Bootstrap-based bias corrections for INAR count time series

C. H. Weiß and C. Jentsch

Abstract

Integer-valued autoregressive (INAR) processes form a very useful class of processes suitable to model time series of counts. Several practically relevant estimators based on INAR data are known to be systematically biased away from their population values, e.g. sample autocovariances, sample autocorrelations, or the dispersion index. We propose to do bias correction for such estimators by using a recently proposed INAR-type bootstrap scheme that is tailor-made for INAR processes, and which has been proven to be asymptotically consistent under general conditions. This INAR bootstrap allows an implementation with and without parametrically specifying the innovations’ distribution. To judge the potential of corresponding bias correction, we compare these bootstraps in simulations to several competitors that include the AR bootstrap and block bootstrap. Finally, we conclude with an illustrative data application.

1. Introduction

During the last decades, researchers paid more and more attention to discrete-valued time series, and especially count time series, which have the range \( \mathbb{N}_0 = \{0, 1, 2, \ldots\} \), became a vivid research area [1]. For count time series with an autoregressive autocorrelation structure, i.e. where the autocorrelation function \( \rho(k) := \text{Corr}[X_t, X_{t-k}] \) follows the well-known Yule-Walker equations

\[ \rho(k) = \sum_{i=1}^{p} \alpha_i \rho(|k-i|) \quad \text{for } k \geq 1 \]

for some autoregressive order \( p \in \mathbb{N} = \{1, 2, \ldots\} \) and corresponding autoregressive parameters \( \alpha_1, \ldots, \alpha_p \), a widely used approach is to consider the family of INAR(p) models according to [2]. These integer-valued autoregressive models, which contain the famous INAR(1) model by [3,4] as a special case for \( p = 1 \), are defined by a recursive scheme looking like the conventional AR(p) recursion, but with the multiplications being replaced by the binomial thinning operator of [5]: If \( X \) is a count random variable and if the thinning
parameter satisfies $\alpha \in [0;1]$, then the random variable $\alpha \circ X$ is said to arise from $X$ by \textit{binomial thinning} if $\alpha \circ X$ has a conditional binomial distribution given the value of $X$, i.e. $\alpha \circ X | X \sim \text{Bin}(X, \alpha)$. Using this probabilistic operator, Du and Li [2] define the \textit{INAR($p$)} \textit{model} by the recursion

$$X_t = \alpha_1 \circ_t X_{t-1} + \cdots + \alpha_p \circ_t X_{t-p} + \epsilon_t,$$

where $\alpha_\bullet := \sum_{j=1}^{p} \alpha_j < 1$. Here, the innovations $(\epsilon_t)$ constitute an independent and identically distributed (i.i.d.) count process having some distribution $G$ with range $\mathbb{N}_0$, with $\epsilon_t$ being independent of $(X_t)_{t < t'}$ and with the thinning operations being performed independently of each other and of $(\epsilon_t)$. In particular, the thinnings $(\alpha_1 \circ_{t+1} X_{t}, \ldots, \alpha_p \circ_{t+p} X_{t})$ given $X_t$ are assumed to be conditionally independent.

For estimating the model parameters of an INAR($p$) process, Yule-Walker (YW) estimation has been considered by [2,6]. Conditional least squares (CLS) estimation has been proposed by [2], and [7] investigated maximum likelihood (ML) estimation (which requires to specify the $\epsilon_t$'s distribution family). Besides these classical approaches, Drost et al. [8] proposed a \textit{semi-parametric maximum likelihood (SPML) estimator}, which only relies on the INAR($p$) model structure, but which, in contrast to the ML approach discussed by [7], does not require a distributional assumption with respect to the innovations $\epsilon_t$. In contrast to YW and CLS estimation, in turn, the SPML approach does not only estimate a couple of innovations’ moments, but the full innovations’ distribution (without assuming a parametric model). Such a distribution-free approach is very welcome in practice, since for model order $p \geq 2$, it is rather difficult to identify an appropriate distribution for the INAR's innovations, because, e.g., in contrast to INAR(1), dispersion properties of the innovations do not carry over to the observations, see [9].

For the particular instance of a Poisson INAR(1) process, i.e. with autoregressive order $p = 1$ and with Poisson-distributed innovations $(\epsilon_t)$, the performance of some estimators (or further statistics being computed from the given time series) has been investigated via simulations by several authors, e.g. [3,10–12]; Bu et al. [7] considered the Poisson INAR(2) model in an analogous manner. Such estimators are known to be asymptotically consistent, but especially for small sample sizes, these authors reported about bias problems in estimation. For the specific Poisson INAR(1) model, it is possible to derive analytically asymptotic bias corrections for many relevant statistics [11,12]. However, the model complexity increases substantially if either the Poisson assumption or the limitation to model order $p = 1$ does not hold. Hence, it is not clear if such closed-form corrections could be derived also for differently distributed innovations or for higher model orders $p$, but this task clearly becomes a lot more involved. Therefore, in this work, we propose a general scheme for bias correction, which relies on the INAR bootstrap scheme recently developed by [13].

Bias correction for univariate AR processes has been put forward by a number of authors. Starting with [14], such approaches (with a focus on bias correction for the AR slope parameter) have been addressed by many authors during the last decades, see e.g. [15–17] for references. These include analytical as well as numerical or simulation-based approaches. In a multivariate context, analytical expressions for the finite-sample bias in estimated vector autoregressive (VAR) models have been developed by [18], see also [17] for references. The first to employ the (residual-based AR) bootstrap for bias correction for the VAR slope coefficient estimation was [15,16], see also [17,19] for discussions.
Whereas [18] derived closed-form expressions for the asymptotic first-order bias of the slope coefficients in the VAR model only without a deterministic time trend, the motivation of [15,16] to use the bootstrap for bias correction was to be able to deal easily with deterministic trends as well. A detailed simulation study for comparing different bias correction approaches in VAR models is provided by [17]. Similarly, in the context of INAR models, analytical expressions for the bias of most relevant statistics are available only for the special case of a Poisson INAR(1) model, so under order \( p = 1 \) and under a parametric assumption on the innovations’ distribution, see Section 3.3 for further details. Hence, the idea to make use of distribution-free bootstrap methods appears to be rather natural to achieve robust bias correction. While most papers address only AR coefficient estimates, we look at different relevant estimators computed from INAR data.

In Section 2, we introduce the general INAR bootstrap framework considered in [13], before corresponding parametric and semi-parametric implementations are discussed. The latter approach crucially depends on the SPML estimator proposed by [8], which is addressed as well. In Section 3, we describe in detail how INAR bootstraps and competitors will be applied for bias correction. In Section 4, we investigate the finite sample performances of such bias corrections for several INAR DGPs. Two illustrative data examples are discussed in Section 5, and we conclude in Section 6. Theoretical details will be deferred to an appendix, and further simulation results to a supplement.

