Electronic Journal of Statistics
Vol. 0 (0000)
ISSN: 1935-7524
DOI: 10.1214/154957804100000000

Supplementary material for

Monte Carlo modified profile likelihood in models for clustered data

Claudia Di Caterina, Giuliana Cortese and Nicola Sartori

Department of Statistical Sciences, University of Padova
Via Cesare Battisti 241, 35121 Padova, Italy
e-mail: dicaterina@stat.unipd.it; *e-mail:* gcortese@stat.unipd.it;
e-mail: sartori@stat.unipd.it

S1. Introduction

Additional simulation results concerning possible applications of the methodology presented in Section 3 are available here. Specifically, Section S2 deals with the autoregressive model for nonstationary normally-distributed panel data, where computation of the MPL is quite cumbersome. The probit regression model with missing response is instead considered in Section S3, which complements the analysis for the logistic link shown in Section 4.3 of the paper. For both link functions, results obtained within the fixed effects simulation setup are also reported. Finally, Section S4 contains supplementary outcomes of simulations for the stratified Weibull regression under right censoring, described in Section 5.4.

S2. Nonstationary normal AR(1) model

S2.1. Setup and background

Let us consider the nonstationary version of the first-order autoregressive model for normal response

$$Y_{it}|Y_{i,t-1} = y_{i,t-1} \sim N(\lambda_i + \rho y_{i,t-1}, \sigma^2), \qquad i = 1, \dots, N, \quad t = 1, \dots, T, \tag{S1}$$

with $y_0 = (y_{10}, \dots, y_{N0})$ vector of unrestricted given initial conditions. The parameter of interest is $\psi = (\rho, \sigma^2) \in \mathbb{R} \times \mathbb{R}^+$ and $\lambda = (\lambda_1, \dots, \lambda_N) \in \mathbb{R}^N$ denotes the nuisance component of fixed effects. The lack of stationarity of the stochastic process Y_{it} in each ith group ($i = 1, \dots, N$) implies the temporal variation of its mean or its autocovariance function, i.e. the covariance of the response with itself at pairs of time points. As a consequence, the autoregressive parameter ρ is left free to equal or exceed unity and the fixed vector y_0 does not need to meet any specific requirement, so that the log-likelihood function is expressed by conditioning on it. In order to facilitate the presentation, both exogenous covariates and further lagged responses $y_{i,t-l}\,(l > 1)$ are excluded from the set of model regressors; however, no additional difficulties would be encountered in applying the proposed methodology otherwise.

The incidental parameters problem occurring in the analogue stationary AR(1) model has been addressed in the statistical literature several times. Particularly, [6] proved that a marginal likelihood for ψ exists and is equivalent to the MPL introduced by Barndorff-Nielsen [3, 4]. Also econometricians showed interest in this issue and one latest proposition to improve ML inference in fixed effects dynamic models for stationary panel data is given by the bias-corrected estimator of [9], specially tailored for macroeconomic settings with $N = O(T)$.

Here, a great deal of attention is paid to the nonstationarity assumption of model (S1). Analytical derivation of $I_{\lambda_i\lambda_i}(\hat{\theta}_\psi;\hat{\theta})$ in this case would be possible but quite tedious. Instead Monte Carlo approximation dramatically reduces the amount of effort demanded to use Severini's modification. Moreover, we are specifically concerned with datasets where T is much smaller than N, i.e. with situations where $l_P(\psi)$ exhibits its worst performance. Estimation of ψ under these conditions was already investigated in the past. For example, inference in autoregressions of order l was thoroughly examined in [8], who obtained an adjusted profile log-likelihood through integration of a recentered score function. For $l = 1$, their solution is essentially the same as that found in [10, Section 3]. From a purely Bayesian perspective, the latter proposed a strategy to integrate out the incidental parameters from the likelihood in order to derive a marginal posterior density with consistent mode for ψ. Such approach is in substance equivalent to that adopted by [7] to make inference on ρ via the frequentist integrated likelihood of Severini [12].

S2.2. Monte Carlo modified profile likelihood

When groups are supposed independent, the log-likelihood of model (S1) conditioned on the initial vector y_0 is

$$l(\theta) = -\sum_{i=1}^{N}\left\{\frac{T}{2}\log\sigma^2 + \frac{1}{2\sigma^2}\sum_{t=1}^{T}(y_{it}-\lambda_i-\rho y_{i,t-1})^2\right\}, \tag{S2}$$

and the score function takes the form

$$l_{\lambda_i}(\theta) = \frac{1}{\sigma^2}\sum_{t=1}^{T}(y_{it}-\lambda_i-\rho y_{i,t-1}), \qquad i = 1,\ldots,N.$$

Solution for λ_i to the equation $l_{\lambda_i}(\theta) = 0$ delivers the constrained ML estimate of λ_i depending just on the autoregressive parameter

$$\hat{\lambda}_{i\psi} = \bar{y}_i - \rho\bar{y}_{i,-1} = \hat{\lambda}_{i\rho}, \tag{S3}$$

where $\bar{y}_i = \sum_{t=1}^{T} y_{it}/T$ and $\bar{y}_{i,-1} = \sum_{t=0}^{T-1} y_{it}/T$. The profile log-likelihood is then obtained by replacement of λ_i with $\hat{\lambda}_{i\rho}$ in expression (S2) for each $i = 1,\ldots,N$.

The first part in Severini's modification term is immediately available, since $j_{\lambda_i\lambda_i}(\hat{\theta}_\psi) = T/\sigma^2$. Yet, the derivation of $I_{\lambda_i\lambda_i}(\hat{\theta}_\psi;\hat{\theta})$ requires more elaboration. The ML estimate of λ_i equals $\hat{\lambda}_i = \hat{\lambda}_{i\hat{\rho}} = \bar{y}_i - \hat{\rho}\bar{y}_{i,-1}$, where

$$\hat{\rho} = \frac{\sum_{i=1}^N \sum_{t=1}^T y_{it} y_{i,t-1} - T\sum_{i=1}^N \bar{y}_i \bar{y}_{i,-1}}{\sum_{i=1}^N \sum_{t=1}^T y_{i,t-1}^2 - T\sum_{i=1}^N \bar{y}_{i,-1}^2}\,. \tag{S4}$$

Therefore, by adding and subtracting the same quantity $\rho\bar{y}_{i,-1}$, one gets

$$\begin{aligned}\hat{\lambda}_i &= \bar{y}_i - \rho\bar{y}_{i,-1} + \rho\bar{y}_{i,-1} - \hat{\rho}\bar{y}_{i,-1} \\ &= \hat{\lambda}_{i\rho} - (\hat{\rho} - \rho)\bar{y}_{i,-1}\,.\end{aligned} \tag{S5}$$

The equation above lets us express the score evaluated at the constrained ML estimate in a convenient way. In particular, we start writing

$$\begin{aligned}l_{\lambda_i}(\hat{\theta}_\psi) &= \frac{1}{\sigma^2}\sum_{t=1}^T \big(y_{it} - \hat{\lambda}_{i\rho} - \rho y_{i,t-1}\big) \\ &= \frac{1}{\sigma^2}\sum_{t=1}^T \big(y_{it} - \hat{\lambda}_{i\rho} + \hat{\lambda}_i - \hat{\lambda}_i - \rho y_{i,t-1} + \hat{\rho} y_{i,t-1} - \hat{\rho} y_{i,t-1}\big)\,,\end{aligned} \tag{S6}$$

where the second equality holds because we simultaneously sum to and subtract from the bracketed part both $\hat{\lambda}_i$ and $\hat{\rho}y_{i,t-1}$. Now, since manipulating (S5) leads to

$$\hat{\lambda}_{i\rho} = \hat{\lambda}_i + (\hat{\rho} - \rho)\bar{y}_{i,-1}\,,$$

by substitution of the latter expression in (S6) it is not hard to obtain

$$\begin{aligned}l_{\lambda_i}(\hat{\theta}_\psi) &= \frac{1}{\sigma^2}\Big\{\sum_{t=1}^T \big(y_{it} - \hat{\lambda}_i - \hat{\rho} y_{i,t-1}\big) + T\big(\hat{\lambda}_i - \hat{\lambda}_{i\rho}\big) + \sum_{t=1}^T (\hat{\rho} - \rho) y_{i,t-1}\Big\} \\ &= \frac{1}{\sigma^2}\big\{\hat{\sigma}^2 l_{\lambda_i}(\hat{\theta}) + T\big(\hat{\lambda}_i - \hat{\lambda}_{i\rho}\big) + T\big(\hat{\rho} - \rho\big)\bar{y}_{i,-1}\big\}\,.\end{aligned}$$

Then, the necessary expected value is a linear function of ρ, and specifically

$$\begin{aligned}I_{\lambda_i\lambda_i}(\hat{\theta}_\psi;\hat{\theta}) &= E_{\hat{\theta}}\big\{l_{\lambda_i}(\hat{\theta}_\psi) l_{\lambda_i}(\hat{\theta})\big\} \\ &= \frac{1}{\sigma^2} E_{\hat{\theta}}\big[\big\{\hat{\sigma}^2 l_{\lambda_i}(\hat{\theta}) + T\big(\hat{\lambda}_i - \hat{\lambda}_{i\rho}\big) + T\big(\hat{\rho} - \rho\big)\overline{Y}_{i,-1}\big\} l_{\lambda_i}(\hat{\theta})\big] \\ &= \frac{1}{\sigma^2}\big\{\hat{\sigma}^2 \widehat{E}_1 + T\big(\hat{\rho} - \rho\big)\widehat{E}_2\big\}\,,\end{aligned} \tag{S7}$$

with $\widehat{E}_1 = E_{\hat{\theta}}\big\{l^2_{\lambda_i}(\hat{\theta})\big\}$ and $\widehat{E}_2 = E_{\hat{\theta}}\big\{\overline{Y}_{i,-1} l_{\lambda_i}(\hat{\theta})\big\}$. Note that expectation (S7) is computed with reference to the distribution $p(y_{it}|x_{it};\hat{\psi},\hat{\lambda}_i)$.

