Supporting information for: "Regression type models for extremal dependence"

L. Mhalla[1] | M. de Carvalho[2] | V. Chavez-Demoulin[3]

[1]Geneva School of Economics and Management, Université de Genève, Genève, Switzerland

[2]School of Mathematics, University of Edinburgh, Edinburgh, UK

[3]HEC Lausanne, Université de Lausanne, Lausanne, Switzerland

1 | AUXILIARY LEMMAS AND PROOFS

Recall the following notations

$$\begin{aligned} \ell(\boldsymbol{\beta}) &= \sum_{i=1}^{n_{\mathbf{r}}} c_i + \log\{h(\mathbf{w}_i;\boldsymbol{\beta})\}, \\ \mathbf{m}(\boldsymbol{\beta}) &= \frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}, \\ \mathbf{m}(\mathbf{w},\boldsymbol{\beta}) &= \frac{\partial \log\{h(\mathbf{w};\boldsymbol{\beta})\}}{\partial \boldsymbol{\beta}}, \end{aligned}$$

where c_i is a constant independent of $\boldsymbol{\beta}$, for $i = 1, \dots, n_{\mathbf{r}}$.
As mentioned in the paper, the penalized log-likelihood estimator (PMLE) $\hat{\boldsymbol{\beta}}$ satisfies the following score equation

$$\mathbf{m}(\boldsymbol{\beta}) - \mathbf{P}(\boldsymbol{\gamma})\boldsymbol{\beta} = \mathbf{0}_{p(1+q\tilde{d})}. \tag{1}$$

We now define

$$\boldsymbol{\phi}(\boldsymbol{\beta}) = \frac{\partial^2 \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}^\top}, \quad \boldsymbol{\phi}(\mathbf{w},\boldsymbol{\beta}) = \frac{\partial^2 \log\{h(\mathbf{w};\boldsymbol{\beta})\}}{\partial \boldsymbol{\beta}\partial \boldsymbol{\beta}^\top}.$$

Based on Assumption (A2), we prove the following two lemmas that will streamline the proof of the first part of Theorem 1, i.e., the weak consistency of $\hat{\boldsymbol{\beta}}$.

Lemma 1 *Let $h(\mathbf{w};\boldsymbol{\beta})$ be continuously differentiable a.e. for $\boldsymbol{\beta} \in \mathbf{B}$. If $\int \sup_{\boldsymbol{\beta}\in\mathbf{B}} \| \partial h(\mathbf{w};\boldsymbol{\beta})/\partial\boldsymbol{\beta} \| \, \mathrm{d}\mathbf{w} < \infty$, then for $\boldsymbol{\beta} \in \mathbf{B}$:*

1. *$\int h(\mathbf{w};\boldsymbol{\beta}) \, \mathrm{d}\mathbf{w}$ is continuously differentiable.*
2. *$\int \partial h(\mathbf{w};\boldsymbol{\beta})/\partial\boldsymbol{\beta} \, \mathrm{d}\mathbf{w} = \partial \int h(\mathbf{w};\boldsymbol{\beta}) \, \mathrm{d}\mathbf{w}/\partial\boldsymbol{\beta}$.*

Proof See Newey and McFadden (1994, Lemma 3.6).

Lemma 2 *If (A2) holds, then*

$$\mathrm{E}\left\{\mathbf{m}(\mathbf{W},\beta_0)\right\} = 0, \quad -\mathrm{E}\left\{\boldsymbol{\phi}(\mathbf{W},\beta_0)\right\} = \mathbf{i}(\beta_0).$$

Proof By Lemma 1, it follows that

$$\mathrm{E}\left\{\mathbf{m}(\mathbf{W},\beta_0)\right\} = \int \mathbf{m}(\mathbf{w},\beta_0)h(\mathbf{w};\beta_0)\,\mathrm{d}\mathbf{w} = \int \frac{\partial h(\mathbf{w};\beta)}{\partial \beta}\Big|_{\beta=\beta_0} \mathrm{d}\mathbf{w} = \frac{\partial \int h(\mathbf{w};\beta)\,\mathrm{d}\mathbf{w}}{\partial \beta}\Big|_{\beta=\beta_0} = 0.$$

