Missing observations in observation-driven time series models

Paolo Gorgi

Co-authors: **F. Blasques and S.J. Koopman**

VU University Amsterdam

May 25th 2018

VRIJE UNIVERSITEIT AMSTERDAM

Outline

1. GAS models with missing data
2. Inconsistency of the setting-to-zero method
3. A consistent indirect inference estimator

GAS models with missing data

GAS models

The class of **Generalized Autoregressive Score** (GAS) models of is

$$y_t \sim p(y_t|f_t;\theta), \quad f_{t+1} = \omega + \beta f_t + \alpha s_t,$$

where **score innovation** s_t is given by

$$s_t = S_t u_t, \quad u_t = \frac{\partial \log p(y_t|f_t;\theta)}{\partial f_t},$$

and S_t is a scaling factor.

GAS models with missing data

- **Missing data in time series** can occur for several reasons:
  - Missing data caused by external events.
  - Unequally spaced time series.
  - Mixed frequency data.

- **How are missing observations handled in practice?**

  GAS models with missing observations are estimated using the **setting-to-zero** method.

  Buccheri et al. (2017), Delle Monache et al. (2016), Lucas et al. (2016), Koopman et al. (2015) and Creal et al. (2014).

  Idea: set the score innovation to zero when a missing observation occurs.

Setting-to-zero method (i)

Missing observations: sample of data $\{y_1, y_2, \ldots, y_T\}$ where

$$\begin{cases} y_t \text{ is observed,} & \text{if } I_t = 1 \\ y_t \text{ is not observed,} & \text{if } I_t = 0. \end{cases}$$

Setting-to-zero method:

- **Step 1:** recover the filtered parameter setting the score to zero when an observation is missing

$$\hat{f}_{t+1}(\theta) = \omega + \beta \hat{f}_t(\theta) + \alpha I_t s_t.$$

- **Step 2:** plug-in the filter into the likelihood (*pseudo likelihood*)

$$\hat{L}_T(\theta) = T^{-1} \sum_{t=1}^{T} I_t \log p(y_t | \hat{f}_t(\theta); \theta),$$

- **Step 3:** obtain the *pseudo ML estimator* $\hat{\theta}_T$ maximizing $\hat{L}_T(\theta)$.

Setting-to-zero method (ii)

The **setting-to-zero** method is:

- Simple to implement.
- Intuitive and it can be justified by some arguments.
- **However**, it leads to inconsistent inference.

In this paper:

- We show that the **pseudo ML is inconsistent** for a local mean GAS model.
- We propose an indirect inference estimator that delivers consistent inference for GAS models with missing data.

Inconsistency of the setting-to-zero method

Local mean GAS models

- Consider the **Gaussian local mean GAS** model:

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2),$$

$$\mu_{t+1} = \omega + \beta\mu_t + \alpha(y_t - \mu_t).$$

  Note: this model is in fact a linear model and it can be estimated by exact ML using the Kalman filter.

- **Setting-to-zero method:** it is straightforward to obtain that the **pseudo ML** estimator of σ^2 is inconsistent.

- We show the non-trivial result that the **pseudo ML** is inconsistent also for α and β.

Inconsistency of pseudo ML

We assume that the missing observation process $\{I_t\}_{t\in\mathbb{Z}}$ is an iid Bernoulli sequence with success probability π.

Theorem (inconsistency of α and β)

The pseudo ML estimator $\hat{\theta}_T$ obtained from the **setting-to-zero** *method for the local mean GAS model is not consistent. In particular, there exists an $\epsilon > 0$ such that*

$$\mathbb{P}\left(\liminf_{T\to\infty} \|\hat{\theta}_T - \theta_0\| > \epsilon\right) = 1,$$

for some $\theta = (\alpha, \beta)$ and some $\pi \in (0,1)$.

- This result shows that even the dependence structure of the model is inconsistently estimated.

Simulation: local mean model

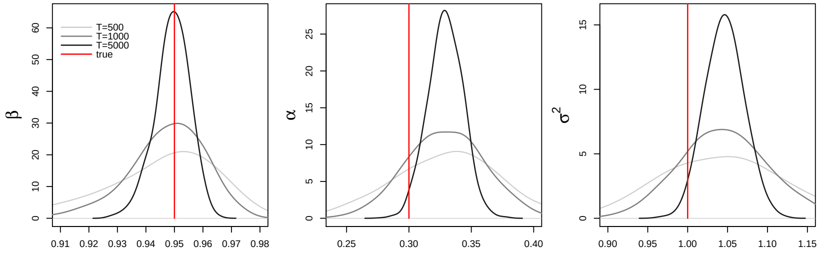

Figure: *Small sample* **distribution of the pseudo ML estimator** *for the Gaussian local mean model. Different sample sizes are considered and* $\pi = 0.75$.