Although possible in principle, the analytical calculation of $\widehat{E}_1$ and $\widehat{E}_2$ is not straightforward. Conversely, estimating $I_{\lambda_i\lambda_i}(\hat{\theta}_\psi;\hat{\theta})$ via Monte Carlo simulation represents an easily implementable solution. Based on what illustrated in

Section 3, the empirical mean for approximating (S7) becomes here

$$I^*_{\lambda_i\lambda_i}(\hat{\theta}_\psi;\hat{\theta}) = \frac{1}{R}\sum_{r=1}^{R}\left[\left\{\frac{1}{\sigma^2}\sum_{t=1}^{T}\left(y^r_{it}-\hat{\lambda}_{i\rho}-\rho y^r_{i,t-1}\right)\right\}\left\{\frac{1}{\hat{\sigma}^2}\sum_{t=1}^{T}\left(y^r_{it}-\hat{\lambda}_i-\hat{\rho} y^r_{i,t-1}\right)\right\}\right], \tag{S8}$$

where $y^r_{it}\,(i=1,\dots,N, t=1,\dots,T)$ is generated by model (S1) with $(\psi,\lambda) = (\hat{\psi},\hat{\lambda})$, but the starting vector is kept unchanged, namely $y^r_0 = y_0$ for each $r = 1,\dots,R$. We acknowledge that, in this specific case, (S7) could be alternatively obtained through analogue Monte Carlo approximations to the expected values $\widehat{E}_1$ and $\widehat{E}_2$, which need to be derived just once because they involve $\hat{\theta}$ only. Nevertheless, for the reasons exposed in Section 3, the general approach detailed by (S8) is not much more costly in terms of computing effort.

S2.3. Computational aspects

The estimate $\hat{\theta}$ can be written in closed form by applying the ordinary least squares method to the linear autoregression with normally distributed errors corresponding to (S1). As a consequence, the ML estimate of σ^2 is expressed by

$$\hat{\sigma}^2 = \sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\left(y_{it}-\hat{\lambda}_i-\hat{\rho}y_{i,t-1}\right)^2}{NT}, \tag{S9}$$

where formulations for $\hat{\rho}$ and $\hat{\lambda}_i$ follow directly from (S4). On the contrary, maximization of $l_{M^*}(\psi)$ for finding the estimate $\hat{\psi}_{M^*} = (\hat{\rho}_{M^*}, \hat{\sigma}^2_{M^*})$ usually has to be performed by means of numerical algorithms, and estimated standard errors are derived using the second derivative of the function at its maximum. Under this particular scenario, it is more convenient to derive $\hat{\sigma}^2_{M^*}$ by evaluation of the explicit constrained estimate

$$\hat{\sigma}^2_{\rho,M^*} = \hat{\sigma}^2_{M^*}(\rho) = \sum_{i=1}^{N}\sum_{t=1}^{T}\frac{\left(y_{it}-\hat{\lambda}_{i\rho}-\rho y_{i,t-1}\right)^2}{N(T-1)}$$

at $\hat{\rho}_{M^*}$, i.e. the scalar solution to the optimization problem with objective function $l^{\rho}_{M^*}(\rho) = l_{M^*}\left(\rho,\hat{\sigma}^2_{\rho,M^*}\right)$. Observe that also $l_P(\psi)$ can be further profiled in order to get $l^\rho_P(\rho) = l_P(\rho,\hat{\sigma}^2_\rho)$, where $\hat{\sigma}^2_\rho$ takes the form equivalent to (S9), but with estimates $\hat{\rho}$ and $\hat{\lambda}_i$ replaced by ρ and $\hat{\lambda}_{i\rho}$ as in (S3), respectively.

According to expression (S7), for values of the autoregressive parameter beyond a certain threshold depending on $\hat{\rho}$ the expectation $I_{\lambda_i\lambda_i}(\hat{\theta}_\psi;\hat{\theta})$ is negative and $l_M(\psi)$ is not computable, similarly to what happens with the integrated likelihood of [7]. In its turn, the approximation $I^*_{\lambda_i\lambda_i}(\hat{\theta}_\psi;\hat{\theta})$ can be smaller than or equal to zero for not too large values of ρ. A potentially undefined modification term poses a problem for the numerical optimization of $l^\rho_{M^*}(\rho)$. In addition, as will emerge from the figures available in Section S2.4, the MCMPL is found to reach its global maximum as $\rho \to +\infty$ for any sample size, in accordance

Table S1
Inference on $\rho = 0.5$ in the nonstationary AR(1) model for panel data. The compared functions are the profile log-likelihood $l_P(\psi)$ and the MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
250	4	$l_P(\psi)$	-0.186	-0.186	0.025	0.187	0.186	0.879	0.000
		$l_{M^*}(\psi)$	0.020	0.018	0.037	0.042	0.028	0.921	0.915
	8	$l_P(\psi)$	-0.114	-0.115	0.018	0.116	0.115	0.921	0.000
		$l_{M^*}(\psi)$	0.002	0.002	0.020	0.020	0.013	0.989	0.942
	16	$l_P(\psi)$	-0.070	-0.070	0.013	0.071	0.070	0.960	0.000
		$l_{M^*}(\psi)$	-0.000	0.000	0.014	0.014	0.009	1.002	0.944
500	4	$l_P(\psi)$	-0.184	-0.183	0.017	0.184	0.183	0.896	0.000
		$l_{M^*}(\psi)$	0.018	0.019	0.025	0.031	0.022	0.952	0.881
	8	$l_P(\psi)$	-0.113	-0.113	0.013	0.114	0.113	0.902	0.000
		$l_{M^*}(\psi)$	0.002	0.002	0.014	0.014	0.010	0.972	0.943
	16	$l_P(\psi)$	-0.069	-0.069	0.009	0.069	0.069	0.983	0.000
		$l_{M^*}(\psi)$	0.000	0.000	0.009	0.009	0.007	1.029	0.959
1000	4	$l_P(\psi)$	-0.187	-0.187	0.013	0.187	0.187	0.879	0.000
		$l_{M^*}(\psi)$	0.019	0.018	0.019	0.026	0.019	0.923	0.795
	8	$l_P(\psi)$	-0.115	-0.115	0.009	0.115	0.115	0.919	0.000
		$l_{M^*}(\psi)$	0.002	0.002	0.010	0.010	0.007	0.987	0.948
	16	$l_P(\psi)$	-0.070	-0.070	0.007	0.070	0.070	0.935	0.000
		$l_{M^*}(\psi)$	0.000	0.000	0.007	0.007	0.005	0.977	0.940

with the various functions for inference on ψ studied in [10], [8] and [7], respectively. On such grounds, we choose to maximize $l^{\rho}_{M^*}(\rho)$ by performing a one-dimensional search in a real bounded interval Υ through the algorithm implemented by the R function `optimize`. Specifically, adopting the same notation of [10], $\Upsilon = (-\rho_l, \rho_u)$ with $\rho_l = \rho_u = 1.5$, since in usual applications the autoregressive parameter is hardly observed to lie outside these extremes. The estimate resulting from the local maximization of $l_{M^*}(\psi)$ is then uniquely defined as $\hat{\psi}_{M^*} = \left(\hat{\rho}_{M^*}, \hat{\sigma}^2_{M^*}\right)$, where $\hat{\rho}_{M^*} = \arg\max_{\rho \in \Upsilon} l^{\rho}_{M^*}(\rho)$ and $\hat{\sigma}^2_{M^*} = \hat{\sigma}^2_{\hat{\rho}_{M^*}, M^*}$. A careful discussion about the conditions under which consistency of this local maximizer is achieved is beyond the scope of the present work. We refer to [8] for further details on the topic.

S2.4. Simulation studies and numerical examples

In this section, the accuracy of the MCMPL in drawing inferences on ψ under the nonstationary normal AR(1) model is assessed with regard to that of the standard profile likelihood through a series of simulations. Two main experiments based on $S = 2000$ iterations are performed, both considering datasets with $T = 4, 8, 16$ and $N = 250, 500, 1000$. The two simulation setups differ only in the true value of the autoregressive parameter used to generate the samples from model (S1): in the first $\rho = 0.5$, while in the second $\rho = 0.9$. The conditional variance of the response variable is $\sigma^2 = 1$ and the fixed effects are

