&= \frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}C_{it}\varepsilon_{it}+\sum_{j\neq i}\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}C_{ijt}\varepsilon_{jt}.
\end{aligned}
$$

First,

$$
\begin{aligned}
\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}C_{it}\varepsilon_{it} &= \frac{1}{\sqrt{T}}\sum_{t=t_i}^{t_{i+1}-1}\left(D_t-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}-\frac{1}{m_t}D_t+\frac{i}{m_t^2}\frac{1}{12}\mathbf{i}_{11}\right)\varepsilon_{it}\\
&+\frac{1}{\sqrt{T}}\sum_{t=t_{i+1}}^{t_{i+2}-1}\left(D_t-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}-\frac{1}{m_t}D_t+\frac{i+1}{m_t^2}\frac{1}{12}\mathbf{i}_{11}\right)\varepsilon_{it}+\\
&+\ldots+\frac{1}{\sqrt{T}}\sum_{t=t_n}^{T}\left(D_t-\frac{1}{m_t}\frac{1}{12}\mathbf{i}_{11}-\frac{1}{m_t}D_t+\frac{n}{m_t^2}\frac{1}{12}\mathbf{i}_{11}\right)\\
&= \left[\frac{1}{\sqrt{T}}\sum_{t=t_i}^{t_{i+1}-1}+\ldots+\frac{1}{\sqrt{T}}\sum_{t=t_n}^{T}\right]\left(D_t-\frac{1}{m_t}D_t\right)\varepsilon_{it}\\
&\Rightarrow N_i\left(0,s_i^2\left(1-r_i-2a_{1i}+a_{2i}\right)\left(\frac{1}{12}I_{11}\right)\right),
\end{aligned}
$$

since, for each fixed j,

$$
\frac{1}{T}\sum_{t=1}^{T}D_tD_{t+j}{}^{\top}\to\begin{cases}\frac{1}{12}I_{11}, & \text{if } j=12k,\\ 0, & \text{if } j\neq 12k\end{cases}.
$$

For $j=1,\ldots,i-1$,

$$
\begin{aligned}
\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}C_{ijt}\varepsilon_{jt} &= \left(\frac{1}{12}\mathbf{i}_{11}\right)\frac{1}{\sqrt{T}}\sum_{t=t_i}^{T}\left[\frac{n}{m_t^2}-\frac{2}{m_t}\right]\varepsilon_{jt}\\
&\to N_j\left(0,\omega_j^2\left(na_{4i}-4na_{3i}+4a_{2i}\right)\left(\frac{1}{12}\mathbf{i}_{11}\right)\left(\frac{1}{12}\mathbf{i}_{11}^{\top}\right)\right),
\end{aligned}
$$

and for $j = i+1, \ldots, n$,

$$\frac{1}{\sqrt{T}} \sum_{t=t_i}^{T} C_{ijt} \varepsilon_{jt} \to N_j \left(0, \omega_j^2 \left(n a_{4j} - 4 n a_{3j} + 4 a_{2j}\right) \left(\frac{1}{12} \mathbf{i}_{11}\right) \left(\frac{1}{12} \mathbf{i}_{11}^{\top}\right)\right).$$

Thus

$$\begin{aligned} e_{A,i} \quad \Rightarrow \quad & N_i \left(0, s_j^2 \left(1 - r_i - 2a_{1i} + a_{2i}\right) \left(\frac{1}{12} I_{11}\right)\right) \\ & + \sum_{j<i} N_j \left(0, \omega_j^2 \left(n a_{4i} - 4 n a_{3i} + 4 a_{2i}\right) \left(\frac{1}{12} \mathbf{i}_{11}\right) \left(\frac{1}{12} \mathbf{i}_{11}^{\top}\right)\right) \\ & + \sum_{j>i} N_j \left(0, \omega_j^2 \left(n a_{4j} - 4 n a_{3j} + 4 a_{2j}\right) \left(\frac{1}{12} \mathbf{i}_{11}\right) \left(\frac{1}{12} \mathbf{i}_{11}^{\top}\right)\right). \end{aligned}$$

