Web Appendix for "Generalizing the Information Content for Stepped Wedge Designs: A Marginal Modeling Approach"

Fan Li*, Jessica Kasza, Elizabeth L. Turner, Paul J. Rathouz, Andrew B. Forbes, John S. Preisser

*fan.f.li@yale.edu

Web Appendix A A Class of Block Correlation Structures for Stepped Wedge Trials

We define a class of general block correlation structures considered in Section 2.1 of the main text. We characterize such a correlation structure based on individual-level outcomes within cluster i. For clarity, we keep the subscript i to indicate the results apply to any cluster i and fully acknowledge that the correlation parameters can even differ between clusters i and i' (such as the structure implied from a random treatment effect model in Hughes et al. (2015) and Hemming et al. (2018)). Assuming the cluster-period size (N) is constant within each cluster i with no period-to-period variation, we define r_{ijj} as the jth within-period correlation, while allowing arbitrary values of the between-period correlations. Specifically, the correlation between two responses measured for two different individuals in the jth and tth periods is r_{ijt} and $r_{ijt} = r_{itj}$. The correlation between two repeated responses measured for the same individual (applicable to a closed-cohort design) in the jth and tth periods is r^*_{ijt} and $r^*_{ijt} = r^*_{itj}$ ($r^*_{ijj} = 1$ by definition); note r^*_{ijt} is not necessarily equal to r_{ijt}. In matrix form, this general block correlation structure for $\mathbf{Y}_i = (\mathbf{Y}'_{i1}, \ldots, \mathbf{Y}'_{iT})'$ can be represented as

$$\begin{bmatrix} (1-r_{i11})\mathbf{I}_N + r_{i11}\mathbf{J}_N & (r^*_{i12}-r_{i12})\mathbf{I}_N + r_{i12}\mathbf{J}_N & \ldots & (r^*_{i1T}-r_{i1T})\mathbf{I}_N + r_{i1T}\mathbf{J}_N \\ (r^*_{i12}-r_{i12})\mathbf{I}_N + r_{i12}\mathbf{J}_N & (1-r_{i22})\mathbf{I}_N + r_{i22}\mathbf{J}_N & \ldots & (r^*_{i2T}-r_{i2T})\mathbf{I}_N + r_{i2T}\mathbf{J}_N \\ (r^*_{i13}-r_{i13})\mathbf{I}_N + r_{i13}\mathbf{J}_N & (r^*_{i23}-r_{i23})\mathbf{I}_N + r_{i23}\mathbf{J}_N & \ldots & (r^*_{i3J}-r_{i3J})\mathbf{I}_N + r_{i3J}\mathbf{J}_N \\ \vdots & \vdots & \ddots & \vdots \\ (r^*_{i1T}-r_{i1T})\mathbf{I}_N + r_{i1T}\mathbf{J}_N & (r^*_{i2T}-r_{i2T})\mathbf{I}_N + r_{i2T}\mathbf{J}_N & \ldots & (1-r_{iTT})\mathbf{I}_N + r_{iTT}\mathbf{J}_N \end{bmatrix}. \tag{1}$$

The above class of general block correlation structure at least includes the following special cases studied in the literature on stepped wedge designs:

- The *simple exchangeable* correlation structure (compound symmetric matrix) is obtained with $r_{ijt} = \alpha_0$ for all $j, t = 1, \ldots, T$, and $r^*_{ijt} = \alpha_0$ for all $|j-t| \geq 1$. This is the correlation matrix implied by the linear mixed model in Hussey and Hughes (2007), and has been used in cross-sectional designs;
- The *nested exchangeable* correlation matrix is obtained with $r_{ijj} = \alpha_0$ for all $j = 1, \ldots, T$, and $r_{ijt} = r^*_{ijt} = \alpha_1$ for all $|j-t| \geq 1$. This is the correlation matrix implied by one version of the Hooper/Girling linear mixed model (Hooper et al., 2016; Girling and Hemming, 2016), and is applicable to cross-sectional designs;

- The *exponential decay* correlation matrix is obtained with $r_{ijj} = \alpha_0$ for all $j = 1, \ldots, T$ and $r_{ijt} = r^*_{ijt} = \alpha_0 \rho^{|j-t|}$ for all $|j - t| \geq 1$. This is the correlation matrix implied by the linear mixed model considered in Kasza et al. (2019), and is applicable to cross-sectional designs;
- The *block exchangeable* correlation matrix is obtained with $r_{ijj} = \alpha_0$ for all $j = 1, \ldots, T$, $r_{ijt} = \alpha_1$ for all $|j - t| \geq 1$ and $r^*_{ijt} = \alpha_2$ for all $|j - t| \geq 1$. This is the correlation matrix studied in Li et al. (2018) and implied by one version of the Hooper/Girling linear mixed model (Hooper et al., 2016; Girling and Hemming, 2016) applicable to closed-cohort designs;
- A more *general proportional decay* correlation matrix is obtained with $r_{ijj} = \alpha_0$ for all $j = 1, \ldots, T$, $r_{ijt} = \alpha_0 \rho^{|j-t|}$ for all $|j - t| \geq 1$, and $r^*_{ijl} = \tau^{|j-t|}$ for all $|j - t| \geq 1$. This is a generalization of the *proportional decay* correlation matrix developed in Li (2020) and Liu et al. (2002) by allowing for different decay rate for the between-individual correlation and within-individual correlation over time;
- Other examples include the correlation structure implied from the random treatment effect model in Hemming et al. (2018) and random coefficient model in Li et al. (2020). A comprehensive set of correlation models previously considered for cross-sectional and closed-cohort stepped wedge design literature were reviewed in Section 3.4 and 3.5 of Li et al. (2020), and are all special cases of the above general block correlation matrix (1).

Web Appendix B Proof of Lemma 1

To show Lemma 1, it is sufficient to show that for each cluster i,

$$N^{-2}\mathbf{E}'\mathbf{V}_i^{-1}\mathbf{E} = (\mathbf{E}'\mathbf{V}_i\mathbf{E})^{-1}. \tag{2}$$

To see why this identity holds, recall that covariance matrix of $\mathbf{Y}_i^*$ is of the form $\mathbf{V}_i = (\mathbf{Q}_{i1} - \mathbf{Q}_{i2}) \otimes \mathbf{I}_N + \mathbf{Q}_{i2} \otimes \mathbf{J}_N$, whose inverse can be written as (Leiva, 2007)

$$\mathbf{V}_i^{-1} = (\mathbf{Q}_{i1} - \mathbf{Q}_{i2})^{-1} \otimes \mathbf{I}_N + \frac{1}{N}\left\{(\mathbf{Q}_{i1} + (N-1)\mathbf{Q}_{i2})^{-1} - (\mathbf{Q}_{i1} - \mathbf{Q}_{i2})^{-1}\right\} \otimes \mathbf{J}_N.$$

Therefore, direct calculation shows

$$\begin{aligned}\mathbf{E}'\mathbf{V}_i^{-1}\mathbf{E} =& N(\mathbf{Q}_{i1} - \mathbf{Q}_{i2})^{-1} + N\left\{(\mathbf{Q}_{i1} + (N-1)\mathbf{Q}_{i2})^{-1} - (\mathbf{Q}_{i1} - \mathbf{Q}_{i2})^{-1}\right\}, \\ =& N(\mathbf{Q}_{i1} + (N-1)\mathbf{Q}_{i2})^{-1}.\end{aligned}$$

On the other hand, we notice that $\mathbf{E}'\mathbf{V}_i\mathbf{E} = N(\mathbf{Q}_{i1} + (N-1)\mathbf{Q}_{i2})$, and therefore (2) holds. By definition, $\overline{\mathbf{V}}_i = N^{-2}\mathbf{E}'\mathbf{V}_i\mathbf{E}$, and by (2), we have $\overline{\mathbf{V}}_i^{-1} = \mathbf{E}'\mathbf{V}_i^{-1}\mathbf{E}$. In addition, we can show

$$\begin{aligned}\mathbf{H}_i'\mathbf{V}_i^{-1}\mathbf{H}_i =& N\mathbf{X}_i'(\mathbf{Q}_{i1} - \mathbf{Q}_{i2})^{-1}\mathbf{X}_i + N\mathbf{X}_i'\left\{(\mathbf{Q}_{i1} + (N-1)\mathbf{Q}_{i2})^{-1} - (\mathbf{Q}_{i1} - \mathbf{Q}_{i2})^{-1}\right\}\mathbf{X}_i, \\ =& N\mathbf{X}_i'(\mathbf{Q}_{i1} + (N-1)\mathbf{Q}_{i2})^{-1}\mathbf{X}_i = \mathbf{X}_i'\overline{\mathbf{V}}_i^{-1}\mathbf{X}_i.\end{aligned}$$

Similar arguments leads to $\mathbf{H}_i'\mathbf{V}_i^{-1}\mathbf{E} = \mathbf{X}_i'\overline{\mathbf{V}}_i^{-1}$. Therefore,

$$var(\widehat{\delta}) = \left\{ \sum_{i=1}^{I} \mathbf{H}_i'\mathbf{V}_i^{-1}\mathbf{H}_i - \left(\sum_{i=1}^{I} \mathbf{H}_i'\mathbf{V}_i^{-1} \right) \mathbf{E} \left(\sum_{i=1}^{I} \mathbf{E}'\mathbf{V}_i^{-1}\mathbf{E} \right)^{-1} \mathbf{E}' \left(\sum_{i=1}^{I} \mathbf{V}_i^{-1}\mathbf{H}_i \right) \right\}^{-1},$$

$$= \left\{ \sum_{i=1}^{I} \mathbf{X}_i'\overline{\mathbf{V}}_i^{-1}\mathbf{X}_i - \left(\sum_{i=1}^{I} \mathbf{X}_i'\overline{\mathbf{V}}_i^{-1} \right) \left(\sum_{i=1}^{I} \overline{\mathbf{V}}_i^{-1} \right)^{-1} \left(\sum_{i=1}^{I} \overline{\mathbf{V}}_i^{-1}\mathbf{X}_i \right) \right\}^{-1},$$

and Lemma 1 holds.

Web Appendix C Proof of Theorem 1 and the Remark

Proof of Theorem 1 We consider the case when one cluster-period, or one cell of the design, is omitted. Define $var(\widehat{\delta})_{[ij]}$ as the variance of treatment effect estimator when cell (i, j) (cluster i period j) is omitted, $\overline{\mathbf{V}}_{[ij]}$ as the $(IT-1) \times (IT-1)$ variance matrix of the *working outcome* means $\overline{\mathbf{Y}}^*$ with the element from cell (i, j) omitted. We further define $\mathbf{Z}_i = (\mathbf{I}_T, \mathbf{X}_i)$ as the $T \times (T+1)$ cluster-period level design matrix for cluster i, and $\mathbf{Z} = (\mathbf{Z}_1', \ldots, \mathbf{Z}_I')'$ as the joint design matrix. Finally, we define $\mathbf{Z}_{[ij]}$ as the matrix $\mathbf{Z}$ with the row corresponding to cluster i and period j omitted.

We rearrange the $IT \times 1$ cell means $\overline{\mathbf{Y}}^* = (\overline{\mathbf{Y}}_1^{*\prime}, \ldots, \overline{\mathbf{Y}}_I^{*\prime})'$ such that the summary statistic corresponding to the omitted cell is the first element, followed by the remaining observations (according to time periods) from the same cluster, and then the observations from other clusters. We write the $IT \times IT$ covariance matrix of the rearranged means as

$$\overline{\mathbf{V}} = \begin{pmatrix} \overline{\nu}_{ij} & \overline{\boldsymbol{\nu}}_{ij}' \\ \overline{\boldsymbol{\nu}}_{ij} & \overline{\mathbf{V}}_{[ij]} \end{pmatrix},$$

where $\overline{\nu}_{ij} = var(\overline{Y}_{ij}^*)$ is the variance of the jth element of $\overline{\mathbf{Y}}_i^*$, and $\overline{\boldsymbol{\nu}}_{ij}$ is the $IT-1$ column vector with (k,t)th element given by $cov(\overline{Y}_{ij}^*, \overline{Y}_{kt}^*)$, with $(k,t) \neq (i,j)$. Due to independence of observations between clusters, we can write $\overline{\boldsymbol{\nu}}_{ij} = (\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)\prime}, \mathbf{0}_{(I-1)T}')'$, where $\mathbf{0}_s$ is the s-vector of zeros. Furthermore, $\overline{\mathbf{V}}_{[ij]}$ is a $(IT-1) \times (IT-1)$ block diagonal structure, with the first block given by the $(T-1) \times (T-1)$ matrix $\overline{\mathbf{V}}_{i[j]}$, the covariance matrix of working outcome means in cluster i with the jth row and column removed, and the remaining blocks given by $cov(\overline{\mathbf{Y}}_k^*)$ for $k \neq i$. We further define $\mathbf{e}_j$ as the T-vector with its jth element as 1 and 0 elsewhere.

Using Lemma 2 from Preisser and Qaqish (1996), we have

$$\begin{aligned}
(\mathbf{Z}_{[ij]}'\overline{\mathbf{V}}_{[ij]}^{-1}\mathbf{Z}_{[ij]})^{-1} =& (\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1} + \frac{1}{\omega_{ij} - \widetilde{q}_{ij}} (\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\widetilde{\mathbf{Z}}_{ij}\widetilde{\mathbf{Z}}_{ij}'(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1} \\
\widetilde{\mathbf{Z}}_{ij} =& \mathbf{Z}_{ij} - \mathbf{Z}_{[ij]}'\overline{\mathbf{V}}_{[ij]}^{-1}\overline{\boldsymbol{\nu}}_{ij} \\
\omega_{ij} =& \overline{\nu}_{ij} - \overline{\boldsymbol{\nu}}_{ij}'\overline{\mathbf{V}}_{[ij]}^{-1}\overline{\boldsymbol{\nu}}_{ij} \\
\widetilde{q}_{ij} =& \widetilde{\mathbf{Z}}_{ij}'(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\widetilde{\mathbf{Z}}_{ij},
\end{aligned}$$

where $\mathbf{Z}_{ij}$ is the row of $\mathbf{Z}$ corresponding to the omitted cell (i, j) (which after the aforementioned re-arrangement is the first row of $\mathbf{Z}$). In the absence of cell omission, we have

$$\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z} = \begin{pmatrix} \sum_{k=1}^{I} \overline{\mathbf{V}}_k^{-1} & \sum_{k=1}^{I} \overline{\mathbf{V}}_k^{-1}\mathbf{X}_k \\ \sum_{k=1}^{I} \mathbf{X}_k'\overline{\mathbf{V}}_k^{-1} & \sum_{k=1}^{I} \mathbf{X}_k'\overline{\mathbf{V}}_k^{-1}\mathbf{X}_k \end{pmatrix} \triangleq \begin{pmatrix} \mathbf{C} & \mathbf{d} \\ \mathbf{d}' & f \end{pmatrix}.$$

It is clear that $var(\widehat{\delta}) = (f - \mathbf{d}'\mathbf{C}^{-1}\mathbf{d})^{-1}$, therefore by block matrix inversion, we have

$$(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1} = \begin{pmatrix} \mathbf{C}^{-1} + var(\widehat{\delta})\mathbf{C}^{-1}\mathbf{d}\mathbf{d}'\mathbf{C}^{-1} & -var(\widehat{\delta})\mathbf{C}^{-1}\mathbf{d} \\ -var(\widehat{\delta})\mathbf{d}'\mathbf{C}^{-1} & var(\widehat{\delta}) \end{pmatrix}.$$

Recall that the design vector can be written as $\mathbf{Z}_{ij} = (\mathbf{e}_j', X_{ij})'$, where $\mathbf{e}_j$ is the $T \times 1$ unit vector with one at the jth position and zero otherwise, and X_{ij} is the treatment indicator corresponding to the omitted cell. We can similarly define $\widetilde{\mathbf{Z}}_{ij} = (\widetilde{\mathbf{e}}_j', \widetilde{X}_{ij})'$ whose explicit expression will become clear in due course. Observe that

$$(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\widetilde{\mathbf{Z}}_{ij} = \begin{pmatrix} \mathbf{C}^{-1}\widetilde{\mathbf{e}}_j + var(\widehat{\delta})\mathbf{C}^{-1}\mathbf{d}\mathbf{d}'\mathbf{C}^{-1}\widetilde{\mathbf{e}}_j - var(\widehat{\delta})\widetilde{X}_{ij}\mathbf{C}^{-1}\mathbf{d} \\ -var(\widehat{\delta})\mathbf{d}'\mathbf{C}^{-1}\widetilde{\mathbf{e}}_j + var(\widehat{\delta})\widetilde{X}_{ij} \end{pmatrix},$$

which then implies that the $(T+1, T+1)$th entry of $(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\widetilde{\mathbf{Z}}_{ij}\widetilde{\mathbf{Z}}_{ij}'(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}$ is $var(\widehat{\delta})^2(\widetilde{X}_{ij} - \mathbf{d}'\mathbf{C}^{-1}\widetilde{\mathbf{e}}_j)^2$. Similarly, we obtain

$$\widetilde{q}_{ij} = \widetilde{\mathbf{Z}}_{ij}'(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\widetilde{\mathbf{Z}}_{ij} = \widetilde{\mathbf{e}}_j'\mathbf{C}^{-1}\widetilde{\mathbf{e}}_j + var(\widehat{\delta})^2(\widetilde{X}_{ij} - \mathbf{d}'\mathbf{C}^{-1}\widetilde{\mathbf{e}}_j)^2.$$

