Supplemental Material

Aggregate Output Measurements

A Common Trend Approach

Martín Almuzara* Gabriele Fiorentini† Enrique Sentana‡

This version: April 2023

[Link to the paper]

SM.A Estimators of unconditional mean and variance

AR(1) example. Consider the following stationary Gaussian AR(1) model:

$$
\begin{aligned}
y_t &= c + r y_{t-1} + u_t, \\
u_t &\overset{iid}{\sim} N(0, s^2)
\end{aligned}
$$

The information matrix for the MLE of $\alpha = (c, r, s^2)$ assuming y_t observable is

$$
\mathcal{I}(\alpha) = \begin{pmatrix} s^{-2} & s^{-2}\mu & 0 \\ s^{-2}\mu & s^{-2}(\mu^2 + \sigma^2) & 0 \\ 0 & 0 & \frac{1}{2}s^{-4} \end{pmatrix},
$$

where $\mu = \mathbb{E}\left[y_{t-1}\right] = c/(1-r)$ and $\sigma^2 = \text{Var}\left(y_{t-1}\right) = s^2/(1-r^2)$.

Consider the following reparameterization: $\alpha \mapsto \theta = (\mu, \rho, \sigma^2)$ where

$$
\begin{aligned}
\mu &= \frac{c}{1-r}, \\
\rho &= r,
\end{aligned}
$$

*Federal Reserve Bank of New York: martin.almuzara@ny.frb.org

†Universitá di Firenze and RCEA: gabriele.fiorentini@unifi.it

‡CEMFI and CEPR: sentana@cemfi.es

$$\sigma^2 = \frac{s^2}{1-r^2},$$

whose inverse is given by

$$\begin{aligned} c &= \mu(1-\rho), \\ r &= \rho, \\ s^2 &= (1-\rho^2)\sigma^2. \end{aligned}$$

Effectively, this amounts to re-writing the Gaussian AR(1) process above as

$$\begin{aligned} (y_t - \mu) &= \rho(y_{t-1} - \mu) + \sqrt{\sigma^2(1-\rho^2)}\varepsilon_t, \\ \varepsilon_t &\overset{iid}{\sim} N(0,1). \end{aligned}$$

The Jacobian of the inverse transformation is

$$\frac{\partial \alpha}{\partial \theta'} = \begin{pmatrix} 1-\rho & -\mu & 0 \\ 0 & 1 & 0 \\ 0 & -2\rho\sigma^2 & 1-\rho^2 \end{pmatrix} = \begin{pmatrix} 1-r & -\frac{c}{1-r} & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{2rs^2}{1-r^2} & 1-r^2 \end{pmatrix}$$

A straightforward application of the chain rule for derivatives implies that the information matrix of the transformed parameters θ will be

$$\tilde{\mathcal{I}}(\theta) = \frac{\partial \alpha'}{\partial \theta}\mathcal{I}(\alpha)\frac{\partial \alpha}{\partial \theta'} = \begin{pmatrix} \frac{1}{s^2}(r-1)^2 & 0 & 0 \\ 0 & \frac{1}{\left(r^2-1\right)^2}\left(r^2+1\right) & -\frac{1}{s^2}r \\ 0 & -\frac{1}{s^2}r & \frac{1}{2s^4}\left(r^2-1\right)^2 \end{pmatrix}$$

whose inverse is

$$\tilde{\mathcal{I}}^{-1}(\theta) = \begin{pmatrix} \frac{s^2}{(1-r)^2} & 0 & 0 \\ 0 & 1-r^2 & 2s^2\frac{r}{1-r^2} \\ 0 & 2s^2\frac{r}{1-r^2} & 2s^4\frac{(1+r^2)}{(1-r^2)^3} \end{pmatrix}$$

Given that the spectral density of y_t at frequency 0 is $s^2/(1-r)^2$, it is clear that the dynamic estimator of μ has the same asymptotic variance as the sample mean of x_t, which coincides with the ML estimator of μ that erroneously imposes that $r = 0$.

