Appendix to `https://doi.org/10.1007/s00362-023-01451-y`

Supplementary material

In this Appendix we present some simulation results that illustrate the asymptotic efficiency of our estimation algorithm SUBEST1. For each data-generating process (DGP) considered, we report experimental outcomes based on 500 replications, except otherwise indicated.

<u>DGP I</u>: The VARMA$_E$ with Kronecker indices $\alpha = (1, 0)$ used in the paper, and defined as:

$$(\mathbf{F_0} + \mathbf{F_1}B)\mathbf{z}_t = (\mathbf{Q_0} + \mathbf{Q_1}B)\mathbf{a}_t, \tag{1}$$

where,

$$\mathbf{F_0} = \mathbf{Q_0} = \begin{bmatrix} 1 & 0 \\ 0.5 & 1 \end{bmatrix}, \quad \mathbf{F_1} = \begin{bmatrix} -0.5 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{Q_1} = \begin{bmatrix} -0.3 & 0.2 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{\Sigma_a} = \begin{bmatrix} 1.0 & 0.4 \\ 0.4 & 1.0 \end{bmatrix}.$$

<u>DGP II</u>: The state-space system obtained from Deistler et al. (1995) :

$$\mathbf{x}_{t+1} = \mathbf{\Phi}\mathbf{x}_t + \mathbf{K}\mathbf{a}_t, \tag{2a}$$

$$\mathbf{z}_t = \mathbf{H}\mathbf{x}_t + \mathbf{a}_t \tag{2b}$$

where,

$$\mathbf{\Phi} = \begin{bmatrix} 0.8 & 0.2 \\ -0.4 & -0.5 \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} 1.5 & 0 \\ -0.2 & -0.8 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{\Sigma_a} = \begin{bmatrix} 1.0 & 0 \\ 0 & 1.0 \end{bmatrix}.$$

<u>DGP III</u>: The state-space system obtained from Deistler et al. (1995) :

$$\mathbf{x}_{t+1} = \mathbf{\Phi}\mathbf{x}_t + \mathbf{K}\mathbf{a}_t, \tag{3a}$$

$$\mathbf{z}_t = \mathbf{H}\mathbf{x}_t + \mathbf{a}_t \tag{3b}$$

where,

$$\mathbf{\Phi} = \begin{bmatrix} 0 & 1 & 0 \\ 0.25 & 0 & 0.25 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} 0.0055 & 0.063 \\ 0.119 & 0.157 \\ 0.674 & 0.0007 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{\Sigma_a} = \begin{bmatrix} 1.0 & 0 \\ 0 & 1.0 \end{bmatrix}.$$

<u>DGP IV</u>: The VARMA$_E$ with Kronecker indices $\alpha = (2,1,1)$ used in the paper, and defined as:

$$(\mathbf{F_0} + \mathbf{F_1}B + \mathbf{F_2}B^2)\mathbf{z}_t = (\mathbf{Q_0} + \mathbf{Q_1}B + \mathbf{Q_2}B^2)\mathbf{a}_t, \tag{4}$$

where,

$$\mathbf{F_0} = \mathbf{Q_0} = \begin{bmatrix} 1 & 0 & 0 \\ 0.4 & 1 & 0 \\ -0.6 & 0 & 1 \end{bmatrix}, \quad \mathbf{F_1} = \begin{bmatrix} -0.7 & 0 & 0 \\ 0.2 & -0.5 & -0.9 \\ 0.4 & 0.3 & -0.2 \end{bmatrix}, \quad \mathbf{F_2} = \begin{bmatrix} 0.3 & -0.2 & 0.5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{Q_1} = \begin{bmatrix} -0.2 & 0.4 & 0.7 \\ 0.6 & -0.3 & -0.4 \\ 0.3 & 1 & -0.8 \end{bmatrix}, \quad \mathbf{Q_2} = \begin{bmatrix} 0.3 & 0.5 & -0.8 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{\Sigma_a} = \begin{bmatrix} 1 & 0.5 & 0.4 \\ 0.5 & 1 & 0.7 \\ 0.4 & 0.7 & 1 \end{bmatrix}.$$