Now, using (A2) (parts 3 and 4) we have that

$$\begin{aligned}\mathrm{E}\left\{\boldsymbol{\phi}(\mathbf{W},\beta_0)\right\} &= \int \boldsymbol{\phi}(\mathbf{w},\beta_0)h(\mathbf{w};\beta_0)\,\mathrm{d}\mathbf{w}\\ &= \int \frac{\partial \mathbf{m}(\mathbf{w},\beta)h(\mathbf{w};\beta)}{\partial \beta^\top}\Big|_{\beta=\beta_0} h(\mathbf{w};\beta_0)\,\mathrm{d}\mathbf{w} - \int \mathbf{m}(\mathbf{w},\beta_0)\mathbf{m}(\mathbf{w},\beta_0)^\top h(\mathbf{w};\beta_0)\,\mathrm{d}\mathbf{w}\\ &= -\int \mathbf{m}(\mathbf{w},\beta_0)\mathbf{m}(\mathbf{w},\beta_0)^\top h(\mathbf{w};\beta_0)\,\mathrm{d}\mathbf{w}.\end{aligned}$$

We now prove the first part of Theorem 1 by proving the following lemma.

Lemma 3 *Let $Q_{\mathbf{a}} = \{\beta \in \mathbf{B} : \beta = \beta_0 + n_{\mathbf{r}}^{-1/2}\mathbf{a}\}$ be the surface of the sphere around β_0 with radius $n_{\mathbf{r}}^{-1/2}\|\mathbf{a}\|$. Then, for every $\varepsilon > 0$, there exists $\mathbf{a}$ such that*

$$\Pr\left\{\sup_{\beta\in Q_{\mathbf{a}}} \ell(\beta,\boldsymbol{\gamma}) < \ell(\beta_0,\boldsymbol{\gamma})\right\} \geq 1-\varepsilon,$$

for $n_{\mathbf{r}}$ large enough.

Proof Let $\beta \in Q_a$, i.e., there exists $\mathbf{a}$ such that $\beta = \beta_0 + n_{\mathbf{r}}^{-1/2}\mathbf{a}$. Applying a second-order Taylor expansion around β_0 of the penalized log-likelihood $\ell(\beta,\boldsymbol{\gamma})$, we have

$$\begin{aligned}\ell(\beta,\boldsymbol{\gamma}) - \ell(\beta_0,\boldsymbol{\gamma}) &= \ell(\beta) - \ell(\beta_0) - \frac{1}{2}\left\{\beta^\top \mathbf{P}(\boldsymbol{\gamma})\beta - \beta_0^\top \mathbf{P}(\boldsymbol{\gamma})\beta_0\right\}\\ &= n_{\mathbf{r}}^{-1/2}\mathbf{m}(\beta_0)^\top \mathbf{a} + \frac{n_{\mathbf{r}}^{-1}}{2}\mathbf{a}^\top \boldsymbol{\phi}(\beta_0)\mathbf{a} - \frac{n_{\mathbf{r}}^{-1}}{2}\mathbf{a}^\top \mathbf{P}(\boldsymbol{\gamma})\mathbf{a} - n_{\mathbf{r}}^{-1/2}\beta_0^\top \mathbf{P}(\boldsymbol{\gamma})\mathbf{a}\\ &\quad + \frac{n_{\mathbf{r}}^{-3/2}}{2}\sum_q\sum_r\sum_s a_q a_r a_s \frac{\partial^3 \ell(\beta)}{\partial \beta_{qrs}}\Big|_{\beta=\beta^*},\end{aligned} \tag{2}$$