To find out the asymptotic variance of the sample variance, we need to obtain the autocorrelation structure of y_t^2, which, given the Gaussian nature of the process, will be that of an AR(1) with autoregressive coefficient r^2. In addition, given that

$$(y_t - \mu)^2 = r^2(y_{t-1} - \mu)^2 + s^2\varepsilon_t^2 + 2rs(y_{t-1} - \mu)\varepsilon_t,$$

the innovation variance of this autoregression would be

$$\text{Var}\left(s^2\varepsilon_t^2 + 2rs(y_{t-1} - \mu)\varepsilon_t\right) = 2s^4 + 4r^2s^2\frac{s^2}{(1-r^2)} = 2s^4\left(1 + \frac{2r^2}{(1-r^2)}\right) = 2s^4\frac{\left(1+r^2\right)}{(1-r^2)},$$

so the spectral density of $(y_t - \mu)^2$ at the frequency 0 will be

$$2s^4\frac{\left(1+r^2\right)}{(1-r^2)}\frac{1}{(1-r^2)^2}.$$

This confirms that the dynamic estimator of σ^2 has the same asymptotic variance as the sample variance of y_t, which coincides with the ML estimator of σ^2 that erroneously assumes that $r = 0$.

The previous example suggests that dynamic misspecification does not matter for the asymptotic distribution of either the unconditional mean or variance estimators when the error is assumed to follow an AR(p) process. As we show next, with MA dynamics the result fails for the unconditional variance. In contrast, it continues to be valid for the unconditional mean, which follows from the fact that the sample mean is the frequency-domain MLE of the first unconditional moment (see, e.g., Grenander and Rosenblatt (1958) and Dzhaparidze (1986) for a formal proof of the equivalence with the time-domain estimator).

MA(1) example. Consider the Gaussian MA(1) model,

$$\begin{aligned} y_t &= \mu + v_t + \theta v_{t-1}, \\ v_t &\overset{iid}{\sim} N(0, \omega^2). \end{aligned}$$

Suppose we are interested in estimating the unconditional mean μ and variance $\sigma^2 = (1+\theta^2)\omega^2$. Let $\hat{\mu}$, $\hat{\theta}$, $\hat{\omega}^2$ and $\hat{\sigma}^2 = (1+\hat{\theta}^2)\hat{\omega}^2$ be the MLEs of the dynamic model and let

$$\tilde{\mu} = T^{-1}\sum_{t=1}^{T} y_t,$$

$$\tilde{\sigma}^2 = T^{-1}\sum_{t=1}^{T}(y_t - \tilde{\mu})^2 (= \tilde{\omega}^2),$$

be the sample mean and variance – equivalent to MLE under the static-model restriction $\theta = 0$. As before, we assume correct specification.

In this context, we show below that $\delta_T = \sqrt{T}(\hat{\sigma}^2 - \tilde{\sigma}^2) = O_p(1)$, unlike in the AR(1) case discussed before, in which $\delta_T = o_p(1)$. In other words, the asymptotic equivalence between the dynamic-model and static-model MLEs breaks down for the unconditional variance. To see this, first consider the asymptotic distribution of the dynamic-model MLE,

$$\sqrt{T}\begin{pmatrix} \hat{\theta} - \theta \\ \hat{\omega}^2 - \omega^2 \end{pmatrix} \Longrightarrow N\left(0_{2\times 1}, \begin{pmatrix} 1-\theta^2 & 0 \\ 0 & 2\omega^4 \end{pmatrix}\right).$$

See e.g. Shephard (1993) for a proof.

The delta method immediately leads to

$$\sqrt{T}\left(\hat{\sigma}^2 - \sigma^2\right) \Longrightarrow N\left(0, 2\omega^4(1 + 4\theta^2 - \theta^4)\right).$$

On the other hand, $(y_t - \mu)^2$ is an MA(1) process with variance $2\omega^4(1 + 2\theta^2 + \theta^4)$ and first-order autocovariance $2\omega^4\theta^2$. Therefore, the static-model MLE of the unconditional mean of this process has asymptotic distribution

$$\sqrt{T}\left(\tilde{\sigma}^2 - \sigma^2\right) \Longrightarrow N\left(0, 2\omega^4(1 + 4\theta^2 + \theta^4)\right).$$

Now, since the asymptotic variance of $\sqrt{T}\left(\tilde{\sigma}^2 - \sigma^2\right)$ strictly exceeds that of $\sqrt{T}\left(\hat{\sigma}^2 - \sigma^2\right)$ unless $\theta = 0$, we can conclude that $\delta_T = O_p(1)$.