2. The INAR bootstrap

Recently, Jentsch and Weiβ [13] proposed an INAR bootstrap scheme, which can be applied to statistics \( T_n = T_n(X) \), where \( X = (X_1, \ldots, X_n) \), belonging to the class of functions of generalized means, i.e. to statistics of the form

\[
T_n = f\left( \frac{1}{n_m} \sum_{t=1}^{n_m} g(X_t, \ldots, X_{t+m-1}) \right),
\]

where \( m \in \mathbb{N} \) (fixed) and \( n_m = n - m + 1 \), and where the functions \( g : \mathbb{R}^m \to \mathbb{R}^d \) and \( f : \mathbb{R}^d \to \mathbb{R} \) have to satisfy some smoothness conditions. Note that the possible types of statistics include e.g. (functions of) the sample mean, sample autocovariances, and sample (partial) autocorrelations.

The general INAR bootstrap scheme proposed by [13] applies to INAR(p) processes (1). Under some mild additional requirements, they were able to prove consistency of the INAR bootstrap. In our application of the INAR bootstrap scheme, we consider both a parametric implementation with \( G = G_\theta \) and corresponding estimator \( \hat{G} = G_{\hat{\theta}} \) as well as a semi-parametric implementation in the sense of [8]. The parametric implementation in general requires the proper choice and estimation of a parametric distribution (family), whereas the semi-parametric version comes without specifying the innovations distribution; see [13] for details. For the parametric case, the most prominent specification is a Poisson INAR model, i.e. \( G_\theta = \text{Poi} (\theta) \), which will be considered also below.

To implement the INAR bootstrap procedure, it is necessary to specify the estimation \( \hat{\alpha}_1, \ldots, \hat{\alpha}_p \) of the INAR coefficients as well as the estimation \( \hat{G} \) of the innovations’ distribution \( G \), which can be done in multiple ways. One such implementation is the fully parametric approach according to [13]:
Parametric INAR Bootstrap

Step 1. Fit an INAR(p) model with \( \epsilon_t \sim G_\theta \) to the data to get (e.g. Yule-Walker) estimates \( \hat{\alpha} = (\hat{\alpha}_1, \ldots, \hat{\alpha}_p) \) and \( \hat{\theta} \).

Step 2. Generate bootstrap observations \( X_t^*, \ldots, X_n^* \) according to

\[
X_t^* = \hat{\alpha}_1 \circ X_{t-1}^* + \cdots + \hat{\alpha}_p \circ X_{t-p}^* + \epsilon_t^*,
\]

where \( \circ \) denotes (mutually independent) bootstrap binomial thinning operations, and where \( (\epsilon_t^*) \) are i.i.d. random variables following the distribution \( G_{\hat{\theta}} \).

Compute statistic (2) for the bootstrap data, i.e. \( T_n^* = T_n(X_1^*, \ldots, X_n^*) \).

Together with some weak assumptions on the parametric innovation distribution \( G \), Jentsch and Weiß [13] proved bootstrap consistency.

As an alternative, we also consider the semi-parametric approach in [13] to implement the INAR bootstrap procedure. It relies on the SPML estimator by [8] for INAR(p) models in (1), i.e. the semi-parametric joint estimation of the INAR coefficients \( \alpha_1, \ldots, \alpha_p \) and the innovation distribution \( G \):

\[
(\hat{\alpha}_{sp}, \hat{G}_{sp}) := (\hat{\alpha}_{sp,1}, \ldots, \hat{\alpha}_{sp,p}, \hat{G}_{sp}(0), \hat{G}_{sp}(1), \hat{G}_{sp}(2), \ldots)
\]

is defined to maximize the conditional log-likelihood function, where the estimated probabilities \( \hat{G}_{sp}(k) \) (which estimate \( G(k) = P(\epsilon_t = k) \)) are equal to 0 for \( k > \max \{X_t \mid t = 1, \ldots, n\} \) and for \( k < \min \{X_t - \sum_{i=1}^p X_{t-i} \mid t = p+1, \ldots, n\} \). The conditional log-likelihood is computed by using the fact that the conditional distribution of \( X_t \) given \( X_{t-1}, \ldots, X_{t-p} \) is a convolution of the binomial distributions \( \text{Bin}(X_{t-1}, \alpha_1), \ldots, \text{Bin}(X_{t-p}, \alpha_p) \) and of \( G \); also see Appendix E.1 in [13] for computational issues.

Semi-parametric INAR Bootstrap

Step 1. Fit semi-parametrically an INAR(p) model to get the estimated INAR coefficients \( \hat{\alpha}_{sp} = (\hat{\alpha}_{sp,1}, \ldots, \hat{\alpha}_{sp,p})' \) and the estimator \( \hat{G}_{sp} = (\hat{G}_{sp}(k), k \in \mathbb{N}_0) \) for \( G \).

Step 2. Generate bootstrap observations \( X_t^*, \ldots, X_n^* \) according to

\[
X_t^* = \hat{\alpha}_{sp,1} \circ X_{t-1}^* + \cdots + \hat{\alpha}_{sp,p} \circ X_{t-p}^* + \epsilon_t^*,
\]

where \( \circ \) denotes (mutually independent) bootstrap binomial thinning operations and \( (\epsilon_t^*) \) are i.i.d. random variables following \( \hat{G}_{sp} \).

Compute statistic (2) for the bootstrap data, i.e. \( T_n^* = T_n(X_1^*, \ldots, X_n^*) \).

With the assumptions specified in Corollary 3.2 in [13], the semi-parametric INAR bootstrap is consistent.

3. INAR Bootstrap for bias correction

In [13], the finite-sample performance of both the parametric and the semi-parametric INAR bootstrap was investigated with respect to confidence intervals for several statistics \( T_n(X) \) in the sense of (2). It turned out that both INAR bootstrap schemes (3) and (4) lead to satisfactory results, where, however, the performance of the parametric INAR bootstrap (3) deteriorated if the required parametric assumption (in this case, a Poisson assumption)
was violated. The semi-parametric INAR bootstrap (4) was robust in these respects, and it also turned out to be superior to some further opponents: a naïve application of the conventional AR bootstrap (by just ignoring the discreteness of counts), a Markov bootstrap (which suffered severely from the huge number of parameters to be estimated), and a circular block bootstrap (the performance of which converged much more slowly with increasing sample size \( n \)). These positive results encourage us to apply the INAR bootstrap schemes for a completely different task, the bias correction of point estimators for some relevant statistics \( T_n(X) \).

We provide the general approach to correct for a bias using the INAR bootstrap proposal, before we discuss the implementation for all statistics under consideration and, in particular, the corresponding centering schemes. Furthermore, we discuss several competitors based on analytical expressions and other resampling techniques.