Now to see the explicit expressions of $\widetilde{\mathbf{e}}_j$ and $\widetilde{X}_{ij}$, we first recall that $\overline{\mathbf{V}}_{[ij]} = blockdiag(\overline{\mathbf{V}}_{i[j]}, \overline{\mathbf{V}}_2, \ldots)$, and $\overline{\boldsymbol{\nu}}_{ij} = (\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)}, \mathbf{0}_{(I-1)T}')'$. Then we have

$$\mathbf{Z}_{[ij]}'\overline{\mathbf{V}}_{[ij]}^{-1}\overline{\boldsymbol{\nu}}_{ij} = \begin{pmatrix} \mathbf{I}_{[j]}'\overline{\mathbf{V}}_{i[j]}^{-1} & \overline{\mathbf{V}}_2^{-1} & \cdots \\ \mathbf{X}_{i[j]}'\overline{\mathbf{V}}_{i[j]}^{-1} & \mathbf{X}_2'\overline{\mathbf{V}}_2^{-1} & \cdots \end{pmatrix} \begin{pmatrix} \overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)} \\ \mathbf{0}_{(I-1)T} \end{pmatrix} =, \begin{pmatrix} \mathbf{I}_{[j]}'\overline{\mathbf{V}}_{i[j]}^{-1} \\ \mathbf{X}_{i[j]}'\overline{\mathbf{V}}_{i[j]}^{-1} \end{pmatrix} \overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)},$$

which directly implies that $\widetilde{\mathbf{e}}_j = \mathbf{e}_j - \mathbf{I}_{[j]}'\overline{\mathbf{V}}_{i[j]}^{-1}\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)}$ and $\widetilde{X}_{ij} = X_{ij} - \mathbf{X}_{i[j]}'\overline{\mathbf{V}}_{i[j]}^{-1}\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)}$, where $\mathbf{I}_{[j]}$ is the identity matrix with the jth row omitted, and $\mathbf{X}_{i[j]}$ is the treatment vector $\mathbf{X}_i$ with the jth row omitted. These algebraic expressions than lead to

$$var(\widehat{\delta})_{[ij]} = var(\widehat{\delta}) + \frac{var(\widehat{\delta})^2}{\omega_{ij} - \widetilde{q}_{ij}}(\widetilde{X}_{ij} - \mathbf{d}'\mathbf{C}^{-1}\widetilde{\mathbf{e}}_j)^2.$$

To further simplify the above expression, we notice that $\overline{\nu}_{ij} = \mathbf{e}_j'\overline{\mathbf{V}}_i\mathbf{e}_j$, $\overline{\mathbf{V}}_{i[j]} = \mathbf{I}_{[j]}\overline{\mathbf{V}}_i\mathbf{I}_{[j]}'$ and $\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)} = \mathbf{I}_{[j]}\overline{\mathbf{V}}_i\mathbf{e}_j$, we can write

$$\begin{aligned}\omega_{ij} =& \overline{\nu}_{ij} - \overline{\boldsymbol{\nu}}_{ij}'\overline{\mathbf{V}}_{[ij]}^{-1}\overline{\boldsymbol{\nu}}_{ij} = \overline{\nu}_{ij} - \overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)\prime}\overline{\mathbf{V}}_{i[j]}^{-1}\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)} = \mathbf{e}_j'\overline{\mathbf{V}}_i(\mathbf{e}_j - \mathbf{I}_{[j]}'\overline{\mathbf{V}}_{i[j]}^{-1}\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)}) \\ =& (\mathbf{e}_j - \mathbf{I}_{[j]}'\overline{\mathbf{V}}_{i[j]}^{-1}\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)})'\overline{\mathbf{V}}_i(\mathbf{e}_j - \mathbf{I}_{[j]}'\overline{\mathbf{V}}_{i[j]}^{-1}\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)}) + (\mathbf{I}_{[j]}'\overline{\mathbf{V}}_{i[j]}^{-1}\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)})'\overline{\mathbf{V}}_i(\mathbf{e}_j - \mathbf{I}_{[j]}'\overline{\mathbf{V}}_{i[j]}^{-1}\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)}) \\ =& \widetilde{\mathbf{e}}_j'\overline{\mathbf{V}}_i\widetilde{\mathbf{e}}_j + \overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)\prime}\overline{\mathbf{V}}_{i[j]}^{-1}\mathbf{I}_{[j]}\overline{\mathbf{V}}_i\mathbf{e}_j - \overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)\prime}\overline{\mathbf{V}}_{i[j]}^{-1}\mathbf{I}_{[j]}\overline{\mathbf{V}}_i\mathbf{I}_{[j]}'\overline{\mathbf{V}}_{i[j]}^{-1}\overline{\boldsymbol{\nu}}_{ij}^{(1:T-1)} = \widetilde{\mathbf{e}}_j'\overline{\mathbf{V}}_i\widetilde{\mathbf{e}}_j.\end{aligned}$$

Therefore, we obtain the formula for the information content as

$$\begin{aligned}IC(i,j) = \frac{var(\widehat{\delta})_{[ij]}}{var(\widehat{\delta})} =& \frac{\widetilde{\mathbf{e}}_j'\overline{\mathbf{V}}_i\widetilde{\mathbf{e}}_j - \widetilde{\mathbf{e}}_j'\mathbf{C}^{-1}\widetilde{\mathbf{e}}_j}{\widetilde{\mathbf{e}}_j'\overline{\mathbf{V}}_i\widetilde{\mathbf{e}}_j - \widetilde{\mathbf{e}}_j'\mathbf{C}^{-1}\widetilde{\mathbf{e}}_j - var(\widehat{\delta})(\widetilde{X}_{ij} - \mathbf{d}'\mathbf{C}^{-1}\widetilde{\mathbf{e}}_j)^2} \\ =& \frac{\widetilde{\mathbf{e}}_j'\left\{\overline{\mathbf{V}}_i - \left(\sum_{k=1}^I \overline{\mathbf{V}}_k^{-1}\right)^{-1}\right\}\widetilde{\mathbf{e}}_j}{\widetilde{\mathbf{e}}_j'\left\{\overline{\mathbf{V}}_i - \left(\sum_{k=1}^I \overline{\mathbf{V}}_k^{-1}\right)^{-1}\right\}\widetilde{\mathbf{e}}_j - var(\widehat{\delta})\left\{\widetilde{X}_{ij} - \left(\sum_{k=1}^I \mathbf{X}_k'\overline{\mathbf{V}}_k^{-1}\right)\left(\sum_{k=1}^I \overline{\mathbf{V}}_k^{-1}\right)^{-1}\widetilde{\mathbf{e}}_j\right\}^2}.\end{aligned}$$

Proof of Remark We wish to prove that $IC(i,j) \geq 1$ for all $i = 1, \dots, I$ and $j = 1, \dots, T$. To see why this holds, we write $\mathbf{C}_{-i} = \sum_{k \neq i}^{I} \overline{\mathbf{V}}_k^{-1}$, then $\left(\sum_{k=1}^{I} \overline{\mathbf{V}}_k^{-1}\right)^{-1} = \left(\overline{\mathbf{V}}_i^{-1} + \mathbf{C}_{-i}\right)^{-1}$. Since the correlation matrix $\mathbf{R}_i$ must be positive definite, the variance $\overline{\mathbf{V}}_i$ must be positive definite. Using the Binomial Inverse Theorem (Henderson and Searle, 1981), we have

$$\overline{\mathbf{V}}_i - \left(\sum_{k=1}^{I} \overline{\mathbf{V}}_k^{-1}\right)^{-1} = \overline{\mathbf{V}}_i - \left(\overline{\mathbf{V}}_i^{-1} + \mathbf{C}_{-i}\right)^{-1} = \overline{\mathbf{V}}_i \mathbf{C}_{-i} \left(\mathbf{C}_{-i} + \mathbf{C}_{-i} \overline{\mathbf{V}}_i \mathbf{C}_{-i}\right)^{-1} \mathbf{C}_{-i} \overline{\mathbf{V}}_i,$$

which is positive definite since we could verify that $\mathbf{C}_{-i} + \mathbf{C}_{-i}\overline{\mathbf{V}}_i\mathbf{C}_{-i}$ is positive definite. Therefore, $\widetilde{\mathbf{e}}_j' \left\{ \overline{\mathbf{V}}_i - \left(\sum_{k=1}^{I} \overline{\mathbf{V}}_k^{-1}\right)^{-1} \right\} \widetilde{\mathbf{e}}_j > 0$ since $\widetilde{\mathbf{e}}_j \neq \mathbf{0}$. Because $IC(i,j)$ is a ratio of two asymptotic variances, it must be positive, and so its denominator must be positive. Finally, $var(\widehat{\delta}) \left\{ \widetilde{X}_{ij} - \left(\sum_{k=1}^{I} \mathbf{X}_k' \overline{\mathbf{V}}_k^{-1}\right) \left(\sum_{k=1}^{I} \overline{\mathbf{V}}_k^{-1}\right)^{-1} \widetilde{\mathbf{e}}_j \right\}^2$ is a product of two non-negative quantities, and must be non-negative. Therefore $IC(i,j) \geq 1$ for all $i = 1, \dots, I$ and $j = 1, \dots, T$.

Web Appendix D Omitting Multiple Cells, an Entire Treatment Sequence or an Entire Period

Precision of the treatment effect estimator when multiple cells are omitted We now extend the expression for $var(\widehat{\delta})_{[ij]}$ in Theorem 1 of the main manuscript when multiple cells are omitted: we denote that variance by $var(\widehat{\delta})_{[o]}$. Assuming that m_i cells are omitted from each cluster i, we can re-arrange the vector of cell means $\overline{\mathbf{Y}}^*$ such that the observations corresponding to the omitted cells are all first, group by clusters, followed by the observations corresponding to the included cells (again group by clusters). This way, the covariance matrix of $\overline{\mathbf{Y}}^*$ can be reconfigured and decomposed as

$$\overline{\mathbf{V}} = \begin{pmatrix} \overline{\mathbf{V}}_o & \boldsymbol{\nu}' \\ \boldsymbol{\nu} & \overline{\mathbf{V}}_{[o]} \end{pmatrix},$$

where $\overline{\mathbf{V}}_o$ is of dimension $\sum_{i=1}^{I} m_i \times \sum_{i=1}^{I} m_i$, $\overline{\mathbf{V}}_{[o]}$ is of dimension $(IT - \sum_{i=1}^{I} m_i) \times (IT - \sum_{i=1}^{I} m_i)$, and $\boldsymbol{\nu}$ is of dimension $(IT - \sum_{i=1}^{I} m_i) \times \sum_{i=1}^{I} m_i$. Due to independence across different clusters, $\overline{\mathbf{V}}_o$ has a block diagonal structure, with each block given by $\overline{\mathbf{V}}_{m_i}$. Likewise, $\overline{\mathbf{V}}_{[o]}$ has a block diagonal structure, with each block given by $\overline{\mathbf{V}}_{[m_i]}$. The inverse can be written by block matrix inversion as

$$\overline{\mathbf{V}}^{-1} = \begin{pmatrix} (\overline{\mathbf{V}}_o - \boldsymbol{\nu}'\overline{\mathbf{V}}_{[o]}^{-1}\boldsymbol{\nu})^{-1} & -(\overline{\mathbf{V}}_o - \boldsymbol{\nu}'\overline{\mathbf{V}}_{[o]}^{-1}\boldsymbol{\nu})^{-1}\boldsymbol{\nu}'\overline{\mathbf{V}}_{[o]}^{-1} \\ -\overline{\mathbf{V}}_{[o]}^{-1}\boldsymbol{\nu}(\overline{\mathbf{V}}_o - \boldsymbol{\nu}'\overline{\mathbf{V}}_{[o]}^{-1}\boldsymbol{\nu})^{-1} & \overline{\mathbf{V}}_{[o]}^{-1} + \overline{\mathbf{V}}_{[o]}^{-1}\boldsymbol{\nu}(\overline{\mathbf{V}}_o - \boldsymbol{\nu}'\overline{\mathbf{V}}_{[o]}^{-1}\boldsymbol{\nu})^{-1}\boldsymbol{\nu}'\overline{\mathbf{V}}_{[o]}^{-1} \end{pmatrix}.$$

Define the covariance block as $\boldsymbol{\nu} = (\mathbf{G}_{m_1}', \dots, \mathbf{G}_{m_I}')$, where each $\mathbf{G}_{m_i}$ is of dimension $m_i \times (IT - \sum_{i=1}^{I} m_i)$, but only contains $T - m_i$ non-zero columns due to between-cluster independence. We will denote the non-zero columns by $\mathbf{G}_{m_i}^{\neq 0}$.

We similarly decompose $\mathbf{Z}$ as $\mathbf{Z} = (\mathbf{Z}_o', \mathbf{Z}_{[o]}')'$, where $\mathbf{Z}_o$ is the $\sum_{i=1}^{I} m_i \times (T+1)$ submatrix corresponding to the omitted cells, and $\mathbf{Z}_{[o]}$ is the remaining $(IT - \sum_{i=1}^{I} m_i) \times (T+1)$

sub-matrix corresponding to the included cells. Some matrix algebra gives

$$\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z} = \mathbf{Z}'_{[o]}\overline{\mathbf{V}}_{[o]}^{-1}\mathbf{Z}_{[o]} + (\mathbf{Z}'_o - \mathbf{Z}'_{[o]}\overline{\mathbf{V}}_{[o]}^{-1}\overline{\boldsymbol{\nu}})(\overline{\mathbf{V}}_o - \overline{\boldsymbol{\nu}}'\overline{\mathbf{V}}_{[o]}^{-1}\overline{\boldsymbol{\nu}})^{-1}(\mathbf{Z}'_o - \mathbf{Z}'_{[o]}\overline{\mathbf{V}}_{[o]}^{-1}\overline{\boldsymbol{\nu}})',$$

which can be equivalently represented by

$$\mathbf{Z}'_{[o]}\overline{\mathbf{V}}_{[o]}^{-1}\mathbf{Z}_{[o]} = \mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z} - \widetilde{\mathbf{Z}}'_o\mathbf{W}^{-1}\widetilde{\mathbf{Z}}_o,$$

where $\widetilde{\mathbf{Z}}'_o = \mathbf{Z}'_o - \mathbf{Z}'_{[o]}\overline{\mathbf{V}}_{[o]}^{-1}\overline{\boldsymbol{\nu}}$, $\mathbf{W} = \overline{\mathbf{V}}_o - \overline{\boldsymbol{\nu}}'\overline{\mathbf{V}}_{[o]}^{-1}\overline{\boldsymbol{\nu}}$. This then suggests that $var(\widehat{\delta})_{[o]}$ is the $(T+1)\times(T+1)$ entry of $(\mathbf{Z}'_{[o]}\overline{\mathbf{V}}_{[o]}^{-1}\mathbf{Z}_{[o]})^{-1} = (\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z} - \widetilde{\mathbf{Z}}'_o\mathbf{W}^{-1}\widetilde{\mathbf{Z}}_o)^{-1}$. Implicitly, here we have assumed the cell deletion still ensures that the concentration matrix $\mathbf{Z}'_{[o]}\overline{\mathbf{V}}_{[o]}^{-1}\mathbf{Z}_{[o]}$ is of full rank and therefore invertible. If the cell deletions engender a rank-deficient concentration matrix $\mathbf{Z}'_{[o]}\overline{\mathbf{V}}_{[o]}^{-1}\mathbf{Z}_{[o]}$, we will replace certain inverse with the *Moore-Penrose generalized inverse* and the calculation can still go through. We elaborate this latter point at the end of Appendix D when an entire period is deleted.