<u>DGP V</u>: The VARMA$_E$ with Kronecker indices $\alpha = (2,1)$ obtained from Poskitt (2016), and defined as:

$$(\mathbf{F_0} + \mathbf{F_1}B + \mathbf{F_2}B^2)\mathbf{z}_t = (\mathbf{Q_0} + \mathbf{Q_1}B + \mathbf{Q_2}B^2)\mathbf{a}_t, \tag{5}$$

where,

$$\mathbf{F_0} = \mathbf{Q_0} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{F_1} = \begin{bmatrix} 0 & 0 \\ 0.201 & -0.491 \end{bmatrix}, \quad \mathbf{F_2} = \begin{bmatrix} -0.438 & -0.843 \\ 0 & 0 \end{bmatrix},$$

$$\mathbf{Q_1} = \begin{bmatrix} 0.300 & 0.644 \\ 0 & 0.348 \end{bmatrix}, \quad \mathbf{Q_2} = \begin{bmatrix} -0.468 & 0 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{\Sigma_a} = \begin{bmatrix} 4.739 & 1.358 \\ 1.358 & 15.381 \end{bmatrix}.$$

In this case we reduced the number of replications to 200 due to the high computational cost of MLE estimation.

DGP I:

Table 1: Mean and standard deviation of the estimates for the free parameters in DGP I relative to MLE.

Metric	T	$f_0 = 0.5$	$f_1 = -0.5$	$q_{11} = -0.3$	$q_{12} = 0.2$
Mean	150	100.71	96.58	95.60	98.83
	300	100.30	98.62	98.92	100.44
	500	100.05	99.18	99.35	100.42
	1000	99.97	99.68	100.08	100.40
	5000	100.00	99.95	100.01	100.31
St. Dev.	150	113.49	109.55	110.50	105.28
	300	102.89	104.94	106.95	103.65
	500	101.23	102.62	103.51	102.36
	1000	100.96	101.62	102.87	100.99
	5000	100.02	100.31	100.69	100.43

Relative values are multiplied by 100 to facilitate the comparison. In the subspace algorithm $p = f$ are chosen as the integer closer to $\log(T)$. We initialized the ML estimation procedure with the true value of the parameters.

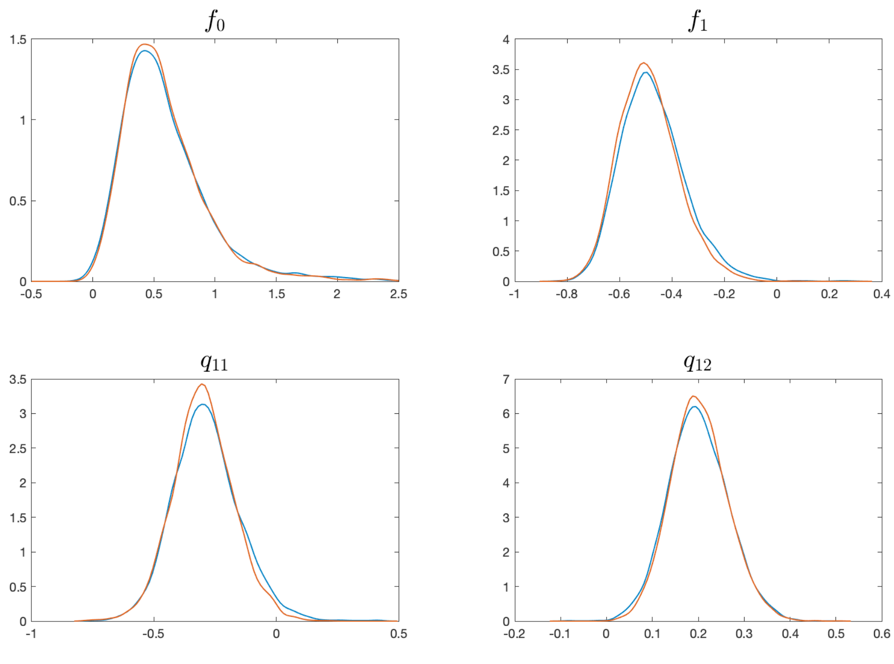


Figure 1: Kernel estimates of empirical distributions of the estimates of DGP I with 5000 replications with sample size $T = 150$. Blue line corresponds to the SUBEST1 estimates. Red line corresponds to MLE.