### 3.1. Bootstrap-based bias correction

In general, let \( \vartheta \) be a parameter of interest, which is estimated from the INAR(p) data \( X \) by the biased estimator \( \hat{\vartheta} \) (computed according to an expression \( T_n(X) \) in the sense of (2)). For the INAR-bootstrap bias correction, we apply either the parametric or the semi-parametric INAR bootstrap. Then, depending on the selected INAR bootstrap implementation, we compute an appropriate centering \( \text{cent}(\hat{\vartheta}^\ast) \), which we obtain from a functional relationship between the population quantity \( \vartheta \) and the model parameters that are estimated for the bootstrap procedure itself (see details below). We execute the bootstrap procedure \( B \) times, where \( B \) is large, to get bootstrap samples \( X^\ast(b) \), \( b = 1, \ldots, B \), where \( X^\ast(b) = (X^\ast_1(b), \ldots, X^\ast_n(b)) \). For each bootstrap sample \( X^\ast \), we compute the centred bootstrap estimates

\[
\hat{\vartheta}^\ast \text{cent} := \hat{\vartheta}^\ast - \text{cent}(\hat{\vartheta}^\ast)
\]

to get in total \( \hat{\vartheta}^\ast \text{cent}, b = 1, \ldots, B \). Finally, we calculate the mean of the centred bootstrap estimates

\[
\overline{\hat{\vartheta}^\ast \text{cent}} := \frac{1}{B} \sum_{b=1}^{B} \hat{\vartheta}^\ast \text{cent}
\]

and define the bias-corrected estimate for \( \vartheta \) by

\[
\hat{\vartheta} \text{corr} := \hat{\vartheta} - \overline{\hat{\vartheta}^\ast \text{cent}}.
\]

### 3.2. Considered statistics and centering schemes

Let \( \bar{X} := (1/n) \sum_{t=1}^{n} X_t \) denote the sample mean, which is an unbiased estimator of the observations’ mean \( \mu \). In our simulations presented in Section 4, we consider the following biased estimators \( \hat{\vartheta} \), which refer to important characteristics of a count process [11,12]:

- \( \hat{\gamma}(0) := (1/n) \sum_{t=1}^{n} (X_t - \bar{X})^2 \); estimator of the observations’ variance \( \sigma^2 \);
- \( \hat{I}_{\text{disp}} := \hat{\gamma}(0)/\bar{X} \); estimator of the observations’ dispersion index \( \sigma^2/\mu \);
- \( \hat{\gamma}(1) := (1/n) \sum_{t=2}^{n} (X_t - \bar{X})(X_{t-1} - \bar{X}) \); estimator of the first-order autocovariance \( \gamma(1) \);
• \( \hat{\rho}(1) := \hat{\gamma}(1)/\hat{\gamma}(0) \); estimator of the first-order autocorrelation \( \rho(1) \);

• \( \bar{X}(1 - \hat{\rho}(1)) \); estimator for \( \mu_X(1 - \rho_X(1)) \), a quantity, which coincides with the innovations' mean \( \mu_\epsilon \) for autoregressive order \( p = 1 \).

To be able to capture the finite sample bias of \( \hat{\theta} \), the use of an appropriate centering scheme is essential when computing the respective bias-corrected estimates \( \hat{\theta}_{\text{corr}} \). These centering schemes are computed based on the respective model assumption, see Section 3.1. For example, the population mean of an INAR(p) model (1), which is estimated by the sample mean \( \bar{X} \), is known to satisfy the functional relationship \( \mu_X := \mu_\epsilon/(1 - \alpha_1 - \ldots - \alpha_p) \). This, in turn, leads to a centering for \( \bar{X}^* \) defined as \( \text{cent}(\bar{X}^*) := \hat{\mu}_{\epsilon; sp}/(1 - \hat{\alpha}_{1;sp} - \ldots - \hat{\alpha}_{p; sp}) \) (semi-parametric implementation); see details below. In the sequel, we provide expressions for centering \( \bar{X}^* \), \( \hat{\gamma}^*(0) \) and \( \hat{\gamma}^*(1) \). Since the remaining estimators are functions thereof, we apply the corresponding function to \( \text{cent}(\bar{X}^*), \ldots \) to obtain an appropriate centering scheme, e.g. we centre \( \hat{\rho}^*(1) := \hat{\gamma}^*(1)/\hat{\gamma}^*(0) \) by \( \text{cent}(\hat{\rho}^*(1)) := \text{cent}(\hat{\gamma}^*(1))/\text{cent}(\hat{\gamma}^*(0)) \).

For the semi-parametric INAR bootstrap scheme (4), we assume an underlying INAR(p) process with specified model order \( p \), but without further distributional assumptions. We obtain the estimated INAR coefficients \( \hat{\alpha}_{1;sp}, \ldots, \hat{\alpha}_{p; sp} \) and innovations' probabilities \( \hat{g}_{sp}(k) \) for \( k = 0, \ldots, \max X \), and using the latter, we compute \( \hat{\mu}_{\epsilon; sp} := \sum_{k=0}^{\max X} X_k \cdot \hat{g}_{sp}(k) \) and \( \hat{\sigma}_{\epsilon; sp}^2 := \sum_{k=0}^{\max X} k^2 \cdot \hat{g}_{sp}(k) - \hat{\mu}_{\epsilon; sp}^2 \). These are used to compute the following centerings:

• \( \text{cent}(\bar{X}^*) := \hat{\mu}_{\epsilon; sp}/(1 - \hat{\alpha}_{1;sp} - \ldots - \hat{\alpha}_{p; sp}) \);

• \( \text{cent}(\hat{\gamma}^*(0)) := \text{cent}(\bar{X}^*) \).

\[
\times \begin{cases} 
\frac{\hat{\alpha}_{1;sp} + \frac{\hat{\sigma}_{\epsilon; sp}^2}{\hat{\mu}_{\epsilon; sp}}}{1 + \hat{\alpha}_{1;sp}} & \text{for } p = 1, \\
1 - \hat{\alpha}_{1;sp}^2 - \hat{\alpha}_{2;sp}^2 + \left( \frac{\hat{\sigma}_{\epsilon; sp}^2}{\hat{\mu}_{\epsilon; sp}} - 1 \right) (1 - \hat{\alpha}_{1;sp} - \hat{\alpha}_{2;sp}) & \text{for } p = 2; \\
1 - \hat{\alpha}_{1;sp}^2 - \hat{\alpha}_{2;sp}^2 - 2\hat{\alpha}_{1;sp}^2 \frac{\hat{\alpha}_{2;sp}}{1 - \hat{\alpha}_{2;sp}} & 
\end{cases}
\]

• \( \text{cent}(\hat{\gamma}^*(1)) := \text{cent}(\hat{\gamma}^*(0)^*) \left( 1 - \frac{1}{n} \right) \times \begin{cases} \frac{\hat{\alpha}_{1;sp}}{\hat{\alpha}_{1;sp}} & \text{for } p = 1, \\
\frac{\hat{\alpha}_{1;sp}}{1 - \hat{\alpha}_{2;sp}} & \text{for } p = 2. 
\end{cases} \)

The corresponding functional relationships can be found, e.g. in [1,9]. Here and in Section 4, we restrict ourselves to model orders \( p \leq 2 \), but centering for higher orders \( p \) could be computed by utilizing the respective Yule-Walker equations, see [2].

