Appendix

A Estimation of the Factor Model Parameters by ML

Bai and Li (2012) apply the EM algorithm on a slightly different static factor model than ours defined in (3.3) to (3.5); i.e., one that fulfils the identification restrictions of the Principal Components Analysis (PCA):

$$\boldsymbol{y}_t = \underline{B}^* \boldsymbol{g}_t^* + \boldsymbol{e}_t \tag{A.1}$$

$$\boldsymbol{g}_t^* \overset{i.i.d}{\sim} \mathcal{N}\left(\mathbf{0}_{k\times 1}, I_k\right), \tag{A.2}$$

$$\boldsymbol{e}_t \overset{i.i.d}{\sim} \mathcal{N}\left(\mathbf{0}_{N\times 1}, \Sigma_2^*\right), \tag{A.3}$$

where I_k is a k dimensional identity matrix. Thus, in the model (A.1)-(A.2), the identification assumptions are that the diagonal elements of $(1/N)\underline{B}^{*\prime}(\Sigma_2^*)^{-1}\underline{B}^*$ are distinct and positive and are arranged in decreasing order. $\underline{B}^*$ is identified up to a column sign change (Bai and Li, 2012).

One may see that while Equation (A.3) is identical to Equation (3.4), meaning that in the PCA identification framework the error inherits the structure of the one in the exact static factor model, the loadings and factors/components differ slightly due to other identification restrictions. The EM algorithm we use in this paper is applied to the model (A.1)-(A.2), however, we ex-post use the rotation procedure described by Bai and Li (2012) to recover the parameters of our static factor model defined in equations (3.3) to (3.5), as described below.

For convenience, we present here the estimation based on the negative log-likelihood of the model specified in equations (A.1), (A.3) and (A.2) that is proportional to:

$$F^{a,1}(\boldsymbol{y}_t, \underline{B}^*, \Sigma_2^*) = \frac{1}{N}\underbrace{\ln\left\{\det\left(\underline{B}^*\underline{B}^{*\prime} + \Sigma_2^*\right)\right\}}_{M^*} + \frac{1}{N}\underbrace{\operatorname{tr}\left[\left(\underline{B}^*\underline{B}^{*\prime} + \Sigma_2^*\right)^{-1} A\right]}_{C^*}, \tag{A.4}$$

with A being the empirical covariance matrix of the data. Thus, the specification in Equation (A.4) is a combination of a concave part (denoted by M^*) and a convex part (denoted by C^*), which usually leads to many local minima in the optimisation (Bai and Li, 2012; Bien and Tibshirani, 2011). We describe below the steps of implementing the EM algorithm of Bai and Li (2012) to circumvent this problem, however not directly to the proportional likelihood in Equation (A.4), but to the "majorized" negative log-likelihood proposed by Bien and Tibshirani

(2011) that approximates the concave part of (A.4), M^* by its tangent hyperplane at each iteration l:

$$\begin{aligned}\tilde{F}_{\ell}^{a,1}(\boldsymbol{y}_t, \underline{B}^*, \Sigma_2^*) = & \frac{1}{N}\ln\left|\det\left(\widehat{\underline{B}_{\ell}^*}\widehat{\underline{B}_{\ell}^*}' + \widehat{\Sigma_{2,\ell}^*}\right)\right| \\ & + \frac{1}{N}\operatorname{tr}\left[2\widehat{\underline{B}_{\ell}^*}'\left(\widehat{\underline{B}_{\ell}^*}\widehat{\underline{B}_{\ell}^*}' + \widehat{\Sigma_{2,\ell}^*}\right)^{-1}(\underline{B}^* - \widehat{\underline{B}_{\ell}^*})\right] \\ & + \frac{1}{N}\operatorname{tr}\left[(\underline{B}^*\underline{B}^{*\prime} + \widehat{\Sigma_{2,\ell}^*})^{-1}A\right], \end{aligned} \tag{A.5}$$

where $\widehat{\underline{B}_{\ell}^*}$ and $\widehat{\Sigma_{2,\ell}^*}$ are the estimators at the ℓ-th iteration of the EM algorithm obtained at S2, described below, yielding, thus, a fully convex optimisation problem. After the algorithm converges, we follow Bai and Li (2012) and rotate $\widehat{\underline{B}^*}$ to obtain $\widehat{B^*}$ and $\widehat{\Sigma_1^*}$, which correspond to the static factor of our interest defined in equations (3.3) to (3.5):

$$\widehat{B^*} = \widehat{\underline{B}^*}\mathcal{Q}\mathcal{W}^{-1}, \tag{A.6}$$

$$\widehat{\Sigma_1^*} = \mathcal{W}\mathcal{W}' \tag{A.7}$$

with $\mathcal{Q}$ stemming from the QR-decomposition of the transpose of the first $k \times k$ block of $\widehat{\underline{B}}^*$ and $\mathcal{W} = \operatorname{diag}(\widehat{\underline{B}^*}\mathcal{Q})$.

The estimators $\widehat{B^*}$, $\widehat{\Sigma_1^*}$ and $\widehat{\Sigma_2^*}$ fulfil the identification conditions of our static factor model defined in equations (3.3) to (3.5) (Bai and Li, 2012).