Finally we analyze the covariance terms:

$$\begin{aligned} & \operatorname{Cov}\left(e_{a,i}, e_{A,i}^{\top}\right) \\ = \ & \left(\frac{1}{T} \sum_{t=t_i}^{T} \sum_{s=t_i}^{T} c_{it} C_{is}^{T}\right) \gamma_i(t-s) + \sum_{j<i} \left(\frac{1}{T} \sum_{t=t_i}^{T} \sum_{s=t_i}^{T} c_{ijt} C_{ijs}^{T}\right) \gamma_j(t-s) + \sum_{j>i} \left(\frac{1}{T} \sum_{t=t_i}^{T} \sum_{s=t_i}^{T} c_{ijt} C_{ijs}^{T}\right) \gamma_j(t-s) \\ \to \ & \sum_{j=i}^{n} \left(r_{j+1} - r_j\right) \left(1 - \frac{1}{j}\right)^2 \frac{1}{12} \mathbf{i}_{11} \omega_i^2 \\ & + \left[\sum_{l=i}^{n} \left(r_{l+1} - r_l\right) \left(\frac{n}{l^2} - \frac{2}{l}\right)^2\right] \frac{1}{12} \mathbf{i}_{11} \sum_{j<i} \omega_j^2 \\ & + \sum_{j>i} \left[\sum_{l=j}^{n} \left(r_{l+1} - r_l\right) \left(\frac{n}{l^2} - \frac{2}{l}\right)^2 \frac{1}{12} \mathbf{i}_{11}\right] \omega_j^2 \\ = \ & \left(1 - r_i - 2a_{1i} + a_{2i}\right) \frac{1}{12} \mathbf{i}_{11}^{\top} \omega_i^2 + \frac{1}{12} \mathbf{i}_{11}^{\top} \left(n^2 a_{4i} - 4 n a_{3i} + 4 a_{2i}\right) \sum_{j<i} \omega_j^2 \\ & + \frac{1}{12} \mathbf{i}_{11} \sum_{j>i} \left[\left(n^2 a_{4j} - 4 n a_{3j} + 4 a_{2j}\right)\right] \omega_j^2, \end{aligned}$$

since

$$\begin{aligned} & \sum_{j<i} \left(\frac{1}{T} \sum_{t=t_i}^{T} \sum_{s=t_i}^{T} c_{ijt} C_{ijs}^{T}\right) \gamma_j(t-s) \\ = \ & \ldots\ldots \\ = \ & \sum_{j<i} \frac{1}{T} \sum_{t=t_i}^{T} \sum_{s=t_i}^{T} \left(n\left(\frac{1}{m_t^2}\right) - \frac{2}{m_t}\right) \left[\frac{n}{m_s^2} - \frac{2}{m_s}\right] \frac{1}{12} \mathbf{i}_{11} \gamma_j(t-s) \\ = \ & \left[\sum_{l=i}^{n} \left(r_{l+1} - r_l\right) \left(\frac{n}{l^2} - \frac{2}{l}\right)^2\right] \frac{1}{12} \mathbf{i}_{11} \sum_{j<i} \omega_j^2, \end{aligned}$$

$$
\begin{aligned}
& \sum_{j>i}\left(\frac{1}{T}\sum_{t=t_j}^{T}\sum_{s=t_i}^{T} c_{ijt}C_{ijs}^{T}\right)\gamma_j(t-s) \\
= & \;\cdots\cdots \\
= & \sum_{j>i}\left\{\frac{1}{T}\sum_{t=t_j}^{T}\sum_{s=t_i}^{T}\left(n\left(\frac{1}{m_t^2}\right)-\frac{2}{m_t}\right)\left[\frac{n}{m_z^2}-\frac{2}{m_z}\right]\frac{1}{12}\mathbf{i}_{11}\gamma_j(t-s)\right\} \\
= & \sum_{j>i}\left\{\sum_{l=j}^{n}(r_{l+1}-r_l)\left(\frac{n}{l^2}-\frac{2}{l}\right)^2\frac{1}{12}\mathbf{i}_{11}\right\}\omega_j^2,
\end{aligned}
$$