SM.B Additional simulation results

Simulation results for designs similar to the ones displayed in the text but with $\rho_1 = \rho_2 = 0$ (instead of $\rho_1 = \rho_2 = 0.85$) are collected below. In particular, we generate simulated data from the distribution $\mathbb{P}$ with $\mu_0 = 3$, $\rho_0 = 0.5$ and $\sigma_0 = 3.25$. We also take $N = 2$ and let R^2 (with $\sigma_1 = \sigma_2$) and $\rho_1 = \rho_2$ vary over the interval $(0, 1)$. We also consider asymmetric designs in which $\rho_1 \neq \rho_2$ too to represent the difference in persistence of measurement errors we find in the data of our application.

Each experiment is based on $n_{\mathrm{MC}} = 2,000$ samples of size $T = 280$ (amounting to 70 years of quarterly data) generated from the data generating process described above.

Tables SM.B.1, SM.B.2 and SM.B.3 together with tables SM.B.4, SM.B.5 and SM.B.6 are analogous to tables 1, 2 and 3 in the text and summarize the sampling distribution of the maximum likelihood estimates discussed in the text. Figure SM.B.1 is analogous to figure 3 in the text and contains weight comparisons for smoothed estimates of the signal; weights are only computed for the symmetric designs. Finally, tables SM.B.7 and SM.B.8 together with tables SM.B.9 and SM.B.10 are analogous to tables 4 and 5 and describe the performance of filtering procedures based on the models that neglect and impose the common trend in levels. The main text contains further details.

TABLE SM.B.1. Monte Carlo simulation for $\rho_1 = \rho_2 = 0$ and $R^2 = 0.30$.

		True	**Differences**	**Two-step**	**Levels**
μ_0	mean	3	3.003	3.003	3.003
	stderr		0.34	0.341	0.34
	corr			1	1
ρ_0	mean	0.5	-0.487	-0.496	0.466
	stderr		0.148	0.158	0.132
	corr			0.663	-0.086
σ_0	mean	3.25	3.393	3.194	3.298
	stderr		0.55	0.639	0.364
	corr			0.917	0.587
ρ_i	mean	0			0
	stderr				0.087
σ_i	mean	7.021	6.91	6.994	6.972
	stderr		0.433	0.451	0.405

NOTES. Number of samples is $n_{\mathrm{MC}} = 2,000$ and sample size is $T = 280$. The bias in $\hat{\rho}$ is to be compared with the theoretical inconsistency $B \approx -1.01$ computed from equation (2) as indicated in the text.

TABLE SM.B.2. Monte Carlo simulation for $\rho_1 = \rho_2 = 0$ and $R^2 = 0.50$.

		True	**Differences**	**Two-step**	**Levels**
μ_0	mean	3	3.002	3.002	3.002
	stderr		0.341	0.341	0.341
	corr			1	1
ρ_0	mean	0.5	0.007	-0.001	0.479
	stderr		0.208	0.186	0.096
	corr			0.933	0.43
σ_0	mean	3.25	3.192	3.218	3.255
	stderr		0.369	0.33	0.255
	corr			0.942	0.788
ρ_i	mean	0			-0.002
	stderr				0.098
σ_i	mean	4.596	4.601	4.587	4.572
	stderr		0.321	0.309	0.283

NOTES. Number of samples is $n_{\mathrm{MC}} = 2,000$ and sample size is $T = 280$. The bias in $\hat{\rho}$ is to be compared with the theoretical inconsistency $B \approx -0.35$ computed from equation (2) as indicated in the text.

TABLE SM.B.3. Monte Carlo simulation for $\rho_1 = \rho_2 = 0$ and $R^2 = 0.85$.

		True	Differences	Two-step	Levels
μ_0	mean	3	3	3.001	3
	stderr		0.342	0.342	0.342
	corr			0.999	0.999
ρ_0	mean	0.5	0.414	0.414	0.489
	stderr		0.071	0.071	0.064
	corr			1	0.95
σ_0	mean	3.25	3.228	3.231	3.236
	stderr		0.19	0.19	0.188
	corr			1	0.986
ρ_i	mean	0			-0.013
	stderr				0.144
σ_i	mean	1.931	1.926	1.922	1.917
	stderr		0.155	0.158	0.158

NOTES. Number of samples is $n_{\mathrm{MC}} = 2,000$ and sample size is $T = 280$. The bias in $\hat{\rho}$ is to be compared with the theoretical inconsistency $B \approx -0.07$ computed from equation (2) as indicated in the text.

