Statistica Sinica **24** (2014), 1075-1095
doi:http://dx.doi.org/10.5705/ss.2012.206

ROTATION SAMPLING FOR FUNCTIONAL DATA

David Degras

DePaul University

Abstract: This paper addresses the survey estimation of a population mean in continuous time. For this purpose we extend the rotation sampling method to functional data. In contrast to conventional rotation designs that select the sample before the survey, our approach randomizes each sample replacement and thus allows for adaptive sampling. Using Markov chain theory, we evaluate the covariance structure and the integrated squared error (ISE) of the related Horvitz-Thompson estimator. Our sampling designs decrease the mean ISE by suitably reallocating the sample across population strata during replacements. They also reduce the variance of the ISE by increasing the frequency or the intensity of replacements. To investigate the benefits of using both current and past measurements in the estimation, we develop a new composite estimator. In an application to electricity usage data, our rotation method outperforms fixed panels and conventional rotation samples. Because of the weak temporal dependence of the data, the composite estimator only slightly improves upon the Horvitz-Thompson estimator.

Key words and phrases: Asymptotic theory, composite estimator, functional data, Horvitz-Thompson estimator, Markov chain, rotation sampling.

1. Introduction

In various monitoring applications, sensor networks generate large volumes of data in continuous time. Due to cost or energy constraints, these collections of functional data (that is, curve data) often cannot be exhaustively observed. Electric utilities, for instance, need to monitor their clients? total consumption in order to adjust the power generation to the system load, to predict future consumption, and to determine pricing policies. However, they cannot access all clients? smart meters at each instant since this would exceed the network transmission capacity and/or incur considerable costs. Under such observation constraints, survey sampling provides competitive solutions for monitoring global parameters. (See Chiky et al. (2008) for a comparison between survey sampling and signal compression approaches.)

Several recent studies explore survey estimation based on functional data. Cardot, Degras, and Josserand (2012) extend the Horvitz-Thompson (HT) estimator to functional data and construct simultaneous confidence bands for the population mean function based on results of Degras (2011). Cardot et al. (2010)

investigate functional principal component analysis in design-based surveys and apply this technique to integrate auxiliary information. Cardot, Goga and Lardin (2013) develop model-assisted survey estimators for functional data. All of these studies rely on fixed panel designs: the same sample is used throughout the survey. Rotation sampling, in contrast, typically yields more accurate estimates of population parameters. This method, which replaces part of the sample at each survey occasion, is widely used in practice (e.g., U.S. Current Population Survey) and has received considerable attention in the literature (e.g., Eckler (1955); Rao and Graham (1964); Wolter (1979); and Lavallée (1995)). However, the available studies rely on modeling frameworks (e.g., discrete time, infinite population and stationary measurements) that are unsuitable for continuous-time monitoring applications and functional data.

In this paper we investigate rotation sampling for functional data. First, we devise sampling designs that replace the sample in part or in full at prespecified times. In contrast to conventional rotation designs that determine the sample prior to the survey, our approach randomizes each sample replacement. The sample can thus be adaptively selected and improved over time based on the observed data. (See Thompson and Seber (1996) for a comparison between adaptive sampling designs and conventional designs.) Second, we study the HT estimator of the mean function in a stratified population. Using Markov chain theory, we derive large-sample approximations for the mean and variance of the integrated squared error (ISE). If sample sizes in each stratum are constant over time, rotation samples and fixed panels have on average the same ISE. However, our rotation designs can reduce the mean ISE by suitably reallocating the sample at each replacement time (Neyman allocation). In addition, rotation samples dramatically decrease the variance of the ISE in comparison to fixed panels. Third, we develop a composite estimation procedure (see e.g., Rao and Graham (1964)) in order to improve upon the HT estimator. The composite estimator is recursively defined in terms of its value at an arbitrary previous instant, the estimated change in the population mean, and the HT estimator. Finally, we apply our sampling strategies to electricity usage data from the Irish CER Smart Metering Project (CER (2011)). For the electricity usage dataset, our rotation designs outperform fixed panels and conventional rotation samples both in terms of estimation accuracy and stability. The composite estimator slightly improves upon the HT estimator.

The paper is organized as follows. In Section 2, we present the modeling framework and the HT estimator. In Section 3, we define the new rotation sampling designs. The mean and covariance of the HT estimator are studied in Section 4. Section 5 gives the main results on the variance of the ISE. Section 6 introduces the composite estimator. The numerical study is described in Section

7. Section 8 provides concluding remarks. The main proofs are gathered in the Appendix. Additional proofs are available online as supplementary material.

2. Statistical Framework

Let $U_N = \{1, \ldots, N\}$ be a finite population and let X_k, $k \in U_N$, be deterministic functions defined on a bounded interval $[0, T]$. We study the estimation of the population mean function

$$\mu_N(t) = \frac{1}{N} \sum_{k \in U_N} X_k(t)$$

based on $X_k(t)$, $k \in s(t)$, where $t \in [0, T]$ and $s(t) \subset U_N$ is a probability sample of fixed size $n(t)$. The collection of samples $s = \{s(t) : t \in [0, T]\}$ can be viewed as a random function from $[0, T]$ to the set $\mathcal{P}(U_N)$ of all subsets of U_N. It is selected from the function space $\mathcal{S} = \{s : [0, T] \to \mathcal{P}(U_N)\}$ according to a probability measure P specified by the statistician. For all $k, l \in U_N$ and $t, t' \in [0, T]$, we denote the first and second order inclusion probabilities by $\pi_k(t) = P(k \in s(t)) = P(\{s \in \mathcal{S} : k \in s(t)\})$ and $\pi_{kl}(t, t') = P(k \in s(t),\, l \in s(t'))$ $= P(\{s \in \mathcal{S} : k \in s(t), l \in s(t')\})$. All expectations are taken with respect to P.

We consider the estimator of Horvitz and Thompson (1952)

$$\hat{\mu}_{ht}(t) = \frac{1}{N} \sum_{k \in U_N} \frac{I_k(t)}{\pi_k(t)} X_k(t), \tag{2.1}$$

where $I_k(t)$ is the sample indicator function of $k \in U_N$ at time t: $I_k(t) = 1$ if $k \in s(t)$ and $I_k(t) = 0$ otherwise. This estimator, which we refer to as the HT estimator, is unbiased for $\mu_N(t)$ and its covariance function is

$$\mathrm{Cov}(\hat{\mu}_{ht}(t), \hat{\mu}_{ht}(t')) = \frac{1}{N^2} \sum_{k,l \in U} \frac{\Delta_{kl}(t, t')}{\pi_k(t)\pi_l(t')} X_k(t) X_l(t'),$$

where $\Delta_{kl}(t, t') = \mathrm{Cov}(I_k(t), I_l(t')) = \pi_{kl}(t, t') - \pi_k(t)\pi_l(t')$.

The estimation accuracy can often be improved by stratifying the population. From now on we drop the subscript in U_N and assume that U is partitioned into strata U_h, $1 \le h \le H$, of size N_h. We denote the strata mean and covariance functions by $\mu_h(t) = (1/N_h) \sum_{k \in U_h} X_k(t)$ and

$$\gamma_h(t, t') = \frac{1}{N_h - 1} \sum_{k \in U_h} \left(X_k(t) - \mu_h(t) \right) \left(X_k(t') - \mu_h(t') \right).$$

Let $n_h(t) = \#(s(t) \cap U_h)$ be the sample size in U_h at time t and $f_h(t) = n_h(t)/N_h$ be the sampling rate. If $s(t)$ is obtained by simple random sampling without replacement (SRSWOR) independently in each U_h, the HT estimator is

$$\hat{\mu}_{ht}(t) = \frac{1}{N} \sum_{h=1}^{H} \frac{1}{f_h(t)} \sum_{k \in U_h} I_k(t) X_k(t) \tag{2.2}$$

and its covariance rewrites as

$$\mathrm{Cov}\big(\hat{\mu}_{ht}(t), \hat{\mu}_{ht}(t')\big) = \frac{1}{N^2} \sum_{h=1}^{H} \frac{1}{f_h(t) f_h(t')} \sum_{k,l \in U_h} \Delta_{kl}(t,t') X_k(t) X_l(t'). \tag{2.3}$$

Remark 1. For simplicity, we assume that the sampled curves X_k are observed in continuous time and without noise. If these curves are observed at discrete times and/or with noise, interpolation or smoothing methods should be applied. In this case the results of this paper still hold under standard interpolation or smoothing conditions. See for example Cardot and Josserand (2011) and Cardot, Degras, and Josserand (2012).