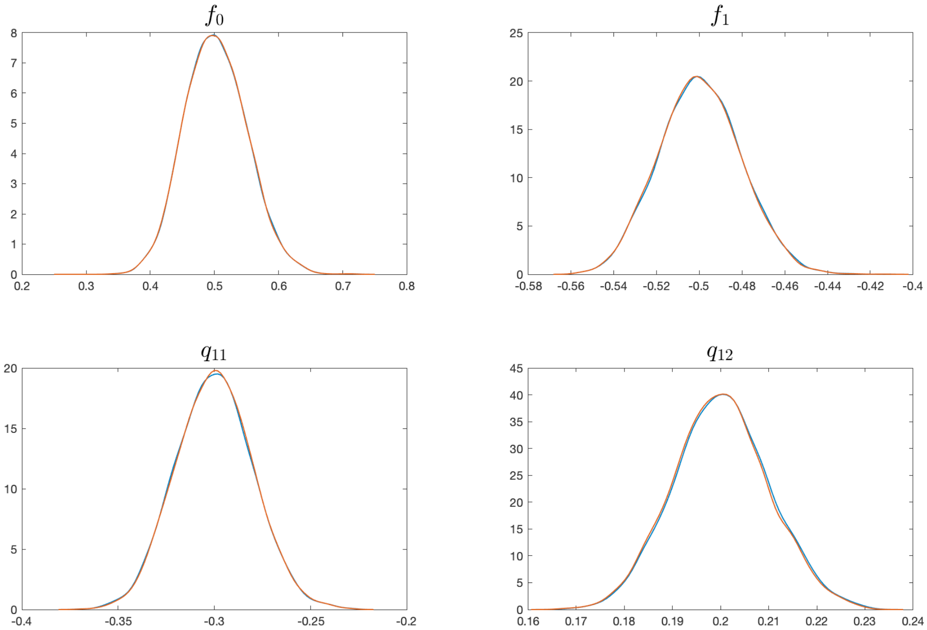


Figure 2: Kernel estimates of empirical distributions of the estimates of DGP I with 5000 replications with sample size $T = 5000$. Blue line corresponds to the SUBEST1 estimates. Red line corresponds to MLE.

Table 2: Simulation results for DGP I

Algorithm	$T = 2000$	$T = 4000$	$T = 8000$	$T = 16000$
M(SUBEST1)	4.1636	4.0026	3.5731	4.0586
M(MLE)	4.1493	4.0091	3.5542	4.0452
CRB	3.8955	3.8955	3.8955	3.8955
D(SUBEST1)	0.0583	0.0259	0.0153	0.0083

The row M(SUBEST1) shows the average of T times the mean-square error of the SUBEST1 estimates for all the parameters of the model, see definition of M in the paper. The row D(SUBEST1) displays the average of T times the mean-square difference between SUBEST1 and MLE estimates for all the parameters of the model, see definition of D in the paper. CRB is the mean of the Cramer-Rao Bounds obtained from the inverse of the Fisher information matrix.