Example: Student-t GAS (i)

- The **inconsistency** of the setting-to-zero method applies to **GAS models in general**.
- **Example:** the conditional volatility Student-t GAS model of Creal et al. (2013) and Harvey (2013) is

$$y_t = \sqrt{h_t}\varepsilon_t, \quad \varepsilon_t \sim t_\nu,$$

$$h_{t+1} = \omega + \beta h_t + \alpha \left[\frac{(\nu+1)y_t^2}{(\nu-2) + y_t^2 h_t^{-1}} - h_t \right]$$

where ω, β, α and ν are parameters to be estimated.

Example: Student-t GAS (ii)

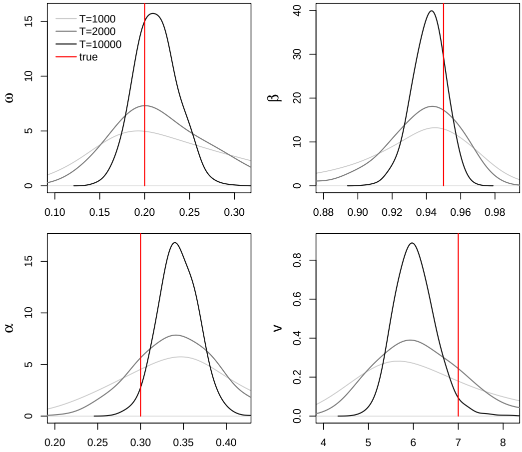

Figure: *Distribution for conditional volatility Student's t model.*

A consistent indirect inference estimator

Indirect Inference estimation (i)

Idea: remove the bias of the setting-to-zero method by indirect inference.

We propose the following **indirect inference estimator**:

- Simulate S series from the GAS model $\{y_{i,t}(\bar{\theta})\}_{t=1}^{T}$, $i = 1, \ldots, S$.
- Introduce missing observations in the simulations $y_{i,t}(\bar{\theta})$ at time points t where the real observations y_t are missing.
- Obtain **pseudo ML** estimator using the setting-to-zero method:

$$\hat{L}_{S,T}(\theta, \bar{\theta}) = \frac{1}{S} \sum_{i=1}^{S} \hat{L}_{i,T}(\theta, \bar{\theta}), \quad \hat{\theta}_{S,T}(\bar{\theta}) = \arg\sup_{\theta \in \Theta} \hat{L}_{S,T}(\theta, \bar{\theta}).$$

- Finally, the **indirect inference** estimator $\tilde{\theta}_{S,T}$ is

$$\tilde{\theta}_{S,T} = \arg\inf_{\bar{\theta} \in \Theta} \left\| \hat{\theta}_{S,T}(\bar{\theta}) - \hat{\theta}_T \right\|.$$

Indirect Inference estimation (ii)

Assumption: data missing at random and missing values process I_t is stationary with $P(I_t = 1) > 0$.

Theorem (asymptotic distribution)

Under some additional conditions, we obtain that

$$\sqrt{T}\left(\tilde{\theta}_{S,T} - \theta_0\right) \xrightarrow{d} N(0, W_S) \quad \text{as} \quad T \to \infty,$$

where

$$W_S := \left(1 + \frac{1}{S}\right)\left[\frac{\partial \theta^*(\theta_0)}{\partial \theta^\top}\right]^{-1} V(\theta_0) \left[\frac{\partial \theta^*(\theta_0)}{\partial \theta}^\top\right]^{-1}$$

where $V(\theta_0)$ *denotes the asymptotic variance*
$V(\theta_0) := \Omega^*(\theta_0)^{-1}(\Sigma^*(\theta_0) - K^*(\theta_0))\Omega^*(\theta_0)^{-1}$.

Simulation study (i)

- We compare the performance of the **indirect inference** estimator with **exact ML** and **pseudo ML**.
- In general, **exact ML** is infeasible but we consider a local mean GAS model for which **exact ML** is available via the Kalman filter.
- **Gaussian local mean GAS** model:

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2),$$

$$\mu_{t+1} = \omega + \beta \mu_t + \alpha (y_t - \mu_t).$$

Note: this model can be shown to be an ARMA(1,1) model.