$$
\begin{aligned}
& \operatorname{Cov}\left(e_{a,i}, e_{A,j}^{\top}\right) \\
= & \left(\frac{1}{T}\sum_{t=t_i}^{T}\sum_{s=t_i}^{T} c_{it}C_{jis}^{T}\right)\gamma_i(t-s)+\left(\frac{1}{T}\sum_{t=t_i}^{T}\sum_{s=t_i}^{T} c_{ijt}C_{js}^{T}\right)\gamma_j(t-s)+\sum_{l\neq i,j}\left(\frac{1}{T}\sum_{t=t_i}^{T}\sum_{s=t_i}^{T} c_{ilt}C_{jls}^{T}\right)\gamma_l(t-s) \\
\to & \sum_{l=\max(i,j)}^{n}(r_{l+1}-r_l)\left(1-\frac{1}{l}\right)\left(\frac{n}{l^2}-\frac{2}{l}\right)\frac{1}{12}\mathbf{i}_{11}\left(\omega_i^2+\omega_j^2\right) \\
& +\sum_{l\neq i,j}\sum_{s=\max(i,j,l)}^{n}(r_{s+1}-r_s)\left(\frac{n}{s^2}-\frac{2}{s}\right)^2\frac{1}{12}\mathbf{i}_{11}\omega_l^2 \\
= & \sum_{l=\max(i,j)}^{n}(r_{l+1}-r_l)\left(\frac{n+2}{l^2}-\frac{2}{l}-\frac{n}{l^3}\right)\frac{1}{12}\mathbf{i}_{11}\left(\omega_i^2+\omega_j^2\right) \\
& +\sum_{l\neq i,j}\sum_{s=\max(i,j,l)}^{n}(r_{s+1}-r_s)\left(\frac{n^2}{s^4}-\frac{4n}{s^3}+\frac{4}{s^2}\right)^2\frac{1}{12}\mathbf{i}_{11}\omega_l^2 \\
= & \left[(n+2)a_{2,\max(i,j)}-2a_{1,\max(i,j)}-na_{3,\max(i,j)}\right]\frac{1}{12}\mathbf{i}_{11}\left(\omega_i^2+\omega_j^2\right) \\
& +\sum_{l\neq i,j}\left(n^2a_{4,\max(i,j,l)}-4na_{3,\max(i,j,l)}+4a_{2,\max(i,j,l)}\right)\frac{1}{12}\mathbf{i}_{11}\omega_l^2.
\end{aligned}
$$

For example, if $i<j$,

$$
\begin{aligned}
& \operatorname{Cov}\left(e_{a,i}, e_{A,j}^{\top}\right) \\
\to & \sum_{l=j}^{n}(r_{l+1}-r_l)\left(1-\frac{1}{l}\right)\left(\frac{n}{l^2}-\frac{2}{l}\right)\frac{1}{12}\mathbf{i}_{11}\left(\omega_i^2+\omega_j^2\right) \\
& +\sum_{l\neq i,j}\sum_{s=\max(i,j,l)}^{n}(r_{s+1}-r_s)\left(\frac{n}{s^2}-\frac{2}{s}\right)^2\frac{1}{12}\mathbf{i}_{11}\omega_l^2 \\
= & \sum_{l=j}^{n}(r_{l+1}-r_l)\left(1-\frac{1}{l}\right)\left(\frac{n}{l^2}-\frac{2}{l}\right)\frac{1}{12}\mathbf{i}_{11}\left(\omega_i^2+\omega_j^2\right) \\
& +\sum_{s=j}^{n}(r_{s+1}-r_s)\left(\frac{n}{s^2}-\frac{2}{s}\right)^2\frac{1}{12}\mathbf{i}_{11}\sum_{l<j}\omega_l^2+\sum_{l>j}\left[\sum_{s=l}^{n}(r_{s+1}-r_s)\left(\frac{n}{s^2}-\frac{2}{s}\right)^2\frac{1}{12}\mathbf{i}_{11}\right]\omega_l^2.
\end{aligned}
$$