TABLE SM.B.4. Monte Carlo simulation for $\rho_1 = 0, \rho_2 = 0.95$ and $R^2 = 0.30$.

		True	Differences	Two-step	Levels
μ_0	mean	3	3.012	3.012	3.012
	stderr		0.347	0.347	0.345
	corr			1	0.99
ρ_0	mean	0.5	-0.127	-0.115	0.479
	stderr		0.355	0.293	0.145
	corr			0.815	0.224
σ_0	mean	3.25	3.161	3.203	3.26
	stderr		0.61	0.579	0.438
	corr			0.939	0.746
ρ_1	mean	0			0.007
	stderr				0.114
σ_1	mean	7.021	6.959	7.006	6.994
	stderr		0.474	0.436	0.412
ρ_2	mean	0.95			0.943
	stderr				0.02
σ_2	mean	7.021	7.102	7.013	6.992
	stderr		0.394	0.379	0.343

NOTES. Number of samples is $n_{\mathrm{MC}} = 2,000$ and sample size is $T = 280$.

TABLE SM.B.5. Monte Carlo simulation for $\rho_1 = 0, \rho_2 = 0.95$ and $R^2 = 0.50$.

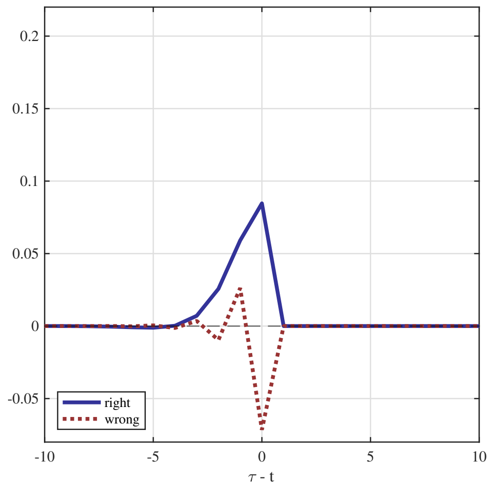
		True	Differences	Two-step	Levels
μ_0	mean	3	3.01	3.01	3.012
	stderr		0.343	0.343	0.343
	corr			1	0.995
ρ_0	mean	0.5	0.304	0.277	0.484
	stderr		0.166	0.153	0.1
	corr			0.987	0.733
σ_0	mean	3.25	3.212	3.231	3.239
	stderr		0.325	0.312	0.275
	corr			0.985	0.866
ρ_1	mean	0			0.044
	stderr				0.221
σ_1	mean	4.596	4.743	4.59	4.599
	stderr		0.305	0.301	0.286
ρ_2	mean	0.95			0.933
	stderr				0.056
σ_2	mean	4.596	4.481	4.596	4.58
	stderr		0.273	0.267	0.25

NOTES. Number of samples is $n_{\mathrm{MC}} = 2,000$ and sample size is $T = 280$.

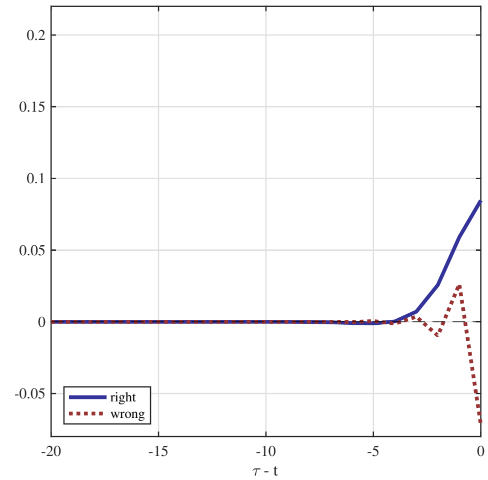
TABLE SM.B.6. Monte Carlo simulation for $\rho_1 = 0, \rho_2 = 0.95$ and $R^2 = 0.85$.