For the parametric INAR bootstrap scheme (3), we additionally assume a distribution family for the innovations \( (\epsilon_t) \). Here, we consider the Poisson INAR(p) model, i.e. we have \( \epsilon_t \sim \text{Poi}(\mu_\epsilon) \) and thus equidispersed innovations, i.e. \( \sigma_{\epsilon}^2/\mu_\epsilon = 1 \). Note, however, that the observations become equidispersed (in fact even Poisson distributed) only for model order \( p = 1 \) [9]. Using the estimators \( \hat{X} \) for the marginal mean \( \mu \) and \( \hat{\alpha}_{1;YW}, \ldots, \hat{\alpha}_{p;YW} \)
(computed from the sample autocorrelation function according to the Yule-Walker equations of the INAR(p) model), we obtain the following centerings:

- \( \text{cent}(\bar{X}^*) := \bar{X} \);
- \( \text{cent}(\hat{\gamma}(0)^*) := \text{cent}(\bar{X}^*) \cdot \begin{cases} 1 & \text{for } p = 1, \\ 1 - \hat{\alpha}_{1;YW}^2 - \hat{\alpha}_{2;YW}^2 & \text{for } p = 2; \end{cases} \)
- \( \text{cent}(\hat{\gamma}(1)^*) := \text{cent}(\hat{\gamma}(0)^*) \left( 1 - \frac{1}{n} \right) \cdot \begin{cases} \hat{\alpha}_{1;YW} & \text{for } p = 1, \\ \hat{\alpha}_{1;YW} & \text{for } p = 2. \end{cases} \)

### 3.3. Possible competitors

In our simulation study presented in Section 4, we compare our above INAR-bootstrap bias corrections also to bias corrections obtained by (naïvely) applying the conventional AR bootstrap as well as the circular block bootstrap (using the corrected version of the automatic block-length selection procedure by [20,21]); for the corresponding centering schemes, we refer to Appendix E.2 in [13]. But in view of the bad experiences by the latter authors, we did not consider again the Markov bootstrap as a competitor.

Furthermore, for model order \( p = 1 \) and assuming an underlying Poisson INAR(1) DGP, asymptotic approximations for the biases of the considered statistics \( \hat{\gamma}(0), \hat{I}_{\text{disp}}, \hat{\gamma}(1), \hat{\rho}(1) \) and \( \bar{X}(1 - \hat{\rho}(1)) \) are available in [11,12]. These are summarized in the centre column of Table 1. Plugging-in Yule-Walker estimates in these expressions instead of the true (but unknown) parameter values, a simple asymptotic bias correction is possible.

We also added the corresponding approximate bias corrections for an AR(1) DGP to Table 1 (right column; the derivation is sketched in Appendix), as this will help us in interpreting the simulation results below. Note that the distribution of the considered AR(1) model is not specified, it is only assumed that relevant moments exist.

| Statistic          | Pois-INAR(1)                  | AR(1)                      |
|--------------------|--------------------------------|-----------------------------|
| \( \hat{\gamma}(0) \) | \( -\frac{1}{n} \left( 1 + \alpha \right) \frac{1 + \alpha}{1 - \alpha} \) | \( -\frac{1}{n} \left( 1 + \alpha \right) \frac{1 + \alpha}{1 - \alpha} \) |
| \( \hat{I}_{\text{disp}} \) | \( -\frac{1}{n} \left( 1 + \alpha \right) \frac{1 + \alpha}{1 - \alpha} \) | \( \frac{1}{n} \left( \frac{\sigma^2 (\sigma^2 - \mu^2)}{\mu^3} \right) \frac{1 + \alpha}{1 - \alpha} - \frac{\tilde{\mu}_3}{\mu^2} \frac{1 + \alpha + \alpha^2}{1 - \alpha} \) |
| \( \hat{\gamma}(1) \) | \( -\frac{1}{n} \left( 1 + \frac{3\alpha}{\mu} \right) \) | \( -\frac{1}{n} \left( 1 + \frac{3\alpha}{\mu} \right) \) |
| \( \hat{\rho}(1) \) | \( -\frac{1}{n} \left( 1 + 3\alpha + \frac{\alpha}{\mu} \right) \) | \( -\frac{1}{n} (1 + 3\alpha) \) |
| \( \bar{X}(1 - \hat{\rho}(1)) \) | \( \frac{1}{n} (1 + 3\alpha) \mu \) | \( \frac{1}{n} (1 + 3\alpha) \mu \) |
4. Performance of bias corrections

To evaluate the performance of the bias corrections presented in Section 3, a comprehensive simulation study was done, where mean and standard deviation of the (bias-corrected) estimates were computed. For each of the considered scenarios, 500 Monte-Carlo samples were simulated, and for each sample, in turn, 500 bootstrap replicates. For cases where the stationary distribution was not directly available, a pre-run of length 100 was used to approximately reach the steady state. Poisson innovations (‘Poi-INAR’) were parametrized as \( \mu \in \{1, 2.5\} \), and negative binomial innovations (‘NB-INAR’) as \( \mu \in \{1, 2.5\} \) and \( \sigma^2 / \mu \in \{1.5, 2.0\} \) (the parameters \( n, \pi \) of the NB\((n, \pi)\)-distribution were computed from the relations \( \sigma^2 / \mu = 1/\pi \) and \( \mu = n((1 - \pi) / \pi) \)). The generated INAR(1) and INAR(2) processes satisfied \( \rho(1) \in \{0.4, 0.8\} \), and for INAR(2) we chose \( \alpha_2 \in \{0.15, 0.3\} \) and \( \alpha_1 = \rho(1) (1 - \alpha_2) \). The considered sample sizes are \( n \in \{100, 250, 500, 1000\} \). All computations were done using MATLAB, the complete results are provided as a Supplementary file. In the sequel, excerpts from these results are shown to illustrate our general findings, where we use the following abbreviations: ‘spINAR’ (semi-parametric INAR bootstrap), ‘pINAR’ (parametric INAR bootstrap), ‘CBB’ (circular block bootstrap), ‘AR’ (AR bootstrap), ‘asymp’ (asymptotic bias correction), ‘no’ (no bias correction).

Let us start with the case, where the data generating process (DGP) follows a Poi-INAR(1) model, and where both types of INAR bootstrap as well as the AR bootstrap correctly assume the autoregressive order \( p = 1 \). As illustrated by the simulated means shown in the left part of Tables 2 (concerning estimator \( \tilde{X} (1 - \hat{\rho}(1)) \)) and 3 (concerning estimator \( \hat{I}_{\text{disp}} \)), any of the considered bias corrections leads to an improved mean, but with the CBB bootstrap being clearly worst. The latter is reasonable since the CBB bootstrap does not make use of any assumptions on the model structure. Since the DGP is Poi-INAR(1), and since both bias corrections ‘pINAR’ and ‘asymp’ rely on this Poisson assumption, it is clear that these corrections work very well in Tables 2 and 3. But notably, the spINAR-correction (which does not use the Poisson assumption) works equally well.