Table S2
Inference on $\sigma^2 = 1$ in the nonstationary AR(1) model for panel data with $\rho = 0.5$. The compared functions are the profile log-likelihood $l_P(\psi)$ and the MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
250	4	$l_P(\psi)$	-0.300	-0.301	0.036	0.303	0.301	0.862	0.000
		$l_{M^*}(\psi)$	0.013	0.011	0.060	0.062	0.041	0.976	0.954
	8	$l_P(\psi)$	-0.147	-0.148	0.029	0.150	0.148	0.938	0.001
		$l_{M^*}(\psi)$	-0.000	-0.001	0.035	0.035	0.024	1.011	0.957
	16	$l_P(\psi)$	-0.071	-0.071	0.022	0.074	0.071	0.964	0.092
		$l_{M^*}(\psi)$	-0.001	-0.001	0.023	0.023	0.016	1.000	0.954
500	4	$l_P(\psi)$	-0.299	-0.299	0.026	0.300	0.299	0.853	0.000
		$l_{M^*}(\psi)$	0.013	0.013	0.043	0.045	0.029	0.973	0.944
	8	$l_P(\psi)$	-0.147	-0.148	0.020	0.148	0.148	0.935	0.000
		$l_{M^*}(\psi)$	-0.000	-0.001	0.024	0.024	0.017	1.008	0.950
	16	$l_P(\psi)$	-0.070	-0.070	0.015	0.072	0.070	0.959	0.006
		$l_{M^*}(\psi)$	-0.000	-0.000	0.017	0.017	0.011	0.995	0.949
1000	4	$l_P(\psi)$	-0.300	-0.299	0.018	0.301	0.299	0.867	0.000
		$l_{M^*}(\psi)$	0.013	0.014	0.030	0.033	0.022	0.988	0.930
	8	$l_P(\psi)$	-0.147	-0.147	0.015	0.147	0.147	0.923	0.000
		$l_{M^*}(\psi)$	0.001	0.000	0.018	0.018	0.012	0.995	0.945
	16	$l_P(\psi)$	-0.070	-0.070	0.011	0.071	0.070	0.976	0.000
		$l_{M^*}(\psi)$	-0.000	-0.000	0.012	0.012	0.008	1.012	0.948

independently drawn from a $N(1, 1)$ distribution, following the example of [10]. In every simulated dataset, all N initial observations in the vector y_0 are fixed equal to zero with no loss of generality, as this is equivalent to interpret each y_{it} as $y_{it} - y_{i0}$ and each λ_i as $\lambda_i - y_{i0}(1 - \rho)$ $(i = 1, \dots, N, t = 1, \dots, T)$ [10]. Lastly, the number of Monte Carlo replicates employed to compute $l_{M^*}(\psi)$ is $R = 500$.

Inferential results of the first study for ρ and σ^2 are displayed in Tables S1 and S2, respectively, using indexes defined in Section 4.3. Similar comments as in [5] can be made. In all configurations, no significant differences between bias and median bias of the same estimator are observed, but the improvement determined by using the MCMPL in this sense is remarkable. Consistently with the theory for independent units [11], the bias does not vary with N but decreases as T increases, whereas the root mean squared error depends on both indexes. Empirical coverage probabilities of 0.95 Wald confidence intervals based on $l_{M^*}(\psi)$ are generally accurate for both components of interest, with larger departures from the nominal level occurring when $T = 4$. Such conspicuous refinements to the poor interval estimation supplied by $l_P(\psi)$ mainly stem from bias reduction. Yet some correction in curvature also takes place, being SE/SD for the MCMPL typically closer to one than for the ordinary profile likelihood.

Tables S3 and S4 illustrate results of the simulation experiment run with a true value of ρ approaching the boundaries of the stationary region $(-1, 1)$, particularly $\rho = 0.9$. Relative behaviours of the two methods for estimating the

TABLE S3
Inference on $\rho = 0.9$ in the nonstationary AR(1) model for panel data. The compared functions are the profile log-likelihood $l_P(\psi)$ and the MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
250	4	$l_P(\psi)$	-0.130	-0.130	0.018	0.131	0.130	0.894	0.000
		$l_{M^*}(\psi)$	0.022	0.021	0.028	0.036	0.024	0.899	0.871
	8	$l_P(\psi)$	-0.051	-0.051	0.008	0.052	0.051	0.922	0.000
		$l_{M^*}(\psi)$	0.003	0.003	0.010	0.010	0.007	0.976	0.933
	16	$l_P(\psi)$	-0.022	-0.023	0.004	0.023	0.023	0.957	0.001
		$l_{M^*}(\psi)$	0.000	0.000	0.005	0.005	0.003	1.003	0.950
500	4	$l_P(\psi)$	-0.128	-0.128	0.013	0.129	0.128	0.905	0.000
		$l_{M^*}(\psi)$	0.021	0.020	0.019	0.028	0.021	0.928	0.774
	8	$l_P(\psi)$	-0.050	-0.050	0.006	0.050	0.050	0.933	0.000
		$l_{M^*}(\psi)$	0.003	0.003	0.007	0.007	0.005	0.980	0.928
	16	$l_P(\psi)$	-0.022	-0.022	0.003	0.022	0.022	0.957	0.000
		$l_{M^*}(\psi)$	0.001	0.001	0.003	0.003	0.002	1.001	0.946
1000	4	$l_P(\psi)$	-0.131	-0.131	0.009	0.131	0.131	0.895	0.000
		$l_{M^*}(\psi)$	0.021	0.021	0.014	0.025	0.021	0.909	0.612
	8	$l_P(\psi)$	-0.051	-0.051	0.004	0.051	0.051	0.923	0.000
		$l_{M^*}(\psi)$	0.003	0.003	0.005	0.006	0.004	0.969	0.884
	16	$l_P(\psi)$	-0.022	-0.022	0.002	0.022	0.022	0.923	0.000
		$l_{M^*}(\psi)$	0.001	0.001	0.002	0.002	0.002	0.968	0.930

component of interest are basically in line with those viewed in the previous study. Perhaps here the general improvements originating from the employment of $l_{M^*}(\psi)$ are somewhat milder than when the autoregressive parameter is farther away from nonstationariety. This observation can be referred both to bias and, mostly, to empirical coverages of confidence intervals for ρ. Nonetheless, the quality of MPL-based inference remains unquestionably higher than that reached through standard ML techniques.

Figure S1 shows $l_P^\rho(\rho)$ and $l_{M^*}^\rho(\rho)$, as defined in Section S2.3, in their relative version. Specifically, the plots are referred to samples generated from model (S1) with $\rho = 0.5, 0.9$, $T = 4$ and $N = 250, 1000$. All panels confirm the results of simulations discussed so far. The maximum of the profile log-likelihood is significantly smaller than the true value of the autoregressive parameter, corresponding to the vertical line. Because of this and the accentuated curvature of $l_P^\rho(\rho)$, such value never belongs to the 0.95 confidence region defined by inversion of the profile likelihood ratio statistic and marked by the horizontal line. Conversely, the local maximization of the MCMPL yields to adequate point and interval estimations of ρ. The unusual trend of $l_{M^*}^\rho(\rho)$, whose global maximizer lies at infinity, was already anticipated in Section S2.3. The absence of restrictions on the initial conditions in y_0 causes in fact the MCMPL to be re-increasing, sometimes already in the stationary parameter region [8].

TABLE S4
Inference on $\sigma^2 = 1$ in the nonstationary AR(1) model for panel data with $\rho = 0.9$. The compared functions are the profile log-likelihood $l_P(\psi)$ and the MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
250	4	$l_P(\psi)$	-0.297	-0.298	0.036	0.299	0.298	0.870	0.000
		$l_{M^*}(\psi)$	0.021	0.017	0.062	0.066	0.043	0.956	0.944
	8	$l_P(\psi)$	-0.144	-0.145	0.029	0.147	0.145	0.945	0.002
		$l_{M^*}(\psi)$	0.002	0.000	0.035	0.035	0.024	1.007	0.956
	16	$l_P(\psi)$	-0.070	-0.070	0.021	0.074	0.070	0.969	0.095
		$l_{M^*}(\psi)$	-0.001	-0.001	0.023	0.023	0.016	1.001	0.951
500	4	$l_P(\psi)$	-0.295	-0.295	0.026	0.297	0.295	0.861	0.000
		$l_{M^*}(\psi)$	0.020	0.019	0.044	0.048	0.032	0.956	0.925
	8	$l_P(\psi)$	-0.144	-0.144	0.020	0.145	0.144	0.943	0.000
		$l_{M^*}(\psi)$	0.001	0.001	0.024	0.025	0.017	1.006	0.954
	16	$l_P(\psi)$	-0.069	-0.069	0.015	0.071	0.069	0.962	0.006
		$l_{M^*}(\psi)$	0.000	0.000	0.017	0.017	0.011	0.993	0.945
1000	4	$l_P(\psi)$	-0.296	-0.296	0.018	0.297	0.296	0.875	0.000
		$l_{M^*}(\psi)$	0.021	0.021	0.031	0.037	0.026	0.972	0.903
	8	$l_P(\psi)$	-0.144	-0.144	0.015	0.144	0.144	0.928	0.000
		$l_{M^*}(\psi)$	0.003	0.002	0.018	0.018	0.012	0.988	0.942
	16	$l_P(\psi)$	-0.070	-0.070	0.011	0.070	0.070	0.981	0.000
		$l_{M^*}(\psi)$	0.000	0.000	0.012	0.012	0.008	1.012	0.952

S3. Probit regression model with missing response

Suppose now that specifications (8)–(10) in Section 4.2 of the paper hold with $F(\cdot) = \Phi(\cdot)$ and $G(\cdot) = \mathrm{logit}^{-1}(\cdot)$, where $\Phi(\cdot)$ is the CDF of the standard normal random variable. Even in probit regressions for complete clustered binary data y_{it} ($i = 1, \ldots, N$, $t = 1, \ldots, T$) an explicit formulation for Severini's adjustment exists, and can be computed on the available units under the MCAR assumption. Denoting by $\phi(\cdot)$ the probability density function of the $N(0, 1)$, expectation (16) in Section 4.2 becomes

$$I_{\lambda_i\lambda_i}(\hat{\theta}; \hat{\theta}_\beta) = \sum_{t\colon y_{it} \in y^{obs}} \frac{\phi\big(\hat{\lambda}_{i\beta} + \beta^{\mathrm{T}} x_{it}\big)\phi\big(\hat{\lambda}_i + \hat{\beta}^{\mathrm{T}} x_{it}\big)}{\big\{1 - \Phi\big(\hat{\lambda}_{i\beta} + \beta^{\mathrm{T}} x_{it}\big)\big\}\Phi\big(\hat{\lambda}_{i\beta} + \beta^{\mathrm{T}} x_{it}\big)}, \qquad i = 1, \ldots, N. \tag{S10}$$