		True	Differences	Two-step	Levels
μ_0	mean	3	3.009	3.009	3.009
	stderr		0.345	0.345	0.344
	corr			0.999	0.998
ρ_0	mean	0.5	0.46	0.451	0.493
	stderr		0.065	0.065	0.061
	corr			0.992	0.968
σ_0	mean	3.25	3.242	3.233	3.256
	stderr		0.199	0.197	0.201
	corr			0.998	0.951
ρ_1	mean	0			0.276
	stderr				0.441
σ_1	mean	1.931	2.05	1.925	1.972
	stderr		0.15	0.167	0.159
ρ_2	mean	0.95			0.774
	stderr				0.289
σ_2	mean	1.931	1.798	1.926	1.865
	stderr		0.151	0.159	0.161

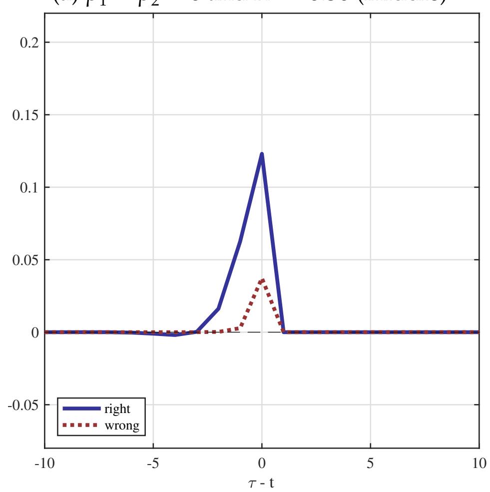
NOTES. Number of samples is $n_{\mathrm{MC}} = 2,000$ and sample size is $T = 280$.



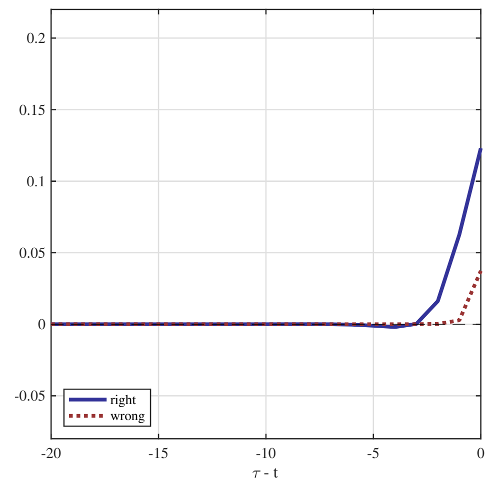
(a) $\rho_1 = \rho_2 = 0$ and $R^2 = 0.30$ (middle)



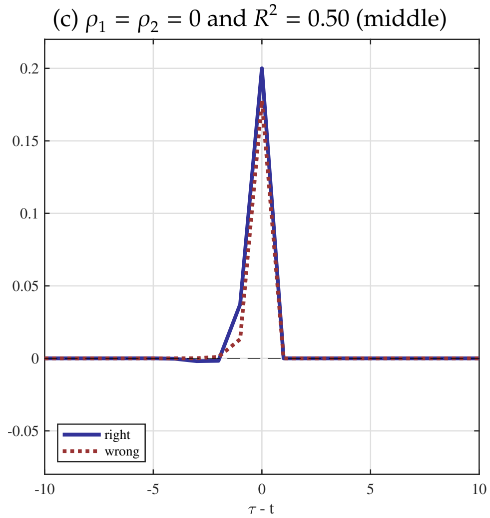
(b) $\rho_1 = \rho_2 = 0$ and $R^2 = 0.30$ (end)



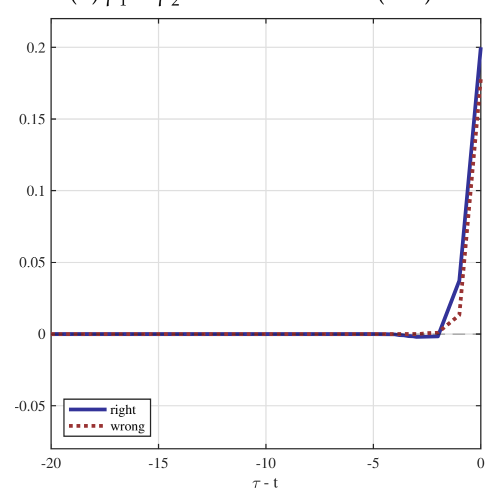
(c) $\rho_1 = \rho_2 = 0$ and $R^2 = 0.50$ (middle)