If $i > j$,

$$\begin{aligned}
& \operatorname{Cov}\left(e_{a,i}, e_{A,j}^{\top}\right) \\
\rightarrow\ & 2\sum_{l=i}^{n}\left(r_{l+1}-r_{l}\right)\left(1-\frac{1}{l}\right)\left(\frac{n}{l^{2}}-\frac{2}{l}\right)\frac{1}{12}\mathbf{i}_{11}\left(\omega_{i}^{2}+\omega_{j}^{2}\right) \\
& +\sum_{l\neq i,j}\sum_{s=\max(i,j,l)}^{n}\left(r_{s+1}-r_{s}\right)\left(\frac{n}{s^{2}}-\frac{2}{s}\right)^{2}\frac{1}{12}\mathbf{i}_{11}\omega_{l}^{2} \\
=\ & \sum_{l=i}^{n}\left(r_{l+1}-r_{l}\right)\left(1-\frac{1}{l}\right)\left(\frac{n}{l^{2}}-\frac{2}{l}\right)\frac{1}{12}\mathbf{i}_{11}\left(\omega_{i}^{2}+\omega_{j}^{2}\right) \\
& +\sum_{l<i}\sum_{s=i}^{n}\left(r_{s+1}-r_{s}\right)\left(\frac{n}{s^{2}}-\frac{2}{s}\right)^{2}\frac{1}{12}\mathbf{i}_{11}\omega_{l}^{2}+\sum_{l>i}\sum_{s=l}^{n}\left(r_{s+1}-r_{s}\right)\left(\frac{n}{s^{2}}-\frac{2}{s}\right)^{2}\frac{1}{12}\mathbf{i}_{11}\omega_{l}^{2}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\operatorname{Cov}\left(e_{a,i}, e_{a,j}^{\top}\right) \rightarrow\ & \sum_{l=\max(i,j)}^{n}\left(r_{l+1}-r_{l}\right)\left(1-\frac{1}{l}\right)\left(\frac{n}{l^{2}}-\frac{2}{l}\right)\left(\omega_{i}^{2}+\omega_{j}^{2}\right) \\
& +\sum_{l\neq i,j}\sum_{s=\max(i,j,l)}^{n}\left(r_{s+1}-r_{s}\right)\left(\frac{n}{s^{2}}-\frac{2}{s}\right)^{2}\omega_{l}^{2} \\
=\ & \left[(n+2)a_{2,\max(i,j)}-2a_{1,\max(i,j)}-na_{3,\max(i,j)}\right]\left(\omega_{i}^{2}+\omega_{j}^{2}\right) \\
& +\sum_{l\neq i,j}\left(n^{2}a_{4,\max(i,j,l)}-4na_{3,\max(i,j,l)}+4a_{2,\max(i,j,l)}\right)\omega_{l}^{2},
\end{aligned}$$

for example, for $i < j$,

$$\begin{aligned}
\operatorname{Cov}\left(e_{a,i}, e_{a,j}^{\top}\right) =\ & \left[(n+2)a_{2,j}-2a_{1,j}-na_{3,j}\right]\left(\omega_{i}^{2}+\omega_{j}^{2}\right) \\
& +\left(n^{2}a_{4,j}-4na_{3,j}+4a_{2,j}\right)\sum_{l<j}\omega_{l}^{2}+\sum_{l>j}\left(n^{2}a_{4,l}-4na_{3,l}+4a_{2,l}\right)\omega_{l}^{2},
\end{aligned}$$