Under these hypotheses, the ith score function may be expressed as

$$l_{\lambda_i}(\theta) = \sum_{t\colon y_{it} \in y^{obs}} \frac{\big\{y_{it} - \Phi\big(\lambda_i + \beta^{\mathrm{T}} x_{it}\big)\big\}\phi\big(\lambda_i + \beta^{\mathrm{T}} x_{it}\big)}{\Phi\big(\lambda_i + \beta^{\mathrm{T}} x_{it}\big)\big\{1 - \Phi\big(\lambda_i + \beta^{\mathrm{T}} x_{it}\big)\big\}}, \qquad i = 1, \ldots, N, \tag{S11}$$

and $j_{\lambda_i\lambda_i}(\theta)$ is readily derived by changing sign to its first derivative with respect to λ_i. Using (S10) and (S11), it is possible to obtain both $l_M(\beta)$ in closed form and $l_{M^*}(\beta)$ as described in (17). In the present probit framework, the formula of the standard profile log-likelihood $l_P(\beta)$ follows in fact from (15) with $\pi_{it} = \Phi(\lambda_i + \beta^{\mathrm{T}} x_{it})$.

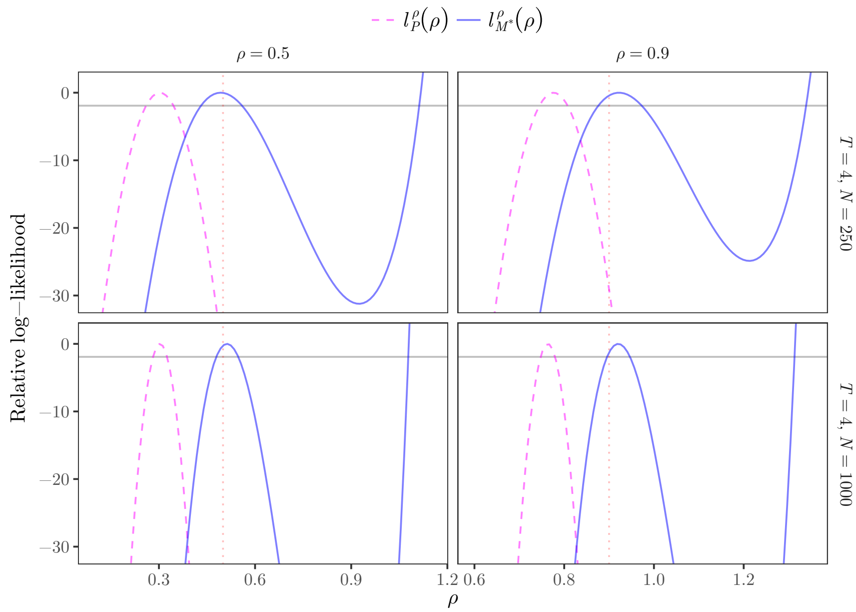

Fig S1. *Relative profile (dashed) and Monte Carlo modified profile (solid) log-likelihoods of four datasets generated under the nonstationary AR(1) model with* $\rho = 0.5, 0.9$, $T = 4$ *and* $N = 250, 1000$. *The dotted vertical line indicates the true value of the autoregressive parameter, while the horizontal line gives the 0.95 confidence intervals for* ρ *based on the likelihood ratio statistics.*

When we conjecture that incompleteness of the data originates from a non-ignorable process, Monte Carlo simulation serves to approximate the unconditional expected value $I_{\lambda_i\lambda_i}(\hat{\varphi}; \hat{\varphi}_\psi)$, whose exact formulation is hardly retrievable. The expression of $l_{M^*}(\psi)$ in the probit setting may be obtained by multiple substitution of $\Phi(\lambda_i + \beta^{\mathrm{T}} x_{it})$ for π_{it}, $\phi(\lambda_i + \beta^{\mathrm{T}} x_{it})$ for f_{it} and $-(\lambda_i + \beta^{\mathrm{T}} x_{it})\phi(\lambda_i + \beta^{\mathrm{T}} x_{it})$ for f'_{it} in equations (11)–(14) of the paper.

The numerical optimization methods employed for the various functions correspond to those of the logistic case. We also recall that exclusion of the non-informative clusters by the dataset must take place prior to the fitting phase.

The basic structure of the studies performed in Section 4.3 is held unchanged, yet here a probit link between the response and the unique predictor is considered. In the first experiment, missing observations are chosen according to an MCAR mechanism with $\gamma_1 = 2.5$; in the second, the true missingness generation process is MNAR with $\gamma_1 = 5$ and $\gamma_2 = 1$. The covariate is again simulated from the $N(-0.35, 1)$ distribution. Exploiting the well-known relation between the logistic and normal distributions [2], in order to obtain data and quantity of informative groups comparable to the logistic setting, the complete simulated samples are generated by fixing $\beta = 1/1.6 = 0.625$ and intercepts λ_i $(i = 1, \ldots, N)$ are independently generated as $N(-0.22, 0.39)$

Table S5

Inference on $\beta = 1/1.6$ in the probit regression for MCAR longitudinal data. The compared functions are the MCAR profile log-likelihood $l_P(\beta)$, Severini's exact MCAR MPL $l_M(\beta)$, and the MCAR MCMPL $l_{M^*}(\beta)$ computed with $R = 500$. Results based on a simulation study with 2000 trials.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\beta)$	0.555	0.473	0.579	0.802	0.478	0.663	0.725
		$l_M(\beta)$	0.142	0.131	0.283	0.317	0.201	1.018	0.967
		$l_{M^*}(\beta)$	0.142	0.128	0.301	0.333	0.203	0.956	0.960
	6	$l_P(\beta)$	0.296	0.268	0.319	0.435	0.280	0.760	0.774
		$l_M(\beta)$	0.074	0.065	0.213	0.226	0.141	0.960	0.956
		$l_{M^*}(\beta)$	0.067	0.057	0.215	0.225	0.140	0.945	0.955
	10	$l_P(\beta)$	0.162	0.147	0.181	0.243	0.158	0.852	0.821
		$l_M(\beta)$	0.039	0.029	0.145	0.151	0.095	0.960	0.940
		$l_{M^*}(\beta)$	0.035	0.025	0.145	0.149	0.093	0.960	0.940
100	4	$l_P(\beta)$	0.460	0.421	0.358	0.582	0.421	0.707	0.585
		$l_M(\beta)$	0.108	0.102	0.197	0.225	0.148	1.005	0.947
		$l_{M^*}(\beta)$	0.104	0.095	0.204	0.229	0.147	0.970	0.941
	6	$l_P(\beta)$	0.284	0.270	0.211	0.354	0.270	0.805	0.636
		$l_M(\beta)$	0.072	0.066	0.147	0.164	0.102	0.985	0.938
		$l_{M^*}(\beta)$	0.064	0.056	0.148	0.161	0.100	0.975	0.941
	10	$l_P(\beta)$	0.155	0.148	0.127	0.200	0.149	0.867	0.703
		$l_M(\beta)$	0.034	0.029	0.103	0.109	0.073	0.971	0.942
		$l_{M^*}(\beta)$	0.030	0.025	0.102	0.107	0.072	0.972	0.944
250	4	$l_P(\beta)$	0.413	0.401	0.208	0.462	0.401	0.754	0.283
		$l_M(\beta)$	0.095	0.091	0.124	0.156	0.105	1.017	0.907
		$l_{M^*}(\beta)$	0.087	0.084	0.127	0.154	0.101	0.990	0.909
	6	$l_P(\beta)$	0.260	0.253	0.130	0.291	0.253	0.818	0.340
		$l_M(\beta)$	0.058	0.055	0.092	0.109	0.074	0.989	0.915
		$l_{M^*}(\beta)$	0.049	0.046	0.092	0.105	0.071	0.982	0.921
	10	$l_P(\beta)$	0.145	0.143	0.077	0.164	0.143	0.904	0.462
		$l_M(\beta)$	0.027	0.025	0.062	0.068	0.045	1.012	0.941
		$l_{M^*}(\beta)$	0.022	0.020	0.062	0.066	0.043	1.014	0.947

The corresponding value of the parameter for x_{it} under a marginal model with generalized estimating equations (GEE) is $\beta_m = \beta/\sqrt{1+\sigma_\lambda^2}$, where σ_λ^2 is the variance of the random effects' distribution [1, Section 9.4.1]. Here $\sigma_\lambda^2 = 0.39$, and therefore $\beta_m = 0.625/\sqrt{1.39} = 0.53$.

Tables S5 and S6 summarize the usual measures of estimation accuracy for β based on $S = 2000$ simulations of the study regarding MCAR data. Relative behaviours of the three MCAR log-likelihoods illustrated by Table S5 do not significantly differentiate from those viewed in Table 1 of Section 4.3 for the logit link. The defective performance of $l_P(\beta)$ is greatly corrected by the adjustment proposed by Severini, from any relevant inferential perspective and for all possible couples (T, N) with $T = 4, 6, 10$ and $N = 50, 100, 250$. Results achieved by $l_M(\beta)$ and $l_{M^*}(\beta)$ are still very similar, thanks to the validity of the MCAR hypothesis.