(d) $\rho_1 = \rho_2 = 0$ and $R^2 = 0.50$ (end)



(e) $\rho_1 = \rho_2 = 0$ and $R^2 = 0.85$ (middle)



(f) $\rho_1 = \rho_2 = 0$ and $R^2 = 0.85$ (end)

FIGURE SM.B.1. Weights of Kalman smoother. Horizontal axis is $\tau - t$; vertical axis is first entry of $\bar{\phi}_{\tau,T}$ (red) and $\phi^*_{\tau,T}$ (blue). Panels (a), (c) and (e) display weights for $t \approx T/2$ (middle), and panels (b), (d) and (f) for $t = T$ (end). The filters are computed using $\mu_0 = 3$, $\rho_0 = 0.50$, $\sigma_0 = 3.25$. Wrong filter uses $\rho_0 + B$ as AR root with B computed from (2).

TABLE SM.B.7. Monte Carlo simulation for $\rho_1 = \rho_2 = 0$ and $t \approx T/2$.

		$\Delta\hat{x}_t$	$\Delta\bar{x}_t$	$\Delta\hat{x}_t^*$	Δx_t^*
$R^2 = 0.30$	RMSE	3.54	3.58	3.05	3.06
	increase	0.37	0.37	0.01	
$R^2 = 0.50$	RMSE	3.24	3.16	2.98	3
	increase	0.19	0.1	0.01	
$R^2 = 0.85$	RMSE	2.88	2.89	2.86	2.88
	increase	0.02	0.01	0.01	

NOTES. Number of samples is $n_{\text{MC}} = 2,000$ and sample size is $T = 280$. Columns "$\Delta\hat{x}_t$" and "$\Delta\bar{x}_t$" refer to the wrong filter at the ML estimates and pseudo true values, respectively. Columns "$\Delta\hat{x}_t^*$" and "Δx_t^*" refer to the right filter at the ML estimates and true values, respectively. Root MSE and increase in MSE as a fraction of the MSE of Δx_t^* are indicated for each filter and R^2.

TABLE SM.B.8. Monte Carlo simulation for $\rho_1 = \rho_2 = 0$ and $t = T$.

		$\Delta\hat{x}_t$	$\Delta\bar{x}_t$	$\Delta\hat{x}_t^*$	Δx_t^*
$R^2 = 0.30$	RMSE	3.51	3.53	3.05	3.05
	increase	0.38	0.37	0.02	
$R^2 = 0.50$	RMSE	3.21	3.12	2.98	2.99
	increase	0.19	0.1	0.01	
$R^2 = 0.85$	RMSE	2.88	2.89	2.87	2.88
	increase	0.02	0.01	0.01	

NOTES. Number of samples is $n_{\text{MC}} = 2,000$ and sample size is $T = 280$. Columns "$\Delta\hat{x}_t$" and "$\Delta\bar{x}_t$" refer to the wrong filter at the ML estimates and pseudo true values, respectively. Columns "$\Delta\hat{x}_t^*$" and "Δx_t^*" refer to the right filter at the ML estimates and true values, respectively. Root MSE and increase in MSE as a fraction of the MSE of Δx_t^* are indicated for each filter and R^2.

TABLE SM.B.9. Monte Carlo simulation for $\rho_1 = 0, \rho_2 = 0.95$ and $t \approx T/2$.

		$\Delta\hat{x}_t$	$\Delta\bar{x}_t$	$\Delta\hat{x}_t^*$	Δx_t^*
$R^2 = 0.30$	RMSE	3.27	3.2	3.02	3.03
	increase	0.18	0.11	0.02	
$R^2 = 0.50$	RMSE	3.06	3.15	2.95	2.95
	increase	0.07	0.14	0.02	
$R^2 = 0.85$	RMSE	2.84	3.05	2.82	2.83
	increase	0.02	0.18	0.01	

NOTES. Number of samples is $n_{\mathrm{MC}} = 2,000$ and sample size is $T = 280$. Columns "$\Delta\hat{x}_t$" and "$\Delta\bar{x}_t$" refer to the wrong filter at the ML estimates and pseudo true values, respectively. Columns "$\Delta\hat{x}_t^*$" and "Δx_t^*" refer to the right filter at the ML estimates and true values, respectively. Root MSE and increase in MSE as a fraction of the MSE of Δx_t^* are indicated for each filter and R^2.