with β^* in the interior of $Q_{\mathbf{a}}$. The terms involving the penalty matrix $\mathbf{P}(\boldsymbol{\gamma})$ converge in probability to 0 due to the vanishing penalty from (A1). By the central limit theorem (based on a Lindeberg-type condition), Lemma 2, and (A2), we have that $n_{\mathbf{r}}^{-1/2}\mathbf{m}(\beta_0) \xrightarrow{d} N(0,\mathbf{i}(\beta_0))$ implying that $|n_{\mathbf{r}}^{-1/2}\mathbf{m}(\beta_0)^\top \mathbf{a}| = O_p(1)\|\mathbf{a}\|$. By the law of large numbers and Lemma 2, we have that $n_{\mathbf{r}}^{-1}\boldsymbol{\phi}(\beta_0) \xrightarrow{p} -\mathbf{i}(\beta_0)$. Hence, applying the continuous mapping theorem, we end up with $\mathbf{a}^\top n_{\mathbf{r}}^{-1}\boldsymbol{\phi}(\beta_0)\mathbf{a}/2 \xrightarrow{p} -\mathbf{a}^\top \mathbf{i}(\beta_0)\mathbf{a}/2 \leq \|\mathbf{a}\|^2\lambda_{\min}/2$, where $\lambda_{\min} > 0$ is the smallest eigenvalue of $\mathbf{i}(\beta_0)$. Finally, Assumption (A2) (last part) implies that the terms $\partial^3\ell(\beta)/\partial\beta_{qrs}|_{\beta=\beta^*} < \infty$ and that, by Cauchy–Schwartz inequality, the remainder term in (2) vanishes in probability (is $O_p(n_{\mathbf{r}}^{-1/2})$). Leaving out the terms vanishing in probability, (2) yields

$$\ell(\boldsymbol{\beta},\boldsymbol{\gamma}) - \ell(\boldsymbol{\beta}_0,\boldsymbol{\gamma}) \leq O_p(1)\|\mathbf{a}\| - \|\mathbf{a}\|^2\lambda_{\min}/2 = T, \quad \boldsymbol{\beta} \in Q_{\mathbf{a}},$$

for $n_{\mathbf{r}}$ large enough. Thus,

$$\Pr\left\{\sup_{\boldsymbol{\beta}\in Q_{\mathbf{a}}} \ell(\boldsymbol{\beta},\boldsymbol{\gamma}) < \ell(\boldsymbol{\beta}_0,\boldsymbol{\gamma})\right\} \geq \Pr(T < 0),$$

implying that for every $\varepsilon > 0$, there exists an $\mathbf{a}$ such that $\Pr\{\sup_{\boldsymbol{\beta}\in Q_{\mathbf{a}}} \ell(\boldsymbol{\beta},\boldsymbol{\gamma}) < \ell(\boldsymbol{\beta}_0,\boldsymbol{\gamma})\} \geq 1-\varepsilon$. Hence, with probability tending to 1, the penalized log-likelihood $\ell(\boldsymbol{\beta},\boldsymbol{\gamma})$ has a local maximum $\hat{\boldsymbol{\beta}}$ in the interior of a sphere around $\boldsymbol{\beta}_0$.

Lemma 3 yields the first part of Theorem 1. Now we move to the second part of Theorem 1. The asymptotic normality of the PMLE $\hat{\boldsymbol{\beta}}$ is derived from a second-order Taylor expansion of the score equation (1) around the true parameter $\boldsymbol{\beta}_0$. The Taylor expansion of (1) yields

$$\mathbf{m}(\boldsymbol{\beta}_0) - \mathbf{P}(\boldsymbol{\gamma})\boldsymbol{\beta}_0 + \{\boldsymbol{\phi}(\boldsymbol{\beta}_0) - \mathbf{P}(\boldsymbol{\gamma})\}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) + \mathbf{r} = \mathbf{0}_{p(1+q\tilde{d})},$$