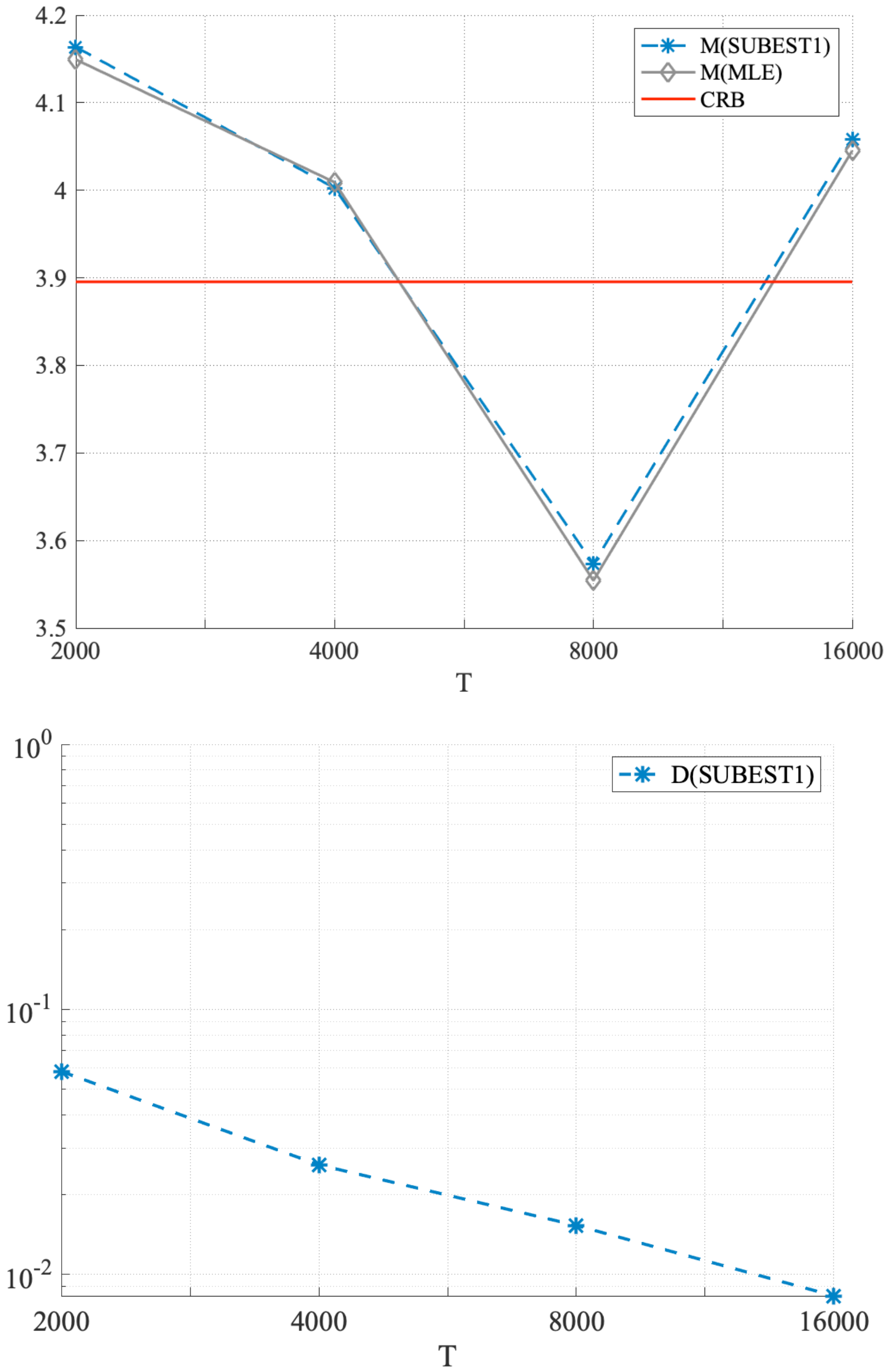


Figure 3: Results of the simulation of DGP I; see Table 2 for further details.

GDP II:

Table 3: Simulation results for DGP II

Algorithm	$T = 2000$	$T = 4000$	$T = 8000$	$T = 16000$
M(SUBEST1)	0.8594	0.8247	0.8104	0.8138
M(MLE)	0.8307	0.8088	0.8034	0.8116
M(ECH)*	1.0040	0.8900	0.9084	0.9376
M(CVA)*	0.8929	0.7745	0.7796	0.8234
CRB	0.8195	0.8195	0.8195	0.8195
D(SUBEST1)	0.0424	0.0234	0.0119	0.0050
D(ECH)*	0.1418	0.1381	0.1460	0.1190
D(CVA)*	0.0224	0.0115	0.0073	0.0038

The row M(SUBEST1) shows the average of T times the mean-square error of the SUBEST1 estimates for all the parameters of the model, see definition of M in the paper. The row D(SUBEST1) displays the average of T times the mean-square difference between SUBEST1 and MLE estimates for all the parameters of the model, see definition of D in the paper. CRB is the mean of the Cramer-Rao Bounds obtained from the inverse of the Fisher information matrix.

*Values for ECH and CVA have been obtained from Deistler et al. (1995).