3. Rotation Designs for Continuous-time Surveys

Rotation sampling has so far been developed for discrete-time surveys. In this section we extend it to the continuous-time framework of functional data. We propose two choices of the probability measure P for selecting the time-varying sample $s = \{s(t) : t \in [0,T]\}$ in a stratified population. These sampling designs, which we refer to as full replacement and partial replacement, share the following features

- The time-varying samples $s_h = \{s(t) \cap U_h : t \in [0,T]\}$, $1 \leq h \leq H$, are independent across strata.
- At time $\tau_0 = 0$, the samples $s_h(\tau_0)$ are obtained by SRSWOR.
- The s_h can be modified at fixed times $0 < \tau_1 < \ldots < \tau_m < T$.

It remains to specify the probability distribution of the discrete processes $\{s_h(\tau_r) : 1 \leq r \leq m\}$ under full and partial replacement.

1. **Full replacement.** For each h, the successive samples $s_h(\tau_r), 1 \leq r \leq m$, are obtained by independent SRSWOR of $n_h(\tau_r)$ units in U_h.
2. **Partial replacement.** For each h, a fraction $\alpha_h \in [0,1]$ of $s_h(t)$ is replaced at each time τ_r, $1 \leq r \leq m$. Given $s_h(\tau_{r-1})$, $s_h(\tau_r)$ is obtained by the following independent operations:
   - select $\alpha_h n_h(\tau_{r-1})$ units in $s_h(\tau_{r-1})$ by SRSWOR and discard them from the sample;
   - select $(n_h(\tau_r) - (1-\alpha_h) n_h(\tau_{r-1}))$ units in $U_h \setminus s_h(\tau_{r-1})$ by SRSWOR and add them to the sample.

By construction, the process $\{s_h(\tau_r) : 0 \le r \le m\}$ is a Markov chain both under full and partial replacement. Full replacement is not a special case of partial replacement with $\alpha_h = 1$: indeed, $s_h(\tau_{r-1})$ and $s_h(\tau_r)$ are independent in the former case whereas they are disjoint (and thus dependent) in the latter. In partial replacement we refer to the α_h as the replacement rates. For simplicity we assume that the α_h are constant over time and that the proposed sample replacements are possible without modifications, which entails that $\alpha_h n_h(\tau_{r-1}) \in \mathbb{N}$ and $n_h(\tau_{r-1}) \le n_h(\tau_r) + \alpha_h n_h(\tau_{r-1}) \le N_h$ for all h, r. Fixed panels correspond to partial replacement with $\alpha_h = 0$ and the $n_h(t)$ constant over time.

We state the probability distribution of $s_h(t)$, using an induction argument on the τ_r for partial replacement (the result holds trivially for full replacement).

Proposition 1. *Assume either the full or the partial replacement design. For all $1 \le h \le H$ and $t \in [0, T]$, the probability distribution of $s_h(t)$ is identical to the SRSWOR of $n_h(t)$ units in U_h.*

4. Covariance of the Horvitz-Thompson Estimator

Here we derive the covariance function (2.3) of the HT estimator (2.2) under the rotation designs, which amounts to determining $\Delta_{kl}(t, t')$ explicitly. Let $\nu(t) = \min\{r : \tau_r \le t\}$ be the number of sample replacements before time t. For $0 \le t < T$, it holds that $\tau_{\nu(t)} \le t < \tau_{\nu(t)+1}$. By convention, we set $\tau_{m+1} = T$ and $\nu(T) = m$. Let $\delta_{..}$ indicate the Krönecker delta.

4.1. Covariance under full replacement

Under the full replacement design, $\hat{\mu}_{ht}(t)$ and $\hat{\mu}_{ht}(t')$ are independent if the sample has been replaced between times t and t'. If no replacement occurred between t and t', $\Delta_{kl}(t, t')$ can be derived from the properties of SRSWOR.

Theorem 1. *Assume the full replacement design. For all $t, t' \in [0, T]$,*

$$\mathrm{Cov}\big(\hat{\mu}_{ht}(t), \hat{\mu}_{ht}(t')\big) = \frac{1}{N} \sum_{h=1}^{H} \frac{N_h}{N} \, \frac{1 - f_h(t)}{f_h(t)} \, \gamma_h(t, t') \, \delta_{\nu(t)\nu(t')} \, .$$

4.2. Covariance under partial replacement

To derive $\Delta_{kl}(t, t')$, it suffices to find $\pi_{kl}(t, t')$ in view of Proposition 1. By definition of SRSWOR, for a given stratum U_h, $\pi_{kk}(t, t')$ and $\pi_{kl}(t, t')$ do not depend on $k, l \in U_h$ $(k \ne l)$. Since $\sum_{k \in U_h} I_k(t) = n_h(t)$, it follows that $E(\sum_k I_k(t) \sum_l I_l(t')) = n_h(t) n_h(t') = N_h \pi_{kk}(t, t') + N_h(N_h - 1)\pi_{kl}(t, t')$, with $k \ne l$ two arbitrary units in U_h. Therefore, it suffices to determine $\pi_{kk}(t, t')$. This in turn reduces to computing $P(k \in s_h(t') | k \in s_h(t))$.

Let D be a subset of U_h. The Markovian nature of $\{s_h(\tau_r) : 0 \leq r \leq m\}$ and the properties of SRSWOR yield the following result.

Lemma 1. *Under the partial replacement design,* $\{s_h(\tau_r) \cap D : 0 \leq r \leq m\}$ *is a Markov chain.*

By setting $D = \{k\}$ in Lemma 1, $\{I_k(\tau_r) : 0 \leq r \leq m\}$ is a Markov chain whose transition probabilities can be found with the Chapman-Kolmogorov equations. To this end, define

$$\lambda_h(t,t') = \prod_{r=\nu(t)+1}^{\nu(t')} \frac{1-\alpha_h - f_h(\tau_r)}{1-f_h(\tau_{r-1})} \tag{4.1}$$

for $0 \leq t \leq t' \leq T$ with $\lambda_h(t,t') = 1$ if $\nu(t) = \nu(t')$. Set to 1 all factors in $\lambda_h(t,t')$ for which $f_h(\tau_{r-1}) = 1$ and extend $\lambda_h(t,t')$ as a symmetric function on $[0,T]^2$.

Lemma 2. *Assume the partial replacement design. For all* U_h, $k \in U_h$, *and* $0 \leq t \leq t' \leq T$,

$$\begin{cases} P\left(k \in s_h(t') \middle| k \in s_h(t)\right) = \left(1 - f_h(t)\right)\lambda_h(t,t') + f_h(t'), \\ P\left(k \in s_h(t') \middle| k \notin s_h(t)\right) = f_h(t') - f_h(t)\lambda_h(t,t'). \end{cases}$$

The lemma is easily proved by induction and, with simple matrix diagonalizations, $\lambda_h(t,t')$ is the product of the eigenvalues of the transition probability matrices of $\{I_k(\tau_r) : 0 \leq r \leq m\}$ between t and t'.

For any two real numbers x, y, write $x \wedge y = \min(x,y)$ and $x \vee y = \max(x,y)$. and Lemma 2,

Theorem 2. *Assume the partial replacement design. For all* $t, t' \in [0,T]$,

$$\operatorname{Cov}\left(\hat{\mu}_{ht}(t), \hat{\mu}_{ht}(t')\right) = \frac{1}{N}\sum_{h=1}^{H} \frac{N_h}{N} \frac{1 - f_h(t \wedge t')}{f_h(t \vee t')} \gamma_h(t,t')\, \lambda_h(t,t').$$

To gain insight into Theorems 1−2, we suppose that the $n_h(t)$ are constant over time. In the case of fixed panels (partial replacement with $\alpha_h = 0$), $\operatorname{Cov}(\hat{\mu}_{ht}(t), \hat{\mu}_{ht}(t')) = N^{-2}\sum_h N_h(f_h^{-1} - 1)\gamma_h(t,t')$ for all t, t'. Under full replacement, the estimator covariance is the same as for fixed panels on the diagonal blocks $[\tau_r, \tau_{r+1}]^2, 1 \leq r \leq m$, and is zero outside these blocks. Under partial replacement, the term $\lambda_h(t,t')$ simplifies to $(1 - \alpha_h/(1-f_h))^{|\nu(t)-\nu(t')|}$. Hence for fixed times t, t', the correlation between $\hat{\mu}_{ht}(t)$ and $\hat{\mu}_{ht}(t')$ decreases as $\alpha_h \in [0, 1-f_h]$ increases (assuming $\gamma_h(t,t') > 0$). If $\alpha_h = 1 - f_h$ for all h, the covariance is the same as under full replacement. For values $\alpha_h > 1 - f_h$, the covariance becomes unstable in the sense that $\lambda_h(t,t')$ changes sign on every block $[\tau_q, \tau_{q+1}] \times [\tau_r, \tau_{r+1}]$. If $\alpha_h \notin \{0, 1-f_h\}$ for all h, $|\mathrm{Cov}(\hat{\mu}_{ht}(t), \hat{\mu}_{ht}(t'))|$ decreases at an exponential rate as $|t-t'|$ increases.