where

$$\mathbf{r} = \frac{1}{2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^\top \frac{\partial^2 \mathbf{m}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^\top}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0), \tag{3}$$

and $\boldsymbol{\beta}^*$ is such that $\|\boldsymbol{\beta}^* - \boldsymbol{\beta}_0\| \leq \|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\|$. Dividing (3) by $n_{\mathbf{r}}^{1/2}$, we obtain

$$\frac{1}{n_{\mathbf{r}}}\{\boldsymbol{\phi}(\boldsymbol{\beta}_0) - \mathbf{P}(\boldsymbol{\gamma}) + \tilde{\mathbf{r}}\}n_{\mathbf{r}}^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) = n_{\mathbf{r}}^{-1/2}\{\mathbf{P}(\boldsymbol{\gamma})\boldsymbol{\beta}_0 - \mathbf{m}(\boldsymbol{\beta}_0)\}, \tag{4}$$

where

$$\tilde{\mathbf{r}} = \frac{1}{2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^\top \frac{\partial^2 \mathbf{m}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^\top}\Big|_{\boldsymbol{\beta}=\boldsymbol{\beta}^*}.$$

The consistency of $\hat{\boldsymbol{\beta}}$ and the assumption on the third order derivative of $\log\{h(\mathbf{w};\boldsymbol{\beta})\}$ in (A2) implies that $\tilde{\mathbf{r}} \xrightarrow{p} 0$. Assumption (A1) implies that the terms involving $\mathbf{P}(\boldsymbol{\gamma})$ vanish in probability. Since $n_{\mathbf{r}}^{-1/2}\mathbf{m}(\boldsymbol{\beta}_0) \xrightarrow{d} N(0, \mathbf{i}(\boldsymbol{\beta}_0))$ and $n_{\mathbf{r}}^{-1}\boldsymbol{\phi}(\boldsymbol{\beta}_0) \xrightarrow{p} -\mathbf{i}(\boldsymbol{\beta}_0)$ (see above), Slutsky's theorem implies that $n_{\mathbf{r}}^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \xrightarrow{d} N(0, \mathbf{i}(\boldsymbol{\beta}_0)^{-1})$ and proves hence the second part of Theorem 1.

2 | EXTREME TEMPERATURE ANALYSIS

This section supplements Section 5.2 in the paper.

Dependence of extreme high winter temperatures

The Dirichlet model of Table 2 is fitted to the pseudo-sample of extreme high temperatures where the angular observations corresponding to a radial component exceeding its 90%, 93%, and 97% quantiles, are considered. Figure

1 displays the fitted smooth effects of time, NAO, and day in season on the extremal coefficient, along with their associated 95% confidence intervals.

Dependence of extreme low winter temperatures

The Dirichlet model of Table 3 is fitted to the pseudo-sample of extreme low temperatures where the angular observations corresponding to a radial component exceeding its 90%, 93%, and 97% quantiles, are considered. Figure 2 displays the fitted smooth effects of time and day in season on the extremal coefficient, along with their associated 95% confidence intervals.

References

Newey, W. K. and McFadden, D. (1994) Large sample estimation and hypothesis testing. In *Handbook of econometrics* (eds. R. F. Engle and D. McFadden), 2111–2245.