Inferences on β drawn using the MNAR functions and the GEE method for the same MCAR samples are displayed by Table S6. In this probit regression setting the empirical properties of $l_{M^*}(\psi)$ are much more favourable than those

Table S6

Inference on $\beta = 1/1.6$ in the probit regression for MCAR longitudinal data. The compared methods are the MNAR profile log-likelihood $l_P(\psi)$, the MNAR MCMPL $l_{M^*}(\psi)$ computed with $R = 500$ and GEE. Results based on a simulation study with 2000 trials.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\psi)$	0.419	0.340	0.643	0.767	0.424	0.630	0.784
		$l_{M^*}(\psi)$	0.121	0.099	0.296	0.320	0.183	0.996	0.981
		GEE	0.026	0.020	0.182	0.184	0.120	0.987	0.955
	6	$l_P(\psi)$	0.213	0.186	0.350	0.409	0.250	0.720	0.809
		$l_{M^*}(\psi)$	0.044	0.031	0.213	0.218	0.136	0.978	0.968
		GEE	0.014	0.010	0.152	0.153	0.099	0.971	0.949
	10	$l_P(\psi)$	0.111	0.109	0.217	0.244	0.156	0.728	0.814
		$l_{M^*}(\psi)$	0.026	0.016	0.144	0.146	0.093	1.001	0.955
		GEE	0.012	0.010	0.112	0.113	0.074	1.000	0.974
100	4	$l_P(\psi)$	0.335	0.302	0.400	0.522	0.332	0.691	0.731
		$l_{M^*}(\psi)$	0.080	0.067	0.196	0.212	0.130	1.029	0.964
		GEE	0.011	0.009	0.127	0.127	0.084	0.993	0.957
	6	$l_P(\psi)$	0.178	0.183	0.266	0.320	0.222	0.649	0.715
		$l_{M^*}(\psi)$	0.035	0.028	0.147	0.151	0.093	1.007	0.955
		GEE	0.009	0.009	0.103	0.103	0.070	1.021	0.954
	10	$l_P(\psi)$	0.101	0.114	0.175	0.202	0.142	0.637	0.737
		$l_{M^*}(\psi)$	0.021	0.016	0.101	0.103	0.067	1.025	0.957
		GEE	0.007	0.003	0.081	0.080	0.053	0.992	0.957
250	4	$l_P(\psi)$	0.283	0.275	0.238	0.370	0.279	0.727	0.598
		$l_{M^*}(\psi)$	0.060	0.054	0.120	0.134	0.087	1.061	0.950
		GEE	0.005	0.003	0.080	0.080	0.054	1.001	0.949
	6	$l_P(\psi)$	0.143	0.167	0.205	0.250	0.197	0.518	0.548
		$l_{M^*}(\psi)$	0.022	0.018	0.091	0.093	0.061	1.020	0.955
		GEE	0.003	0.001	0.065	0.065	0.043	1.004	0.951
	10	$l_P(\psi)$	0.079	0.108	0.154	0.173	0.130	0.447	0.563
		$l_{M^*}(\psi)$	0.013	0.011	0.061	0.063	0.040	1.061	0.962
		GEE	0.001	0.001	0.049	0.049	0.034	1.008	0.950

of the corresponding unmodified function. This is well reflected by all bias and Wald coverage indicators. We observe that, contrary to expectations, for fixed T the bias of the ML estimator decreases with N. In addition, the MNAR log-likelihoods seem to supply better point estimation but less trustworthy confidence intervals compared to their MCAR counterparts. Comments pertaining the performance of the approach based on GEE are along the same lines as those made about Table 2 in Section 4.3 of the paper, dedicated to the logistic regression. Even with the probit link, $l_{M^*}(\psi)$ is more reliable than GEE as the group size grows up and generally succeeds in detecting the underlying ignorable missingness process, which represents a reduced form of the full MNAR model presupposed by that MCMPL.

An account of the last simulation experiment is given in Table S7, which is referred to incomplete datasets with MNAR units. Similarly to what emerged by Table 3 in the paper under the MNAR logistic framework, $l_{M^*}(\psi)$ appears to retain higher inferential precision than the analytical MPL which ignores the missingness model, the only exception being the case with $T = 4$. Thus, when groups are not extremely small, taking the true nonignorable missing-

TABLE S7

Inference on $\beta = 1/1.6$ in the probit regression for MNAR longitudinal data. The compared methods are the MNAR profile log-likelihood $l_P(\psi)$, Severini's exact MCAR MPL $l_M(\beta)$, the MNAR MCMPL $l_{M^*}(\psi)$ computed with $R = 500$ and GEE. Results based on a simulation study with 2000 trials.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\psi)$	0.301	0.235	0.623	0.692	0.383	0.687	0.863
		$l_M(\beta)$	0.008	-0.011	0.349	0.349	0.230	0.987	0.963
		$l_{M^*}(\psi)$	0.012	-0.022	0.405	0.405	0.233	0.857	0.964
		GEE	-0.135	-0.142	0.204	0.244	0.175	1.002	0.891
	6	$l_P(\psi)$	0.163	0.131	0.351	0.387	0.230	0.775	0.884
		$l_M(\beta)$	-0.072	-0.091	0.245	0.255	0.175	0.973	0.928
		$l_{M^*}(\psi)$	-0.041	-0.063	0.253	0.256	0.167	0.956	0.941
		GEE	-0.140	-0.146	0.164	0.215	0.162	1.015	0.854
	10	$l_P(\psi)$	0.079	0.087	0.244	0.256	0.155	0.714	0.865
		$l_M(\beta)$	-0.113	-0.116	0.163	0.199	0.146	0.986	0.870
		$l_{M^*}(\psi)$	-0.035	-0.042	0.166	0.170	0.116	0.990	0.937
		GEE	-0.143	-0.144	0.124	0.190	0.148	1.025	0.784
100	4	$l_P(\psi)$	0.254	0.240	0.390	0.466	0.287	0.729	0.834
		$l_M(\beta)$	-0.018	-0.024	0.237	0.238	0.157	0.976	0.939
		$l_{M^*}(\psi)$	-0.028	-0.033	0.240	0.242	0.161	0.974	0.940
		GEE	-0.142	-0.149	0.134	0.195	0.154	1.062	0.829
	6	$l_P(\psi)$	0.140	0.134	0.257	0.293	0.187	0.742	0.827
		$l_M(\beta)$	-0.074	-0.085	0.169	0.185	0.131	0.997	0.920
		$l_{M^*}(\psi)$	-0.050	-0.063	0.171	0.178	0.124	1.002	0.941
		GEE	-0.145	-0.151	0.114	0.184	0.153	1.046	0.749
	10	$l_P(\psi)$	0.053	0.089	0.236	0.242	0.132	0.519	0.791
		$l_M(\beta)$	-0.119	-0.119	0.116	0.166	0.127	0.994	0.788
		$l_{M^*}(\psi)$	-0.037	-0.042	0.119	0.124	0.085	0.990	0.928
		GEE	-0.150	-0.151	0.088	0.174	0.151	1.018	0.598
250	4	$l_P(\psi)$	0.197	0.190	0.229	0.302	0.208	0.781	0.769
		$l_M(\beta)$	-0.046	-0.048	0.147	0.154	0.107	1.008	0.933
		$l_{M^*}(\psi)$	-0.063	-0.066	0.146	0.159	0.113	1.022	0.924
		GEE	-0.150	-0.150	0.085	1.173	0.150	1.056	0.607
	6	$l_P(\psi)$	0.035	0.099	0.262	0.264	0.163	0.427	0.727
		$l_M(\beta)$	-0.094	-0.094	0.103	0.139	0.102	1.027	0.853
		$l_{M^*}(\psi)$	-0.070	-0.071	0.105	0.127	0.088	1.017	0.890
		GEE	-0.156	-0.158	0.070	0.171	0.158	1.044	0.412
	10	$l_P(\psi)$	-0.022	0.081	0.282	0.283	0.130	0.260	0.665
		$l_M(\beta)$	-0.132	-0.132	0.069	0.149	0.132	1.051	0.545
		$l_{M^*}(\psi)$	-0.045	-0.045	0.070	0.083	0.057	1.060	0.912
		GEE	-0.161	-0.163	0.053	0.170	0.163	1.053	0.179

data mechanism into consideration via Monte Carlo simulation is practically translated into improved bias and coverage properties of the estimator based on the MNAR MCMPL. Note that the performance of the MCAR MCMPL is not displayed, like for the logit link function, since it does not significantly differ from that of Severini's $l_M(\beta)$. Furthermore, as expected, the GEE approach is found inconsistent when fitting a binary regression with MNAR response. In closing, the use of Monte Carlo approximation to compute the MNAR MPL proves to be generally recommendable in the presence of binary missing observations even under the probit link assumption, thanks to its robustness to the true mechanism of missingness.

Finally, we consider a slightly different simulation setting, in which the incidental parameters are generated in such a way that they are intrinsically correlated with the covariate in the model. Specifically, we assume $\lambda_i = \sum_{t=1}^{T} x_{it}/T + u_i$, where $u_i \sim N(0,1)$ $(i = 1,\dots,N)$. In this case, a random effects model would not be correctly specified and thus neither a corresponding marginal model would be available for comparing inference based on GEE. Tables S8–S10 show for the logistic link function that results based on the fixed effects approach and achieved via the MCMPL are, as expected, as accurate as those obtained under the alternative simulations in Section 4.3 of the paper. The same pattern of performance can be found in Tables S11–S13 concerning the probit specification.