4.3. Mean integrated squared error

To measure the accuracy of an estimator $\hat{\mu}_N$ of μ_N over $[0,T]$, we use the Integrated Squared Error

$$\mathrm{ISE} = \int_0^T \left(\hat{\mu}_N(t) - \mu_N(t)\right)^2 dt.$$

As seen in Sections 2−3, the HT estimator (2.2) is unbiased and, when the sample sizes $n_h(t)$ are constant over time, its variance function is the same under the full and partial replacement designs (in particular, for fixed panels). Therefore the HT estimator has the same mean integrated squared error

$$\mathrm{MISE} = \int_0^T E\left(\hat{\mu}_N(t) - \mu_N(t)\right)^2 dt$$

under both designs. On the other hand, in comparison to fixed panels, the full and partial replacement designs can reduce the MISE by using suitable time-varying sample sizes $n_h(t)$. Specifically, the variance of $\hat{\mu}_{ht}(\tau_r)$, $1 \le r \le m$, is minimal when $n_h(\tau_r)$ is chosen according to the Neyman allocation rule: $n_h(\tau_r) = c_r N_h \sqrt{\gamma_h(\tau_r, \tau_r)}$ with the constant c_r such that $\sum_h n_h(\tau_r) = n(\tau_r)$ (see e.g., Fuller (2009, p.21) for more details). Note that in practice, $\gamma_h(\tau_r, \tau_r)$ is unknown and must be estimated from the data.

5. Asymptotic Results for the ISE

To determine the variance of the ISE for the HT estimator (2.2) under the full and partial replacement designs, we first write

$$\begin{aligned}\mathrm{Var}(\mathrm{ISE}) &= \iint_{[0,T]^2} \mathrm{Cov}\left(\left\{\hat{\mu}_{ht}(t) - \mu_N(t)\right\}^2, \left\{\hat{\mu}_{ht}(t') - \mu_N(t')\right\}^2\right) dt dt' \\ &= \frac{1}{N^4} \iint_{[0,T]^2} \sum_{i,j,k,l \in U} \frac{\Delta_{ijkl}(t,t')}{\pi_i(t)\pi_j(t)\pi_k(t')\pi_l(t')} X_i(t)X_j(t)X_k(t')X_l(t')\, dt dt',\end{aligned}$$

where

$$\Delta_{ijkl}(t,t') = \mathrm{Cov}\big(\{I_i(t) - \pi_i(t)\}\{I_j(t) - \pi_j(t)\}, \{I_k(t') - \pi_k(t')\}\{I_l(t') - \pi_l(t')\}\big).$$

Based on the independence of samples across strata, it can be shown that

$$\begin{aligned}\mathrm{Var}(\mathrm{ISE}) &= \frac{1}{N^4} \sum_{h=1}^{H} \iint_{[0,T]^2} \sum_{i,j,k,l \in U_h} \frac{\Delta_{ijkl}(t,t')}{f_h^2(t) f_h^2(t')} X_i(t)X_j(t)X_k(t')X_l(t') dt dt' \\ &+ \frac{2}{N^4} \sum_{h \neq h'} \iint_{[0,T]^2} \sum_{i,k \in U_h} \frac{\Delta_{ik}(t,t')}{f_h(t) f_h(t')} X_i(t)X_k(t') \sum_{j,l \in U_{h'}} \frac{\Delta_{jl}(t,t')}{f_{h'}(t) f_{h'}(t')} X_j(t)X_l(t') dt.\end{aligned} \tag{5.1}$$

The variance (5.1) can be computed exactly if the $n_h(t)$ are constant over time, but requires large-sample approximations otherwise.

5.1. Asymptotic framework

We let the strata sizes N_h, sample sizes $n_h(t)$, replacement rates α_h and number of replacements m depend on the population size N and let $N \to \infty$. The parameters $m, n_h(t), N_h$ go to infinity with N while the number of strata and the observation period $[0,T]$ stay fixed. We make the following assumptions.

(A1) The curves X_k, $k \geq 1$, are integrable and uniformly bounded on $[0,T]$.

(A2) $\int_0^{\tau_r} g(t)dt = \frac{r}{m+1}$, $0 \leq r \leq m+1$, where g is a continuous, positive, and bounded function on $(0,T)$.

(A3) For all h, the sampling rate function f_h converges uniformly on $[0,T]$ to a continuous, positive limit function as $N \to \infty$.

(A4) For all h, the covariance function γ_h converges uniformly on $[0,T]^2$ to a continuous limit as $N \to \infty$.

(A5) $m = o(\min_h(N_h))$ as $N \to \infty$.

The condition $N_h \to \infty$ is not restrictive as, typically, small strata U_h are fully observed and do not contribute to the estimation error. (A1) allows discontinuity jumps in the individual curves X_k, while (A4) requires that the strata covariance functions can be uniformly approximated by continuous functions; this entails that at any time t, only a negligible fraction of the $X_k(t)$ have discontinuity jumps. This is needed under full replacement to approximate the covariance $\gamma_h(t,t')$ by the variance $\gamma_h(t,t)$ around the diagonal $\{t = t'\}$. (A2) ensures that the replacement times are regularly spaced. (A3) requires positive sampling rate functions, which is necessary for the consistent estimation of μ_N. (A5) is needed under partial replacement to approximate certain transition probabilities.

5.2. Intermediate results

Let $\tilde{X}_k(t) = X_k(t) - \mu_h(t)$ for $k \in U_h$ and $1 \leq h \leq H$. Simple algebra yields

$$\begin{aligned}
&\sum_{i,j,k,l \in U_h} \Delta_{ijkl}(t,t')\, X_i(t)X_j(t)X_k(t')X_l(t') \\
&\quad = \sum_{i,j,k,l \in U_h} E\left(I_i(t)I_j(t)I_k(t')I_l(t')\right) \tilde{X}_i(t)\tilde{X}_j(t)\tilde{X}_k(t')\tilde{X}_l(t') \qquad (5.2) \\
&\qquad - N_h^2\, f_h(t)f_h(t')\,(1-f_h(t))\left(1-f_h(t')\right)\gamma_h(t,t)\,\gamma_h(t',t')\,.
\end{aligned}$$

The sum on the right-hand side of (5.2) can be developed using the properties of SRSWOR. Let $a_N \sim b_N$ denote the asymptotic equivalence of real sequences (a_N) and (b_N).

Proposition 2. *Assume either the full or the partial replacement design and (A1). Let i^*, j^*, k^*, l^* be distinct units in a given stratum U_h. Then*

$$\sum_{i,j,k,l \in U_h} E\left(I_i(t)I_j(t)I_k(t')I_l(t')\right) \tilde{X}_i(t)\tilde{X}_j(t)\tilde{X}_k(t')\tilde{X}_l(t') \sim \left(C_1(t,t')\,\gamma_h(t,t)\gamma_h(t',t') + C_2(t,t')\gamma_h^2(t,t')\right) N_h^2$$

uniformly in $t, t' \in [0,T]$ as $N_h \to \infty$, where

$$\begin{aligned} C_1(t,t') &= E\left(I_{i^*}(t)I_{k^*}(t')\right) - E\left(I_{i^*}(t)I_{j^*}(t)I_{k^*}(t')\right) - E\left(I_{i^*}(t)I_{k^*}(t')I_{l^*}(t')\right) \\ &\quad + E\left(I_{i^*}(t)I_{j^*}(t)I_{k^*}(t')I_{l^*}(t')\right), \\ C_2(t,t') &= 2\,E\left(I_{i^*}(t)I_{i^*}(t')I_{k^*}(t)I_{k^*}(t')\right) - 4\,E\left(I_{i^*}(t)I_{i^*}(t')I_{j^*}(t)I_{k^*}(t')\right) \\ &\quad + 2\,E\left(I_{i^*}(t)I_{j^*}(t)I_{k^*}(t')I_{l^*}(t')\right). \end{aligned}$$

Under the full replacement design, the functions C_1 and C_2 can be expressed in terms of $f_h(t)$ and $f_h(t')$ and Var(ISE) can readily be computed. Under partial replacement, an additional result is required. Let k and l be two distinct units in a stratum U_h. Applying Lemma 1 to $D = \{k,l\}$ and using the Chapman-Kolmogorov equations and large-sample approximations, we obtain the following transition probabilities.