To proceed, recall that $\overline{\mathbf{V}}_{[o]}$ is block diagonal, so we can write

$$\mathbf{W} = \begin{pmatrix} \overline{\mathbf{V}}_{m_1} - \mathbf{G}^{\neq 0}_{m_1}\overline{\mathbf{V}}_{[m_1]}^{-1}(\mathbf{G}^{\neq 0}_{m_1})' & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \overline{\mathbf{V}}_{m_I} - \mathbf{G}^{\neq 0}_{m_I}\overline{\mathbf{V}}_{[m_I]}^{-1}(\mathbf{G}^{\neq 0}_{m_I})' \end{pmatrix}.$$

Meanwhile, we can write

$$\widetilde{\mathbf{Z}}_o = \begin{pmatrix} \widetilde{\mathbf{T}}_{m_1} & \widetilde{\mathbf{X}}_{m_1} \\ \vdots & \vdots \\ \widetilde{\mathbf{T}}_{m_I} & \widetilde{\mathbf{X}}_{m_I} \end{pmatrix},$$

therefore $\widetilde{\mathbf{Z}}'_o\mathbf{W}^{-1}\widetilde{\mathbf{Z}}_o$ can be written as

$$\begin{pmatrix} \sum_{i=1}^I \widetilde{\mathbf{T}}'_{m_i}\left(\overline{\mathbf{V}}_{m_i} - \mathbf{G}^{\neq 0}_{m_i}\overline{\mathbf{V}}_{[m_i]}^{-1}(\mathbf{G}^{\neq 0}_{m_i})'\right)^{-1}\widetilde{\mathbf{T}}_{m_i} & \sum_{i=1}^I \widetilde{\mathbf{T}}'_{m_i}\left(\overline{\mathbf{V}}_{m_i} - \mathbf{G}^{\neq 0}_{m_i}\overline{\mathbf{V}}_{[m_i]}^{-1}(\mathbf{G}^{\neq 0}_{m_i})'\right)^{-1}\widetilde{\mathbf{X}}_{m_i} \\ \sum_{i=1}^I \widetilde{\mathbf{X}}'_{m_i}\left(\overline{\mathbf{V}}_{m_i} - \mathbf{G}^{\neq 0}_{m_i}\overline{\mathbf{V}}_{[m_i]}^{-1}(\mathbf{G}^{\neq 0}_{m_i})'\right)^{-1}\widetilde{\mathbf{T}}_{m_i} & \sum_{i=1}^I \widetilde{\mathbf{X}}'_{m_i}\left(\overline{\mathbf{V}}_{m_i} - \mathbf{G}^{\neq 0}_{m_i}\overline{\mathbf{V}}_{[m_i]}^{-1}(\mathbf{G}^{\neq 0}_{m_i})'\right)^{-1}\widetilde{\mathbf{X}}_{m_i} \end{pmatrix},$$

Recall that

$$\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z} = \begin{pmatrix} \sum_{i=1}^I \overline{\mathbf{V}}_i^{-1} & \sum_{i=1}^I \overline{\mathbf{V}}_i^{-1}\mathbf{X}_i \\ \sum_{i=1}^I \mathbf{X}'_i\overline{\mathbf{V}}_i^{-1} & \sum_{i=1}^I \mathbf{X}'_i\overline{\mathbf{V}}_i^{-1}\mathbf{X}_i \end{pmatrix} \triangleq \begin{pmatrix} \mathbf{C} & \mathbf{d} \\ \mathbf{d}' & f \end{pmatrix}.$$

The lower-right element of $(\mathbf{Z}'_{[o]}\overline{\mathbf{V}}_{[o]}^{-1}\mathbf{Z}_{[o]})^{-1} = (\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z} - \widetilde{\mathbf{Z}}'_o\mathbf{W}^{-1}\widetilde{\mathbf{Z}}_o)^{-1}$ is given by $(\mathbf{F}_1 - \mathbf{F}_2\mathbf{F}_3^{-1}\mathbf{F}'_2)^{-1}$, where

$$\mathbf{F}_1 = \sum_{i=1}^I \mathbf{X}'_i\overline{\mathbf{V}}_i^{-1}\mathbf{X}_i - \sum_{i=1}^I \widetilde{\mathbf{X}}'_{m_i}\left(\overline{\mathbf{V}}_{m_i} - \mathbf{G}^{\neq 0}_{m_i}\overline{\mathbf{V}}_{[m_i]}^{-1}(\mathbf{G}^{\neq 0}_{m_i})'\right)^{-1}\widetilde{\mathbf{X}}_{m_i}$$

$$\mathbf{F}_2 = \sum_{i=1}^I \mathbf{X}'_i\overline{\mathbf{V}}_i^{-1} - \sum_{i=1}^I \widetilde{\mathbf{X}}'_{m_i}\left(\overline{\mathbf{V}}_{m_i} - \mathbf{G}^{\neq 0}_{m_i}\overline{\mathbf{V}}_{[m_i]}^{-1}(\mathbf{G}^{\neq 0}_{m_i})'\right)^{-1}\widetilde{\mathbf{T}}_{m_i}$$

$$\mathbf{F}_3 = \sum_{i=1}^I \overline{\mathbf{V}}_i^{-1} - \sum_{i=1}^I \widetilde{\mathbf{T}}'_{m_i}\left(\overline{\mathbf{V}}_{m_i} - \mathbf{G}^{\neq 0}_{m_i}\overline{\mathbf{V}}_{[m_i]}^{-1}(\mathbf{G}^{\neq 0}_{m_i})'\right)^{-1}\widetilde{\mathbf{T}}_{m_i}.$$

Recall the expression of $var(\widehat{\delta})$ from Lemma 1, and plugging in the expressions for $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$, we arrive at the expression

$$\begin{aligned}\frac{1}{var(\widehat{\delta})_{[o]}} =& \frac{1}{var(\widehat{\delta})} + \left(\sum_{i=1}^{I}\mathbf{X}_i'\overline{\mathbf{V}}_i^{-1}\right)\left(\sum_{i=1}^{I}\overline{\mathbf{V}}_i^{-1}\right)^{-1}\left(\sum_{i=1}^{I}\overline{\mathbf{V}}_i^{-1}\mathbf{X}_i\right)\\
&-\sum_{i=1}^{I}\widetilde{\mathbf{X}}_{m_i}'\left(\overline{\mathbf{V}}_{m_i}-\mathbf{G}_{m_i}^{\neq 0}\overline{\mathbf{V}}_{[m_i]}^{-1}(\mathbf{G}_{m_i}^{\neq 0})'\right)^{-1}\widetilde{\mathbf{X}}_{m_i}\\
&-\left\{\sum_{i=1}^{I}\mathbf{X}_i'\overline{\mathbf{V}}_i^{-1}-\sum_{i=1}^{I}\widetilde{\mathbf{X}}_{m_i}'\left(\overline{\mathbf{V}}_{m_i}-\mathbf{G}_{m_i}^{\neq 0}\overline{\mathbf{V}}_{[m_i]}^{-1}(\mathbf{G}_{m_i}^{\neq 0})'\right)^{-1}\widetilde{\mathbf{T}}_{m_i}\right\}\\
&\times\left\{\sum_{i=1}^{I}\overline{\mathbf{V}}_i^{-1}-\sum_{i=1}^{I}\widetilde{\mathbf{T}}_{m_i}'\left(\overline{\mathbf{V}}_{m_i}-\mathbf{G}_{m_i}^{\neq 0}\overline{\mathbf{V}}_{[m_i]}^{-1}(\mathbf{G}_{m_i}^{\neq 0})'\right)^{-1}\widetilde{\mathbf{T}}_{m_i}\right\}^{-1}\\
&\times\left\{\sum_{i=1}^{I}\mathbf{X}_i'\overline{\mathbf{V}}_i^{-1}-\sum_{i=1}^{I}\widetilde{\mathbf{X}}_{m_i}'\left(\overline{\mathbf{V}}_{m_i}-\mathbf{G}_{m_i}^{\neq 0}\overline{\mathbf{V}}_{[m_i]}^{-1}(\mathbf{G}_{m_i}^{\neq 0})'\right)^{-1}\widetilde{\mathbf{T}}_{m_i}\right\}'.\end{aligned}$$

Information content of each treatment sequence Suppose that the treatment sequence i is omitted from the study, and we let $var(\widehat{\delta})_{[i,\bullet]}$ denote the variance of the treatment effect estimator after omitting sequence i. Then we can re-arrange the vector of cell means such that the omitted elements precede the included elements. This re-arrangement allows us to partition $\mathbf{Z} = (\mathbf{Z}_i', \mathbf{Z}_{[i]}')'$ and

$$\overline{\mathbf{V}} = \begin{pmatrix} \overline{\mathbf{V}}_i & \mathbf{0} \\ \mathbf{0} & \overline{\mathbf{V}}_{[i]} \end{pmatrix},$$

which then gives $\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z} = \mathbf{Z}_i'\overline{\mathbf{V}}_i^{-1}\mathbf{Z}_i + \mathbf{Z}_{[i]}'\overline{\mathbf{V}}_{[i]}^{-1}\mathbf{Z}_{[i]}$. Therefore the covariance matrix of the marginal mean regression estimators can be found as

$$\begin{aligned}(\mathbf{Z}_{[i]}'\overline{\mathbf{V}}_{[i]}^{-1}\mathbf{Z}_{[i]})^{-1} =& (\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z} - \mathbf{Z}_i'\overline{\mathbf{V}}_i^{-1}\mathbf{Z}_i)^{-1}\\
=& (\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1} + (\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\mathbf{Z}_i'\{\overline{\mathbf{V}}_i - \mathbf{Z}_i(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\mathbf{Z}_i'\}^{-1}\mathbf{Z}_i(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}.\end{aligned}$$

Recall that

$$(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1} = \begin{pmatrix} \mathbf{C}^{-1} + var(\widehat{\delta})\mathbf{C}^{-1}\mathbf{d}\mathbf{d}'\mathbf{C}^{-1} & -var(\widehat{\delta})\mathbf{C}^{-1}\mathbf{d} \\ -var(\widehat{\delta})\mathbf{d}'\mathbf{C}^{-1} & var(\widehat{\delta}) \end{pmatrix},$$

where the relevant notation is defined in Web Appendix C. Because we can partition $\mathbf{Z}_i = (\mathbf{I}, \mathbf{X}_i)$, we have

$$(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\mathbf{Z}_i' = \begin{pmatrix} \mathbf{C}^{-1} + var(\widehat{\delta})\mathbf{C}^{-1}\mathbf{d}\mathbf{d}'\mathbf{C}^{-1} - var(\widehat{\delta})\mathbf{C}^{-1}\mathbf{d}\mathbf{X}_i \\ -var(\widehat{\delta})\mathbf{d}'\mathbf{C}^{-1} + var(\widehat{\delta})\mathbf{X}_i \end{pmatrix}.$$

Then

$$\overline{\mathbf{V}}_i - \mathbf{Z}_i(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\mathbf{Z}_i' = \overline{\mathbf{V}}_i - \mathbf{C}^{-1} - var(\widehat{\delta})(\mathbf{C}^{-1}\mathbf{d} - \mathbf{X}_i)(\mathbf{C}^{-1}\mathbf{d} - \mathbf{X}_i)' = \mathbf{\Lambda}_i - var(\widehat{\delta})\mathbf{M}_i\mathbf{M}_i',$$

where $\mathbf{\Lambda}_i = \overline{\mathbf{V}}_i - \mathbf{C}^{-1}$ and

$$\mathbf{M}_i = \mathbf{C}^{-1}\mathbf{d} - \mathbf{X}_i = \left(\sum_{i=1}^{I}\overline{\mathbf{V}}_i^{-1}\right)^{-1}\left(\sum_{i=1}^{I}\overline{\mathbf{V}}_i^{-1}\mathbf{X}_i\right) - \mathbf{X}_i.$$

Using the Woodbury matrix identity, the inverse is given by

$$\{\overline{\mathbf{V}}_i - \mathbf{Z}_i(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\mathbf{Z}_i'\}^{-1} = \boldsymbol{\Lambda}_i^{-1} + \boldsymbol{\Lambda}_i^{-1}\mathbf{M}_i\left(\frac{1}{var(\widehat{\delta})} - \mathbf{M}_i'\boldsymbol{\Lambda}_i^{-1}\mathbf{M}_i\right)\mathbf{M}_i'\boldsymbol{\Lambda}_i^{-1}.$$

Because $var(\widehat{\delta})_{[i,\bullet]} = var(\widehat{\delta})_{[i,\bullet]}$ is given by the $(T+1, T+1)$th element of $(\mathbf{Z}_{[i]}'\overline{\mathbf{V}}_{[i]}^{-1}\mathbf{Z}_{[i]})^{-1}$, we can write

$$\begin{aligned}
var(\widehat{\delta})_{[i,\bullet]} =& var(\widehat{\delta}) + \{-var(\widehat{\delta})\mathbf{d}'\mathbf{C}^{-1} + var(\widehat{\delta})\mathbf{X}_i\}' \\
&\times \left\{\boldsymbol{\Lambda}_i^{-1} + \boldsymbol{\Lambda}_i^{-1}\mathbf{M}_i\left(\frac{1}{var(\widehat{\delta})} - \mathbf{M}_i'\boldsymbol{\Lambda}_i^{-1}\mathbf{M}_i\right)\mathbf{M}_i'\boldsymbol{\Lambda}_i^{-1}\right\}\{-var(\widehat{\delta})\mathbf{d}'\mathbf{C}^{-1} + var(\widehat{\delta})\mathbf{X}_i\} \\
=& var(\widehat{\delta}) + var(\widehat{\delta})^2\mathbf{M}_i'\left\{\boldsymbol{\Lambda}_i^{-1} + \boldsymbol{\Lambda}_i^{-1}\mathbf{M}_i\left(\frac{1}{var(\widehat{\delta})} - \mathbf{M}_i'\boldsymbol{\Lambda}_i^{-1}\mathbf{M}_i\right)\mathbf{M}_i'\boldsymbol{\Lambda}_i^{-1}\right\}\mathbf{M}_i.
\end{aligned}$$

Some algebra then gives

$$IC(i,\bullet) = \frac{var(\widehat{\delta})_{[i,\bullet]}}{var(\widehat{\delta})} = 1 + var(\widehat{\delta})\left\{\frac{\mathbf{M}_i'\boldsymbol{\Lambda}_i^{-1}\mathbf{M}_i}{1 - var(\widehat{\delta})\mathbf{M}_i'\boldsymbol{\Lambda}_i^{-1}\mathbf{M}_i}\right\},$$

which equals to equation (4) in the main manuscript after algebraic simplification.

Information content of each time period We denote the variance of the treatment effect estimator when period j is omitted by $var(\widehat{\delta})_{[\bullet,j]}$, whose expression will be derived as follows. We denote the cell means from period j across all clusters as $\overline{\mathbf{Y}}_j^*$ and then re-arrange the collection of all cell means such that $\overline{\mathbf{Y}}^* = (\overline{\mathbf{Y}}_j^{*'}, \overline{\mathbf{Y}}_{[j]}^{*'})'$. The variance of $\overline{\mathbf{Y}}^*$ is then partitioned into

$$\overline{\mathbf{V}} = \begin{pmatrix} \overline{\mathbf{V}}_j & \mathbf{G}_j' \\ \mathbf{G}_j & \overline{\mathbf{V}}_{[j]} \end{pmatrix}.$$

Due to between-cluster independence, $\overline{\mathbf{V}}_j$ has a block diagonal structure, and $\overline{\mathbf{V}}_{[j]}$ also has a block diagonal structure, and each block is given by $\overline{\mathbf{V}}_i$ with the jth row and column omitted. Further, we can write $\mathbf{G}_j = (\mathbf{G}_{j1}', \dots, \mathbf{G}_{jI}')$, where each $\mathbf{G}_{ji}$ is of dimension $1 \times I(T-1)$. Of note, only certain columns of $\mathbf{G}_{ji}$ contain non-zero elements, which we collectively denote by $\mathbf{G}_{ji}^{\neq 0}$. We also partition $\mathbf{Z} = (\mathbf{Z}_j', \mathbf{Z}_{[j]}')'$. Uniquely, when an entire period is deleted, one element in $\boldsymbol{\theta}$ corresponding to the jth period effect becomes unestimable, and therefore $var(\widehat{\delta})_{[\bullet,j]}$ should be obtained as the *Moore-Penrose generalized inverse* of the rank-deficient concentration matrix $\mathbf{Z}_{[j]}'\overline{\mathbf{V}}_{[j]}^{-1}\mathbf{Z}_{[j]}$. Invoking a version of the Woodbury matrix identity with rank-deficient matrix (Theorem 2.1 in Deng (2011)), we obtain

$$(\mathbf{Z}_{[j]}'\overline{\mathbf{V}}_{[j]}^{-1}\mathbf{Z}_{[j]})^+ = (\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1} + (\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\widetilde{\mathbf{Z}}_j'(\mathbf{W}_j - \widetilde{\mathbf{Q}}_j)^+\widetilde{\mathbf{Z}}_j(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1},$$

where '+' denotes in the Moore-Penrose generalized inverse, $\widetilde{\mathbf{Z}}_j = \mathbf{Z}_j - \mathbf{G}_j'\overline{\mathbf{V}}_{[j]}^{-1}\mathbf{Z}_{[j]}$, $\mathbf{W}_j = \overline{\mathbf{V}}_j - \mathbf{G}_j'\overline{\mathbf{V}}_{[j]}^{-1}\mathbf{G}_j$ and $\widetilde{\mathbf{Q}}_j = \widetilde{\mathbf{Z}}_j(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\widetilde{\mathbf{Z}}_j'$. Now observe that

$$\mathbf{W}_j = \begin{pmatrix} \overline{\mathbf{V}}_{1j} - \mathbf{G}_{j1}^{\neq 0}\overline{\mathbf{V}}_{1[j]}^{-1}(\mathbf{G}_{j1}^{\neq 0})' & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \overline{\mathbf{V}}_{Ij} - \mathbf{G}_{Ij}^{\neq 0}\overline{\mathbf{V}}_{I[j]}^{-1}(\mathbf{G}_{Ij}^{\neq 0})' \end{pmatrix}.$$

We further partition $\widetilde{\mathbf{Z}}_j = (\widetilde{\mathbf{E}}_j, \widetilde{\mathbf{X}}_j)$, where

$$\widetilde{\mathbf{E}}_j = \begin{pmatrix} \mathbf{e}'_j - \mathbf{G}^{\neq 0}_{j1}\overline{\mathbf{V}}^{-1}_{1[j]}\mathbf{I}_{[j]} \\ \vdots \\ \mathbf{e}'_j - \mathbf{G}^{\neq 0}_{jI}\overline{\mathbf{V}}^{-1}_{I[j]}\mathbf{I}_{[j]} \end{pmatrix}, \qquad \widetilde{\mathbf{X}}_j = \begin{pmatrix} X_{1j} - \mathbf{G}^{\neq 0}_{j1}\overline{\mathbf{V}}^{-1}_{1[j]}\mathbf{X}_{1[j]} \\ \vdots \\ X_{Ij} - \mathbf{G}^{\neq 0}_{jI}\overline{\mathbf{V}}^{-1}_{I[j]}\mathbf{X}_{I[j]} \end{pmatrix},$$

where $\mathbf{X}_{i[j]}$ is $\mathbf{X}_i$ with the jth element omitted, and $\mathbf{I}_{[j]}$ is the identity matrix $\mathbf{I}_{J\times J}$ with the jth row omitted.