Simulation study (ii)

		$\pi = 0.40$			$\pi = 0.60$		
		β	α	σ^2	β	α	σ^2
$T = 500$	**PML**	0.028	0.120	0.204	0.023	0.075	0.121
	ML	0.029	0.060	0.109	0.023	0.047	0.087
	II	0.025	0.064	0.118	0.020	0.051	0.091
$T = 1000$	**PML**	0.018	0.108	0.187	0.016	0.065	0.105
	ML	0.018	0.042	0.078	0.016	0.033	0.058
	II	0.014	0.045	0.084	0.013	0.035	0.062

Table: *Mean squared error (MSE) of estimators obtained from 500 simulations with* $S = 10$. *The true parameter vector is* $\theta = (0.95, 0.3, 1)^\top$.

Application to S&P500

VRIJE UNIVERSITEIT AMSTERDAM

We perform an experiment to access the performance of the **Indirect Inference** and **pseudo ML** estimator with a real dataset.

- **Dataset:** daily log-returns of the **S&P500 stock index** from January 2000 to December 2016.
- **Model:** conditional variance **Student-t GAS model**.
- **Experiment:**
  - Estimate the parameter $\theta = (\omega, \beta, \alpha, \nu)^\top$ using the full dataset.
  - Remove observations using a Bernoulli process with success probability π and estimate the model using this new dataset.
  - Repeat previous point multiple times and compare the estimates with missing data with full-sample estimate.

Pseudo ML

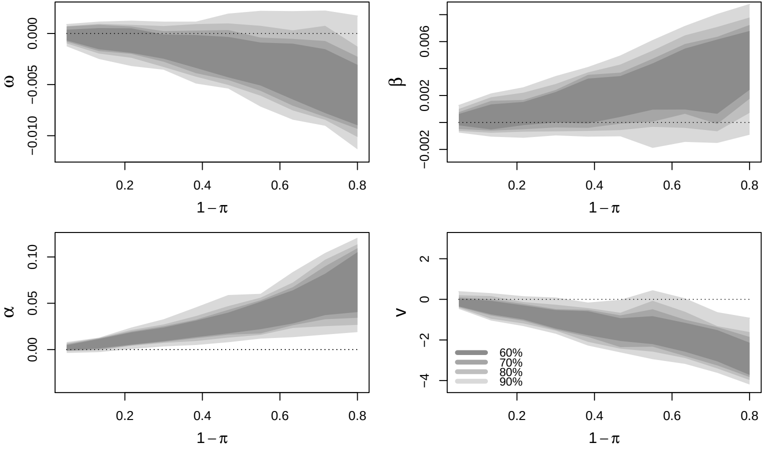

Figure: *Bias of the pseudo ML estimator wrt full sample estimator for different* π. *Grey areas represent confidence bounds of the bias.*

Indirect Inference

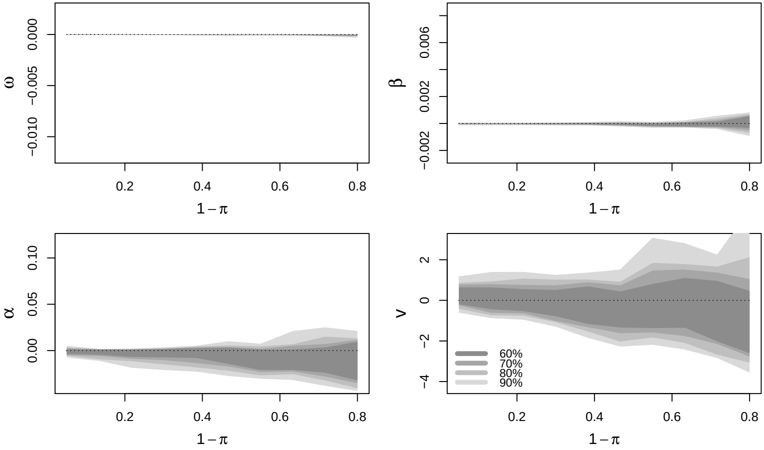

Figure: *Bias of the indirect inference estimator wrt full sample estimator for different π. Grey areas represent confidence bounds of the bias.*

Conclusion

- We prove that the **setting-to-zero method** for estimating GAS models with missing data leads to inconsistent inference.
- We propose an indirect inference estimator that delivers consistent parameter estimates.
- The proposed estimator shows comparable accuracy to the infeasible exact ML estimator.