Proposition 3. *Assume the partial replacement design and* (A3)−(A5). *For all $0 \le t \le t' \le T$, as $N \to \infty$,*

$$\begin{cases} P\left(k,l \in s_h(t') \middle| k,l \in s_h(t)\right) \sim \left[\,(1-f_h(t))\lambda_h(t,t') + f_h(t')\,\right]^2, \\ P\left(k,l \in s_h(t') \middle| k \in s_h(t),\, l \notin s_h(t)\right) \\ \quad \sim \left[\,-f_h(t)\,(1-f_h(t))\,\lambda_h^2(t,t') + f_h(t')\,(1-2\,f_h(t))\,\lambda_h(t) + f_h^2(t')\right], \\ P\left(k,l \in s_h(t') \middle| k,l \notin s_h(t)\right) \sim \left[\,(1-f_h(t))\,\lambda_h(t,t') - (1-f_h(t'))\,\right]^2. \end{cases}$$

5.3. Variance of the integrated squared error

We can now state the main results.

Theorem 3. *Consider the HT estimator* (2.2) *based on the full replacement design. Assume* (A1), (A2), (A3) and (A4). *Then as $N \to \infty$,*

$$\mathrm{Var}\,(\mathrm{ISE}) \sim \frac{2}{mN^2} \int_0^T \Big(\sum_{h=1}^H \frac{N_h}{N}\, \frac{1-f_h(t)}{f_h(t)}\, \frac{1}{g(t)}\, \gamma_h(t,t) \Big)^2 dt.$$

Theorem 4. *Consider the HT estimator* (2.2) *based on the partial replacement design. Assume* (A1), (A2), (A3), *and* (A5). *Then as* $N \to \infty$,

$$\mathrm{Var}\,(\mathrm{ISE}) \sim \frac{2}{N^2} \iint_{[0,T]^2} \Big(\sum_{h=1}^{H} \frac{N_h}{N} \, \frac{1 - f_h(t)}{f_h(t')} \, \lambda_h(t,t') \, \gamma_h(t,t') \Big)^2 dtdt'.$$

Under additional assumptions, it is possible to find an asymptotic expression for $\lambda_h(t,t')$. Let G be an antiderivative of the density g in (A2).

Corollary 1. *Assume the conditions of Theorem* 4 *and suppose that* (i) *the sample sizes* $n_h(t)$ *are constant over time, and* (ii) $\lim_{N\to\infty} (\alpha_h m/(1-f_h)) = c_h < \infty$ *exists. Then as* $N \to \infty$,

$$\mathrm{Var}\,(\mathrm{ISE}) \sim \frac{2}{N^2} \iint_{[0,T]^2} \Big(\sum_{h=1}^{H} \frac{N_h}{N} \frac{1-f_h}{f_h} \exp\big(-c_h \left|G(t) - G(t')\right|\big) \, \gamma_h(t,t') \Big)^2 dtdt'.$$

Condition (ii) is reasonable since $\alpha_h m/T$ is the average sample replacement rate per unit time, which in practice stays bounded. The symmetry of α_h and m conforms to intuition, since multiplying either of these parameters by a given integer produces the same total of replaced units.

Under the assumptions of Theorems 3−4 and Corollary 1, we can compare the full and partial replacement designs in terms of variability of the ISE. As in Section 4 we include fixed panels as a special case of partial replacement where $\alpha_h = c_h = 0$. In comparison to fixed panels, partial replacement with $c_h > 0$ induces an exponentially decreasing function in Var(ISE). The decrease rate is larger when c_h is large and the data are highly positively correlated. In comparison to partial replacements, the full replacement design divides the order of Var(ISE) by a factor m, which massively stabilizes the estimation performance.

Remark 2. If the survey's goals include evaluating $I_N = \int_0^T \mu_N(t)dt$, then $\hat{I}_N = \int_0^T \hat{\mu}_{ht}(t)dt$ provides an unbiased estimator whose variance can be deduced from the previous results. As above, in comparison to fixed panels, partial replacement of the sample reduces $\mathrm{Var}(\hat{I}_N)$ by an exponentially decreasing function and full replacement divides the order of $\mathrm{Var}(\hat{I}_N)$ by a factor m.

6. Composite Estimation

The HT estimator (2.2) of $\mu_N(t)$ is only based on current observations X_k, $k \in s(t)$. The estimation can likely be improved by using past data in addition to current ones. Following the principle of composite estimation (e.g., Eckler (1955), we utilize the partial replacement design of Section 3 and recursively define a new estimator $\hat{\mu}_c(t)$ as a linear combination of $\hat{\mu}_{ht}(t)$ and $\hat{\mu}_c(t-\delta)$ plus the estimated change in μ_N between $t-\delta$ and t, where $\delta > 0$ is a lag parameter to be specified.

Let $0 \le t \le t' \le T$. If $|\nu(t) - \nu(t')| \le 1$, the estimator

$$\widehat{\Delta}\mu_N(t,t') = \frac{1}{N} \sum_{k \in U} \frac{I_k(t) I_k(t')}{\pi_{kk}(t,t')} \left(X_k(t') - X_k(t)\right) \tag{6.1}$$

of the level change $\Delta\mu_N(t,t') = \mu_N(t') - \mu_N(t)$ is unbiased. If $|\nu(t) - \nu(t')| \ge 2$, this estimator is extended as

$$\widehat{\Delta}\mu_N(t,t') = \widehat{\Delta}\mu_N(t, \tau_{\nu(t)+1}) + \sum_{r=\nu(t)+2}^{\nu(t')} \widehat{\Delta}\mu_N(\tau_{r-1}, \tau_r) + \widehat{\Delta}\mu_N(\tau_{\nu(t')}, t'). \tag{6.2}$$

The composite estimator is

$$\hat{\mu}_c(t) = \begin{cases} \hat{\mu}_{ht}(t), & 0 \le t < \tau_1, \\ Q\,\hat{\mu}_{ht}(t) + (1-Q)\left(\hat{\mu}_c(t-\delta) + \widehat{\Delta}\mu_N(t-\delta, t)\right), & \tau_1 \le t \le T, \end{cases} \tag{6.3}$$

where $Q \in [0,1]$ must be specified and, by convention, $\hat{\mu}_c(t) = \hat{\mu}_c(0)$ and $\widehat{\Delta}\mu_N(t,t') = \widehat{\Delta}\mu_N(0,t')$ if $t < 0 \le t'$. Note that if $\alpha_h = 0$ with sample sizes $n_h(t)$ constant over time, if $\delta = 0$, or if $Q = 1$, then $\hat{\mu}_c(t)$ reduces to $\hat{\mu}_{ht}(t)$. The composite estimator is thus a shrinkage estimator whose parameters α_h, δ, and Q determine the relative importance of past and present data in the estimation.

7. Numerical Study

Here we examine the numerical performances of the HT estimator (2.2) and composite estimator (6.3) based on the sampling designs of Section 3. We use electricity consumption data from the Irish CER Smart Metering Project conducted in 2009-10 (CER (2011)). During the project, smart meter readings (in kW) were collected every 30mn for $N = 6445$ residential and business customers. (The data are available by request at `www.ucd.ie/issda/data/commissionforenergyregulation/`.) We focus on one month of data (8/17/2009-9/17/2009) and set the sampling rate to 5% ($n = 322$) for the whole period. Customer electricity curves and the population mean curve are displayed in Figures 1 and 2. We stratify the population according to the type of contract (see Table 1) and replace the sample every 12 hours so that $\tau_r = 12r$, $1 \le r \le m$ (in hours) with $m = 61$ replacements.

The study investigates three factors in the estimation: rotation design (full, partial, or conventional), sample allocation (proportional, optimal, or adaptive), and estimator (HT or composite). The conventional rotation design consists in specifying a rotation pattern, and then randomly permuting the population

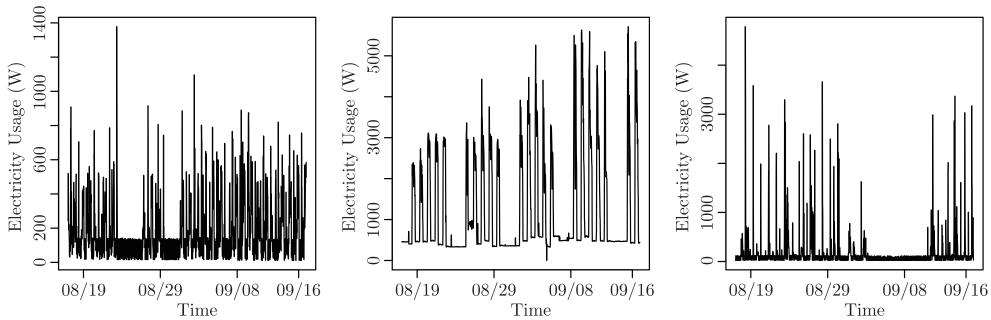

Figure 1. Sample electricity curves.