TABLE SM.B.10. Monte Carlo simulation for $\rho_1 = 0, \rho_2 = 0.95$ and $t = T$.

		$\Delta\hat{x}_t$	$\Delta\bar{x}_t$	$\Delta\hat{x}_t^*$	Δx_t^*
$R^2 = 0.30$	RMSE	3.31	3.31	3.15	3.15
	increase	0.16	0.1	0.02	
$R^2 = 0.50$	RMSE	3.12	3.28	3.07	3.09
	increase	0.06	0.12	0.01	
$R^2 = 0.85$	RMSE	2.96	3.22	2.94	2.96
	increase	0.01	0.17	0.01	

NOTES. Number of samples is $n_{\mathrm{MC}} = 2,000$ and sample size is $T = 280$. Columns "$\Delta\hat{x}_t$" and "$\Delta\bar{x}_t$" refer to the wrong filter at the ML estimates and pseudo true values, respectively. Columns "$\Delta\hat{x}_t^*$" and "Δx_t^*" refer to the right filter at the ML estimates and true values, respectively. Root MSE and increase in MSE as a fraction of the MSE of Δx_t^* are indicated for each filter and R^2.

SM.C Long-run objects

Formally, by a long-run object we mean henceforth a weighted average $X = \sum_{t=1}^{T} \omega_t \Delta x_t$, where the weights $\omega_{1:T}$ satisfy $\|\omega_{1:T}\| = \sqrt{\sum_{t=1}^{T} \omega_t^2} = O\left(1/\sqrt{T}\right)$. And, of course, $\omega_t \geq 0$ and $\sum_{t=1}^{T} \omega_t = 1$. To be precise, we ask that $\sqrt{T}\omega_t = \tilde{\omega}(t/T)$ where $\tilde{\omega} : [0,1] \to \mathbb{R}$ is of bounded variation and $\int_0^1 \tilde{\omega}^2(s)ds = O(1)$. As an example, consider writing the average growth rate of economic activity for the 2010's decade in a sample running from 1950 to 2019 as X with $\omega_t \propto \mathbb{1}\{\text{decade}(t) = 2010\}$. As mentioned in the body of the paper, neglecting the common trend affects inferences about long-run objects by inflating measures of their uncertainty, such as standard errors or confidence intervals.

As in our discussion of signal extraction for short-run objects, we abstract from estimation uncertainty by using pseudo-true parameter values for the misspecified model and true values for the correctly specified one. Let $Y = \sum_{t=1}^{T} \omega_t \Delta y_t$ and $V = \sum_{t=1}^{T} \omega_t \Delta v_t$ for a given set of weights $\omega_{1:T}$. The measurement equation (1) delivers

$$Y = X1_{N\times 1} + V.$$

We are interested in the problem of constructing a confidence interval for X. To keep the exposition simple, we will condition on X, which effectively treats X as a fixed quantity rather than as a latent variable.[1] Theorem 1 in Müller and Watson (2017) implies that under the misspecified model at the pseudo-true values,[2]

$$(Y - X1_{N\times 1})\,|X \implies N\left[0_{N\times 1}, \tilde{\Omega}^2 \operatorname{diag}\left(\sigma_{1:N}^2\right)\right],$$

where $\tilde{\Omega}^2 = \int_{-\infty}^{\infty} \left| \int_0^1 e^{i\lambda s} \tilde{\omega}(s)\, ds \right|^2 d\lambda$, with $\tilde{\omega}(s) = \sqrt{T}\omega_{\lfloor sT \rfloor}$, $\lfloor sT \rfloor$ the integer part of sT and $\omega_{1:T}$ the weights used to construct X, Y and V. As a consequence, a (pointwise asymptotic) level-$(1-\alpha)$ confidence interval for X based on this approximation will be

$$\overline{\text{CI}}_\alpha = \left[\sum_{i=1}^{N} (\bar{\sigma}^2/\sigma_i^2) Y_i \pm \Phi^{-1}(\alpha)\tilde{\Omega}\bar{\sigma} \right],$$

[1] Our model implies an unconditional distribution for X that smoothing calculations would exploit in constructing confidence intervals, but it appears from our simulation evidence below that this alternative approach would not critically modify our results.