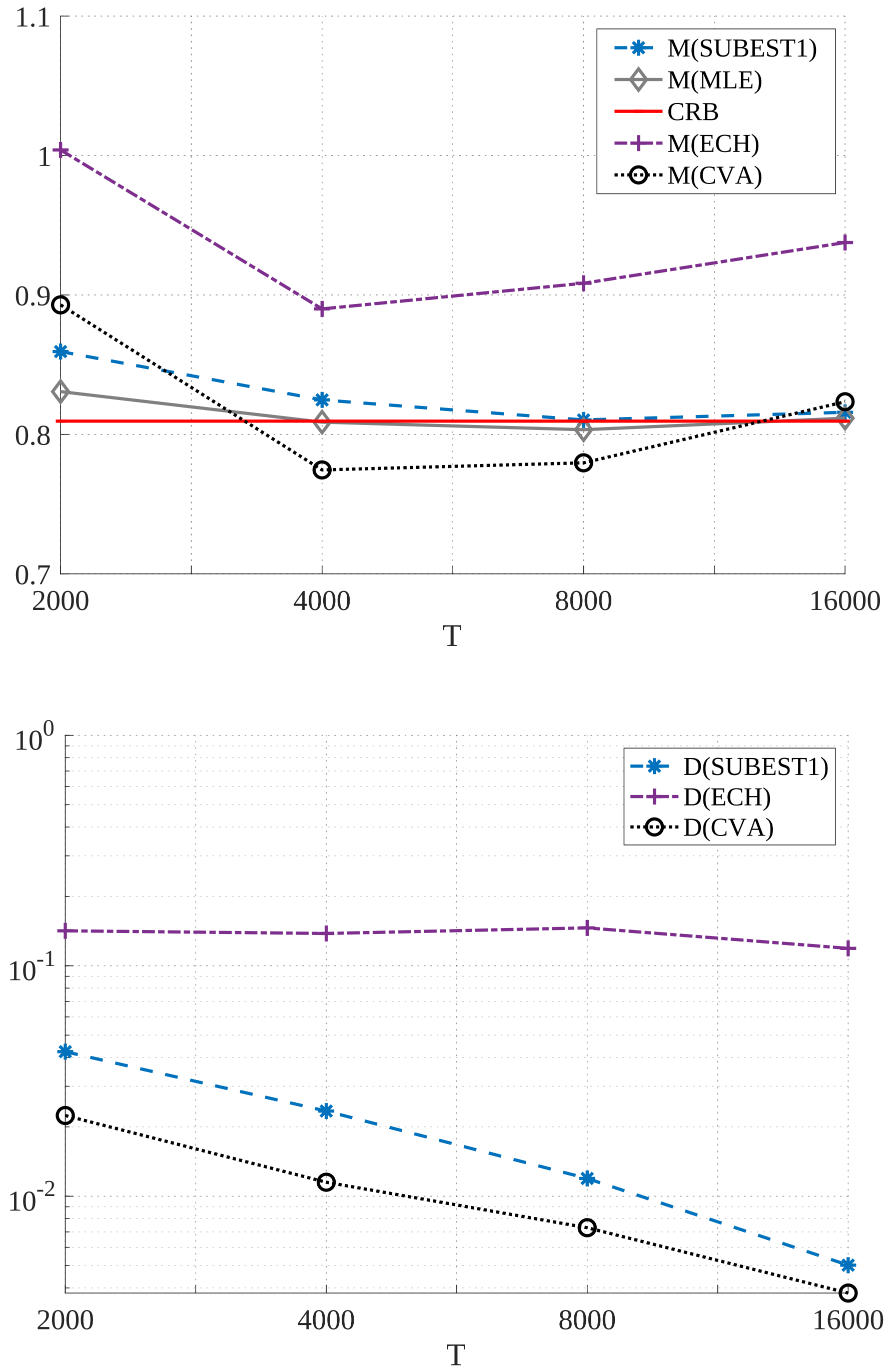


Figure 4: Results of the simulation of DGP II. Values for ECH and CVA have been obtained from Deistler et al. (1995); see Table 3 for further details.

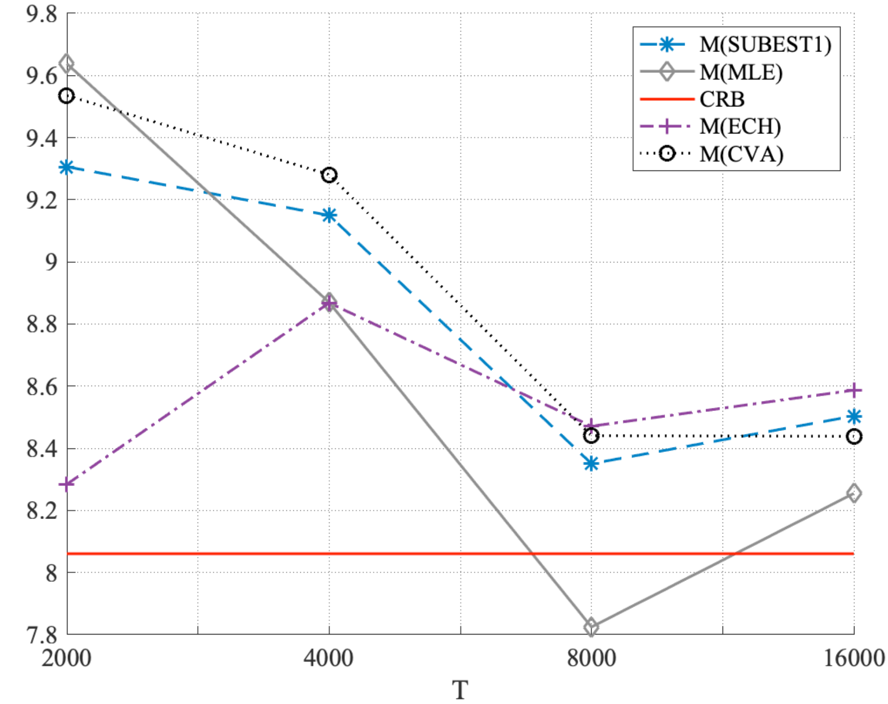
DGP III:

Table 4: Simulation results for DGP III

Algorithm	$T = 2000$	$T = 4000$	$T = 8000$	$T = 16000$
M(SUBEST1)	0.8594	0.8247	0.8104	0.8138
M(MLE)	0.8307	0.8088	0.8034	0.8116
M(ECH)*	1.0040	0.8900	0.9084	0.9376
M(CVA)*	0.8929	0.7745	0.7796	0.8234
CRB	0.8195	0.8195	0.8195	0.8195
D(SUBEST1)	0.0424	0.0234	0.0119	0.0050
D(ECH)*	0.1418	0.1381	0.1460	0.1190
D(CVA)*	0.0224	0.0115	0.0073	0.0038

The row M(SUBEST1) shows the average of T times the mean-square error of the SUBEST1 estimates for all the parameters of the model, see definition of M in the paper. The row D(SUBEST1) displays the average of T times the mean-square difference between SUBEST1 and MLE estimates for all the parameters of the model, see definition of D in the paper. CRB is the mean of the Cramer-Rao Bounds obtained from the inverse of the Fisher information matrix.

*Values for ECH and CVA have been obtained from Deistler et al. (1995).



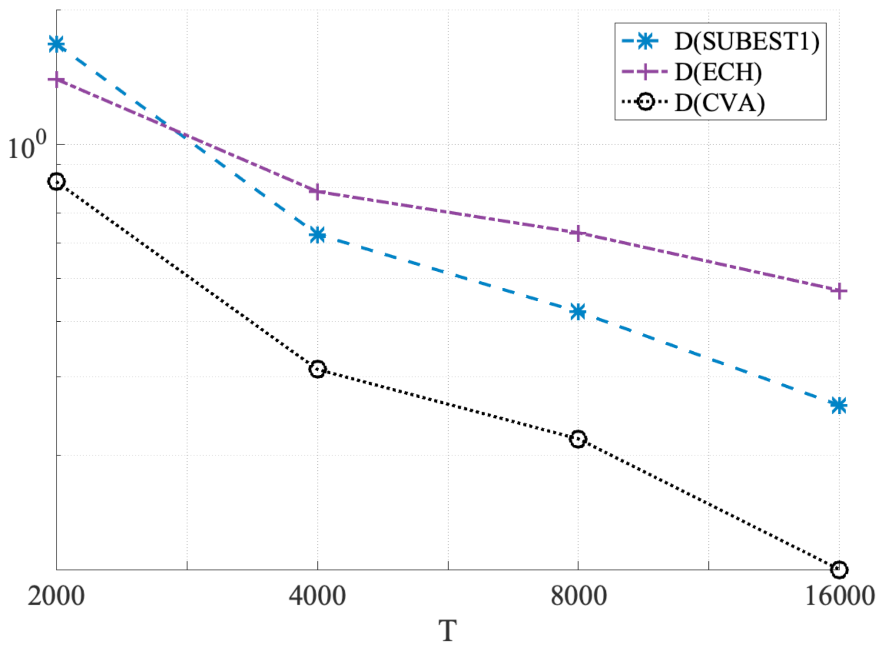


Figure 5: Results of the simulation of DGP III. Values for ECH and CVA have been obtained from Deistler et al. (1995); see Table 4 for further details.

DGP IV:

Table 5: Simulation results for DGP IV

Algorithm	$T = 2000$	$T = 4000$	$T = 8000$	$T = 16000$
M(SUBEST1)	0.8594	0.8247	0.8104	0.8138
M(MLE)	0.8307	0.8088	0.8034	0.8116
CRB	0.8195	0.8195	0.8195	0.8195
D(SUBEST1)	0.0424	0.0234	0.0119	0.0050

The row M(SUBEST1) shows the average of T times the mean-square error of the SUBEST1 estimates for all the parameters of the model, see definition of M in the paper. The row D(SUBEST1) displays the average of T times the mean-square difference between SUBEST1 and MLE estimates for all the parameters of the model, see definition of D in the paper. CRB is the mean of the Cramer-Rao Bounds obtained from the inverse of the Fisher information matrix.