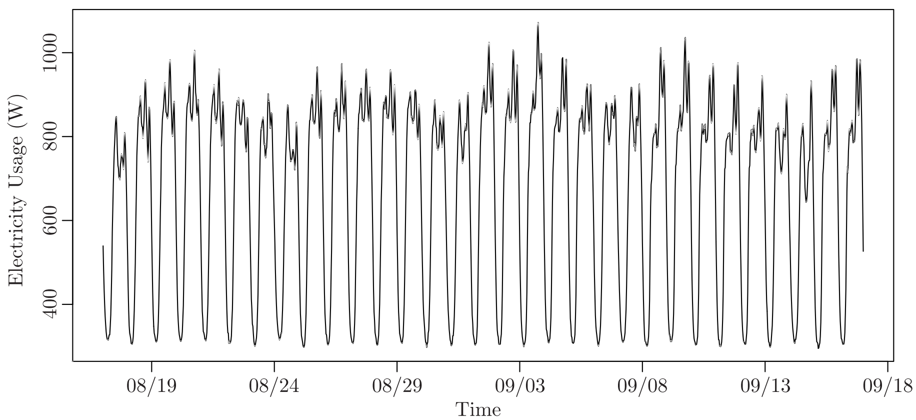

Figure 2. Mean electricity consumption in the population.

Table 1. Population strata. SME denotes Small-to-Medium Enterprises.

Stratum	Residential	SME	Other
Size	4225	485	1735

labels (see e.g., Rao and Graham (1964)). The partial replacement- and conventional rotation designs adopt the same replacement rate $\alpha \in \{0, 0.1, 0.2, \ldots, 1\}$ in each stratum U_h, $1 \leq h \leq 3$. In proportional allocation the sample sizes are $n_h = (N_h/N)n$ rounded to the nearest integer. Optimal (Neyman) allocation uses sample sizes $n_h(\tau_r) = c_r N_h \sqrt{\gamma_h(\tau_r, \tau_r)}$ with the constant c_r such that $\sum_h n_h(\tau_r) = n$. Figure 3 illustrates the difference between these allocations. Since the strata variances $\gamma_h(\tau_r, \tau_r)$ are unknown in practice, optimal allocation is infeasible; we use it as a benchmark. Adaptive allocation replaces the strata variances by estimates $\hat{\gamma}_h(\tau_r, \tau_r)$ in the optimal allocation method. We define $\hat{\gamma}_h(\tau_r, \tau_r)$ as the sample variance of the $X_k(t)$, $k \in s_h(t)$, at the last observation time t before τ_r. In composite estimation we used the parameter values $\alpha \in \{0.1, 0.2, \ldots, 1\}$, $\delta \in \{0.5, 1, 6, 12, 24\}$ (in hours), and $Q \in \{0, 0.1, 0.2, \ldots, 1\}$. For each combination of factors (sampling design, sample allocation and esti-

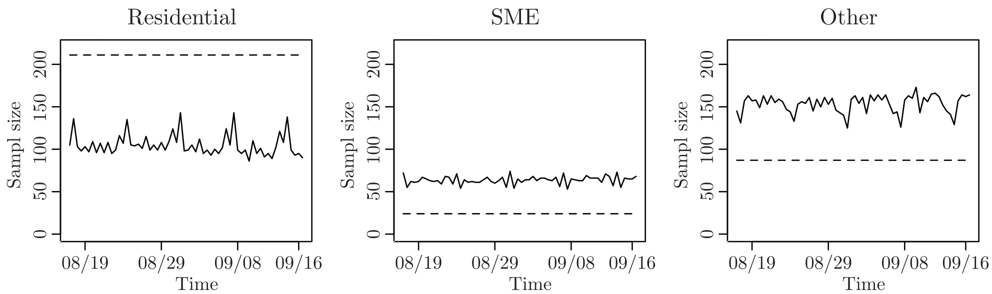

Figure 3. Proportional (dashes) and optimal (solid line) sample allocation.

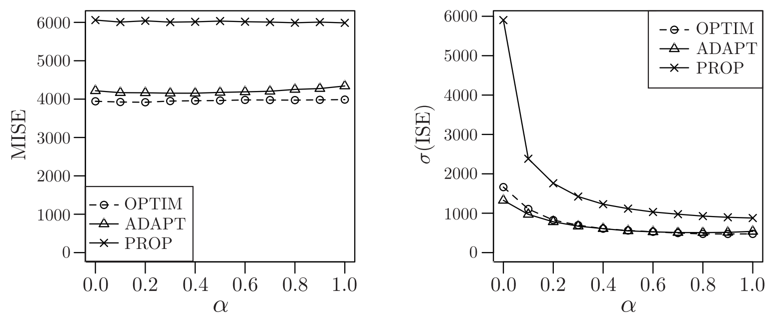

Figure 4. Comparison of proportional (PROP), adaptive (ADAPT) and optimal (OPTIM) sample allocation using the Horvitz-Thompson estimator. The mean and standard deviation of the ISE are displayed in terms of the replacement rate α

mator) and parameter values, we replicated the following sequence 10^4 times: generate the time-varying sample $s = \{s(t) : t \in [0, T]\}$ by Monte Carlo simulation, compute the estimator and determine its integrated squared error (ISE). The numerical study was implemented in the R language and each replicate ran in 0.05s to 0.2s, depending on the sample and estimator used.

The mean and standard deviation of the ISE for the HT estimator are shown in Figure 4. This figure compares the three types of sample allocation under the partial replacement design. Numerical results for the full replacement design are nearly identical to partial replacement with $\alpha = 1$. In line with Section 4.3 and Theorem 4, for a given allocation, the mean integrated squared error (MISE) is the same for all α while the standard deviation σ(ISE) is a decreasing function of α. Unsurprisingly, proportional allocation gives far less accurate results than optimal allocation. Adaptive allocation is comparable to optimal allocation in terms of MISE, with a relative efficiency between 91% and 95% across the range of α. Regarding σ(ISE), adaptive allocation is superior to optimal allocation for $\alpha \leq 0.6$. This result is not contradictory given that the optimal (Neyman) allocation is only optimal for the MISE and for fixed sample sizes (note that adaptive allocation requires random sample sizes).

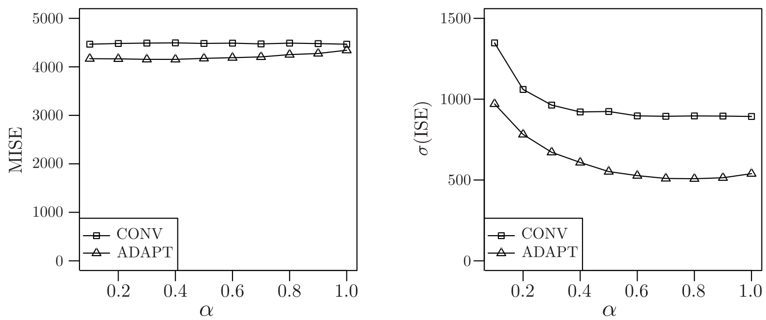

Figure 5. Comparison of conventional rotation sampling (CONV) to our adaptive rotation designs (ADAPT) using the Horvitz-Thompson estimator. The mean value and standard deviation of the ISE are displayed in terms of the replacement rate α.

We can compare conventional rotation sampling to our rotation designs, using again the HT estimator. Under the conventional design, we employed proportional allocation before τ_1 and took $n_h \propto N_h(\int_0^{\tau_1} \hat{\gamma}_h(t,t)dt)^{1/2}$ afterwards, with $\sum_h n_h = n$ and $\hat{\gamma}_h(t,t)$ the sample variance of the $X_k(t)$, $k \in s_h(t)$. The sample sizes n_h approximately minimize the MISE over $[0, \tau_1]$ and yield a reasonable estimator $\hat{\mu}_{ht}(t)$ for $t \geq \tau_1$ provided the strata variances $\gamma_h(t,t)$ do not vary excessively with respect to each other. For our rotation designs, we used the adaptive sample allocation described earlier. As Figure 5 shows, our rotation designs improve upon conventional rotation sampling by 3% to 8% for the MISE and by 25% to 45% for σ(ISE) across the range of α. We also tried using the customers' monthly consumption to improve sample allocation under the conventional design. (This auxiliary information is readily available since customers are billed monthly.) However, it did not increase the performances of conventional rotation.