Finally, let us look at the AR bootstrap. Although it has some corrective effect on estimator \( \hat{I}_{\text{disp}} \) in Table 3, it is clearly worse than ‘(s)pINAR’ and ‘asymp’, which is plausible as the DGP is not an AR process. In particular, we can see in Table 1 that the bias of the dispersion index based on AR data has a very different form in comparison to the bias when computed from INAR data. For the autocorrelation, the AR bootstrap is not consistent; see Appendix D in [13]. However, the bias expressions in Table 1 are very similar and differ only by a minor additional term for the INAR case, such that the AR bootstrap will work also well here. Furthermore, as the AR bootstrap will be generally consistent for the sample mean even for INAR data, it is not surprising that AR bootstrapping shows a good performance

| \( n \) | spINAR | pINAR | CBB | AR | asymp | no | spINAR | pINAR | CBB | AR | asymp | no |
|-------|--------|-------|-----|----|-------|----|--------|-------|-----|----|-------|----|
| 100   | 1.051  | 1.054 | 1.093| 1.054| 1.059 | 1.221| 0.362  | 0.368 | 0.370| 0.370| 0.365 | 0.419|
| 250   | 1.026  | 1.027 | 1.039| 1.025| 1.028 | 1.094| 0.225  | 0.225 | 0.229| 0.227| 0.226 | 0.243|
| 500   | 1.009  | 1.009 | 1.012| 1.009| 1.009 | 1.043| 0.148  | 0.148 | 0.149| 0.148| 0.148 | 0.153|
| 1000  | 1.002  | 1.002 | 1.003| 1.002| 1.002 | 1.019| 0.100  | 0.100 | 0.100| 0.100| 0.100 | 0.102|
Table 3. DGP Poi-INAR(1) with $\mu_\epsilon = 1, \alpha = 0.8$. Estimator $\hat{l}_{\text{disp}}$, true value $l_{\text{disp}} = \sigma^2 / \mu = 1.000$.

| n     | spINAR | pINAR | CBB | AR | asymp | no     | RMSE of estimates $\hat{l}_{\text{disp}}$ |
|-------|--------|-------|-----|----|-------|--------|-----------------------------------------|
| 100   | 0.996  | 0.995 | 0.967 | 0.976 | 0.999 | 0.921 | 0.298 |
| 250   | 0.990  | 0.991 | 0.981 | 0.986 | 0.991 | 0.958 | 0.189 |
| 500   | 0.999  | 0.999 | 0.996 | 0.996 | 0.999 | 0.981 | 0.139 |
| 1000  | 1.001  | 1.001 | 0.999 | 0.999 | 1.001 | 0.992 | 0.097 |

Figure 1. DGP Poi-INAR(1) with $\mu_\epsilon = 1, \alpha = 0.8$. Estimators $\bar{X} (1 - \hat{\rho}(1))$ (left) and $\hat{l}_{\text{disp}}$ (right).

in correcting the bias of $\bar{X} (1 - \hat{\rho}(1))$ as well, see Table 2. Actually, the corresponding bias approximations in Table 1 are even identical.

Besides the bias of an estimator, it is also important to look at the mean squared error (or the square-root thereof, hereafter abbreviated as ‘RMSE’), because this measure is not only sensitive to the bias but also to the dispersion of the estimator. A reduction in bias often tends to be associated with an increase in the variance of an estimator. Hence, it will not be generally possible to reduce the (R)MSE as well. Looking again at Tables 2 and 3 (respective right part), we see that the RMSE is either also improved by using an INAR (or asymptotic) bias correction (Table 2), or it is only slightly increased as in Table 3. But compared to the strong improvement in bias, the small increase in RMSE appears to be acceptable for practice, which is also supported by the box plots shown in Figure 1: the uncorrected estimates (‘no’) look misplaced with respect to the true parameter value, but there is no visible difference in dispersion. The reason for a possible increase in RMSE is an increase in the standard deviation (SD) of the estimators, see Table 4.

Let us return to Table 3, where we concluded that the AR bootstrap performs worse than the INAR bootstraps because it corrects for a differing bias according to Table 1. But the picture is different in Table 5 about the variance estimator $\hat{\gamma}(0)$. Note that the expression derived by [22] implies that asymptotically, all AR(1)-like DGP$s lead to the same asymptotic bias for $\hat{\gamma}(0)$, namely $-(1/n)((1 + \alpha)/(1 - \alpha)) \sigma^2$ (with $\sigma^2 = \mu$ in the Poisson case), also see Table 1. Hence, although we see some deviation for small sample size $n = 100$, INAR and AR bootstraps tend to be very similar for larger $n$ in correcting the bias for $\hat{\gamma}(0)$.
Table 4. DGP Poi-INAR(1) with $\mu_e = 1, \alpha = 0.8$. Estimators $\hat{X} (1 - \hat{\rho}(1))$ and $\hat{I}_{disp}$.

| n   | spINAR | pINAR | CBB   | AR    | asymp | no   | spINAR | pINAR | CBB   | AR    | asymp | no   |
|-----|--------|-------|-------|-------|-------|-----|--------|-------|-------|-------|-------|-----|-----|
| 100 | 0.359  | 0.364 | 0.358 | 0.366 | 0.360 | 0.357| 0.298  | 0.300 | 0.307 | 0.302 | 0.302 | 0.281|
| 250 | 0.224  | 0.224 | 0.225 | 0.225 | 0.224 | 0.224| 0.188  | 0.190 | 0.194 | 0.194 | 0.191 | 0.184|
| 500 | 0.147  | 0.148 | 0.148 | 0.148 | 0.148 | 0.147| 0.139  | 0.139 | 0.142 | 0.141 | 0.139 | 0.137|
| 1000| 0.100  | 0.100 | 0.100 | 0.100 | 0.100 | 0.100| 0.097  | 0.097 | 0.098 | 0.098 | 0.097 | 0.097|

Table 5. DGP Poi- or NB-INAR(1) with $\mu_e = 1, \alpha = 0.8$. Estimator $\hat{\gamma}(0)$, true values $\gamma(0) = \sigma^2 = 5.000$ (Poi-INAR(1), left part) and 7.778 (NB-INAR(1), right part), respectively.

| n   | Poi-INAR(1), i.e. $\sigma^2/\mu_e = 1$; | NB-INAR(1) with $\sigma^2/\mu_e = 1$; |
|-----|----------------------------------------|----------------------------------------|
|     | means of estimates $\hat{\gamma}(0)$   | means of estimates $\hat{\gamma}(0)$   |
|-----|----------------------------------------|----------------------------------------|
| 100 | 4.945  | 7.739  | 4.945  | 7.772  |
| 250 | 4.923  | 7.772  | 4.925  | 7.753  |
| 500 | 4.991  | 7.806  | 4.979  | 7.772  |
| 1000| 5.009  | 7.807  | 5.010  | 7.785  |

Table 6. DGP Poi-INAR(2) with $\mu_e = 1, \rho(1) = 0.8, \alpha_2 = 0.15$. Estimator $\hat{I}_{disp}$, true value $I_{disp} = \sigma^2/\mu = 1.464$.

| n   | Falsely assuming $p = 1$, $\gamma(0)$ | Correctly assuming $p = 2$, $\gamma(0)$ |
|-----|---------------------------------------|----------------------------------------|
|     | means of estimates $\hat{I}_{disp}$    | means of estimates $\hat{I}_{disp}$    |
|-----|---------------------------------------|----------------------------------------|
| 100 | 1.375  | 1.440  | 1.379  | 1.443  | 1.382  | 1.438 | 1.382  | 1.438 | 1.305  | 1.407  | 1.305 |
| 250 | 1.432  | 1.467  | 1.435  | 1.449  | 1.438  | 1.449 | 1.438  | 1.449 | 1.401  | 1.452  | 1.401 |
| 500 | 1.431  | 1.451  | 1.432  | 1.443  | 1.438  | 1.443 | 1.438  | 1.443 | 1.416  | 1.441  | 1.416 |
| 1000| 1.456  | 1.467  | 1.457  | 1.464  | 1.459  | 1.463 | 1.459  | 1.463 | 1.449  | 1.463  | 1.449 |

But even more interesting is a comparison of both parts of Table 5: in the left part, the DGP is Poi-INAR(1) as assumed by ‘pINAR’ and ‘asymp’, whereas this is violated in the right part about an NB-INAR(1) DGP. As a result, these bias corrections perform clearly worse for the NB-INAR(1) DGP, while ‘spINAR’ still successfully corrects the bias and is clearly best among all approaches.