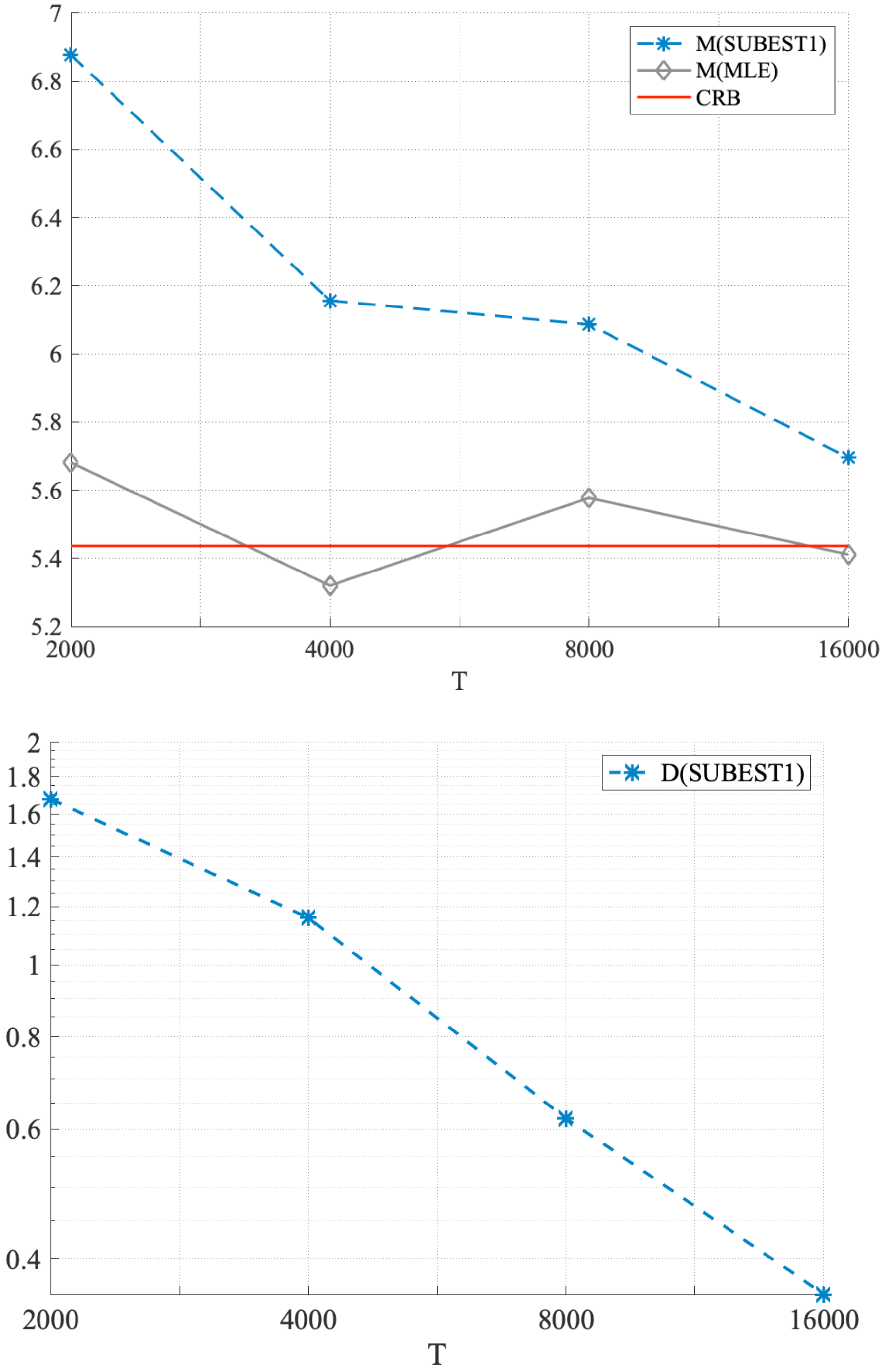


Figure 6: Results of the simulation of DGP IV; see Table 5 for further details.

DGP V:

Table 6: Simulation results for DGP V

Algorithm	$T = 2000$	$T = 4000$	$T = 8000$	$T = 16000$
M(SUBEST1)	6.5128	6.5687	4.8594	4.3226
M(MLE)	2.8926	3.4417	3.0581	3.0915
CRB	3.1154	3.1154	3.1154	3.1154
D(SUBEST1)	4.9566	3.9884	2.4740	1.1657

The row M(SUBEST1) shows the average of T times the mean-square error of the SUBEST1 estimates for all the parameters of the model, see definition of M in the paper. The row D(SUBEST1) displays the average of T times the mean-square difference between SUBEST1 and MLE estimates for all the parameters of the model, see definition of D in the paper. CRB is the mean of the Cramer-Rao Bounds obtained from the inverse of the Fisher information matrix.