Recall that

$$(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1} = \begin{pmatrix} \mathbf{C}^{-1} + var(\widehat{\delta})\mathbf{C}^{-1}\mathbf{d}\mathbf{d}'\mathbf{C}^{-1} & -var(\widehat{\delta})\mathbf{C}^{-1}\mathbf{d} \\ -var(\widehat{\delta})\mathbf{d}'\mathbf{C}^{-1} & var(\widehat{\delta}) \end{pmatrix},$$

where the relevant notation is defined in Web Appendix C. We have

$$(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\widetilde{\mathbf{Z}}'_j = \begin{pmatrix} \mathbf{C}^{-1}\widetilde{\mathbf{E}}'_j + var(\widehat{\delta})\mathbf{C}^{-1}\mathbf{d}\{\mathbf{d}'\mathbf{C}^{-1}\widetilde{\mathbf{E}}'_j - \widetilde{\mathbf{X}}'_j\} \\ -var(\widehat{\delta})\{\mathbf{d}'\mathbf{C}^{-1}\widetilde{\mathbf{E}}'_j - \widetilde{\mathbf{X}}'_j\} \end{pmatrix}.$$

We can obtain

$$\begin{aligned}\widetilde{\mathbf{Q}}_j = \widetilde{\mathbf{Z}}_j(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\widetilde{\mathbf{Z}}'_j =& \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\widetilde{\mathbf{E}}'_j + var(\widehat{\delta})\widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\{\mathbf{d}'\mathbf{C}^{-1}\widetilde{\mathbf{E}}'_j - \widetilde{\mathbf{X}}'_j\} - var(\widehat{\delta})\widetilde{\mathbf{X}}_j\{\mathbf{d}'\mathbf{C}^{-1}\widetilde{\mathbf{E}}'_j - \widetilde{\mathbf{X}}'_j\} \\ =& \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\widetilde{\mathbf{E}}'_j + var(\widehat{\delta})\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}'.\end{aligned}$$

Using Corollary 2.1 in Deng (2011), we can further express the generalized inverse as

$$\begin{aligned}(\mathbf{W}_j - \widetilde{\mathbf{Q}}_j)^{+} =& \left[\widetilde{\mathbf{W}}_j - var(\widehat{\delta})\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}'\right]^{+} \\ =& \widetilde{\mathbf{W}}^{+}_j + \gamma\widetilde{\mathbf{W}}^{+}_j\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}'\widetilde{\mathbf{W}}^{+}_j,\end{aligned}$$

where $\widetilde{\mathbf{W}}_j = \mathbf{W}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\widetilde{\mathbf{E}}'_j$,

$$\gamma = \frac{var(\widehat{\delta})}{1 - var(\widehat{\delta})\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}'\widetilde{\mathbf{W}}^{+}_j\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}}.$$

These expressions then suggest that the $(T+1, T+1)$th entry of $(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}\widetilde{\mathbf{Z}}'_j(\mathbf{W}_j - \widetilde{\mathbf{Q}}_j)^{+}\widetilde{\mathbf{Z}}_j(\mathbf{Z}'\overline{\mathbf{V}}^{-1}\mathbf{Z})^{-1}$ is given by

$$var(\widehat{\delta})^2\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}'\widetilde{\mathbf{W}}^{+}_j\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}\left[1 + \gamma\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}'\widetilde{\mathbf{W}}^{+}_j\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}\right].$$

The variance of treatment effect omitting the jth period is

$$\begin{aligned}var(\widehat{\delta})_{[\bullet,j]} =& var(\widehat{\delta}) + var(\widehat{\delta})^2\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}'\widetilde{\mathbf{W}}^{+}_j\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\} \\ &\times\left[1 + \gamma\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}'\widetilde{\mathbf{W}}^{+}_j\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}\right].\end{aligned}$$

This then gives

$$\begin{aligned}IC(\bullet, j) = \frac{var(\widehat{\delta})_{[\bullet,j]}}{var(\widehat{\delta})} =& 1 + var(\widehat{\delta})\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}'\widetilde{\mathbf{W}}^{+}_j\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\} \\ &\times\left[1 + \gamma\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}'\widetilde{\mathbf{W}}^{+}_j\{\widetilde{\mathbf{X}}_j - \widetilde{\mathbf{E}}_j\mathbf{C}^{-1}\mathbf{d}\}\right] = \frac{\gamma}{var(\widehat{\delta})}.\end{aligned}$$

Web Appendix E Explicit Expressions for Information Content in Example 1 and Example 2

Block Exchangeable Correlation Structure Using the inverse expression of $\mathbf{R}$ provided in Li et al. (2018), we can obtain

$$\overline{\mathbf{V}}^{-1} = \phi^{-1}\left\{\frac{N}{\lambda_3}\mathbf{I}_T - \frac{N(\lambda_4-\lambda_3)}{T\lambda_3\lambda_4}\mathbf{J}_T\right\} = \phi^{-1}\{a_T + b\mathbf{J}_T\}.$$

Notice that because $\overline{\mathbf{V}}$ is no longer cluster-dependent,

$$\mathbf{M}_i = \frac{1}{I}\left(\sum_{k=1}^{I}\mathbf{X}_k\right) - \mathbf{X}_i = \frac{1}{I}\begin{pmatrix} 0 \\ 1 \\ \vdots \\ T-1 \end{pmatrix} - \begin{pmatrix} X_{i1} \\ X_{i2} \\ \vdots \\ X_{iT} \end{pmatrix} = \frac{1}{I}c - d. \tag{3}$$

Note that in a standard stepped wedge design, the vector d contains 1 from the $(i+1)$th position to the Tth position and zero elsewhere. Under a variance-stabilizing link function, the information content of each sequence futher simplifies to

$$IC(i,\bullet) = 1 + \frac{Ivar(\widehat{\delta})\mathbf{M}_i'\overline{\mathbf{V}}^{-1}\mathbf{M}_i}{(I-1) - Ivar(\widehat{\delta})\mathbf{M}_i'\overline{\mathbf{V}}^{-1}\mathbf{M}_i}. \tag{4}$$

After some simplification algebra, we arrive at $I\mathbf{M}_i'\overline{\mathbf{V}}^{-1}\mathbf{M}_i = \phi^{-1}\left\{aW + bU^2 + (a+Ib)i(i-T)\right\}$, where $W = \sum_{j=1}^{T}(\sum_{i=1}^{I}X_{ij})^2 = I(I+1)(2I+1)/6$ and $U = \sum_{i=1}^{I}\sum_{j=1}^{T}X_{ij} = I(I+1)/2$ are design constants in a standard stepped wedge design. Plugging in the values for a and b, we then have

$$aW + bU^2 + (a+Ib)i(i-T) = \frac{N}{T\lambda_3\lambda_4}\left\{\lambda_4(TW - U^2 + i(i-T)I) + \lambda_3(U^2 + i(i-T)I^2)\right\},$$

which then leads to the expression

$$IC(i,\bullet) = \frac{I(I-1)T\phi/N}{I(I-1)T\phi/N - var(\widehat{\delta})\left\{(TW-U^2+i(i-T)I)/\lambda_3 + (U^2+i(i-T)I^2)/\lambda_4\right\}}.$$

Proportional Decay Correlation Structure Using the inverse expression of $\mathbf{R}$ established in Li (2020), we can obtain

$$\overline{\mathbf{V}}^{-1} = \frac{N\phi^{-1}}{\xi_2(1-\rho^2)}(\mathbf{I}_T + \rho^2\mathbf{C}_2 - \rho\mathbf{C}_1),$$

where $\mathbf{C}_2 = (0,1,\ldots,1,0)$ and $\mathbf{C}_1$ is a $T\times T$ tridiagonal matrix with zeros on the main diagonal and ones of the two subdiagonals. Furthermore, equation (3) still applies here (thus we define c and d accordingly), and therefore we have

$$\begin{aligned}\mathbf{M}_i'\overline{\mathbf{V}}^{-1}\mathbf{M}_i =& \frac{N\phi^{-1}}{\xi_2(1-\rho^2)}\times\frac{1}{I^2}(c'c+\rho^2c'\mathbf{C}_2c-\rho c'\mathbf{C}_1c) + \frac{N\phi^{-1}}{\xi_2(1-\rho^2)}(d'd+\rho^2d'\mathbf{C}_2d -' \mathbf{C}_1d)\\ &- \frac{N\phi^{-1}}{\xi_2(1-\rho^2)}\times\frac{2}{I}(c'd+\rho^2c'\mathbf{C}_2d-\rho c'\mathbf{C}_1d).\end{aligned}$$

Notice that

$$c'c = \sum_{j=1}^{T}\left(\sum_{i=1}^{I} X_{ij}\right)^2 = W = \frac{1}{6}I(I+1)(2I+1)$$

$$c'\mathbf{C}_1 c = 2\sum_{j=1}^{T-1}\left(\sum_{i=1}^{I} X_{ij}\right)\left(\sum_{i=1}^{I} X_{i,j+1}\right) = 2W - I(I+1) = \frac{2}{3}I(I+1)(I-1)$$

which gives

$$\begin{aligned} c'c + \rho^2 c'\mathbf{C}_2 c - \rho c'\mathbf{C}_1 c &= -I^2\rho^2 + (1+\rho^2)W - 2\rho\left(W - \frac{1}{2}I(I+1)\right) \\ &= \rho I(I+1) - \rho^2 I^2 + (1-\rho)^2 W. \end{aligned}$$

In addition, the term in the second parenthesis

$$\begin{aligned} d'd + \rho^2 d'\mathbf{C}_2 d -' \mathbf{C}_1 d &= (X_{i1} + X_{iT}) + (1+\rho^2)\sum_{j=2}^{T-1} X_{ij} - 2\rho\sum_{j=2}^{T-1} X_{ij}X_{i,j+1} \\ &= -\rho^2(X_{i1} + X_{iT}) + (1+\rho^2)\sum_{j=1}^{T} X_{ij} - 2\rho\sum_{j=2}^{T-1} X_{ij}X_{i,j+1} \\ &= -\rho^2 + (1+\rho^2)(T-i) - 2\rho(T-i-1). \end{aligned}$$

Finally, the term in the third parenthesis is

$$c'd + \rho^2 c'\mathbf{C}_2 d - \rho c'\mathbf{C}_1 d = -\rho^2\left(X_{i1}\sum_{k=1}^{I} X_{k1} + X_{iT}\sum_{k=1}^{I} X_{kT}\right) + (1+\rho^2)c'd - \rho c'\mathbf{C}_1 d$$

$$= -I\rho^2 + (1+\rho^2)\left\{\frac{1}{2}I(I+1) - \frac{1}{2}i(i-1)\right\} - \rho\left[\sum_{j=1}^{T-1}\left(\sum_{k=1}^{I} X_{kj}\right)X_{i,j+1} + \sum_{j=2}^{T}\left(\sum_{k=1}^{I} X_{kj}\right)X_{i,j-1}\right],$$

where each of summations can be simplifies in a standard stepped wedge design as

$$\begin{aligned} &\sum_{j=1}^{T-1}\left(\sum_{k=1}^{I} X_{kj}\right)X_{i,j+1} = \sum_{j=i}^{T-1}\left(\sum_{k=1}^{i-1} X_{kj}\right) + \sum_{j=i}^{T-1}\left(\sum_{k=i}^{I} X_{kj}\right) \\ &= (T-i)(i-1) + (0+1+\ldots+(T-i-1)) = \frac{1}{2}(T-i)(T+i-3) \end{aligned}$$

and

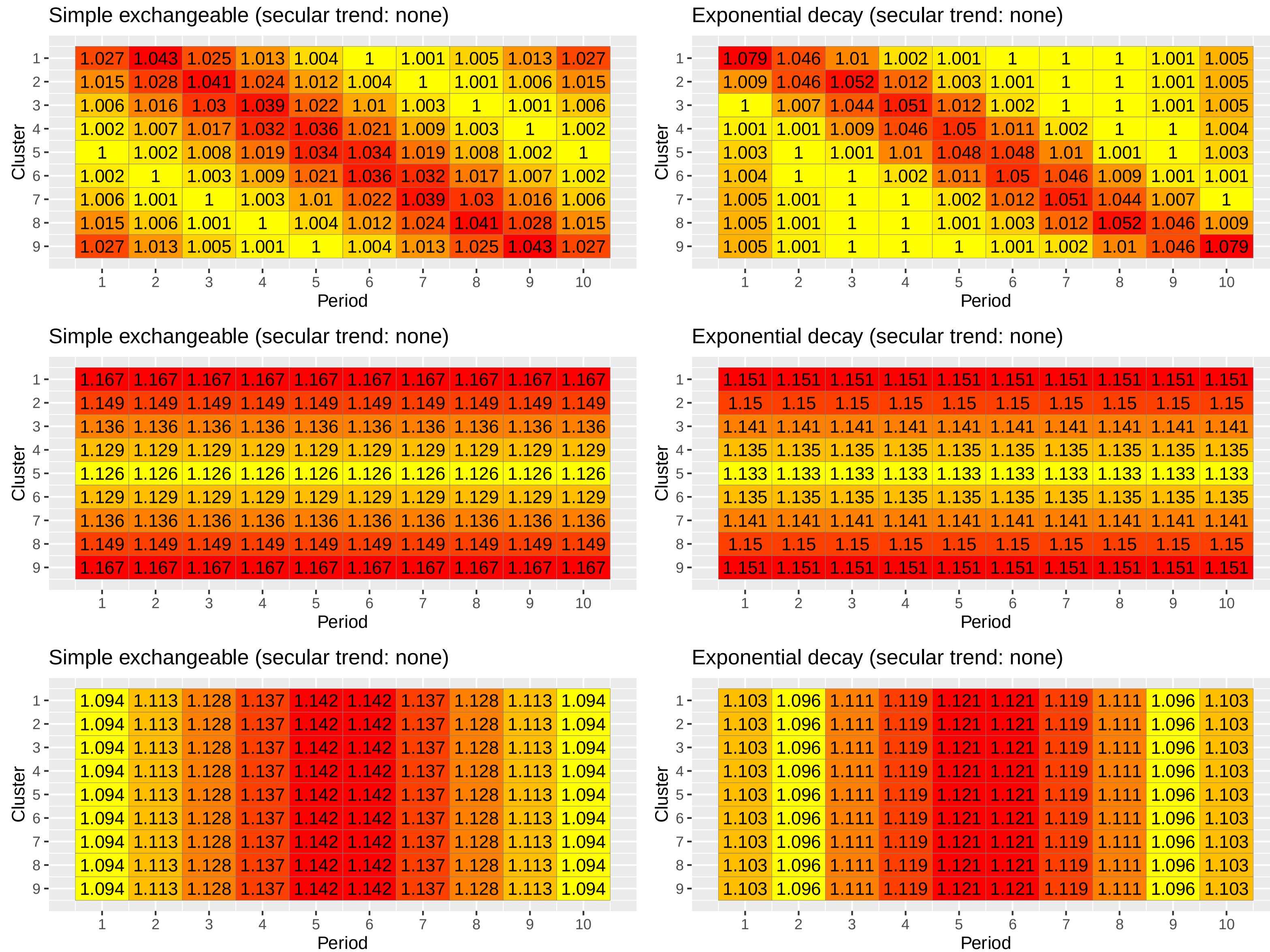
$$\begin{aligned} &\sum_{j=2}^{T}\left(\sum_{k=1}^{I} X_{kj}\right)X_{i,j-1} = \sum_{j=i+2}^{T}\left(\sum_{k=1}^{i+1} X_{kj}\right) + \sum_{j=i+2}^{T}\left(\sum_{k=i+2}^{I} X_{kj}\right) \\ &= (T-i-1)(i+1) + (0+1+\ldots+(T-i-2)) = \frac{1}{2}(T-i-1)(T+i). \end{aligned}$$