$$\begin{aligned}
\operatorname{Cov}\left(e_{A,i}, e_{A,j}^{\top}\right) \rightarrow\ & \sum_{l=\max(i,j)}^{n}\left(r_{l+1}-r_{l}\right)\left(1-\frac{1}{l}\right)\left(\frac{n}{l^{2}}-\frac{2}{l}\right)\frac{1}{12^{2}}\mathbf{i}_{11}\mathbf{i}_{11}^{\top}\left(\omega_{i}^{2}+\omega_{j}^{2}\right) \\
& +\sum_{s=\max(i,j,l)}^{n}\left(r_{s+1}-r_{s}\right)\left(\frac{n}{s^{2}}-\frac{2}{s}\right)^{2}\frac{1}{12^{2}}\mathbf{i}_{11}\mathbf{i}_{11}^{\top}\sigma^{2} \\
=\ & \left[(n+2)a_{2,\max(i,j)}-2a_{1,\max(i,j)}-na_{3,\max(i,j)}\right]\frac{1}{12^{2}}J_{11}\left(\omega_{i}^{2}+\omega_{j}^{2}\right) \\
& +\sum_{l\neq i,j}\left(n^{2}a_{4,\max(i,j,l)}-4na_{3,\max(i,j,l)}+4a_{2,\max(i,j,l)}\right)\frac{1}{12^{2}}J_{11}\sigma^{2}.
\end{aligned}$$

Thus, let

$$\begin{aligned}
\delta_i &= (1 - r_i - 2a_{1i} + a_{2i}) \\
f_i &= (n+2)a_{2,i} - 2a_{1,i} - na_{3,i} \ldots\ldots f_i\left(\omega_l^2 + \omega_i^2\right),\ l < i \\
\lambda_i &= \left(n^2 a_{4,i} - 4na_{3,i} + 4a_{2,i}\right) \ldots\ldots \lambda_i \sum_{j \neq l, j < i} \omega_j^2 + \sum_{j > i} \lambda_j \omega_j^2,\ l < i
\end{aligned}$$

the covariance matrix of $e_T = (e_a^\top, e_A^\top)^\top$ is given by

$$\begin{bmatrix} \mathrm{Var}(e_a) & \mathrm{Cov}(e_a, e_A) \\ \mathrm{Cov}(e_A, e_a) & \mathrm{Var}(e_A) \end{bmatrix},$$

where

$$\begin{aligned}
&\mathrm{Var}\left(e_a\right) \\
= &\begin{bmatrix} \delta_1\omega_1^2 & & \\ & \delta_i\omega_i^2 & \\ & & \delta_n\omega_n^2 \end{bmatrix} \\
&+ \begin{bmatrix} \sum_{j=2}^n \lambda_j\omega_j^2 & f_i\left(\omega_1^2 + \omega_i^2\right) + \lambda_i \sum_{j\neq 1, j<i} \omega_j^2 + \sum_{j>i} \lambda_j\omega_j^2 & f_n\left(\omega_1^2 + \omega_n^2\right) + \lambda_n \sum_{j\neq 1, j<n} \omega_j^2 \\ & \lambda_i \sum_{j=1}^{i-1} \omega_j^2 + \sum_{j=i+1}^n \lambda_j\omega_j^2 & f_n\left(\omega_n^2 + \omega_i^2\right) + \lambda_n \sum_{j\neq i, j<n} \omega_j^2 \\ & & \lambda_n \sum_{j=1}^{n-1} \omega_j^2 \end{bmatrix} \\
= &\ \Omega_n + A_n,
\end{aligned}$$