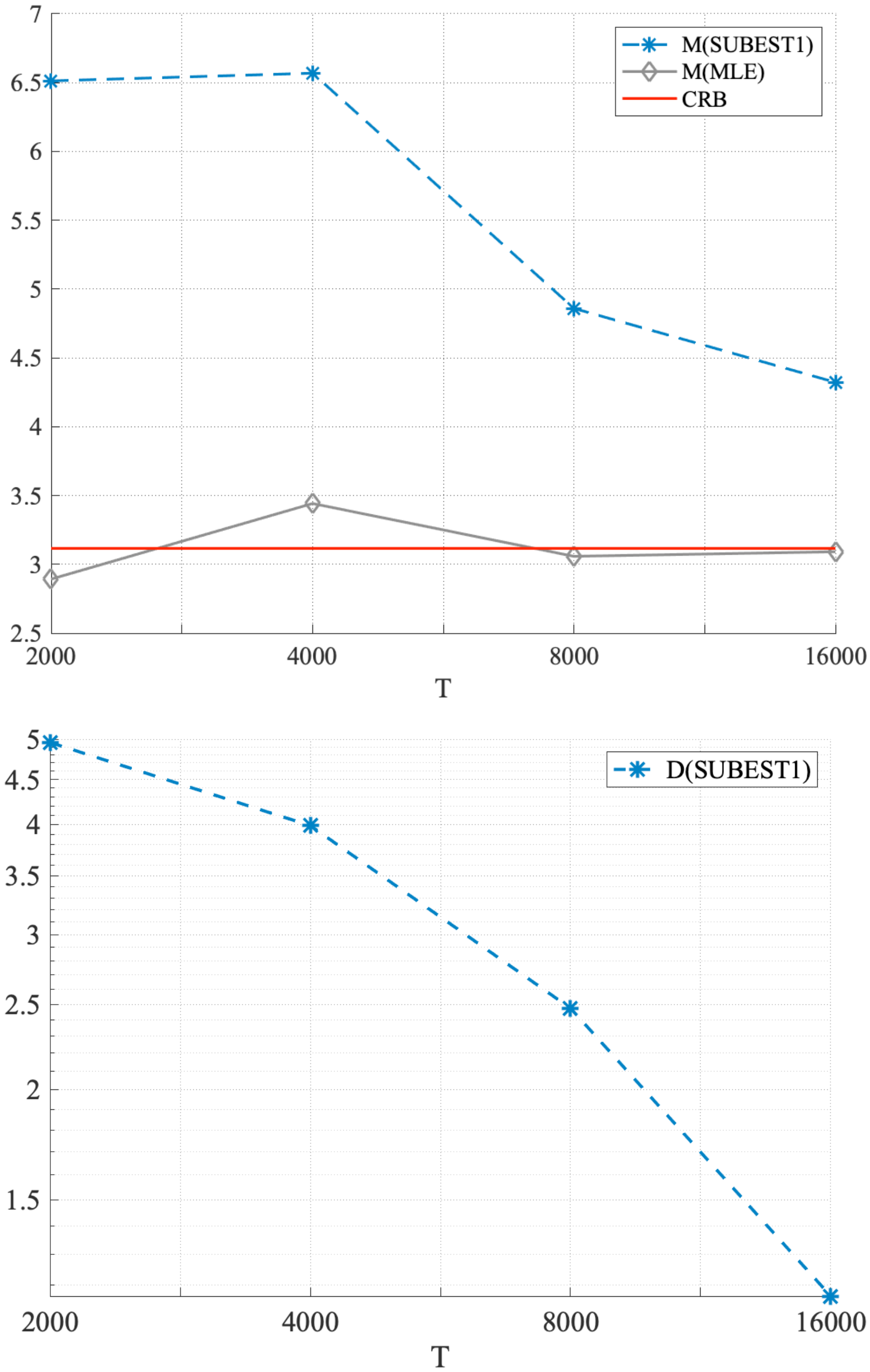


Figure 7: Results of the simulation of DGP IV; see Table 6 for further details.

References

Deistler, M., Peternell, K., and Scherrer, W. (1995). Consistency and relative efficency of subspace methods. *Automatica*, 31, 12:1865–1875.

Poskitt, D. S. (2016). Vector autoregressive moving average identification for macroeconomic modeling: A new methodolgy. *Journal of Econometrics*, 192:468–484.