S4. Weibull regression model for right-censored data

Tables S14–S17 summarize inference on $\beta = (\beta_1, \beta_2)$ obtained through the simulation studies described in Section 5.4 of the paper. In particular, Tables S14 and S15 refer to the first experiment considering an average censoring probability $P_c = 0.2$, while Tables S16 and S17 refer to the case $P_c = 0.4$. The results attest the sufficient adequacy of $l_P(\psi)$ in drawing conclusions about β. However, due to better estimation of the standard errors of $\hat{\beta}_{M^*} = (\hat{\beta}_{1M^*}, \hat{\beta}_{2M^*})$, the MCMPL is still superior in terms of appropriateness of confidence intervals. Hence, one can conclude that the Monte Carlo adjustment is valuable to further improve the quality of standard ML procedures even when making inference on the regression coefficients under the Weibull modelling framework of Section 5.3.

References

[1] Agresti, A. (2015). *Foundations of linear and generalized linear models*. John Wiley & Sons.

[2] Amemiya, T. (1981). Qualitative response models: A survey. *Journal of Economic Li-terature 19*, 1483–1536.

[3] Barndorff-Nielsen, O. E. (1980). Conditionality resolutions. *Biometrika 67*, 293–310.

[4] Barndorff-Nielsen, O. E. (1983). On a formula for the distribution of the maximum likelihood estimator. *Biometrika 70*, 343–365.

[5] Bartolucci, F., R. Bellio, A. Salvan, and N. Sartori (2016). Modified profile likelihood for fixed-effects panel data models. *Econometric Reviews 35*, 1271–1289.

[6] Cruddas, A., N. Reid, and D. Cox (1989). A time series illustration of approximate conditional likelihood. *Biometrika 76*, 231–237.

[7] De Bin, R., N. Sartori, and T. Severini (2015). Integrated likelihoods in models with stratum nuisance parameters. *Electronic Journal of Statistics 9*, 1474–1491.

Table S8

Inference on $\beta = 1$ in the logistic regression for MCAR longitudinal data. The compared functions are the MCAR profile log-likelihood $l_P(\beta)$, Severini's exact MCAR MPL $l_M(\beta)$, and the MCAR MCMPL $l_{M^*}(\beta)$ computed with $R = 500$. Results based on a simulation study with 2000 trials, where $\lambda_i = \sum_{t=1}^{T} x_{it}/T + u_i$, $u_i \sim N(0, 1)$.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\beta)$	0.872	0.679	1.232	1.509	0.716	0.575	0.806
		$l_M(\beta)$	0.209	0.160	0.841	0.867	0.344	0.601	0.960
		$l_{M^*}(\beta)$	0.211	0.163	0.828	0.854	0.347	0.611	0.960
	6	$l_P(\beta)$	0.473	0.426	0.571	0.741	0.460	0.765	0.829
		$l_M(\beta)$	0.108	0.097	0.383	0.398	0.255	0.944	0.949
		$l_{M^*}(\beta)$	0.110	0.099	0.383	0.399	0.256	0.944	0.950
	10	$l_P(\beta)$	0.227	0.201	0.308	0.382	0.237	0.892	0.877
		$l_M(\beta)$	0.036	0.018	0.250	0.252	0.163	0.993	0.956
		$l_{M^*}(\beta)$	0.037	0.020	0.250	0.253	0.161	0.992	0.956
100	4	$l_P(\beta)$	0.715	0.639	0.631	0.954	0.641	0.723	0.674
		$l_M(\beta)$	0.146	0.129	0.345	0.375	0.239	0.995	0.955
		$l_{M^*}(\beta)$	0.148	0.131	0.346	0.376	0.240	0.993	0.955
	6	$l_P(\beta)$	0.430	0.397	0.376	0.571	0.400	0.808	0.725
		$l_M(\beta)$	0.088	0.073	0.262	0.276	0.174	0.971	0.950
		$l_{M^*}(\beta)$	0.089	0.073	0.262	0.277	0.174	0.971	0.948
	10	$l_P(\beta)$	0.213	0.198	0.216	0.304	0.206	0.902	0.812
		$l_M(\beta)$	0.027	0.016	0.176	0.178	0.116	1.002	0.954
		$l_{M^*}(\beta)$	0.028	0.017	0.176	0.178	0.117	1.002	0.954
250	4	$l_P(\beta)$	0.625	0.595	0.370	0.727	0.595	0.764	0.435
		$l_M(\beta)$	0.117	0.109	0.221	0.250	0.160	0.995	0.934
		$l_{M^*}(\beta)$	0.118	0.110	0.221	0.250	0.160	0.994	0.932
	6	$l_P(\beta)$	0.395	0.388	0.228	0.456	0.388	0.817	0.454
		$l_M(\beta)$	0.071	0.067	0.162	0.177	0.116	0.971	0.930
		$l_{M^*}(\beta)$	0.072	0.068	0.162	0.177	0.117	0.970	0.929
	10	$l_P(\beta)$	0.214	0.210	0.134	0.253	0.210	0.906	0.583
		$l_M(\beta)$	0.027	0.024	0.109	0.113	0.075	1.003	0.950
		$l_{M^*}(\beta)$	0.027	0.024	0.109	0.113	0.075	1.003	0.948

[8] Dhaene, G. and K. Jochmans (2014). Likelihood inference in an autoregression with fixed effects. *Econometric Theory First View*, 1–38.

[9] Dhaene, G. and K. Jochmans (2016). Bias-corrected estimation of panel vector autoregressions. *Economics Letters 145*, 98–103.

[10] Lancaster, T. (2002). Orthogonal parameters and panel data. *Review of Economic Studies 69*, 647–666.

[11] Sartori, N. (2003). Modified profile likelihoods in models with stratum nuisance parameters. *Biometrika 90*, 533–549.

[12] Severini, T. A. (2007). Integrated likelihood functions for non-Bayesian inference. *Biometrika 94*, 529–542.

Table S9

Inference on $\beta = 1$ in the logistic regression for MCAR longitudinal data. The compared methods are the MNAR profile log-likelihood $l_P(\psi)$, the MNAR MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials, where $\lambda_i = \sum_{t=1}^{T} x_{it}/T + u_i$, $u_i \sim N(0,1)$.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\psi)$	0.567	0.408	1.235	1.359	0.713	0.599	0.822
		$l_{M^*}(\psi)$	-0.021	-0.030	0.720	0.720	0.454	0.816	0.890
	6	$l_P(\psi)$	0.357	0.326	0.615	0.711	0.435	0.757	0.859
		$l_{M^*}(\psi)$	0.004	0.018	0.452	0.452	0.275	0.913	0.934
	10	$l_P(\psi)$	0.195	0.178	0.336	0.389	0.242	0.874	0.890
		$l_{M^*}(\psi)$	0.014	0.004	0.277	0.278	0.182	0.992	0.957
100	4	$l_P(\psi)$	0.467	0.422	0.727	0.864	0.537	0.687	0.789
		$l_{M^*}(\psi)$	-0.064	-0.046	0.470	0.475	0.303	0.903	0.899
	6	$l_P(\psi)$	0.323	0.293	0.408	0.520	0.331	0.800	0.815
		$l_{M^*}(\psi)$	-0.004	-0.004	0.313	0.313	0.194	0.951	0.952
	10	$l_P(\psi)$	0.185	0.178	0.233	0.298	0.195	0.892	0.847
		$l_{M^*}(\psi)$	0.009	0.005	0.191	0.191	0.128	1.013	0.959
250	4	$l_P(\psi)$	0.362	0.350	0.436	0.567	0.393	0.730	0.732
		$l_{M^*}(\psi)$	-0.140	-0.097	0.357	0.384	0.226	0.857	0.875
	6	$l_P(\psi)$	0.289	0.284	0.248	0.381	0.290	0.805	0.674
		$l_{M^*}(\psi)$	-0.011	-0.006	0.188	0.189	0.117	0.972	0.950
	10	$l_P(\psi)$	0.124	0.118	0.153	0.197	0.135	0.911	0.844
		$l_{M^*}(\psi)$	-0.037	-0.042	0.127	0.132	0.091	1.010	0.940

Table S10

Inference on $\beta = 1$ in the logistic regression for MNAR longitudinal data. The compared methods are the MNAR profile log-likelihood $l_P(\psi)$, Severini's exact MCAR MPL $l_M(\beta)$, the MNAR MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials, where $\lambda_i = \sum_{t=1}^{T} x_{it}/T + u_i$, $u_i \sim N(0,1)$.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\psi)$	0.330	0.190	1.207	1.251	0.621	0.638	0.892
		$l_M(\beta)$	-0.062	-0.121	0.758	0.761	0.431	0.793	0.944
		$l_{M^*}(\psi)$	-0.106	-0.147	0.826	0.833	0.449	0.750	0.933
	6	$l_P(\psi)$	0.217	0.163	0.592	0.631	0.366	0.819	0.926
		$l_M(\beta)$	-0.164	-0.186	0.439	0.469	0.316	0.941	0.898
		$l_{M^*}(\psi)$	-0.061	-0.082	0.432	0.436	0.279	0.979	0.945
	10	$l_P(\psi)$	0.126	0.106	0.346	0.368	0.233	0.916	0.923
		$l_M(\beta)$	-0.234	-0.250	0.283	0.367	0.279	1.002	0.850
		$l_{M^*}(\psi)$	-0.036	-0.052	0.287	0.290	0.193	1.014	0.958
100	4	$l_P(\psi)$	0.291	0.228	0.670	0.730	0.412	0.755	0.886
		$l_M(\beta)$	-0.107	-0.133	0.418	0.431	0.291	0.951	0.922
		$l_{M^*}(\psi)$	-0.106	-0.123	0.431	0.444	0.283	0.969	0.936
	6	$l_P(\psi)$	0.191	0.164	0.395	0.439	0.269	0.865	0.905
		$l_M(\beta)$	-0.193	-0.211	0.293	0.350	0.260	0.996	0.879
		$l_{M^*}(\psi)$	-0.071	-0.087	0.294	0.302	0.207	1.013	0.947
	10	$l_P(\psi)$	0.125	0.116	0.244	0.274	0.181	0.925	0.919
		$l_M(\beta)$	-0.242	-0.250	0.197	0.312	0.255	1.025	0.750
		$l_{M^*}(\psi)$	-0.035	-0.040	0.202	0.205	0.139	1.025	0.955
250	4	$l_P(\psi)$	0.202	0.172	0.418	0.464	0.293	0.775	0.859
		$l_M(\beta)$	-0.160	-0.178	0.274	0.317	0.231	0.942	0.872
		$l_{M^*}(\psi)$	-0.142	-0.153	0.281	0.315	0.223	0.972	0.918
	6	$l_P(\psi)$	0.159	0.162	0.238	0.286	0.197	0.879	0.869
		$l_M(\beta)$	-0.215	-0.213	0.180	0.280	0.217	0.996	0.761
		$l_{M^*}(\psi)$	-0.086	-0.083	0.177	0.197	0.132	1.026	0.917
	10	$l_P(\psi)$	0.124	0.118	0.153	0.197	0.135	0.911	0.844
		$l_M(\beta)$	-0.258	-0.259	0.123	0.286	0.259	1.010	0.458
		$l_{M^*}(\psi)$	-0.037	-0.042	0.127	0.132	0.091	1.010	0.940