$$
\begin{aligned}
&\operatorname{Cov}(e_a, e_A) \\
= &\begin{bmatrix} \delta_1\omega_1^2 & & \\ & \delta_i\omega_i^2 & \\ & & \delta_n\omega_n^2 \end{bmatrix} \otimes \frac{1}{12}\mathbf{i}_{11}^\top \\
+ &\begin{bmatrix} \sum_{j=2}^n \lambda_j\omega_j^2 & f_i\left(\omega_1^2+\omega_i^2\right)+\lambda_i\sum_{l<i}\omega_l^2+\sum_{l>i}\lambda_l\omega_l^2 & f_n\left(\omega_1^2+\omega_n^2\right)+\lambda_n\sum_{l<n}\omega_l^2 \\ & \lambda_i\sum_{j<i}\omega_j^2+\sum_{j=i+1}^n\lambda_j\omega_j^2 & f_n\left(\omega_n^2+\omega_i^2\right)+\lambda_n\sum_{l<n}\omega_l^2 \\ & & \lambda_n\sum_{j<n}\omega_j^2 \end{bmatrix} \otimes \frac{1}{12}\mathbf{i}_{11}^\top \\
= &\left[\Omega_n + A_n\right] \otimes \frac{1}{12}\mathbf{i}_{11}^\top,
\end{aligned}
$$

$$
\begin{aligned}
&\operatorname{Var}(e_A) \\
= &\begin{bmatrix} \delta_1 s_1^2 & & \\ & \delta_i s_i^2 & \\ & & \delta_n s_n^2 \end{bmatrix} \otimes \frac{1}{12}I_{11} \\
+ &\begin{bmatrix} \sum_{j=2}^n \lambda_j\omega_j^2 & f_i\left(\omega_1^2+\omega_i^2\right)+\lambda_i\sum_{j=2}^{i-1}\omega_j^2+\sum_{l>i}\lambda_l\omega_l^2 & f_n\left(\omega_1^2+\omega_n^2\right)+\lambda_n\sum_{j=2}^{n-1}\omega_j^2 \\ & \lambda_i\sum_{j=1}^{i-1}\omega_j^2+\sum_{j=i+1}^n\lambda_j\omega_j^2 & f_n\left(\omega_n^2+\omega_i^2\right)+\lambda_n\sum_{j=2}^{i-1}\omega_j^2 \\ & & \lambda_n\sum_{j=1}^{n-1}\omega_j^2 \end{bmatrix} \otimes \frac{1}{12^2}J_{11} \\
= &S_n \otimes \frac{1}{12}I_{11} + A_n \otimes \frac{1}{12^2}J_{11}.
\end{aligned}
$$

Thus the covariance matrix of the stochastic term is

$$
\Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix} = \begin{bmatrix} \Omega_n + A_n & \left[\Omega_n + A_n\right] \otimes \frac{1}{12}\mathbf{i}_{11}^\top \\ \left[\Omega_n + A_n\right] \otimes \frac{1}{12}\mathbf{i}_{11}^\top & S_n \otimes \frac{1}{12}I_{11} + A_n \otimes \frac{1}{12^2}J_{11} \end{bmatrix}.
$$

■

References

[1] AHN, H., and J.L. POWELL (1993). "Estimation of Censored Selection Models with a Nonparametric Selection Mechanism." *Journal of Econometrics*, 58, 3-30.

[2] BICKEL, P. J., C. A. J. KLAASSEN, Y. RITOV, J. A. WELLNER (1993). Efficient and Adaptive Inference in Semiparametric Models, Monograph, Johns Hopkins University Press, Baltimore. Reprinted (1998), Springer, New York.

[3] CAMPBELL, S.D., and F.X. DIEBOLD (2005). "Weather Forecasting for Weather Derivatives." *Journal of the American Statistical Association*, 100, 6-16.

[4] CHEN, H. (1988): "Convergence Rates for Parametric Components in a Partly Linear Model," *Annals of Statistics*, 16, 136-146.

[5] ENGLE, R. F., C. W. J. GRANGER, J. RICE, and A. WEISS (1986). "Semiparametric Estimates of the Relationship Between Weather and Electricity Sales." *Journal of the American Statistical Association*, 81, 310-320.

[6] FAN, J. (1992)."Design-Adaptive Nonparametric Regression." *Journal of the American Statistical Association*, 87, 998-1004.