[2] In fact, we only need the limit variance calculations from Müller and Watson (2017) since normality is in our case the result of V being a linear combination of normally distributed random variables under $\mathbb{P}$ and $\mathcal{P}$.

where Φ is the standard normal CDF. In contrast, under the true data generating process,

$$T\left(Y - X1_{N\times 1}\right)|X \implies N\left[0_{N\times 1}, \tilde{\Omega}^2 \operatorname{diag}\left(\Sigma^2_{1:N}\right)\right],$$

where $\Sigma_i^2 = \sigma_i^2(1+\rho_i)(1-\rho_i)^{-2}/2$ is the long-run variance of v_{it}. Therefore, the level-$(1-\alpha)$ confidence interval for X based on this approximation will be

$$\mathrm{CI}^*_\alpha = \left[\sum_{i=1}^{N}(\tilde{\Sigma}^2/\Sigma_i^2)Y_i \pm \Phi^{-1}(\alpha)\tilde{\Omega}\frac{\tilde{\Sigma}}{T}\right]$$

with $\tilde{\Sigma}^2 = \left[\sum_{i=1}^{N}(1/\Sigma_i^2)\right]^{-1}$. Hence, it follows that as $T \to \infty$,

$$\frac{\text{length}(\mathrm{CI}^*_\alpha)}{\text{length}(\overline{\mathrm{CI}}_\alpha)} = \frac{\bar{\Sigma}}{T\bar{\sigma}} \to 0.$$

In other words, the confidence interval based on the misspecified model is arbitrarily long compared to the optimal interval. Given that $\tilde{\Omega}\bar{\sigma}$ is the standard error a researcher who believes the misspecified model $\mathcal{P}$ is correct would report, our calculations suggest that "putative" measures of uncertainty of smoothed estimates of long-run objects tend to overstate the actual uncertainty about them.[3]

We should mention an alternative inference approach is available when $\Sigma_{1:N}$ is unknown and must be estimated.[4] Let $\hat{\Sigma}^2_\gamma$ be an estimate of the long-run variance of a linear combination of measurement errors $\sum_{i=1}^{N}\gamma_i\Delta v_{it}$ for weights $\gamma = \gamma_{1:N}$ adding up to 1. Then,

$$\left(\sum_{i=1}^{N}\gamma_i Y_i - X\right)|X \implies N(0, \tilde{\Omega}^2 \ \operatorname{plim}(\hat{\Sigma}_\gamma)).$$

Therefore, the level-$(1-\alpha)$ confidence interval for X based on this approximation will be

$$\widehat{\mathrm{CI}}_\alpha = \left[\sum_{i=1}^{N}\gamma_i Y_i \pm \Phi^{-1}(\alpha)\tilde{\Omega}\hat{\Sigma}_\gamma\right].$$

The interval $\widehat{\mathrm{CI}}_\alpha$ will tend to zero for large T as CI^*_α does.

[3] Although here we focus on a situation with no estimation uncertainty, which allows us to reduce the inference problem by focusing on the sufficient statistics Y, unreported simulation experiments confirm the same patterns for Kalman smoother calculations evaluated at maximum likelihood estimates.

[4] We thank an anonymous referee for this suggestion.

References

Dzhaparidze, K. (1986): *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series*, Springer Verlag.

Grenander, U. and M. Rosenblatt (1958): "Statistical spectral analysis of time series arising from stationary stochastic processes," *Matematika*, 2, 123–144.

Müller, U. K. and M. W. Watson (2017): "Low-frequency econometrics," in *Advances in Economics and Econometrics: Eleventh World Congress of the Econometric Society*, ed. by B. Honoré, A. Pakes, M. Piazzesi, and L. Samuelson, Cambridge University Press, vol. 2.

Shephard, N. (1993): "Distribution of the ML estimator of an MA(1) and a local level model," *Journal of Econometrics*, 9, 377–401.