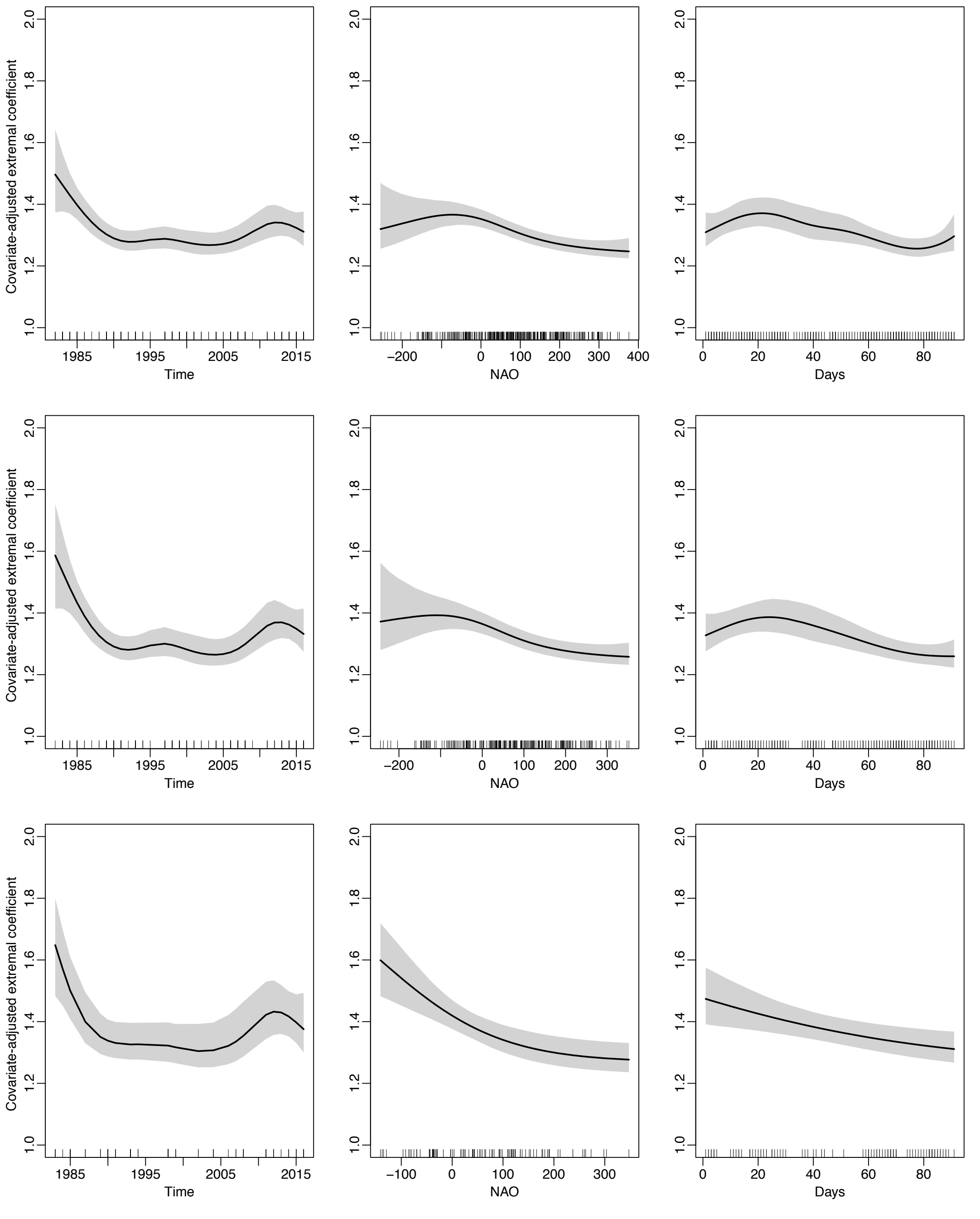

FIGURE 1 Fitted smooth effects for the extremal coefficient under the Dirichlet model of Table 2 along with their associated 95% (pointwise) asymptotic confidence bands. Different radial thresholds are considered: the 90% quantile (top), the 93% quantile (middle), and the 97% quantile (bottom).