Table S11

Inference on $\beta = 1/1.6$ in the probit regression for MCAR longitudinal data. The compared functions are the MCAR profile log-likelihood $l_P(\beta)$, Severini's exact MCAR MPL $l_M(\beta)$, and the MCAR MCMPL $l_{M^*}(\beta)$ computed with $R = 500$. Results based on a simulation study with 2000 trials, where $\lambda_i = \left(\sum_{t=1}^{T} x_{it}/T + u_i\right)/1.6$, $u_i \sim N(0,1)$.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\beta)$	0.592	0.484	0.693	0.911	0.495	0.605	0.752
		$l_M(\beta)$	0.148	0.138	0.304	0.338	0.222	1.026	0.967
		$l_{M^*}(\beta)$	0.152	0.132	0.336	0.369	0.225	0.931	0.955
	6	$l_P(\beta)$	0.321	0.296	0.343	0.470	0.305	0.757	0.784
		$l_M(\beta)$	0.087	0.081	0.227	0.243	0.155	0.961	0.954
		$l_{M^*}(\beta)$	0.080	0.073	0.229	0.242	0.152	0.949	0.950
	10	$l_P(\beta)$	0.162	0.146	0.180	0.242	0.157	0.904	0.831
		$l_M(\beta)$	0.037	0.027	0.145	0.150	0.097	1.015	0.959
		$l_{M^*}(\beta)$	0.033	0.022	0.144	0.148	0.095	1.015	0.960
100	4	$l_P(\beta)$	0.498	0.449	0.393	0.634	0.450	0.698	0.577
		$l_M(\beta)$	0.123	0.114	0.206	0.240	0.157	1.030	0.953
		$l_{M^*}(\beta)$	0.121	0.108	0.219	0.250	0.157	0.966	0.946
	6	$l_P(\beta)$	0.295	0.279	0.223	0.370	0.279	0.809	0.650
		$l_M(\beta)$	0.075	0.069	0.154	0.171	0.107	0.995	0.944
		$l_{M^*}(\beta)$	0.066	0.060	0.154	0.168	0.106	0.986	0.944
	10	$l_P(\beta)$	0.149	0.141	0.128	0.196	0.143	0.902	0.749
		$l_M(\beta)$	0.029	0.024	0.104	0.108	0.068	1.011	0.956
		$l_{M^*}(\beta)$	0.025	0.020	0.103	0.106	0.068	1.012	0.957
250	4	$l_P(\beta)$	0.429	0.414	0.221	0.483	0.414	0.763	0.322
		$l_M(\beta)$	0.103	0.100	0.130	0.166	0.115	1.041	0.914
		$l_{M^*}(\beta)$	0.095	0.090	0.133	0.164	0.109	1.011	0.911
	6	$l_P(\beta)$	0.270	0.263	0.135	0.302	0.263	0.817	0.329
		$l_M(\beta)$	0.063	0.060	0.095	0.114	0.077	0.991	0.913
		$l_{M^*}(\beta)$	0.054	0.052	0.095	0.110	0.072	0.983	0.920
	10	$l_P(\beta)$	0.147	0.146	0.080	0.167	0.146	0.897	0.472
		$l_M(\beta)$	0.027	0.027	0.065	0.070	0.047	1.003	0.939
		$l_{M^*}(\beta)$	0.022	0.022	0.064	0.068	0.046	1.005	0.943

Table S12

Inference on $\beta = 1/1.6$ in the probit regression for MCAR longitudinal data. The compared methods are the MNAR profile log-likelihood $l_P(\psi)$, the MNAR MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials, where $\lambda_i = \left(\sum_{t=1}^{T} x_{it}/T + u_i\right)/1.6$, $u_i \sim N(0,1)$.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\psi)$	0.422	0.331	0.760	0.869	0.466	0.573	0.786
		$l_{M^*}(\psi)$	0.137	0.107	0.346	0.372	0.202	0.924	0.977
	6	$l_P(\psi)$	0.233	0.213	0.371	0.438	0.279	0.733	0.827
		$l_{M^*}(\psi)$	0.056	0.049	0.224	0.231	0.144	0.994	0.965
	10	$l_P(\psi)$	0.113	0.120	0.215	0.243	0.160	0.781	0.844
		$l_{M^*}(\psi)$	0.027	0.018	0.143	0.145	0.093	1.066	0.970
100	4	$l_P(\psi)$	0.359	0.318	0.446	0.573	0.362	0.670	0.729
		$l_{M^*}(\psi)$	0.099	0.085	0.208	0.230	0.140	1.037	0.967
	6	$l_P(\psi)$	0.180	0.180	0.266	0.321	0.218	0.695	0.746
		$l_{M^*}(\psi)$	0.033	0.025	0.150	0.154	0.098	1.041	0.962
	10	$l_P(\psi)$	0.095	0.109	0.178	0.201	0.136	0.664	0.774
		$l_{M^*}(\psi)$	0.017	0.011	0.101	0.103	0.067	1.067	0.970
250	4	$l_P(\psi)$	0.282	0.274	0.257	0.381	0.281	0.740	0.646
		$l_{M^*}(\psi)$	0.069	0.065	0.126	0.143	0.090	1.092	0.949
	6	$l_P(\psi)$	0.145	0.173	0.216	0.260	0.205	0.509	0.542
		$l_{M^*}(\psi)$	0.025	0.021	0.094	0.097	0.062	1.021	0.944
	10	$l_P(\psi)$	0.098	0.116	0.136	0.168	0.130	0.538	0.581
		$l_{M^*}(\psi)$	0.013	0.012	0.063	0.064	0.044	1.062	0.963

Table S13

Inference on $\beta = 1/1.6$ in the probit regression for MNAR longitudinal data. The compared methods are the MNAR profile log-likelihood $l_P(\psi)$, Severini's exact MCAR MPL $l_M(\beta)$, the MNAR MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials, where $\lambda_i = \left(\sum_{t=1}^T x_{it}/T + u_i\right)/1.6$, $u_i \sim N(0,1)$.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\psi)$	0.309	0.217	0.802	0.860	0.436	0.592	0.851
		$l_M(\beta)$	0.011	-0.013	0.456	0.456	0.265	0.846	0.957
		$l_{M^*}(\psi)$	0.012	-0.028	0.517	0.517	0.265	0.749	0.957
	6	$l_P(\psi)$	0.167	0.134	0.369	0.405	0.244	0.799	0.898
		$l_M(\beta)$	-0.069	-0.082	0.266	0.275	0.186	0.964	0.924
		$l_{M^*}(\psi)$	-0.045	-0.060	0.264	0.268	0.180	0.986	0.945
	10	$l_P(\psi)$	0.070	0.077	0.247	0.256	0.153	0.745	0.875
		$l_M(\beta)$	-0.118	-0.123	0.167	0.204	0.148	1.027	0.889
		$l_{M^*}(\psi)$	-0.043	-0.052	0.168	0.173	0.118	1.039	0.948
100	4	$l_P(\psi)$	0.261	0.220	0.429	0.503	0.299	0.728	0.829
		$l_M(\beta)$	-0.014	-0.023	0.257	0.258	0.171	0.986	0.950
		$l_{M^*}(\psi)$	-0.025	-0.037	0.258	0.259	0.169	0.993	0.952
	6	$l_P(\psi)$	0.130	0.132	0.270	0.300	0.192	0.752	0.846
		$l_M(\beta)$	-0.085	-0.092	0.177	0.197	0.140	1.019	0.918
		$l_{M^*}(\psi)$	-0.061	-0.070	0.179	0.189	0.131	1.026	0.937
	10	$l_P(\psi)$	0.034	0.078	0.247	0.249	0.133	0.519	0.837
		$l_M(\beta)$	-0.126	-0.128	0.118	0.173	0.135	1.038	0.808
		$l_{M^*}(\psi)$	-0.046	-0.051	0.121	0.129	0.089	1.031	0.930
250	4	$l_P(\psi)$	0.192	0.172	0.262	0.325	0.207	0.761	0.792
		$l_M(\beta)$	-0.051	-0.059	0.170	0.177	0.121	0.974	0.930
		$l_{M^*}(\psi)$	-0.071	-0.081	0.170	0.184	0.130	0.984	0.921
	6	$l_P(\psi)$	0.081	0.115	0.221	0.235	0.151	0.551	0.768
		$l_M(\beta)$	-0.100	-0.100	0.108	0.147	0.106	1.024	0.844
		$l_{M^*}(\psi)$	-0.078	-0.079	0.110	0.135	0.093	1.025	0.885
	10	$l_P(\psi)$	0.019	0.087	0.251	0.252	0.119	0.314	0.708
		$l_M(\beta)$	-0.138	-0.141	0.074	0.157	0.141	1.013	0.531
		$l_{M^*}(\psi)$	-0.049	-0.050	0.075	0.089	0.064	1.034	0.901