Plugging in these terms and after extensive algebra, we obtain

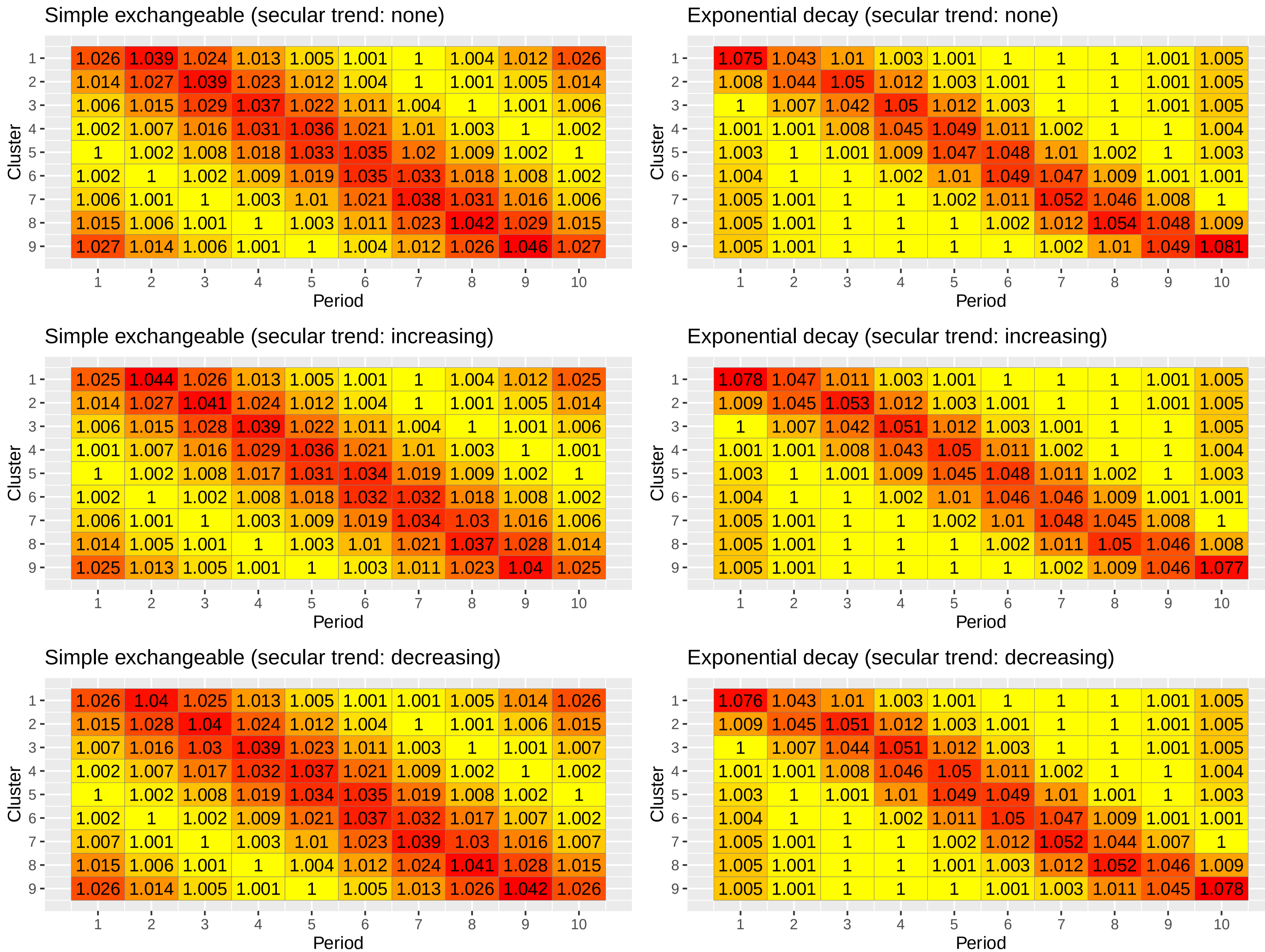
$$I\mathbf{M}_i'\overline{\mathbf{V}}^{-1}\mathbf{M}_i = \frac{N\phi^{-1}}{I\xi_2(1-\rho^2)}\left\{\rho I(I-1) + (1-\rho)^2 W - I(1-\rho)^2 i(T-i)\right\},$$

which then gives the expression of $IC(i,\bullet)$ in Section 3.1.2.

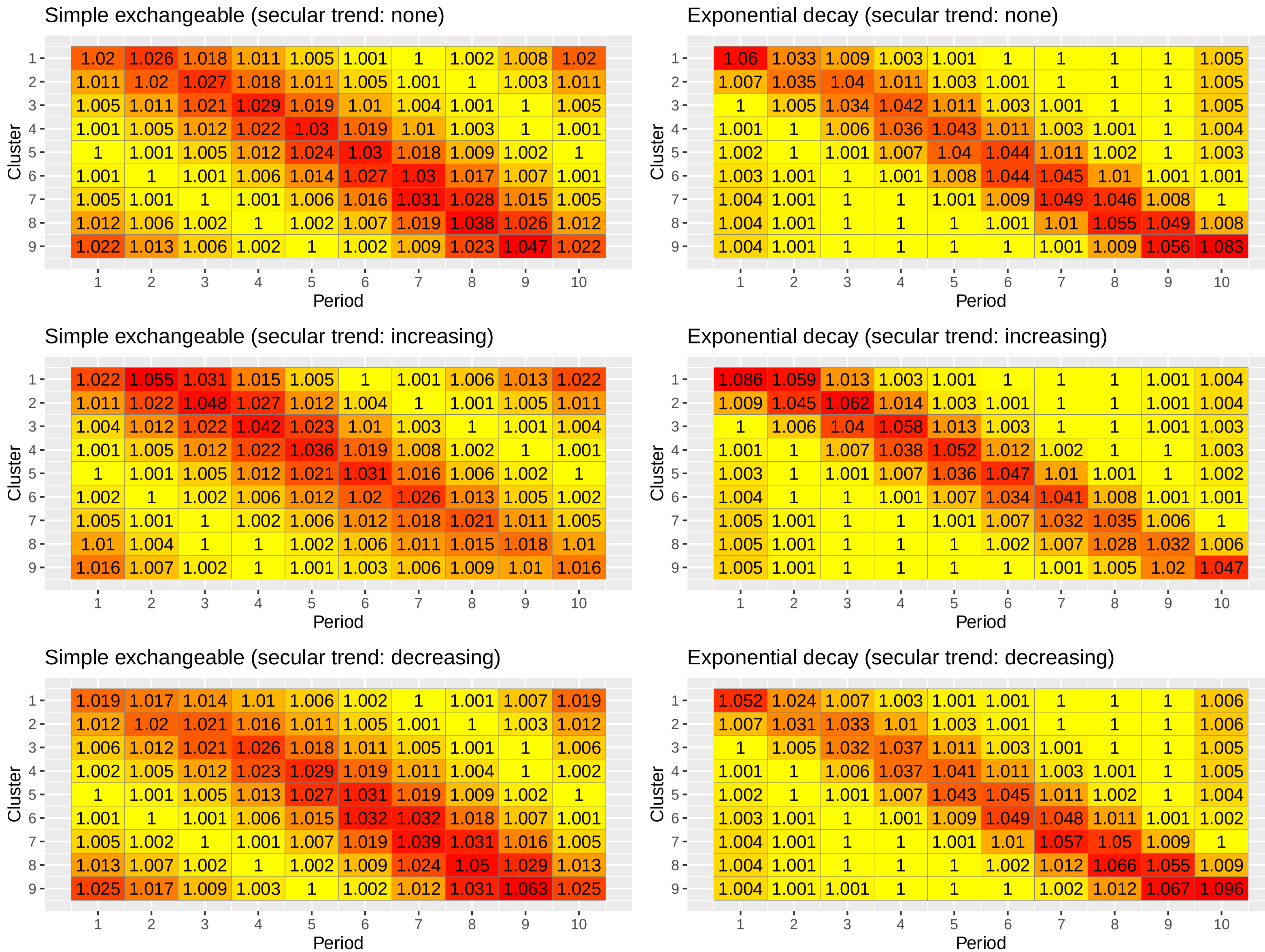
Web Appendix F Web Figures



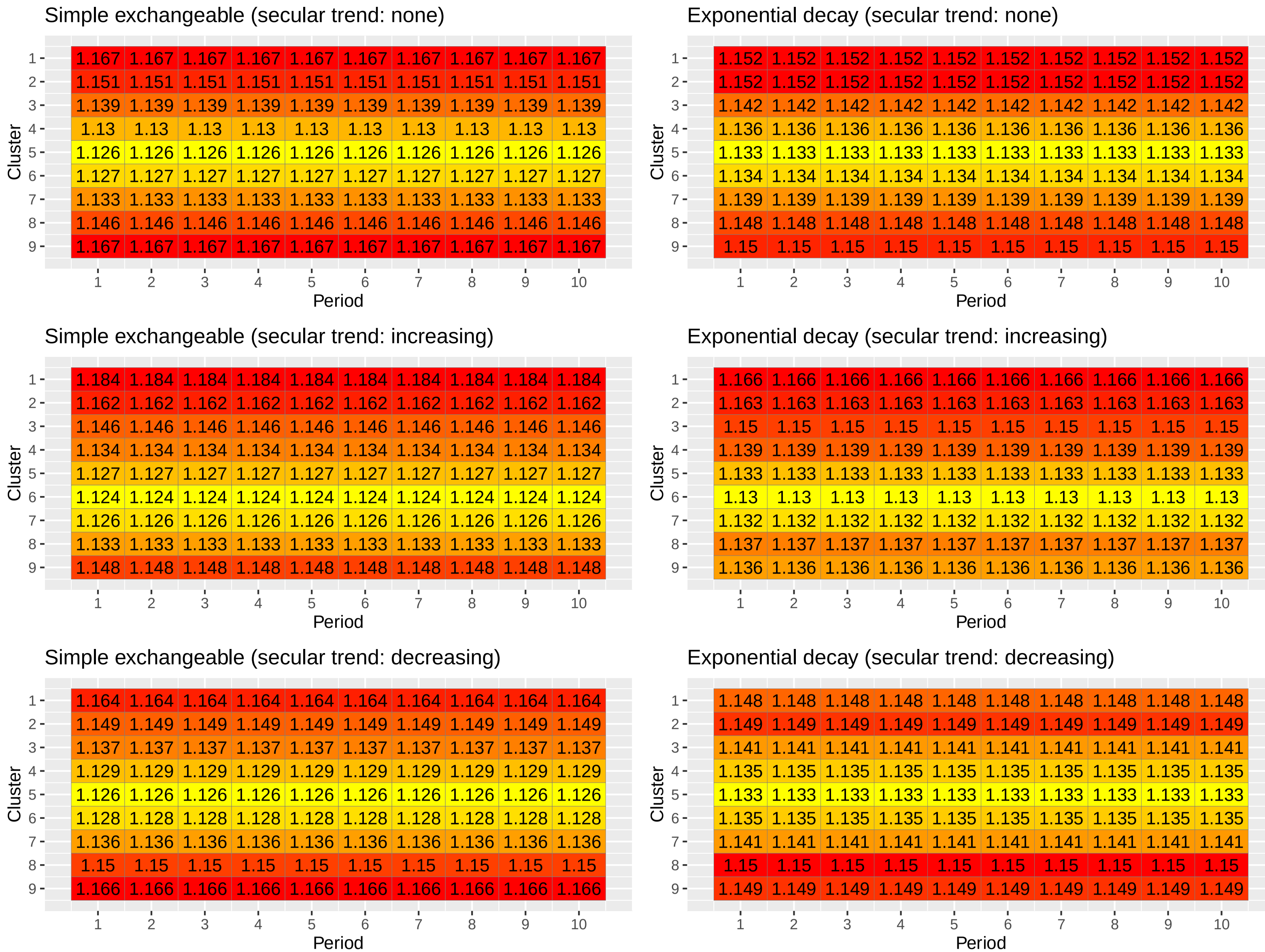
Web Figure 1: Information content of cells, sequences and periods under two correlation models, for arbitrary outcomes with a variance-stabilizing link in the marginal mean model.



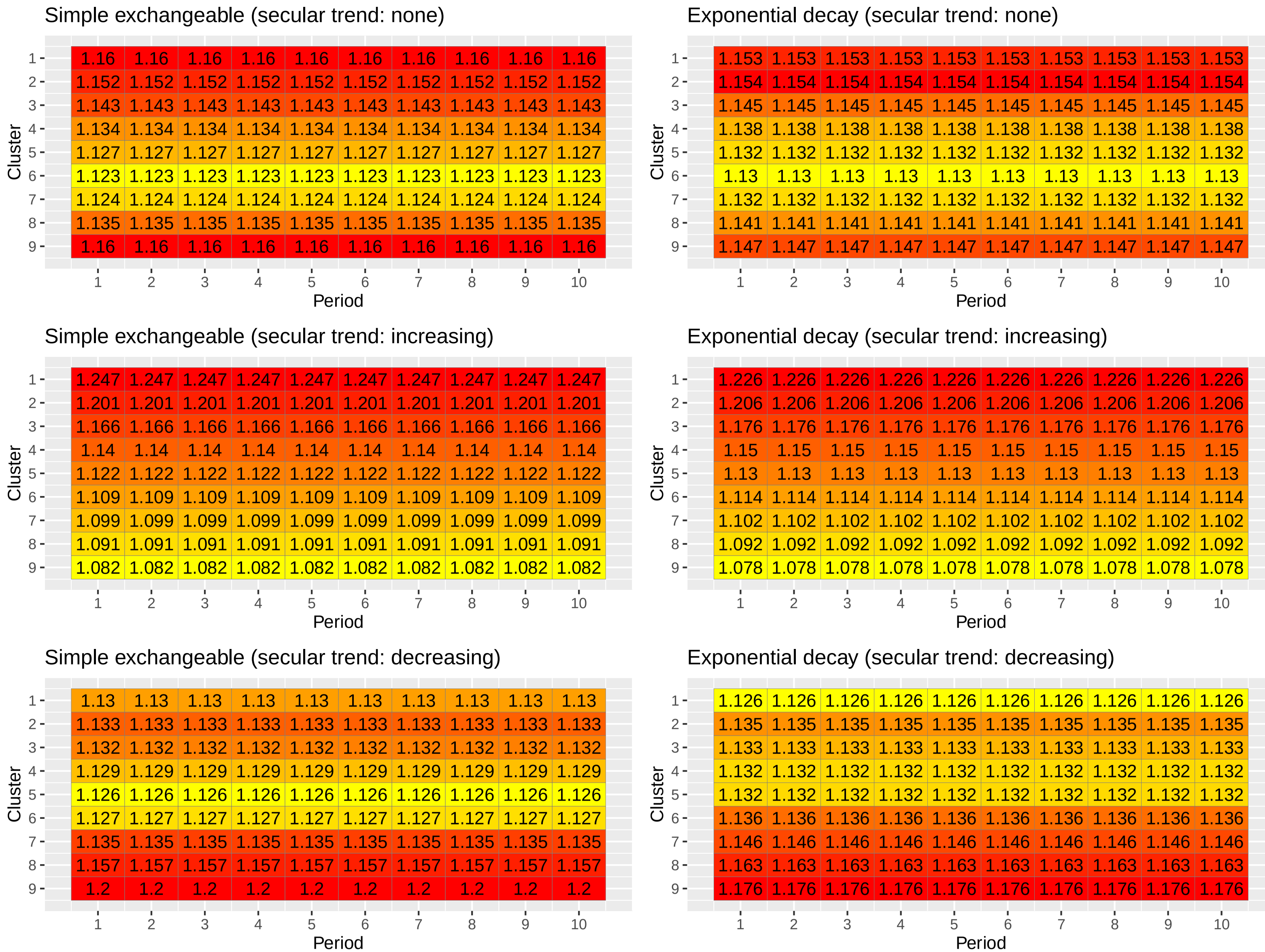
Web Figure 2: Information content of cells under two correlation models and three secular trends with a binary outcome and logistic link. The baseline prevalence is $e^{\beta_1}/(1+e^{\beta_1}) = 50\%$.