Table 2 presents results for the composite estimator $\hat{\mu}_c(t)$, defined in reference to a previous value $\hat{\mu}_c(t-\delta)$. As δ increases, α should decrease and Q should increase in order to obtain the optimal MISE. In other words, if $\hat{\mu}_c(t)$ is defined with respect to a distant past $t-\delta$, then the sample should be closer to a fixed panel so that the change $\Delta\mu_N(t-\delta,t)$ is estimated more reliably. A large weight Q should also be placed on the current estimator $\hat{\mu}_{ht}(t)$ since for large δ, the estimation of $\Delta\mu_N(t-\delta,t)$ is not very accurate. In comparison to the HT estimator (see Figure 4), the composite estimator brings no significant improvement (3% at most for the MISE and 4% at most for σ(ISE) at any α value) although it uses more data. This is due to the overall weakness of the temporal dependence in electricity usage (see Figure 6). Since past data provide little information about current consumption, the composite estimator cannot greatly improve upon the HT estimator. In another numerical study with more strongly

Table 2. Composite estimation: optimal MISE and parameters α, Q in terms of δ.

δ	α_{opt}	Q_{opt}	min(MISE)
0.5h	0.5	0.3	3872
1h	0.5	0.4	3859
6h	0.5	0.8	3850
12h	0.4	0.8	3848
24h	0.2	1	3925

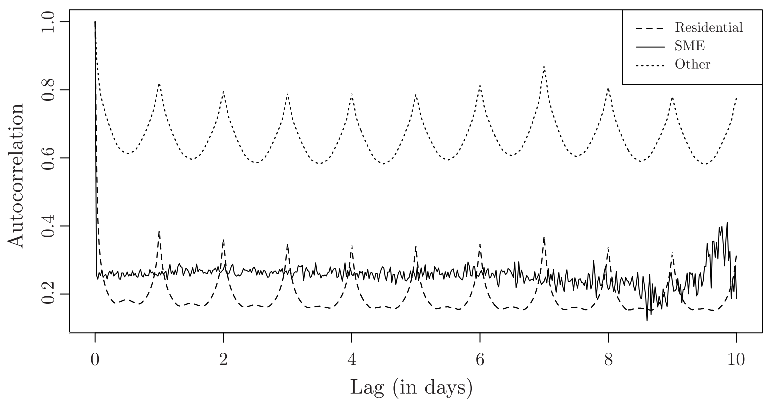

Figure 6. Temporal dependence in electricity usage. The displayed functions are the average autocorrelations $\Delta \mapsto \int_0^{T-\Delta} \rho_h(t, t+\Delta)dt$, where Δ is the time lag and ρ_h is the correlation function of stratum U_h.

correlated data (not reported here), the composite estimator clearly dominated the HT estimator.

The results of this study have been seen to hold qualitatively for a range of numbers of replacements m. The full replacement design produced excellent estimates of $\int_0^T \mu_N(t)dt$ (see Remark 2) with a relative error of 0.5%.

8. Discussion

We have devised rotation sampling designs for functional data. These survey designs are well suited to sensor network applications such as monitoring energy usage, internet traffic, and TV/radio audiences. Unlike conventional rotation designs that specify the sample before the survey, our methodology allows for adaptive sampling. Theoretical results and the numerical study with electricity usage show that our approach produces better survey estimates than fixed panels and conventional rotation samples. The proposed composite estimator enhances the Horvitz-Thompson estimator by integrating both past and current data. Although both estimators gave comparable results in the numerical study, the composite estimator is likely superior in the presence of stronger data correlation.

The present work can be extended in several directions. Our rotation designs can be modified to accommodate cluster-, multistage- or PPS sampling, which could improve upon stratified sampling and SRSWOR. In addition to adaptive sample allocation, other adaptive rules could increase estimation accuracy. For example, in order to maximize the information content of the sample, the replacement rates could be based on the balance between longitudinal and cross-sectional variations in recent measurements. A theoretical study of the composite estimator would facilitate statistical inference, help select the parameters δ and Q, and enable comparisons with the HT estimator. Incorporating auxiliary information would profitably expand our approach. To this end, a comparison of design-based and model-assisted methods is required to determine the most efficient integration scheme.

Acknowledgement

This research was supported in part by the Statistical and Applied Mathematical Sciences Institute (SAMSI). The author thanks Hervé Cardot (Université de Bourgogne), the Editors, and the referees for their helpful comments that led to substantial improvements in the manuscript.

Appendix

Proof of Proposition 2.

The sum under study can be decomposed as $\sum_{\ell=1}^{4} A_\ell(t,t')$, where

$$A_\ell(t,t') = \sum_{\substack{i,j,k,l \in U_h \\ \mathcal{C}_{ijkl}=\ell}} E\left(I_i(t)I_j(t)I_k(t')I_l(t')\right) \tilde{X}_i(t)\tilde{X}_j(t)\tilde{X}_k(t')\tilde{X}_l(t')$$

and $\mathcal{C}_{ijkl} = \#\{i,j,k,l\}$. To compute the A_ℓ, we derive $E\left(I_i(t)I_j(t)I_k(t')I_l(t')\right)$ based on the properties of SRSWOR and develop sums $\sum \tilde{X}_i(t)\tilde{X}_j(t)\tilde{X}_k(t')\tilde{X}_l(t')$ using the identity $\sum_{k\in U_h} \tilde{X}_k(t) = 0$. Let i^*, j^*, k^*, l^* be distinct units in U_h. To lighten the notation, we omit the subscript $k \in U_h$ in the sums to follow.

We begin with

$$A_1(t,t') = E\big(I_{i^*}(t)I_{i^*}(t')\big) \sum \tilde{X}_k^2(t)\tilde{X}_k^2(t'). \tag{A.1}$$

The term $A_2(t,t')$ can be expressed as

$$\begin{aligned} A_2(t,t') = {} & E\big(I_{i^*}(t)I_{k^*}(t')\big)\Big[(N_h-1)^2\gamma_h(t,t)\gamma_h(t',t') - \sum \tilde{X}_k^2(t)\tilde{X}_k^2(t')\Big] \\ & +2E\big(I_{i^*}(t)I_{i^*}(t')I_{k^*}(t)I_{k^*}(t')\big)\Big[(N_h-1)^2\gamma_h^2(t,t') - \sum \tilde{X}_k^2(t)\tilde{X}_k^2(t')\Big] \\ & -2\Big[E\big(I_{i^*}(t)I_{i^*}(t')I_{k^*}(t')\big)+E\big(I_{i^*}(t)I_{k^*}(t)I_{k^*}(t')\big)\Big] \sum \tilde{X}_k^2(t)\tilde{X}_k^2(t'). \end{aligned} \tag{A.2}$$

Next, it can be shown that

$$\begin{aligned}A_3(t,t') = &\Big[E\big(I_{i^*}(t)I_{j^*}(t)I_{k^*}(t')\big) + E\big(I_{i^*}(t)I_{k^*}(t')I_{l^*}(t')\big)\Big] \\ &\times\Big[-(N_h-1)^2\,\gamma_h(t,t)\gamma_h(t',t') + 2\sum \tilde{X}_k^2(t)\tilde{X}_k^2(t')\Big] \\ &+4\,E\big(I_{i^*}(t)I_{i^*}(t')I_{j^*}(t)I_{k^*}(t')\big) \\ &\times\Big[-(N_h-1)^2\,\gamma_h^2(t,t') + 2\sum \tilde{X}_k^2(t)\tilde{X}_k^2(t')\Big]. \qquad (A.3)\end{aligned}$$

To compute $A_4(t,t')$, use the decomposition

$$\sum_{i,j,k,l\in U_h} \tilde{X}_i(t)\tilde{X}_j(t)\tilde{X}_k(t')\tilde{X}_l(t') = \sum_{\ell=1}^{4} \sum_{\substack{i,j,k,l\in U_h \\ \mathcal{C}_{ijkl}=\ell}} \tilde{X}_i(t)\tilde{X}_j(t)\tilde{X}_k(t')\tilde{X}_l(t')$$

together with (A.1)-(A.3) to obtain

$$\begin{aligned}A_4(t,t') = &E\big(I_{i^*}(t)I_{j^*}(t)I_{k^*}(t')I_{l^*}(t')\big)\Big[(N_h-1)^2\,\gamma_h(t,t)\gamma_h(t',t') \\ &+2\,(N_h-1)^2\,\gamma_h^2(t,t') - 6\sum \tilde{X}_k^2(t)\tilde{X}_k^2(t')\Big]. \qquad (A.4)\end{aligned}$$

The proof is completed by gathering (A.1)−(A.4) and observing that all terms involving $\sum \tilde{X}_k^2(t)\tilde{X}_k^2(t')$ are of negligible order $\mathcal{O}(N_h)$ thanks to (A1).