Table S14

Inference on $\beta_1 = -1$ in the stratified Weibull regression model for right-censored survival data and probability of censoring $P_c = 0.2$. The compared functions are the profile log-likelihood $l_P(\psi)$ and the MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\psi)$	-0.005	-0.002	0.122	0.123	0.081	0.823	0.894
		$l_{M^*}(\psi)$	0.004	0.005	0.121	0.121	0.080	0.970	0.937
	6	$l_P(\psi)$	-0.003	-0.003	0.096	0.096	0.066	0.880	0.906
		$l_{M^*}(\psi)$	0.002	0.001	0.095	0.095	0.066	0.983	0.941
	10	$l_P(\psi)$	0.001	0.000	0.072	0.072	0.051	0.921	0.935
		$l_{M^*}(\psi)$	0.003	0.004	0.072	0.072	0.052	0.982	0.951
100	4	$l_P(\psi)$	-0.010	-0.008	0.087	0.087	0.059	0.822	0.894
		$l_{M^*}(\psi)$	-0.002	-0.001	0.086	0.086	0.058	0.971	0.943
	6	$l_P(\psi)$	-0.004	-0.004	0.068	0.068	0.046	0.885	0.915
		$l_{M^*}(\psi)$	0.001	0.001	0.067	0.067	0.045	0.988	0.948
	10	$l_P(\psi)$	-0.004	-0.003	0.051	0.052	0.035	0.912	0.925
		$l_{M^*}(\psi)$	-0.002	-0.001	0.051	0.051	0.035	0.976	0.942
250	4	$l_P(\psi)$	-0.006	-0.007	0.054	0.054	0.036	0.845	0.893
		$l_{M^*}(\psi)$	0.002	0.002	0.053	0.053	0.035	0.995	0.943
	6	$l_P(\psi)$	-0.006	-0.006	0.043	0.043	0.029	0.893	0.917
		$l_{M^*}(\psi)$	-0.000	-0.001	0.042	0.042	0.028	1.000	0.949
	10	$l_P(\psi)$	-0.002	-0.002	0.031	0.031	0.022	0.946	0.939
		$l_{M^*}(\psi)$	0.001	0.000	0.031	0.031	0.021	1.012	0.959

Table S15

Inference on $\beta_2 = 1$ in the Weibull regression model for in the stratified Weibull regression model for right-censored survival data and probability of censoring $P_c = 0.2$. The compared functions are the profile log-likelihood $l_P(\psi)$ and the MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\psi)$	0.010	0.009	0.078	0.079	0.053	0.833	0.901
		$l_{M^*}(\psi)$	0.003	0.002	0.078	0.078	0.053	0.976	0.948
	6	$l_P(\psi)$	0.006	0.004	0.054	0.055	0.038	0.887	0.920
		$l_{M^*}(\psi)$	0.002	0.000	0.054	0.054	0.037	0.989	0.951
	10	$l_P(\psi)$	0.002	0.003	0.039	0.039	0.026	0.934	0.932
		$l_{M^*}(\psi)$	-0.000	-0.000	0.039	0.039	0.026	0.999	0.946
100	4	$l_P(\psi)$	0.007	0.007	0.052	0.053	0.035	0.820	0.891
		$l_{M^*}(\psi)$	-0.001	-0.001	0.051	0.051	0.036	0.971	0.943
	6	$l_P(\psi)$	0.005	0.004	0.040	0.040	0.027	0.876	0.911
		$l_{M^*}(\psi)$	-0.000	-0.000	0.039	0.039	0.027	0.981	0.945
	10	$l_P(\psi)$	0.003	0.003	0.028	0.028	0.018	0.936	0.926
		$l_{M^*}(\psi)$	0.001	0.000	0.028	0.028	0.018	1.002	0.944
250	4	$l_P(\psi)$	0.009	0.009	0.034	0.035	0.024	0.824	0.886
		$l_{M^*}(\psi)$	-0.000	-0.001	0.033	0.033	0.023	0.979	0.944
	6	$l_P(\psi)$	0.004	0.004	0.025	0.025	0.017	0.873	0.900
		$l_{M^*}(\psi)$	-0.001	-0.002	0.024	0.024	0.016	0.981	0.941
	10	$l_P(\psi)$	0.002	0.002	0.018	0.018	0.013	0.910	0.926
		$l_{M^*}(\psi)$	-0.000	-0.000	0.018	0.018	0.013	0.972	0.948

Table S16

Inference on $\beta_1 = -1$ in the stratified Weibull regression model for right-censored survival data and probability of censoring $P_c = 0.4$. The compared functions are the profile log-likelihood $l_P(\psi)$ and the MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\psi)$	-0.014	-0.014	0.149	0.150	0.100	0.793	0.876
		$l_{M^*}(\psi)$	0.001	0.003	0.146	0.146	0.096	0.974	0.940
	6	$l_P(\psi)$	-0.008	-0.008	0.116	0.116	0.078	0.860	0.908
		$l_{M^*}(\psi)1$	0.002	0.004	0.115	0.115	0.078	0.984	0.947
	10	$l_P(\psi)$	-0.001	0.000	0.085	0.085	0.059	0.917	0.930
		$l_{M^*}(\psi)$	0.004	0.006	0.084	0.084	0.059	0.995	0.951
100	4	$l_P(\psi)$	-0.016	-0.016	0.107	0.108	0.073	0.786	0.878
		$l_{M^*}(\psi)$	-0.001	0.000	0.104	0.104	0.072	0.971	0.949
	6	$l_P(\psi)$	-0.009	-0.007	0.082	0.082	0.056	0.867	0.909
		$l_{M^*}(\psi)$	0.002	0.003	0.081	0.081	0.055	0.996	0.949
	10	$l_P(\psi)$	-0.006	-0.006	0.061	0.062	0.040	0.902	0.915
		$l_{M^*}(\psi)$	-0.002	-0.000	0.061	0.061	0.040	0.981	0.943
250	4	$l_P(\psi)$	-0.014	-0.013	0.066	0.067	0.045	0.815	0.881
		$l_{M^*}(\psi)$	0.002	0.002	0.064	0.064	0.042	1.004	0.948
	6	$l_P(\psi)$	-0.010	-0.009	0.051	0.052	0.035	0.873	0.897
		$l_{M^*}(\psi)$	0.001	0.002	0.051	0.051	0.033	1.007	0.945
	10	$l_P(\psi)$	-0.005	-0.004	0.038	0.038	0.025	0.925	0.922
		$l_{M^*}(\psi)$	0.000	0.001	0.038	0.038	0.025	1.005	0.951

Table S17

Inference on $\beta_2 = 1$ in the stratified Weibull regression model for right-censored survival data and probability of censoring $P_c = 0.4$. The compared functions are the profile log-likelihood $l_P(\psi)$ and the MCMPL $l_{M^*}(\psi)$ computed with $R = 500$. Results based on a simulation study with 2000 trials.

N	T	Method	B	MB	SD	RMSE	MAE	SE/SD	0.95 CI
50	4	$l_P(\psi)$	0.018	0.014	0.099	0.100	0.065	0.801	0.877
		$l_{M^*}(\psi)$	0.003	-0.000	0.097	0.097	0.064	0.982	0.946
	6	$l_P(\psi)$	0.011	0.010	0.070	0.071	0.048	0.855	0.907
		$l_{M^*}(\psi)$	0.002	0.001	0.069	0.069	0.047	0.983	0.946
	10	$l_P(\psi)$	0.004	0.002	0.048	0.048	0.031	0.932	0.928
		$l_{M^*}(\psi)$	-0.001	-0.002	0.048	0.048	0.030	1.010	0.948
100	4	$l_P(\psi)$	0.015	0.014	0.066	0.067	0.045	0.799	0.880
		$l_{M^*}(\psi)$	0.000	-0.002	0.064	0.064	0.045	0.985	0.946
	6	$l_P(\psi)$	0.010	0.011	0.050	0.051	0.035	0.851	0.899
		$l_{M^*}(\psi)$	0.000	0.000	0.049	0.049	0.034	0.977	0.944
	10	$l_P(\psi)$	0.006	0.006	0.034	0.035	0.024	0.935	0.929
		$l_{M^*}(\psi)$	0.001	0.001	0.034	0.034	0.023	1.015	0.953
250	4	$l_P(\psi)$	0.017	0.017	0.042	0.045	0.030	0.810	0.864
		$l_{M^*}(\psi)$	0.000	0.000	0.041	0.041	0.027	1.011	0.950
	6	$l_P(\psi)$	0.009	0.008	0.031	0.032	0.021	0.851	0.896
		$l_{M^*}(\psi)$	-0.002	-0.003	0.030	0.030	0.020	0.982	0.943
	10	$l_P(\psi)$	0.005	0.005	0.022	0.023	0.016	0.911	0.919
		$l_{M^*}(\psi)$	-0.000	-0.001	0.022	0.022	0.015	0.986	0.942