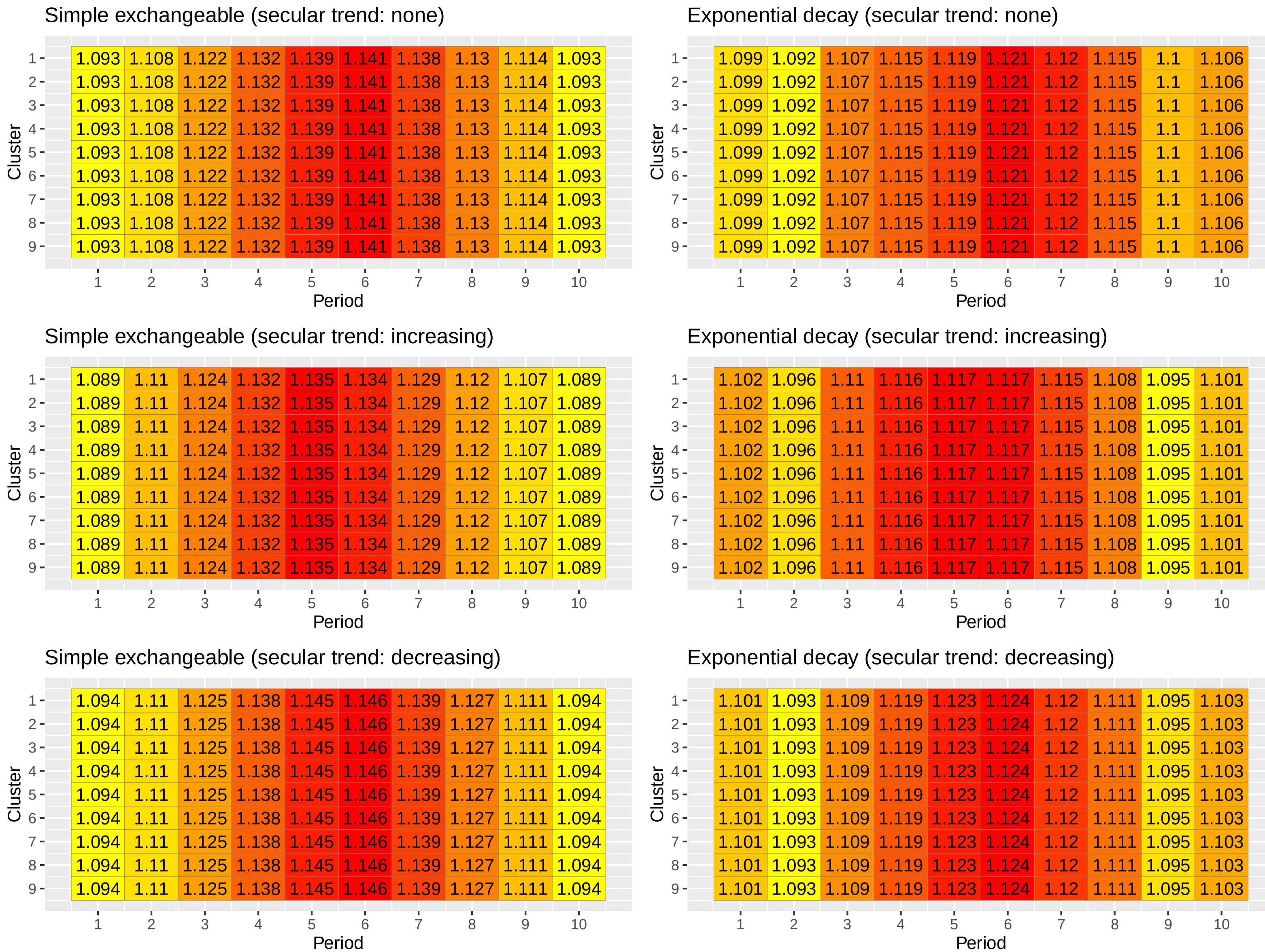
Web Figure 3: Information content of cells under two correlation models and three secular trends with a binary outcome and logistic link. The baseline prevalence is $e^{\beta_1}/(1+e^{\beta_1}) = 80\%$.



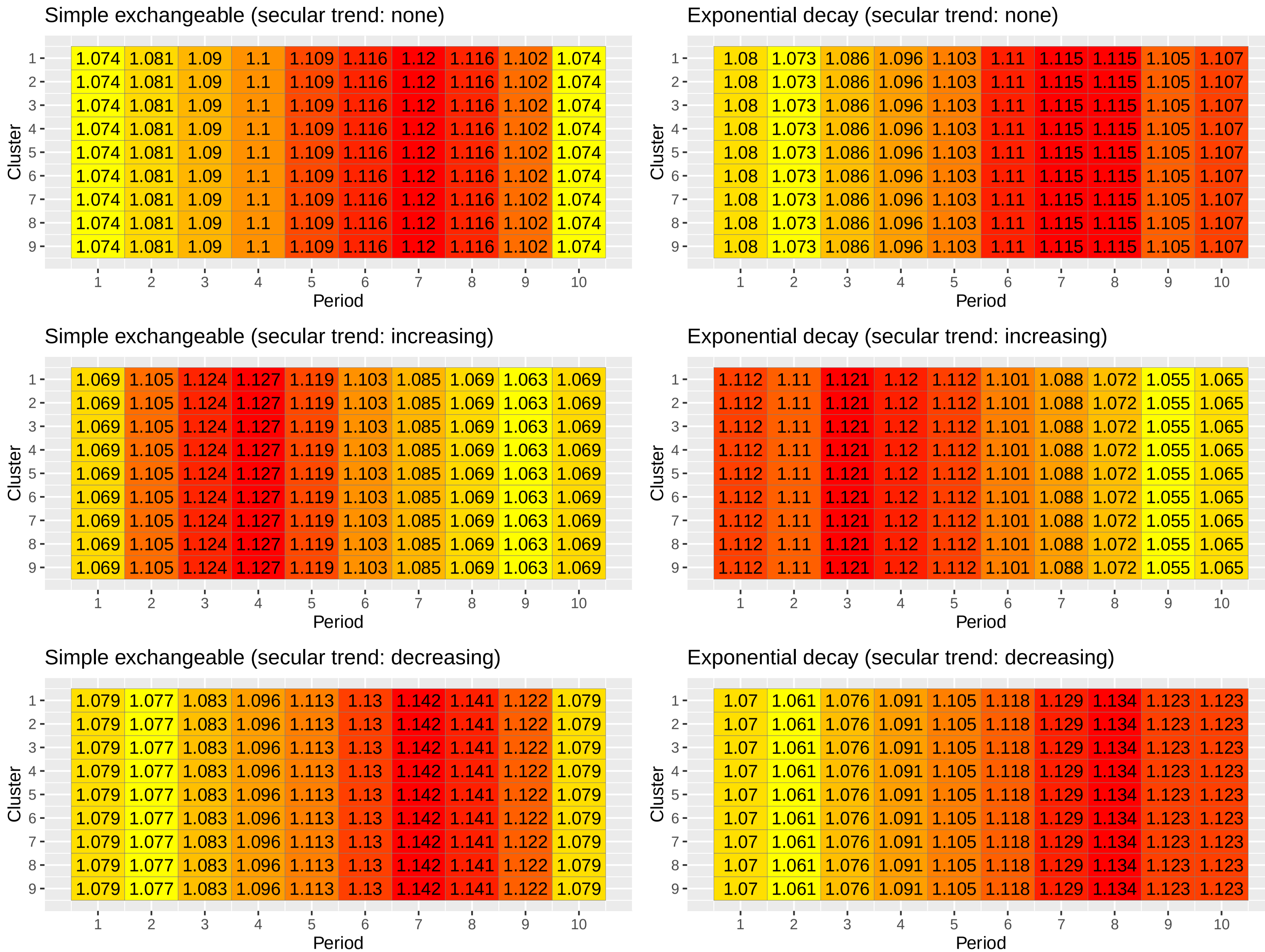
Web Figure 4: Information content of sequences under two correlation models and three secular trends with a binary outcome and logistic link. The baseline prevalence is $e^{\beta_1}/(1+e^{\beta_1}) = 50\%$.



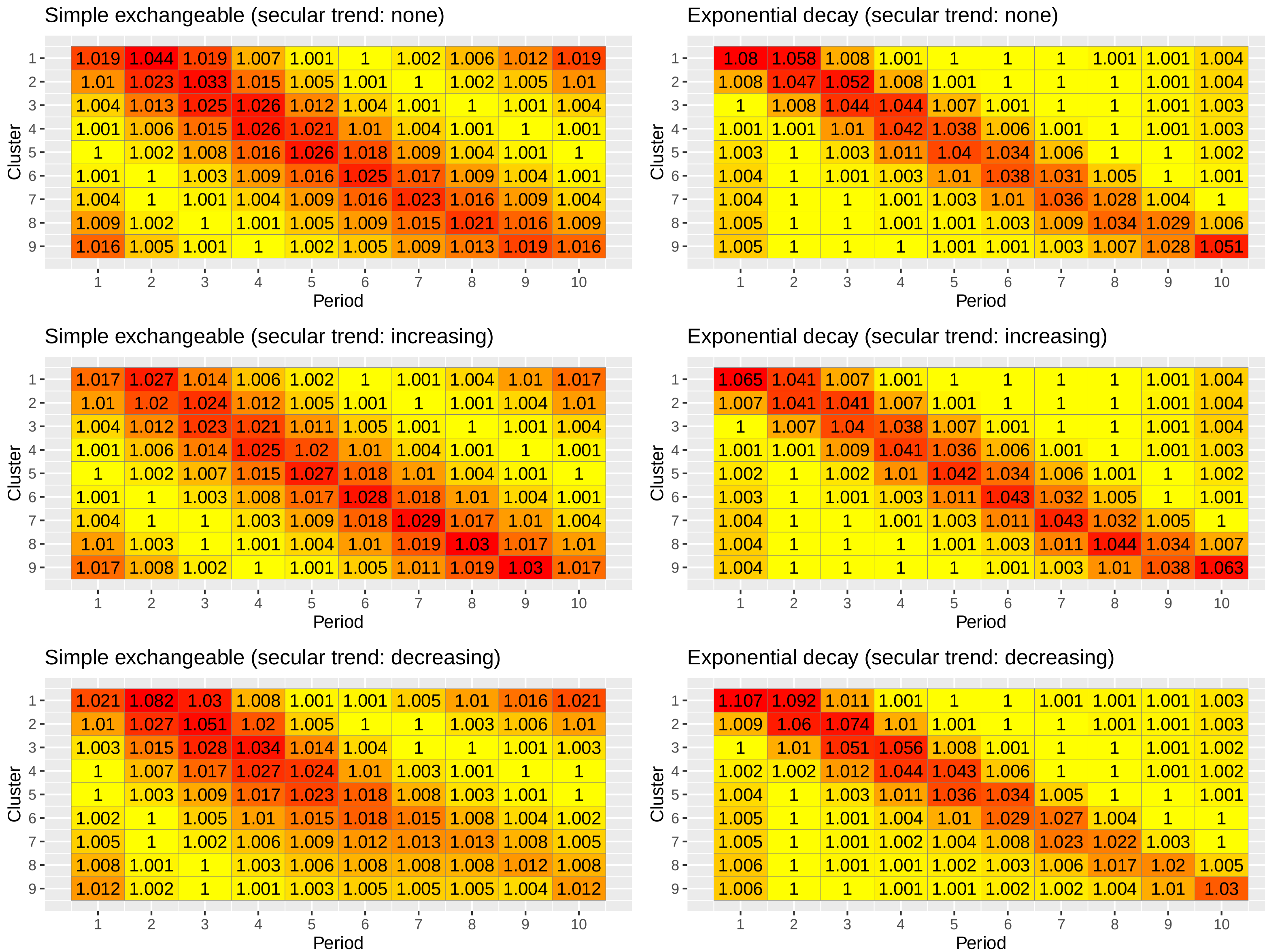
Web Figure 5: Information content of sequences under two correlation models and three secular trends with a binary outcome and logistic link. The baseline prevalence is $e^{\beta_1}/(1 + e^{\beta_1}) = 80\%$.



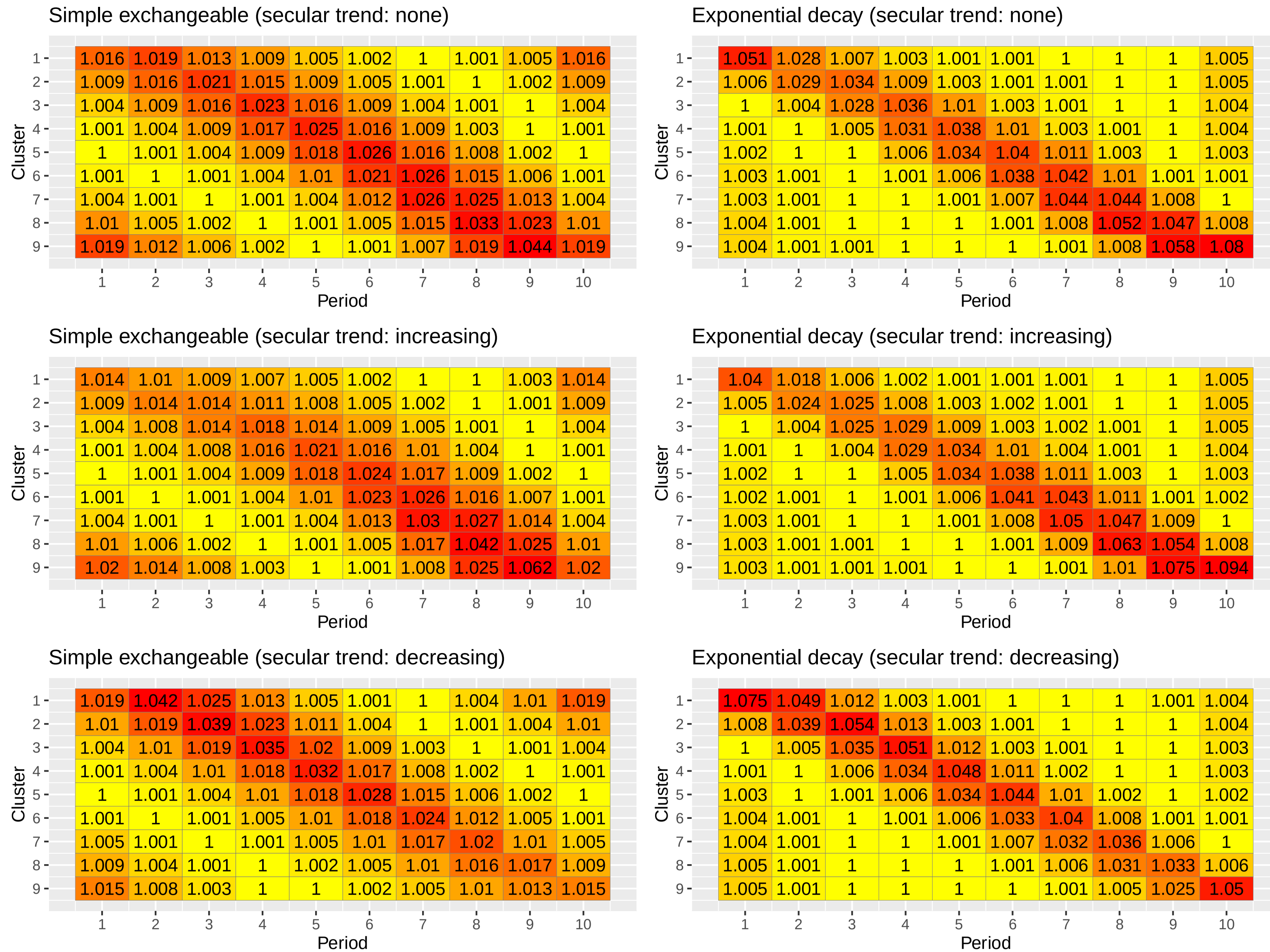
Web Figure 6: Information content of periods under two correlation models and three secular trends with a binary outcome and logistic link. The baseline prevalence is $e^{\beta_1}/(1+e^{\beta_1}) = 50\%$.