Next, we turn our attention to autoregressive order $p = 2$, i.e. the DGP is an INAR(2) process. Nevertheless, it may happen that we falsely assume order $p = 1$ when applying our techniques for bias correction. The result of such a model underfitting becomes clear from Table 6, where the dispersion index is estimated by either falsely assuming $p = 1$, or correctly assuming $p = 2$. Note that the considered Poi-INAR(2) process is marginally overdispersed despite its Poisson innovations. Furthermore, if assuming model order $p > 1$, it makes no sense anymore to consider the asymptotic correction as this relies on order $p = 1$. Looking at Table 6, it becomes clear that both INAR bootstrap schemes suffer from an underfitted model (although they are still superior to not doing a correction), i.e. for practice, it is important to not choose the model order of the INAR(p) DGP too small. If the bootstrap schemes correctly assume $p = 2$, it can again be observed that the AR bootstrap is worse than the INAR bootstraps, for the reasons discussed above.
Finally, we consider the estimation of the first-order autocorrelation, both for the Poi-INAR(1) and Poi-INAR(2) DGP. Tables 7 and 8 (which refer to the medium autocorrelation level \( \rho(1) = 0.4 \)) show that any of the bias corrections leads to an improvement compared to the case without correction, namely a reduced bias without increasing the RMSE. The INAR and AR bootstrap schemes perform similarly well, which is reasonable in view of our above discussion about the asymptotic bias formulae in Table 1. Comparing Tables 7 and 8, we see that the corrected estimates for \( \rho(1) \) show less bias and less RMSE for the lower model order \( p = 1 \), which is reasonable since we have one parameter less to be estimated. Analogous observations can be made for the highly correlated case \( \rho(1) = 0.8 \), which is illustrated by box plots in Figure 2.
5. Illustrative data examples

We illustrate the considered bias corrections with two data examples taken from the book by [1]. The first time series (length \( n = 267 \)) gives daily counts of downloads of a TEX-editor and was shown in Section 2.5 of [1] to be of \( \text{INAR}(1) \)-type with a strong degree of overdispersion. The second time series (length \( n = 370 \)) consisting of counts of gold particles over time, in contrast, appears to be of \( \text{INAR}(2) \)-type with a nearly equidispersed marginal distribution, see Example 3.1.6 in [1]. This can also be seen from Table 9, which shows some results from SPML estimation according to [8]: for the downloads counts, the estimate for the second autoregressive parameter deviates only slightly from 0 (thus also supporting a first-order autoregression), and the innovations’ variance is much larger than the mean (within the fitted model), whereas the gold particles counts have nearly equidispersed innovations (as it would be the case for a Poisson \( \text{INAR} \) model) and a clearly non-zero value for \( \hat{\alpha}_{sp,2} \). Certainly, we do not know the true data-generating mechanism behind these time series, but these observations will help us in interpreting the subsequent findings.

Let us concentrate on the estimation of the variance \( \hat{\gamma}(0) \), the dispersion index \( \hat{I}_{\text{disp}} \), and the first-order autocorrelation \( \hat{\rho}(1) \) in the sequel, which are meaningful and relevant quantities for both model orders \( p = 1 \) and \( p = 2 \). For bias correction, we applied the bootstrap schemes presented in Section 3 (now we managed even 10,000 bootstrap replications), where the optimal block length for the block bootstrap [20,21] was calculated as 6 (downloads) and 33 (gold particles), respectively. For model order \( p = 1 \), we also applied the asymptotic bias corrections described in Section 3.3.

Results for the downloads counts are shown in Table 10. It becomes clear that all methods for bias correction have a similar effect on the autocorrelation \( \hat{\rho}(1) \), which is not surprising in view of the discussions in Section 4. It is also plausible that there is nearly no difference between model orders \( p = 1 \) and \( p = 2 \), since the first-order model is included as a boundary case within the second-order one (overfitting increases computational costs, whereas underfitting deteriorates the performance of bias correction, remember the discussion of Table 6).

### Table 9. Results of SPML estimation.

| Download counts: | Gold particles counts: |
|------------------|-------------------------|
| Model            | \( \hat{\alpha}_{sp,1} \) | \( \hat{\alpha}_{sp,2} \) | \( \hat{\mu}_{\epsilon; sp} \) | \( \hat{\sigma}^2_{\epsilon; sp} \) | Model            | \( \hat{\alpha}_{sp,1} \) | \( \hat{\alpha}_{sp,2} \) | \( \hat{\mu}_{\epsilon; sp} \) | \( \hat{\sigma}^2_{\epsilon; sp} \) |
| \( p = 1 \)      | 0.186                   | 1.925                   | 6.176                     | \( p = 1 \)                  | 0.545           | 0.712                   | 0.741                     |
| \( p = 2 \)      | 0.166                   | 0.043                   | 1.877                     | \( p = 2 \)                  | 0.487           | 0.156                   | 0.558                     | 0.571                     |

### Table 10. Bias-corrected estimates for downloads counts.

| \( \hat{\gamma}(0) \) | Model       | \( \hat{I}_{\text{disp}} \) | Model | \( \hat{\rho}(1) \) | Model |
|-----------------------|-------------|-----------------------------|-------|---------------------|-------|
| Bias corr.            | \( p = 1 \) | \( p = 2 \) Bias corr.     | \( p = 1 \) | \( p = 2 \) Bias corr. | \( p = 1 \) | \( p = 2 \) |
| spINAR                | 7.539       | 7.533                       | spINAR| 3.145               | 3.149             | spINAR| 0.251 | 0.252 |
| plINAR                | 7.518       | 7.524                       | plINAR| 3.132               | 3.134             | plINAR| 0.252 | 0.252 |
| CBB                   | 7.554       | 7.554                       | CBB   | 3.158               | 3.115             | CBB   | 0.250 |
| AR                    | 7.539       | 7.556                       | AR    | 3.097               | 3.115             | AR    | 0.252 | 0.252 |
| asymp                 | 7.521       | asymp                       | asymp | 3.133               | 3.127             | asymp | 0.252 |
| no                    | 7.506       | no                          |       | 3.127               | no               |       | 0.245 |
Table 11. Bias-corrected estimates for gold particles counts.