Proof of Proposition 3. Let k,l be distinct units in a stratum U_h. The Markov chain $\{(I_k+I_l)(\tau_r),\, r=0,\ldots,m\}$, which counts how many units among k,l are present in the sample at the successive replacement times, has three possible states: 0, 1, and 2. For $1\le r\le m$, the transition probability matrix

$$\mathbf{P}_r = \big(P\left((I_k+I_l)(\tau_r)=j-1\big|(I_k+I_l)(\tau_{r-1})=i-1\right)\big)_{1\le i,j\le 3}$$

can be represented as

$$\mathbf{P}_r = \mathbf{P}_r^* + \mathbf{E}_r, \qquad (A.5)$$

where

$$\mathbf{P}_r^* = \begin{pmatrix} (1-\beta_r)^2 & 2\,(1-\beta_r)\,\beta_r & \beta_r^2 \\ \alpha_h\,(1-\beta_r) & \alpha_h\beta_r + (1-\alpha_h)\,(1-\beta_r) & (1-\alpha_h)\,\beta_r \\ \alpha_h^2 & 2\,(1-\alpha_h)\,\alpha_h & (1-\alpha_h)^2 \end{pmatrix},$$

$\beta_r = P\left(k\in s_h(\tau_r)|k\notin s_h(\tau_{r-1})\right) = \left(f_h(\tau_r)-(1-\alpha_h)f_h(\tau_{r-1})\right)/\left(1-f_h(\tau_{r-1})\right)$, and $\alpha_h = P\left(k\notin s_h(\tau_r)|k\in s_h(\tau_{r-1})\right)$. The matrix $\mathbf{E}_r$ is asymptotically negligible in comparison to $\mathbf{P}_r^*$. Indeed, (A5) guarantees that $\max_{r=1,\ldots,m}\|\mathbf{E}_r\| = \mathcal{O}\left(1/N_h\right)$

as $N \to \infty$, where $\|\cdot\|$ denotes an arbitrary matrix norm. For simplicity we use the spectral norm $\|\mathbf{A}\| = \sup_{\mathbf{x}\neq 0} (\mathbf{x}'\mathbf{A}'\mathbf{A}\mathbf{x}/\mathbf{x}'\mathbf{x})^{1/2}$.

The transition probability matrices $\mathbf{P}_r$ have unit spectral norm. Using the binomial formula, the triangle inequality, and $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$ for compatible matrices $\mathbf{A}$ and $\mathbf{B}$, it follows that

$$\begin{aligned}\left\| \prod_{r=\nu(t)+1}^{\nu(t')} \mathbf{P}_r - \prod_{r=\nu(t)+1}^{\nu(t')} \mathbf{P}_r^* \right\| &\leq \sum_{r=1}^{\nu(t')-\nu(t)-1} \binom{\nu(t')-\nu(t)-1}{r} \left(\max_{q=\nu(t)+1,\ldots,\nu(t')} \|\mathbf{E}_q\| \right)^r \\ &\leq \left(1 + \max_{1\leq r\leq m} \|\mathbf{E}_r\|\right)^m - 1 \\ &= \mathcal{O}\left(m \max_{1\leq r\leq m} \|\mathbf{E}_r\|\right) \qquad \text{(A.6)}\end{aligned}$$

uniformly in $0 \leq t \leq t' \leq T$. With (A5),

$$\prod_{r=\nu(t)+1}^{\nu(t')} \mathbf{P}_r = (1+o(1)) \prod_{r=\nu(t)+1}^{\nu(t')} \mathbf{P}_r^* . \qquad \text{(A.7)}$$

We now study the simpler product $\prod_{r=\nu(t)+1}^{\nu(t')} \mathbf{P}_r^*$. Without loss of generality, set $\nu(t) = 0$ and $\nu(t') = r$. Let $\mathbf{Q}_r = \prod_{l=1}^{r} \mathbf{P}_l^*$ and write $[\mathbf{A}]_{ij}$ for the (i,j)th coefficient of a matrix $\mathbf{A}$. It remains to compute $[\mathbf{Q}_r]_{13}$, $[\mathbf{Q}_r]_{23}$, and $[\mathbf{Q}_r]_{33}$.

Since $\mathbf{Q}_r$ is a transition probability matrix and $[\mathbf{P}_r^*]_{2j} = [\mathbf{P}_r^*]_{1j}^{1/2}\,[\mathbf{P}_r^*]_{3j}^{1/2}$ for $j = 1,3$, it can be shown by induction that $[\mathbf{Q}_r]_{11}^{1/2} + [\mathbf{Q}_r]_{13}^{1/2} = 1$ for $1 \leq r \leq m$. As a result, $[\mathbf{Q}_r]_{13}^{1/2} = (1-\beta_r-\alpha_h)\,[\mathbf{Q}_{r-1}]_{13}^{1/2} + \beta_r$ and

$$[\mathbf{Q}_r]_{13}^{1/2} - f_h(\tau_r) = \frac{1-\alpha_h-f_h(\tau_r)}{1-f_h(\tau_{r-1})} \left([\mathbf{Q}_{r-1}]_{13}^{1/2} - f_h(\tau_{r-1})\right). \qquad \text{(A.8)}$$

Since $[\mathbf{Q}_1]_{13}^{1/2} - f_h(\tau_1) = -f_h(\tau_0)(1-\alpha_h-f_h(\tau_1))/(1-f_h(\tau_0))$, iterating (A.8), we obtain the identity $[\mathbf{Q}_r]_{13}^{1/2} = f_h(\tau_r) - f_h(\tau_0)\,\lambda_h(\tau_0,\tau_r)$. Similarly, we show that $[\mathbf{Q}_r]_{33}^{1/2} = (1-f_h(\tau_0))\,\lambda_h(\tau_0,\tau_r) + f_h(\tau_r)$. The expressions of $[\mathbf{Q}_r]_{13}$ and $[\mathbf{Q}_r]_{33}$ can be checked by induction.

Finally we turn to $[\mathbf{Q}_r]_{23}$. The total probability formula yields

$$\begin{aligned}P\left(k,l\in s(\tau_r)\right) = &P\left(k,l\in s(\tau_r)\middle|k,l\in s(0)\right) P\left(k,l\in s(0)\right) \\ &+2\,P\left(k,l\in s(\tau_r)\middle|k\in s(0),\, l\notin s(0)\right) P\left(k\in s(0),\, l\notin s(0)\right) \\ &+P\left(k,l\in_h s_h(t')\middle|k,l\notin s(0)\right) P\left(k,l\notin s(0)\right). \qquad \text{(A.9)}\end{aligned}$$

Based on (A.7) we obtain

$$f_h(\tau_r)^2 \sim \left(f_h(\tau_0)^2\,[\mathbf{Q}_r]_{11} + 2\,f_h(\tau_0)\,(1-f_h(\tau_0))\,[\mathbf{Q}_r]_{21} + (1-f_h(\tau_0))^2\,[\mathbf{Q}_r]_{31}\right). \qquad \text{(A.10)}$$

The proof is completed by expressing $[\mathbf{Q}_r]_{13}$ and $[\mathbf{Q}_r]_{33}$ in (A.10).

Proof of Theorem 3. We start by finding the asymptotic expressions for $C_1(t,t')$ and $C_2(t,t')$ in Proposition 2. In view of (A3), the properties of SRSWOR and the independence of $s_h(\tau_r)$, $1 \le r \le m$, under full replacement,

$$\begin{cases} C_1(t,t') \sim f_h(t)f_h(t')\,(1-f_h(t))\,(1-f_h(t')) \\ C_2(t,t') \sim 2\,f_h(t)f_h(t')\,(1-f_h(t))\,(1-f_h(t'))\,\delta_{\nu(t)\nu(t')} \end{cases}$$

uniformly in $t,t' \in [0,T]$ as $N \to \infty$. Hence, the last term in (5.2) cancels out with the term in $C_1(t,t')$ of Proposition 2.