[7] FAN, J. and I.GIJBELS (1992). "Variable Bandwidth and Local Linear Regression Smoothers." *Annals of Statistics*, 20, 2008-2036.

[8] GAO, J. and K. HAWTHORNE (2006). "Semiparametric Estimation and Testing of the Trend of Temperature Series." *Econometrics Journal* 9, 332-355.

[9] HALL, P., and C. HEYDE (1980). *Martingale Limit Theory and its Application.* New York, Academic Press.

[10] HART, J.D. (1991). "Kernel Regression Estimation with Time Series Errors." *Journal of the Royal Statistical Society, Series B.* 53, 173-187.

[11] HECKMAN, N.E. (1986). "Spline Smoothing in a Partly Linear Model." *Journal of the Royal Statistical Society, Series B*, 48, 244-248.

[12] HOOGSTRATE, A.J., F.C PALM, G.A GERARD (2000). "Pooling in Dynamic Panel-Data Models: An Application to Forecasting GDP Growth Rates." *Journal of Business and Economic Statistics*, Vol. 18, 3, pp. 274-83.

[13] ISSLER, J.V and L.R. LIMA (2009). "A Panel Data Approach to Economic Forecasting: The Bias-corrected Average Forecast," *Journal of Econometrics*, Vol. 152, pp. 153-164.

[14] MARRON, J.S. and M.P.WAND (1992). "Exact Mean Integrated Squared Error." *Annals of Statistics*, 20, 712-736.

[15] MILIONIS, A.E and T.D DAVIES (1994). "Box-Jenkins Univariate Modeling for Climatological Time Series Analysis: An Application to the Monthly Activity of Temperature Inversions." *International Journal of Climatology* 14 569-579.

[16] MÜLLER, H. G. (1988). *Nonparametric Regression Analysis of Longitudinal Data.* Lecture Notes in Statistics, Vol. 46. Heidelberg/New York: Springer-Verlag.

[17] NEWEY, W.K., J.L. POWELL, J.R. WALKER, (1990). "Semiparametric Estimation of Selection Models: Some Empirical Results." *American Economic Review,* 80, 324-328.

[18] RICE, J. (1986). "Convergence Rates for Partially Splined Models." *Statistics and Probability Letters*, 4, 203-208.

[19] ROBINSON, P. M. (1983). "Nonparametric Estimation for Time Series Models." *Journal of Time Series Analysis* 4, 185-208.

[20] ROBINSON, P. M. (1988). "Root-N-Consistent Semiparametric Regression." *Econometrica*, 56, 931-954.

[21] ROBINSON, P. M. (2008). "Nonparametric Trending Regression with Cross- sectional Dependence." Working paper available at http://www.ihs.ac.at/etsast/presentations/Robinson.pdf

[22] SEVERINI, T.A and W.H. WONG (1992). "Profile Likelihood and Conditionally Parametric Models." *The Annals of Statistics,* 20, 1768-1802.

[23] SILVERMAN, B. (1986). *Density Estimation for Statistics and Data Analysis.* London, Chapman and Hall.

[24] SPECKMAN, P. (1988). "Kernel Smoothing in Partial Linear Models." *Journal of the Royal Statistical Society Series B*, 50, 413-436.

[25] STOCK, J. H. (1989). "Nonparametric Policy Analysis." *Journal of the American Statistical Association*, 84, 567-576.

[26] STOCK, J. H. (1991). "Nonparametric Policy Analysis: An Application to Estimating Hazardous Waste Cleanup Benefits." in *Nonparametric and Semiparametric Methods in Econometrics and Statistics.* Eds Barnett, Powell, and Tauchen.

[27] STOKER, T.M, (1989). "Tests of Average Derivative Constraints." *Review of Economic Studies,* 56, 535-552.

[28] STOKER, T.M, (1993). "Smoothing Bias in Derivative Estimation." *Journal of the American Statistical Association,* 88, 855-863.

[29] WIKIPEDIA. *Climate Change.* (2009) http://en.wikipedia.org/wiki/Climate_change