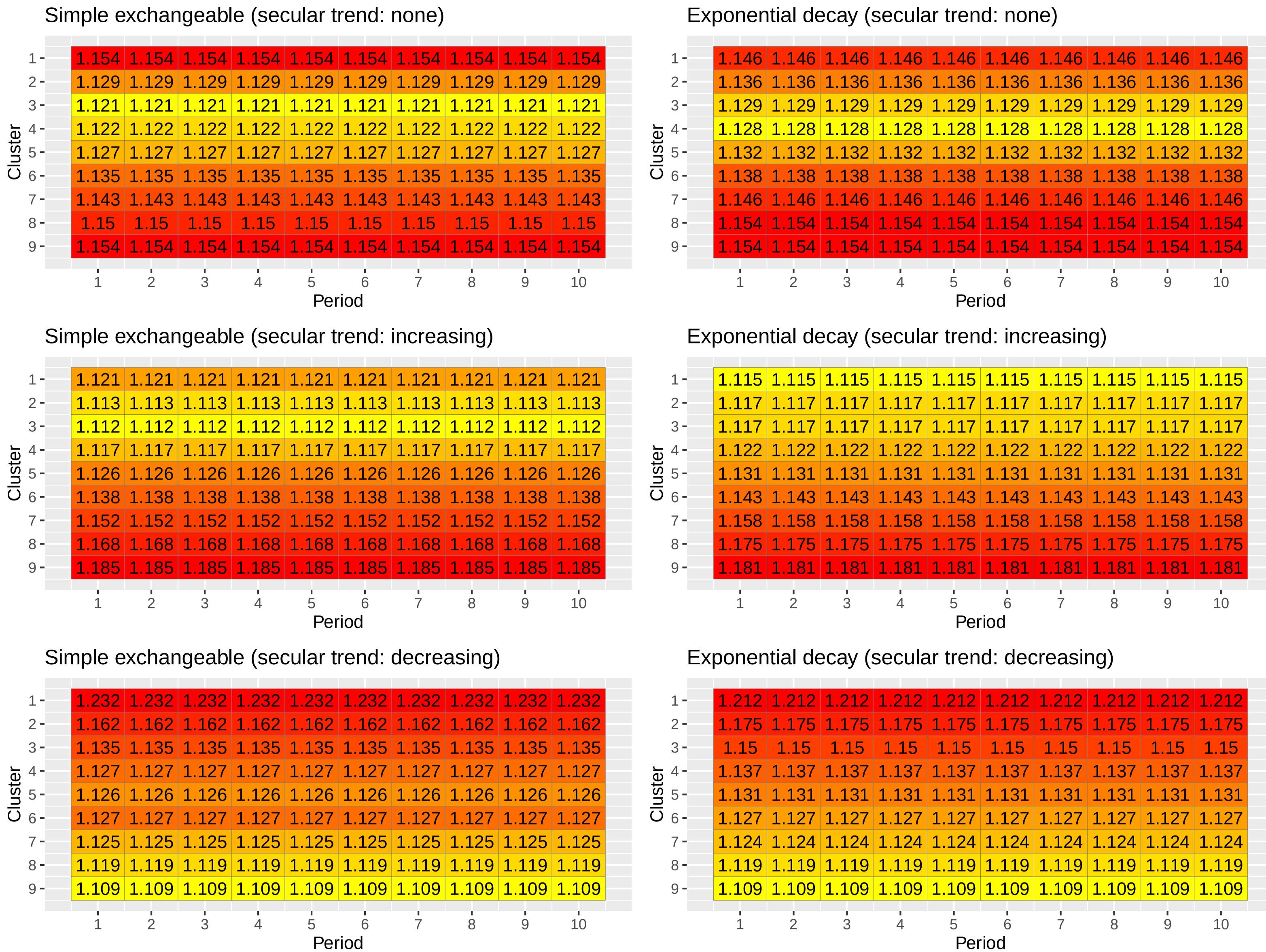
Web Figure 7: Information content of periods under two correlation models and three secular trends with a binary outcome and logistic link. The baseline prevalence is $e^{\beta_1}/(1+e^{\beta_1}) = 80\%$.



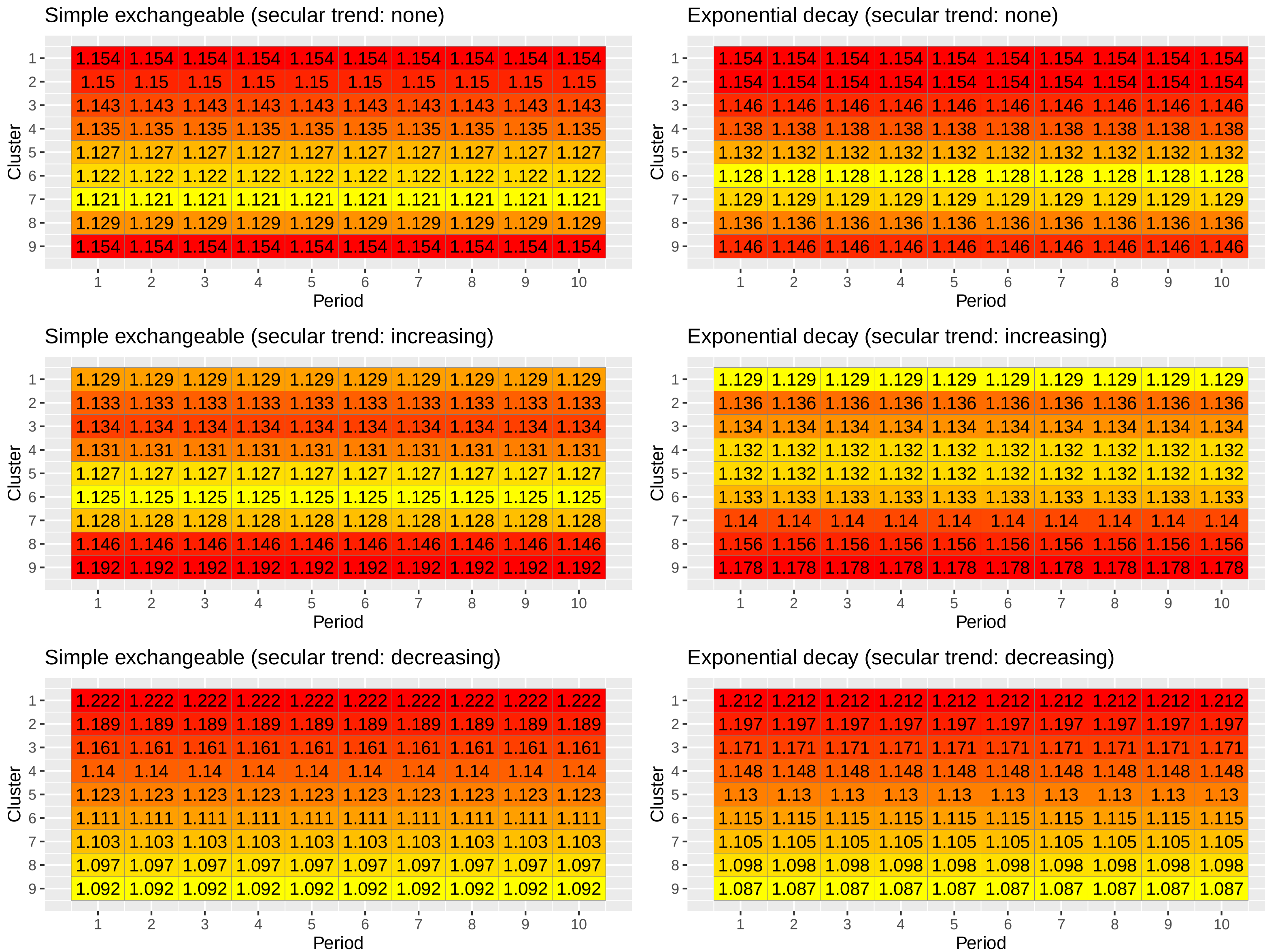
Web Figure 8: Information content of cells under two correlation models and three secular trends with a count outcome and log link. The baseline event rate is $e^{\beta_1} = 1.5$ and the treatment effect in log rate ratio is $\delta = \log(2)$.



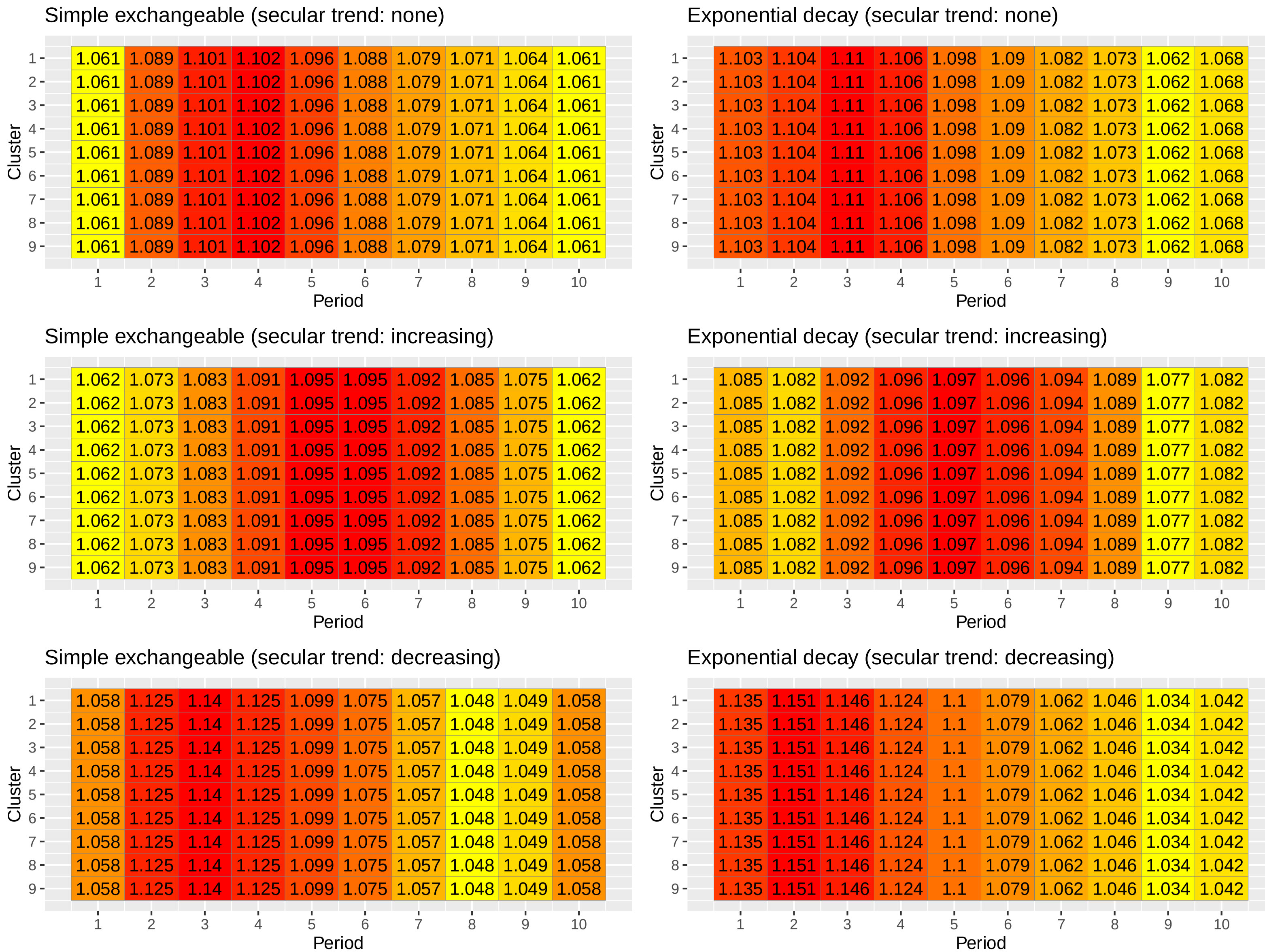
Web Figure 9: Information content of cells under two correlation models and three secular trends with a count outcome and log link. The baseline event rate is $e^{\beta_1} = 1.5$ and the treatment effect in log rate ratio is $\delta = \log(1/2)$.



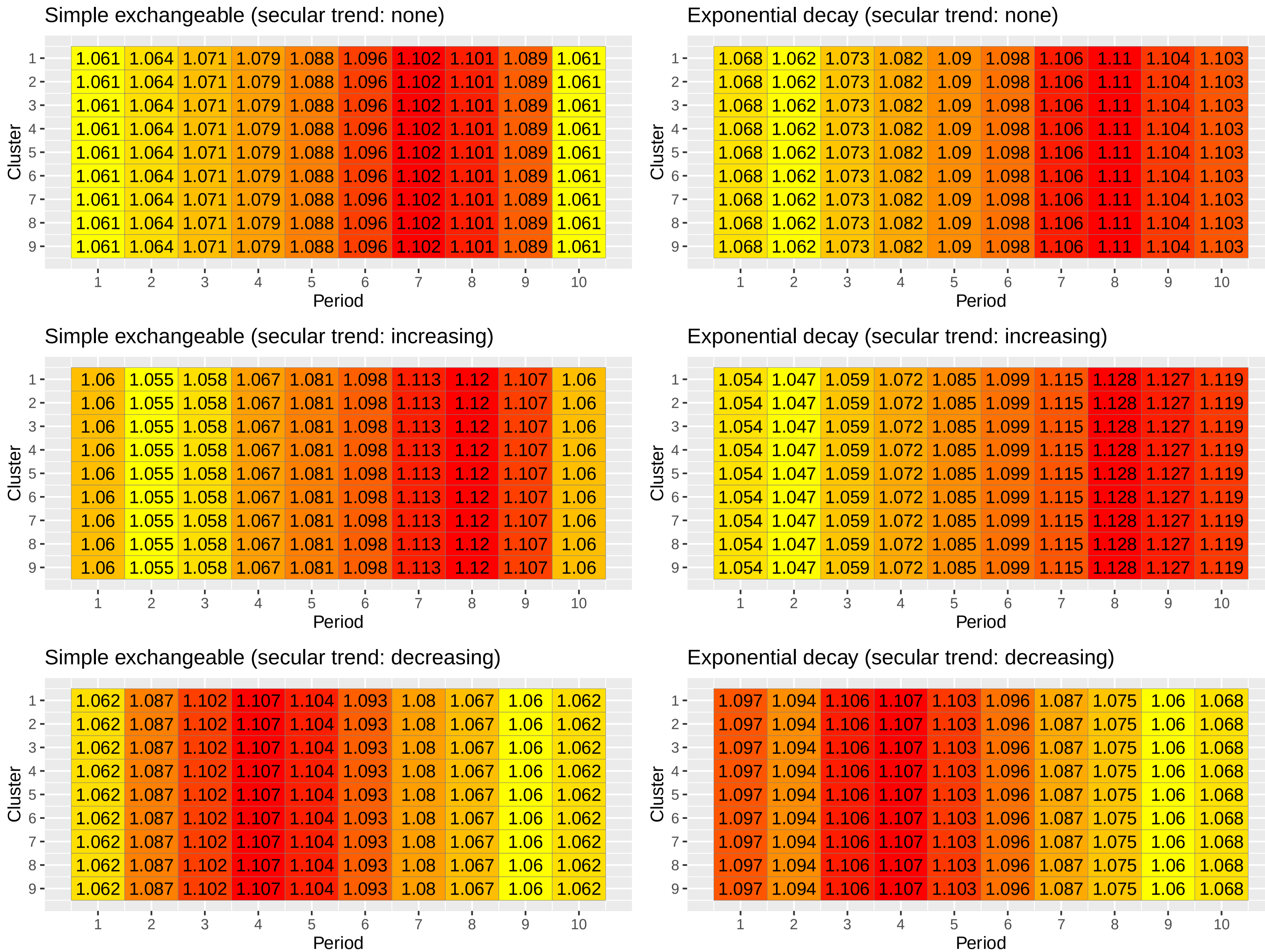
Web Figure 10: Information content of sequences under two correlation models and three secular trends with a count outcome and log link. The baseline event rate is $e^{\beta_1} = 1.5$ and the treatment effect in log rate ratio is $\delta = \log(2)$.