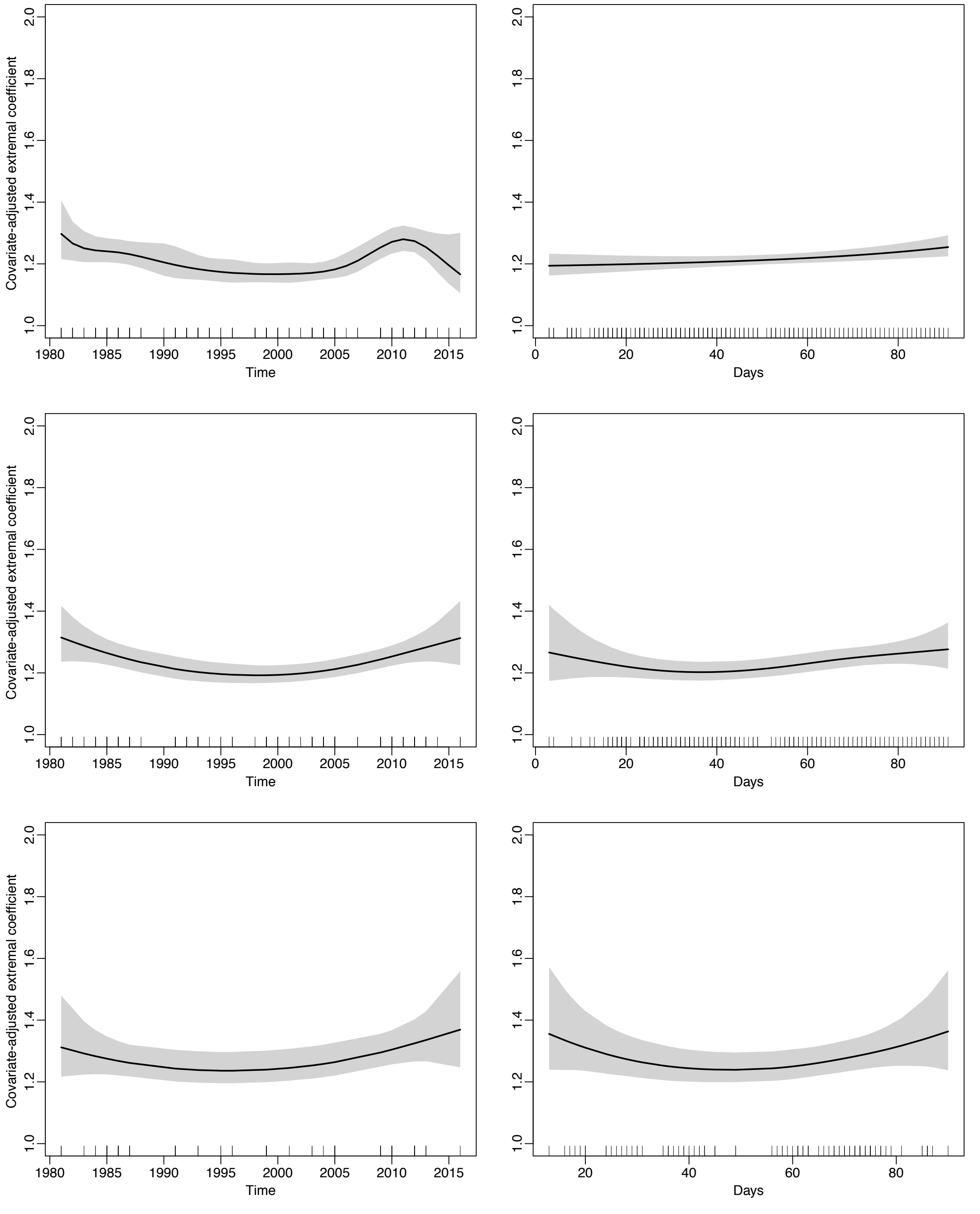

FIGURE 2 Fitted smooth effects for the extremal coefficient under the Dirichlet model of Table 3 along with their associated 95% (pointwise) asymptotic confidence bands. Different radial thresholds are considered: the 90% quantile (top), the 93% quantile (middle), and the 97% quantile (bottom).