|       | Model |       | Model |       | Model |
|-------|-------|-------|-------|-------|-------|
|       | \( \hat{\gamma}(0) \) |       | \( \hat{I}_{\text{disp}} \) |       | \( \hat{\rho}(1) \) |
| Bias corr. | \( p = 1 \) | \( p = 2 \) | Bias corr. | \( p = 1 \) | \( p = 2 \) | Bias corr. | \( p = 1 \) | \( p = 2 \) |
| spINAR | 1.663 | 1.670 | spINAR | 1.071 | 1.079 | spINAR | 0.581 | 0.587 |
| pINAR  | 1.663 | 1.675 | pINAR  | 1.072 | 1.082 | pINAR  | 0.582 | 0.589 |
| CBB    | 1.687 |      | CBB    | 1.091 |      | CBB    | 0.591 |      |
| AR     | 1.663 | 1.673 | AR     | 1.077 | 1.079 | AR     | 0.581 | 0.586 |
| asymp  | 1.663 |      | asymp  | 1.072 |      | asymp  | 0.582 |      |
| no     | 1.647 |      | no     | 1.062 |      | no     | 0.574 |      |

Looking at the variance \( \hat{\gamma}(0) \) and dispersion index \( \hat{I}_{\text{disp}} \), it is apparent that the methods ‘pINAR’ and ‘asymp’ have a smaller corrective effect than ‘spINAR’ and ‘CBB’. This is plausible since both rely on a parametric Poisson assumption, which appears to be severely violated by the data.

Let us now discuss the results for the gold particles counts, see Table 11. For these data, there is strong evidence for a second-order autoregressive structure, also see Example 3.1.6 in [1]. Hence, it is not surprising that the INAR bootstrap schemes show visible deviations if being based on model order \( p = 1 \) or \( p = 2 \). The difference between ‘spINAR’ and ‘pINAR’, however, is rather small, which can be explained by the fact that the gold particles counts are close to equidispersion, i.e. a Poisson INAR(2) model appears to be a reasonable choice for the data.

6. Conclusions

We propose to use INAR-type bootstrap procedures for finite sample bias correction of several practically relevant statistics computed from INAR data. In general, we found that bias correction works quite well in simulations. We compare the potential of INAR bootstrap bias correction to other competitors including bias correction exploiting asymptotic approximations as well as other resampling techniques as AR bootstrap and block bootstrap. The INAR bootstraps tend to outperform their competitors in simulations. In particular, the block bootstrap converges more slowly and performs worse for small sample sizes. The AR bootstrap shows good performance for some statistics, but is clearly worse for others. We were able to explain these phenomena by making use of asymptotic approximations of the bias in special cases. Furthermore, we found that the parametric INAR bootstrap and the asymptotic bias correction, which both rely on Poisson assumptions, performed worse whenever these assumptions were violated. In such cases, the semi-parametric INAR bootstrap were robust to non-Poisson innovations and still led to good results.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

C.H. Weiβ http://orcid.org/0000-0001-8739-6631
C. Jentsch http://orcid.org/0000-0001-7824-1697
Appendix. Derivation of AR(1) Bias Corrections from Table 1 (Sketch)

The approach to derive asymptotic approximations for the bias of moment-based statistics described in [11,12] can be immediately adapted to the case of an AR(1) DGP. More precisely, let us denote the AR(1) recursion as $U_t = \alpha \cdot U_{t-1} + \varepsilon_t$, where moments of some order $> 4$ are assumed to exist, but where we do not further specify the distribution of $U_t$ and $\varepsilon_t$, respectively. We use the notations $\mu$ for the mean of $U_t$, $\sigma^2$ for the variance, and $\overline{\mu}_k$ for the $k$th central moment. If referring to the respective
moments of $\varepsilon_t$, we add the index ‘$e$’. Note that the following relations hold:

\[
\mu = \frac{\mu_e}{1 - \alpha}, \quad \sigma^2 = \frac{\sigma_e^2}{1 - \alpha^2}, \quad \tilde{\mu}_3 = \frac{\tilde{\mu}_3,e}{1 - \alpha^3}, \quad \tilde{\mu}_4 = \frac{\tilde{\mu}_4,e}{1 - \alpha^4} + \frac{6\alpha^2\sigma_e^4}{(1 - \alpha^4)(1 - \alpha^2)}. \quad (A1)
\]

Furthermore, we denote the $U_t$’s joint moments by

\[
\mu(s_1, \ldots, s_{r-1}) := E[U_t \cdot U_{t+s_1} \cdots U_{t+s_{r-1}}] \quad (A2)
\]

with $0 \leq s_1 \leq \ldots \leq s_{r-1}$ and $r \in \mathbb{N}$. With analogous arguments as in [11,12], it follows that

\[
\frac{1}{\sqrt{T}} \sum_{t=1}^{T} Y_t \xrightarrow{D} N(0, \Sigma) \quad \text{with } \Sigma = (\sigma_{ij}) \text{ given by} \quad (A3)
\]

\[
\sigma_{ij} = E[Y_{0,j}Y_{0,j}] + \sum_{k=1}^{\infty} (E[Y_{0,j}Y_{k,j}] + E[Y_{k,i}Y_{0,j}]),
\]

where

\[
Y_t := (U_t - \mu, U_t^2 - \mu(0), U_tU_{t-1} - \mu(1))^T \quad (A4)
\]

with $\mu(k) = \alpha^k \sigma^2 + \mu^2$ for all $k \in \mathbb{N}_0$.

To compute the $\sigma_{ij}$ from (A3), the joint moments (A2) need to be evaluated also for $r = 3,4$. This is done by solving recurrence equations derived from the AR(1)’s conditional moments,

\[
E[U_t^k | U_{t-1}, \ldots] = \sum_{j=0}^{k} \binom{k}{j} \alpha^j E[\varepsilon_t^{k-j}] \cdot U_{t-1}^j,
\]

and the law of total expectation applied to (A2). For $r = 3$, these are the equations

\[
\begin{align*}
\mu(k,l) &= E[U_tU_{t+k}E[U_{t+l} | U_{t+l-1}, \ldots]] = \ldots \\
&= \alpha^{l-k} \mu(k,k) + \mu(k)(1 - \alpha^{l-k}), \\
\mu(k,k) &= E[U_tE[U_{t+k}^2 | U_{t+k-1}, \ldots]] \\
&= \alpha^2 \mu(k-1, k-1) + 2\alpha(1 - \alpha)\mu(k-1) + \mu(\sigma_e^2 + \mu_e^2).
\end{align*}
\]