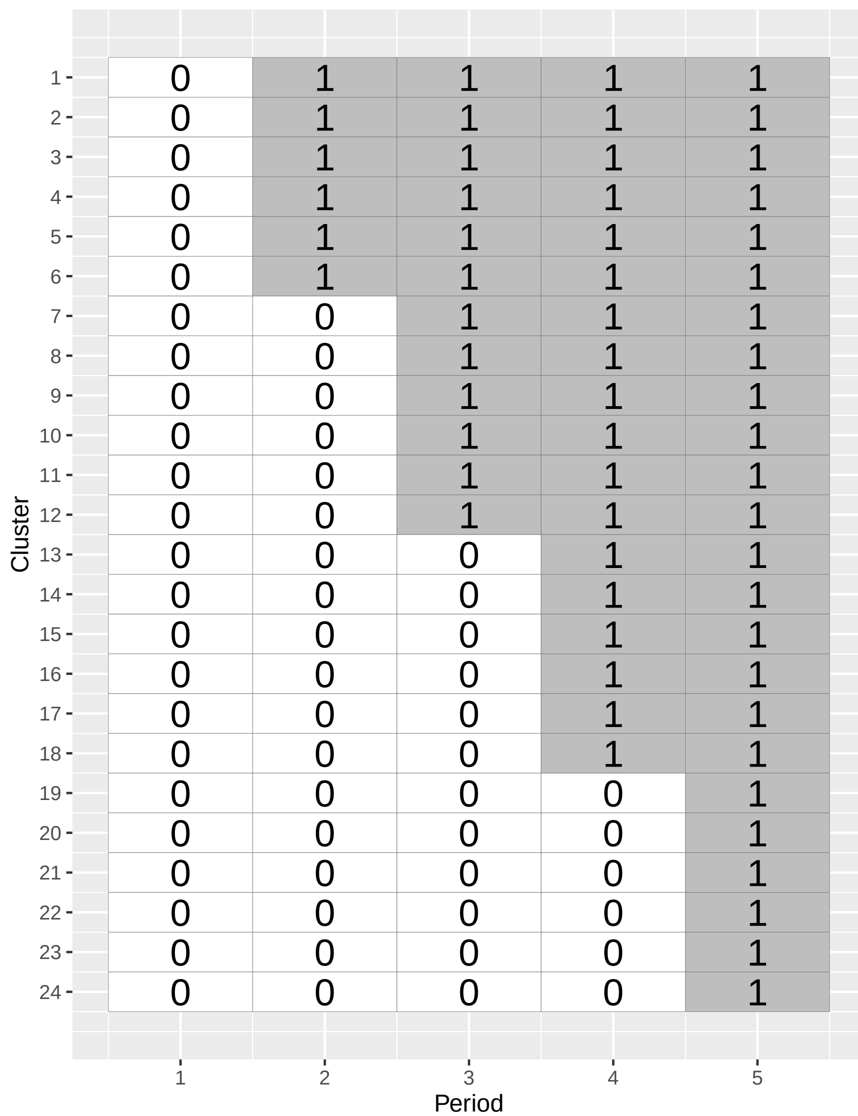
Web Figure 11: Information content of sequences under two correlation models and three secular trends with a count outcome and log link. The baseline event rate is $e^{\beta_1} = 1.5$ and the treatment effect in log rate ratio is $\delta = \log(1/2)$.



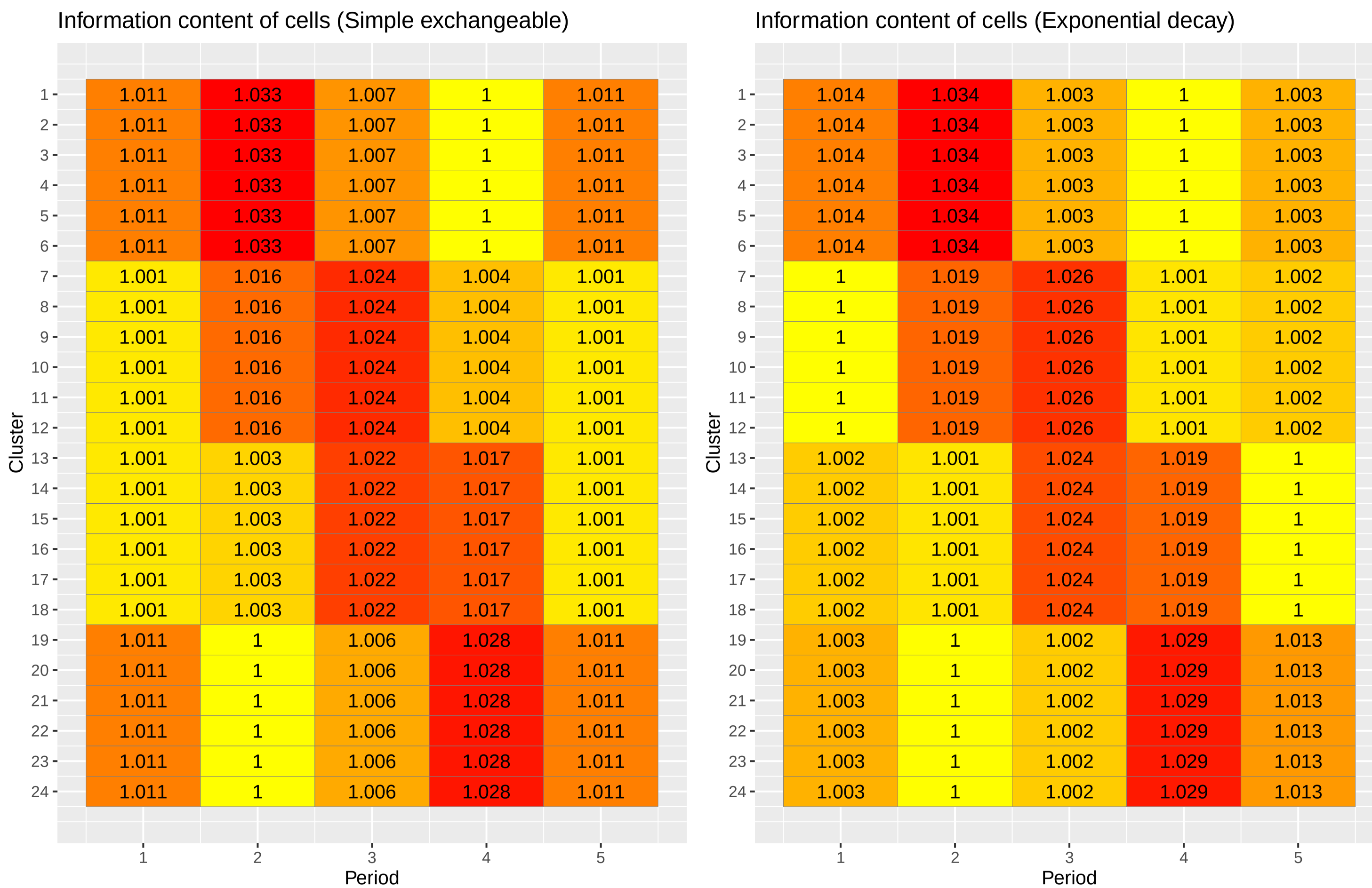
Web Figure 12: Information content of periods under two correlation models and three secular trends with a count outcome and log link. The baseline event rate is $e^{\beta_1} = 1.5$ and the treatment effect in log rate ratio is $\delta = \log(2)$.



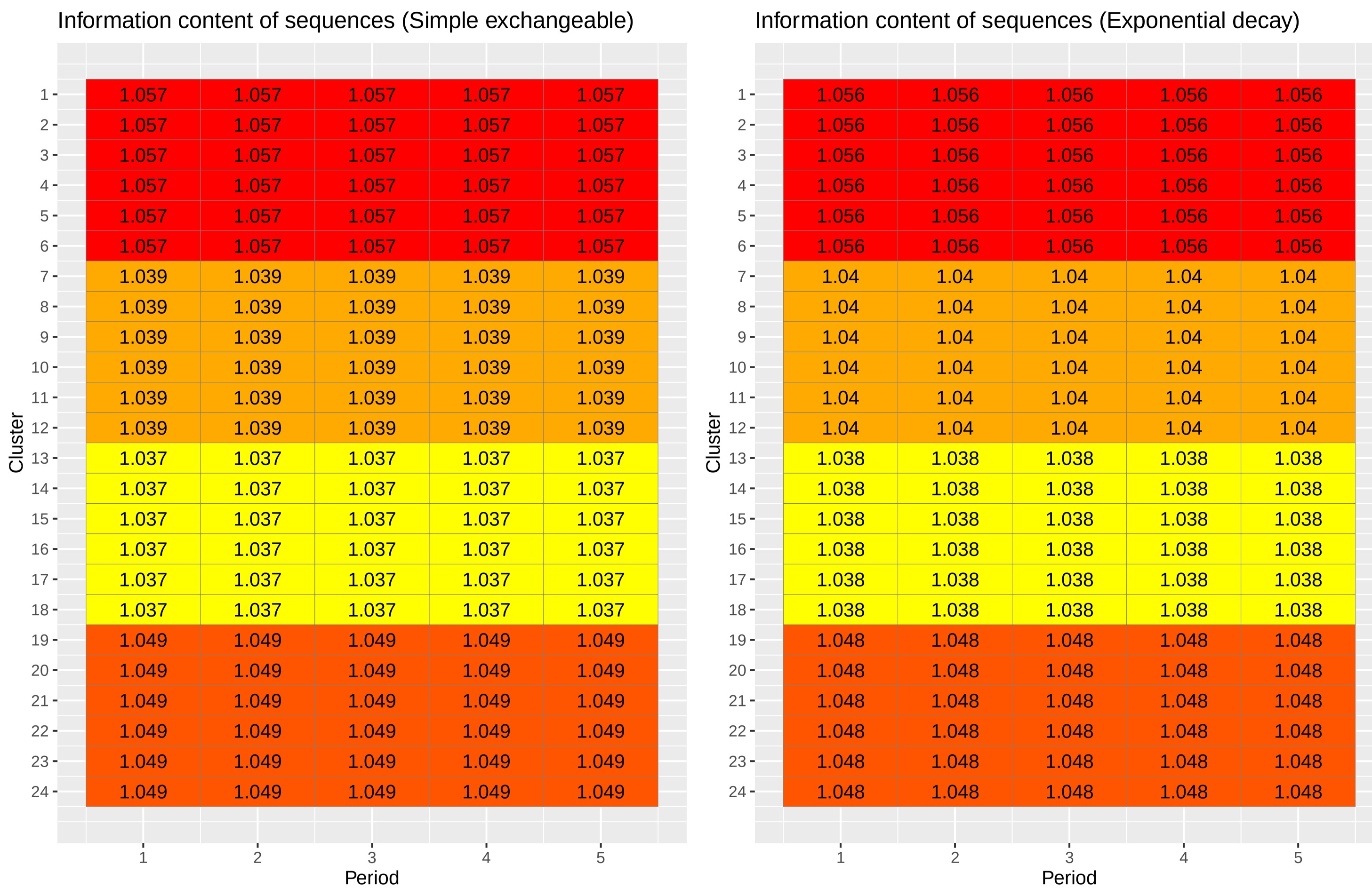
Web Figure 13: Information content of periods under two correlation models and three secular trends with a count outcome and log link. The baseline event rate is $e^{\beta_1} = 1.5$ and the treatment effect in log rate ratio is $\delta = \log(1/2)$.



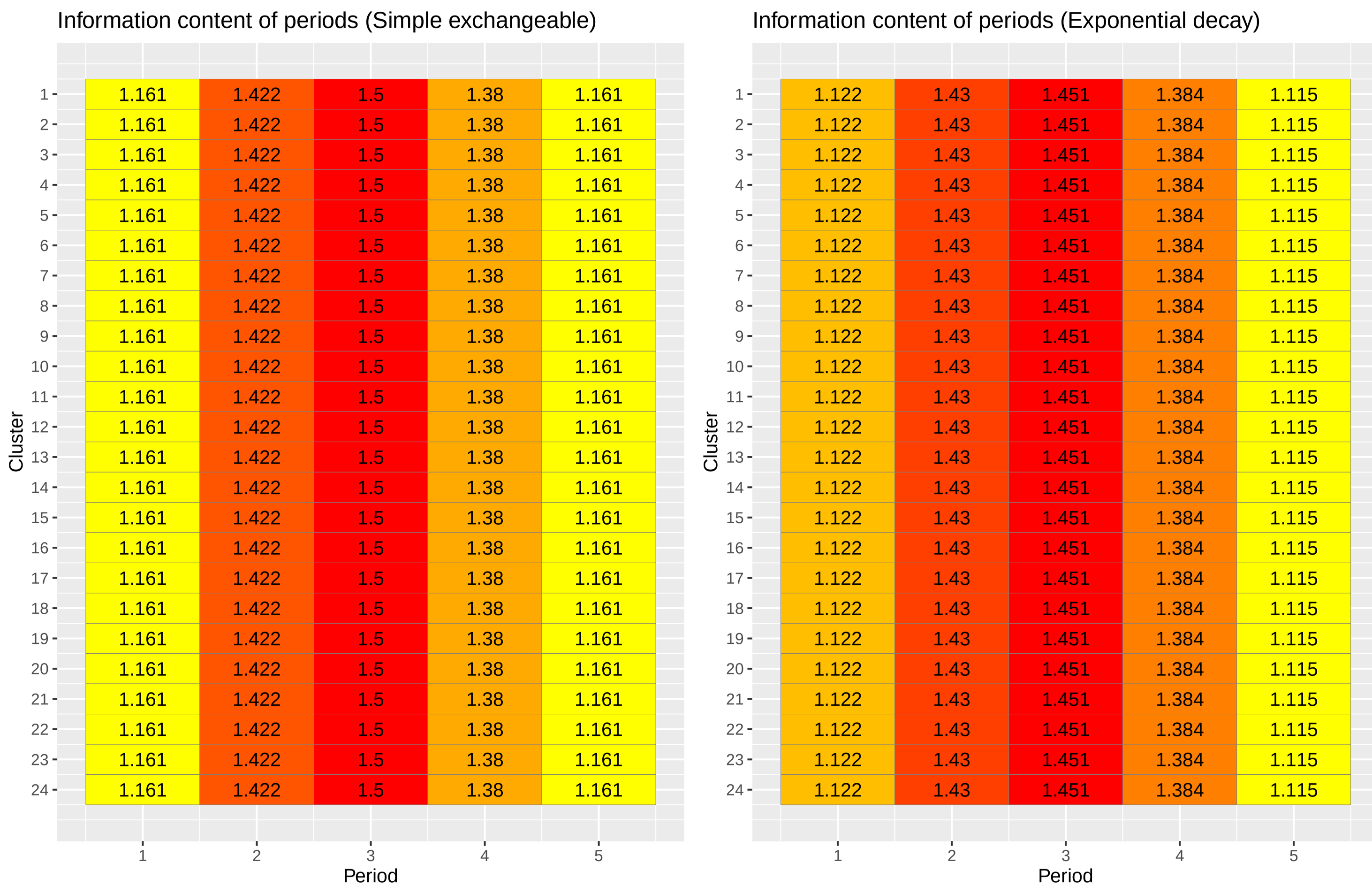
Web Figure 14: Cluster-by-period treatment design matrix of the Washington State EPT study. Each row represents a cluster and each column represents a period, with '1' in each cell denoting treatment and '0' denoting control.



Web Figure 15: Information content of each cluster-period cell in the Washington State EPT trial.



Web Figure 16: Information content of each cluster in the Washington State EPT trial.



Web Figure 17: Information content of each period in the Washington State EPT trial.

References

Deng, C. Y. (2011). A generalization of the sherman–morrison–woodbury formula. *Applied Mathematics Letters* **24,** 1561–1564.

Girling, A. J. and Hemming, K. (2016). Statistical efficiency and optimal design for stepped cluster studies under linear mixed effects models. *Statistics in Medicine* **35,** 2149–2166.

Hemming, K., Taljaard, M., and Forbes, A. (2018). Modeling clustering and treatment effect heterogeneity in parallel and stepped-wedge cluster randomized trials. *Statistics in Medicine* **37,** 883–898.

Henderson, H. V. and Searle, S. R. (1981). On deriving the inverse of a sum of matrices. *Siam Review* **23,** 53–60.

Hooper, R., Teerenstra, S., de Hoop, E., and Eldridge, S. (2016). Sample size calculation for stepped wedge and other longitudinal cluster randomised trials. *Statistics in Medicine* **35,** 4718–4728.

Hughes, J. P., Granston, T. S., and Heagerty, P. J. (2015). Current issues in the design and analysis of stepped wedge trials. *Contemporary Clinical Trials* **45,** 55–60.

Hussey, M. A. and Hughes, J. P. (2007). Design and analysis of stepped wedge cluster randomized trials. *Contemporary Clinical Trials* **28,** 182–191.

Kasza, J., Hemming, K., Hooper, R., Matthews, J. N., and Forbes, A. B. (2019). Impact of non-uniform correlation structure on sample size and power in multiple-period cluster randomised trials. *Statistical Methods in Medical Research* **28,** 703–716.

Leiva, R. (2007). Linear discrimination with equicorrelated training vectors. *Journal of multivariate analysis* **98,** 384–409.

Li, F. (2020). Design and analysis considerations for cohort stepped wedge cluster randomized trials with a decay correlation structure. *Statistics in Medicine* **39,** 438–455.

Li, F., Hughes, J. P., Hemming, K., Taljaard, M., Melnick, E. R., and Heagerty, P. J. (2020). Mixed-effects models for the design and analysis of stepped wedge cluster randomized trials: An overview. *Statistical Methods in Medical Research* **30,** 612–639.

Li, F., Turner, E. L., and Preisser, J. S. (2018). Sample size determination for GEE analyses of stepped wedge cluster randomized trials. *Biometrics* **74,** 1450–1458.

Liu, A., Shih, W. J., and Gehan, E. (2002). Sample size and power determination for clustered repeated measurements. *Statistics in Medicine* **21,** 1787–1801.

Preisser, J. S. and Qaqish, B. F. (1996). Deletion diagnostics for generalised estimating equations. *Biometrika* **83,** 551–562.