Writing $\mathrm{Var}(\mathrm{ISE}) = (2/N^2)\iint_{[0,T]^2} \phi_N(t,t')dtdt'$, we deduce that

$$\phi_N(t,t') \sim \Big(\sum_{h=1}^{H} \frac{N_h}{N}\,\frac{1-f_h(t)}{f_h(t')}\,\delta_{\nu(t)\nu(t')}\gamma_h(t,t') \Big)^2 \tag{A.11}$$

uniformly in $t,t' \in [0,T]$ as $N \to \infty$. Using (A2)-(A4), the Mean Value Theorem, integral approximations, and a change of variable, we obtain

$$\begin{aligned}
&\mathrm{Var}(\mathrm{ISE}) \\
&\sim \frac{2}{N^2}\sum_{h,h'} \frac{N_h N_{h'}}{N^2} \sum_{r=1}^{m+1} \frac{(1-f_h(\tau_r))\,(1-f_{h'}(\tau_r))}{f_h(\tau_r)f_{h'}(\tau_r)} \iint_{[\tau_{r-1},\tau_r]^2} \gamma_h(t,t')\gamma_{h'}(t,t')dtdt' \\
&\sim \frac{2}{N^2}\sum_{h,h'} \frac{N_h N_{h'}}{N^2} \sum_{r=1}^{m+1} \frac{(1-f_h(\tau_r))\,(1-f_{h'}(\tau_r))}{f_h(\tau_r)f_{h'}(\tau_r)} \left(\tau_r-\tau_{r-1}\right)^2 \gamma_h(\tau_r,\tau_r)\gamma_{h'}(\tau_r,\tau_r) \\
&\sim \frac{2}{N^2}\sum_{h,h'} \frac{N_h N_{h'}}{N^2} \sum_{r=1}^{m+1} \frac{(1-f_h(\tau_r))\,(1-f_{h'}(\tau_r))}{f_h(\tau_r)f_{h'}(\tau_r)} \frac{1}{m^2\,g(\tau_r)^2}\,\gamma_h(\tau_r,\tau_r)\gamma_{h'}(\tau_r,\tau_r) \\
&\sim \frac{2}{mN^2}\sum_{h,h'} \frac{N_h N_{h'}}{N^2} \int_0^T \frac{(1-f_h(t))\,(1-f_{h'}(t))}{f_h(t)f_{h'}(t)}\,\frac{1}{g(t)}\,\gamma_h(t,t)\gamma_{h'}(t,t)dt\,.
\end{aligned}$$

Proof of Theorem 4. This result is established along the lines of Theorem 3. In view of Lemmas 1−2 and Proposition 3, it holds for all $0 \le t \le t' \le T$ that, as $N \to \infty$,

$$\begin{aligned}
E\left(I_{i^*}(t)I_{k^*}(t')\right) &= P\left(k^* \in s(t') \middle| i^*, k^* \in s(t)\right) P\left(i^*, k^* \in s(t)\right) \\
&\quad + P\left(k^* \in s(t') \middle| i^* \in s(t), k^* \notin s(t)\right) P\left(i^* \in s(t),\, k^* \notin s(t)\right) \\
&= P\left(k^* \in s(t') \middle| k^* \in s(t)\right) P\left(i^*, k^* \in s(t)\right) \\
&\quad + P\left(k^* \in s(t') \middle| k^* \notin s(t)\right) P\left(i^* \in s(t),\, k^* \notin s(t)\right) \\
&\sim \left[\,(1-f_h(t))\,\lambda_h(t,t') + f_h(t')\,\right] f_h^2(t) \\
&\quad + \left[\,f_h(t') - f_h(t)\lambda_h(t,t')\,\right] f_h(t)\,(1-f_h(t)) \\
&= f_h(t)f_h(t').
\end{aligned}$$

By symmetry this equality holds for all $t, t' \in [0,T]$. Similarly, we find that

$$\begin{cases} E\left(I_{i^*}(t)I_{j^*}(t)I_{k^*}(t')\right) & \sim f_h^2(t)f_h(t'), \\ E\left(I_{i^*}(t)I_{k^*}(t')I_{l^*}(t')\right) & \sim f_h(t)f_h^2(t'), \\ E\left(I_{i^*}(t)I_{j^*}(t)I_{k^*}(t')I_{l^*}(t')\right) & \sim f_h^2(t)f_h^2(t'), \\ E\left(I_{i^*}(t)I_{i^*}(t')I_{k^*}(t)I_{k^*}(t')\right) & \sim \left[\,(1-f_h(t))\lambda_h(t,t') + f_h(t')\,\right]^2 f_h^2(t), \\ E\left(I_{i^*}(t)I_{i^*}(t')I_{j^*}(t)I_{k^*}(t')\right) & \sim \left[\,(1-f_h(t))\,\lambda_h(t,t') + f_h(t')\,\right] f_h^2(t)f_h(t'). \end{cases}$$

Therefore

$$\begin{cases} C_1(t,t') \sim f_h(t)f_h(t')\,(1-f_h(t))\,(1-f_h(t')) \\ C_2(t,t') \sim 2\,f_h^2(t)\,(1-f_h(t))^2\,\lambda_h^2(t,t') \end{cases}.$$

Combining this result and Proposition 2, it follows from (5.2) that

$$\sum_{i,j,k,l\in U_h} \Delta_{ijkl}(t,t')\,X_i(t)X_j(t)X_k(t')X_l(t') \sim 2\,f_h^2(t)\,(1-f_h(t))^2\,\lambda_h^2(t,t')\gamma_h^2(t,t')N_h^2\,. \tag{A.12}$$

Based on Theorem 2, the last term in (5.1) is

$$\sum_{i,k\in U_h} \frac{\Delta_{ik}(t,t')}{f_h(t)f_h(t')}\,X_i(t)X_k(t') = N_h\,\frac{1-f_h(t)}{f_h(t')}\,\gamma_h(t,t')\,\lambda_h(t,t'). \tag{A.13}$$

Writing $\mathrm{Var(ISE)} = (2/N^2)\iint_{[0,T]^2}\phi_N(t,t')dtdt'$, (A.12)−(A.13) imply that

$$\phi_N(t,t') \sim \Big(\sum_{h=1}^{H}\frac{N_h}{N}\,\frac{1-f_h(t)}{f_h(t')}\,\gamma_h(t,t')\,\lambda_h(t,t')\Big)^2 \tag{A.14}$$

for all $t, t' \in [0,T]$ as $N \to \infty$. To apply the Dominated Convergence Theorem, it suffices to check that the ϕ_N, $N \geq 1$, are uniformly bounded on $[0,T]^2$. The right-hand side of (A.14) has a finite number of terms. In view of (A3)−(A4), $(1-f_h(t))/f_h(t')$ and $\gamma_h(t,t')$ are uniformly bounded with respect to $t, t' \in [0,T]$ and N. Finally, $|\lambda_h(t,t')| \leq 1$ as a product of eigenvalues of transition probability matrices, which concludes the proof.

Proof of Corollary 1. It suffices to show that $\lambda_h(t,t') \sim \exp\left(-c_h\,|G(t)-G(t')|\right)$ as $N \to \infty$. Condition (i) yields $\lambda_h(t,t') = (1-\alpha_h/(1-f_h))^{|\nu(t)-\nu(t')|}$ and $|\nu(t)-\nu(t')| = (m+1)\,\left|G(\tau_{\nu(t)}) - G(\tau_{\nu(t')})\right| \sim m\,|G(t)-G(t')|$ by (A2). Condition (ii) and the approximation $\ln(1+x) \sim x$ as $x \to 0$ provide the conclusion.

References

Cardot, H., Chaouch, M., Goga, C. and Labruère, C. (2010). Properties of design-based functional principal components analysis. *J. Statist. Plann. Inference* **140**, 75-91.

Cardot, H., Degras, D. and Josserand, E. (2012). Confidence bands for Horvitz-Thompson estimators using sampled noisy functional data. *Bernoulli.* Forthcoming, also available at `http://arxiv.org/abs/1105.2135`.

Cardot, H., Goga, C. and Lardin, P. (2013). Uniform convergence and asymptotic confidence bands for model-assisted estimators of the mean of sampled functional data. *Electron. J. Statist.* **7**, 562-596.

Cardot, H. and Josserand, E. (2011). Horvitz-Thompson estimators for functional data: asymptotic confidence bands and optimal allocation for stratified sampling. *Biometrika* **98**(2011), 107-118.

Chiky, R., Cubillé, J., Dessertaine, A., Hébrail, G. and Picard, M.-L. (2008). Échantillonnage spatio-temporel de flux de données distribués. In *EGC'08* (Edited by F. Guillet and B. Trousse, B.), 169-180.

Commission for Energy Regulation (2011). Smart metering information paper 4 (CER 11/080). Dublin, Ireland. `http://www.cer.ie`.

Degras, D. (2011). Simultaneous confidence bands for nonparametric regression with functional data. *Statist. Sinica* **21**, 1735-1765.

Eckler, A. R. (1955). Rotation sampling. *Ann. Math. Statist.* **26**, 664-685.

Fuller, W. A. (2009). *Sampling Statistics.* Wiley, Hoboken, N.J.

Horvitz, D. G. and Thompson, D. J. (1952). A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.* **47**, 663-685.

Lavallée, P. (1995). Cross-sectional weighting of longitudinal surveys of individuals and households using the weight share method. *Surv. Methodol.* **21**, 25-32.

Rao, J. N. K. and Graham, J. E. (1964). Rotation designs for sampling on repeated occasions. *J. Amer. Statist. Assoc.* **59**, 492-509.

Thompson, S. K.; Seber, G. A. (1996). *Adaptive Sampling.* Wiley, New York.

Wolter, K. M. (1979). Composite estimation in finite populations. *J. Amer. Statist. Assoc.* **74**, 604-613.

Mathematical Sciences, DePaul University, Chicago, IL 60604, USA.

E-mail: ddegrasv@depaul.edu

(Received July 2012; accepted July 2013)