Solving these equations (and analogously for $r = 4$), one finally obtains

\[
\begin{align*}
\mu(k,l) &= \alpha^{l-k} \left(2\sigma^2 \mu \alpha^k + \sigma^2 \mu + \tilde{\mu}_3 \alpha^{2k} + \mu^3\right) + \mu \left(1 - \alpha^{l-k}\right) \left(\sigma^2 \alpha^k + \mu^2\right), \\
\mu(k,l,m) &= \sigma^4 \left(\alpha^{m-l+k} + 2\alpha^{m+l-k} - 3\alpha^{m+l+k}\right) + \tilde{\mu}_4 \alpha^{m+l+k} + \mu^4 \\
&+ \sigma^2 \mu^2 \left(\alpha^k + \alpha^{l-k} + \alpha^{m-k} + \alpha^l + \alpha^{m-l} + \alpha^m\right) \\
&+ \tilde{\mu}_3 \mu \left(\alpha^{k+l} + \alpha^{m+l-2k} + \alpha^{m+k} + \alpha^{m+l}\right), \quad (A5)
\end{align*}
\]
where $0 \leq k \leq l \leq m$. These moment formulae are now applied to evaluate (A3):

$$
\sigma_{11} = E[Y_{0,1}^2] + 2 \sum_{k=1}^{\infty} E[Y_{0,1} Y_{k,1}]
= \mu(0) - \mu^2 + 2 \sum_{k=1}^{\infty} (\mu(k) - \mu^2),
$$

$$
\sigma_{12} = \sum_{k=1}^{\infty} E[Y_{0,1} Y_{k,2}] + \sum_{k=0}^{\infty} E[Y_{k,1} Y_{0,2}]
= \sum_{k=1}^{\infty} (\mu(k, k) - \mu \mu(0)) + \sum_{k=0}^{\infty} (\mu(0, k) - \mu(0)),
$$

$$
\sigma_{22} = E[Y_{0,2}^2] + 2 \sum_{k=1}^{\infty} E[Y_{0,2} Y_{k,2}]
= \mu(0, 0, 0) - \mu(0)^2 + 2 \sum_{k=1}^{\infty} (\mu(0, k, k) - \mu(0)^2),
$$

$$
\sigma_{13} = \sum_{k=1}^{\infty} E[Y_{0,1} Y_{k,3}] + \sum_{k=0}^{\infty} E[Y_{k,1} Y_{0,3}]
= \sum_{k=1}^{\infty} (\mu(k - 1, k) - \mu \mu(1)) + \sum_{k=0}^{\infty} (\mu(1, k + 1) - \mu(1)),
$$

$$
\sigma_{23} = \sum_{k=1}^{\infty} E[Y_{0,2} Y_{k,3}] + \sum_{k=0}^{\infty} E[Y_{k,2} Y_{0,3}]
= \sum_{k=1}^{\infty} (\mu(0, k - 1, k) - \mu(0) \mu(1)) + \sum_{k=0}^{\infty} (\mu(1, k + 1, k + 1) - \mu(0) \mu(1)),
$$

$$
\sigma_{33} = E[Y_{0,3}^2] + 2 \sum_{k=1}^{\infty} E[Y_{0,3} Y_{k,3}]
= \mu(0, 1, 1) - \mu(1)^2 + 2 \sum_{k=1}^{\infty} (\mu(1, k, k + 1) - \mu(1)^2),
$$

leading to

$$
\sigma_{11} = \sigma^2 \frac{1 + \alpha}{1 - \alpha},
$$

$$
\sigma_{12} = 2\sigma^2 \mu \frac{1 + \alpha}{1 - \alpha} + \bar{\mu}_3 \frac{1 + \alpha + \alpha^2}{1 - \alpha^2},
$$

$$
\sigma_{22} = 4\sigma^2 \mu^2 \frac{1 + \alpha}{1 - \alpha} + (\bar{\mu}_4 - \sigma^4) \frac{1 + \alpha^2}{1 - \alpha^2} + 4\bar{\mu}_3 \mu \frac{1 + \alpha + \alpha^2}{1 - \alpha^2},
$$

$$
\sigma_{13} = 2\sigma^2 \mu \frac{1 + \alpha}{1 - \alpha} + \bar{\mu}_3 \alpha \frac{1 + \alpha + \alpha^2}{1 - \alpha^2},
$$

(A6)
\[
\begin{align*}
\sigma_{23} &= 2\sigma^4 \alpha + 4\sigma^2 \mu^2 \frac{1 + \alpha}{1 - \alpha} + (\bar{\mu}_4 - \sigma^4) \alpha \frac{1 + \alpha}{1 - \alpha^2} + 2\bar{\mu}_3 \mu \frac{1 + \alpha + \alpha^2}{1 - \alpha}, \\
\sigma_{33} &= \sigma^4 (1 + 3\alpha^2) + 4\sigma^2 \mu^2 \frac{1 + \alpha}{1 - \alpha} + (\bar{\mu}_4 - \sigma^4) \alpha^2 \frac{1 + \alpha^2}{1 - \alpha^2} + 4\bar{\mu}_3 \mu \alpha \frac{1 + \alpha + \alpha^2}{1 - \alpha^2}.
\end{align*}
\]

To derive the asymptotic properties of the statistics \(\hat{\gamma}(0), \hat{I}_{\text{disp}}, \hat{\gamma}(1), \hat{\rho}(1)\) and \(\bar{X}(1 - \hat{\rho}(1))\), define the functions \(g_i : \mathbb{R}^3 \to \mathbb{R}\) for \(i = 1, \ldots, 5\) as

\[
\begin{align*}
g_1(y_1, y_2, y_3) &= y_2 - y_1^2, \quad g_2(y_1, y_2, y_3) = \frac{y_2 - y_1^2}{y_1}, \quad g_3(y_1, y_2, y_3) = y_3 - y_1^2, \quad g_4(y_1, y_2, y_3) = \frac{y_3 - y_1^2}{y_2 - y_1^2}, \quad g_5(y_1, y_2, y_3) = y_1 \left(1 - \frac{y_3 - y_1^2}{y_2 - y_1^2}\right),
\end{align*}
\]

such that \(g_i(\mu(0), \mu(1))\) equals \(\gamma(0), I_{\text{disp}}, \gamma(1), \rho(1)\) and \(\mu(1 - \rho(1))\) for \(i = 1, \ldots, 5\). Application of the Delta method to (A3), (A6) and (A7) shows the asymptotic normality of \(\hat{\gamma}(0), \hat{I}_{\text{disp}}, \hat{\gamma}(1), \hat{\rho}(1)\) and \(\bar{X}(1 - \hat{\rho}(1))\), and closed-form expressions for the asymptotic (co-)variances could be obtained.

But here, we are interested in a bias correction for these statistics, so we use a second-order Taylor expansion for each function \(g_i(y_1, y_2, y_3)\). In a first step, the Hessians of \(g_i\) are computed, leading to \(H_{g_i}(y_1, y_2, y_3)\) for \(i = 1, \ldots, 5\). These are evaluated in \((\mu, \mu(0), \mu(1))\), leading to the matrices \(H_i = H_{g_i}(\mu, \mu(0), \mu(1))\). Then, abbreviating \(Z_n := (1/\sqrt{n}) \sum_{t=1}^n Y_t\), we compute the bias corrections

\[
nE \left[ g_i(\bar{X}, \bar{\mu}(0), \bar{\mu}(1)) - g_i(\mu, \mu(0), \mu(1)) \right] \approx E \left[ \frac{1}{2} (Z_n) \top H_i Z_n \right]
\]

\[
\approx \frac{1}{2} h_{i,11} \sigma_{11} + \frac{1}{2} h_{i,22} \sigma_{22} + \frac{1}{4} h_{i,33} \sigma_{33} + h_{i,12} \sigma_{12} + h_{i,13} \sigma_{13} + h_{i,23} \sigma_{23}.
\]

This leads to the formulae shown in the right column of Table 1.