B Estimation of the ARSV Parameters by EMM

The ARSV parameters $\boldsymbol{\theta}_2 = (\varphi_1, \ldots, \varphi_{k+N}, \sigma_{\eta,1}, \ldots, \sigma_{\eta,k+N})'$ are estimated in pairs, i.e., $\boldsymbol{\theta}_{2,m} = (\varphi_m, \sigma_{\eta,m})'$ by making use of the auxiliary GARCH(1,1) estimation applied equation-by-equation to each element of $\hat{\boldsymbol{x}}_m = (\hat{x}_{1,m}, \ldots, \hat{x}_{T,m})'$ with $m = 1, \ldots, k+N$. Thus, the auxiliary log-likelihood is proportional to:

$$\mathcal{L}_m^{a,2}(\hat{\boldsymbol{x}}_m; \boldsymbol{\beta}_{2,m}) = -\frac{1}{T}\sum_{t=1}^{T}\left(\ln\left(\delta_{t,m}^2(\boldsymbol{\beta}_{2,m})\right) + \frac{\hat{x}_{t,m}^2}{\delta_{t,m}^2(\boldsymbol{\beta}_{2,m})}\right), \tag{B.1}$$

where $\delta_{t,m}^2$ is defined in Equation (3.11) and it is a function of $\boldsymbol{\beta}_{2,m}$. The single contributions, i.e., at time t, to the auxiliary model score vector, are obtained by taking the derivative of the log-likelihood function in (B.1) with respect to the parameter vector $\boldsymbol{\beta}_{2,m}$ at time t

$$\begin{aligned} q_{t,m}(\hat{x}_{t,m}, \boldsymbol{\beta}_{2,m}) &= \frac{\partial \mathcal{L}_m^{a,2}(\hat{\boldsymbol{x}}_m; \boldsymbol{\beta}_{2,m})}{\partial \boldsymbol{\beta}_{2,m}}. & \text{(B.2)} \\ &= \frac{1}{\delta_{t,m}^2(\boldsymbol{\beta}_{2,m})} \left(\frac{\partial \delta_{t,m}^2(\boldsymbol{\beta}_{2,m})}{\partial \boldsymbol{\beta}_{2,m}} \right) \left(\frac{\hat{x}_{t,m}^2}{\delta_{t,m}^2(\boldsymbol{\beta}_{2,m})} - 1 \right), & \text{(B.3)} \end{aligned}$$

where

$$\frac{\partial \delta_{t,m}^2(\boldsymbol{\beta}_{2,m})}{\partial \boldsymbol{\beta}_{2,m}} = \begin{pmatrix} \hat{x}_{t-1,m}^2 \\ \delta_{t-1,m}^2(\boldsymbol{\beta}_{2,m}) \end{pmatrix} - \hat{\sigma}_m^2 + \alpha_{2,m} \frac{\partial \delta_{t-1,m}^2(\boldsymbol{\beta}_{2,m})}{\partial \boldsymbol{\beta}_{2,m}}. \tag{B.4}$$

The score vector is then given by taking the mean over the single contributions in Equation (B.4):

$$\mathcal{Q}_m(\hat{\boldsymbol{x}}_m; \boldsymbol{\beta}_{2,m}) = \frac{1}{T} \sum_{t=1}^{T} q_{t,m}(\hat{x}_{t,m}, \boldsymbol{\beta}_{2,m}). \tag{B.5}$$

As described below, the EMM estimation for each of the pair of parameters $\boldsymbol{\theta}_{2,m} = (\varphi_m, \sigma_{\eta,m})'$ requires H number of simulated series $\boldsymbol{y}_t$. Section 3 provides a detailed discussion on how to choose H for our study.

Thus, the EMM estimation of the ARSV parameters can be parallelised over $m = 1, \ldots, k + N$ and consists of the following steps:

A0: Set $m = 1$ and choose the m-th time series of observations from $\hat{\boldsymbol{x}}_t = (\hat{g}_{t,1}, \ldots, \hat{g}_{t,k}, \hat{e}_{t,1}, \ldots, \hat{e}_{t,N})'$.

A1: Get the PML estimate $\hat{\boldsymbol{\beta}}_{2,m}$ of the GARCH(1,1) model based on $\mathcal{L}_m^{a,2}(\hat{\boldsymbol{x}}_m; \boldsymbol{\beta}_{2,m})$.

A2: Evaluate the auxiliary model score at $\hat{\boldsymbol{x}}_m$ and $\hat{\boldsymbol{\beta}}_{2,m}$, $\mathcal{Q}_m(\hat{\boldsymbol{x}}_m; \hat{\boldsymbol{\beta}}_{2,m})$.

A3: Choose a vector of initial values $\boldsymbol{\theta}_2$, i.e., $\boldsymbol{\theta}_2^0$.

A4: Simulate $T \times H$ observations $\tilde{\boldsymbol{y}}_t(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_2^0)$ from the MFSV data generating process defined in Section 2 by using $\hat{\boldsymbol{\theta}}_1$ obtained at the first step and $\boldsymbol{\theta}_2^0$.

A5: Compute $\tilde{\boldsymbol{x}}(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_2^0)$ from equations (3.8) and (3.9), by replacing $\boldsymbol{y}_t$ with its' simulated counterpart $\tilde{\boldsymbol{y}}_t(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_2^0)$ and by fixing Π^* in Equation (3.8) for all iterations and all m's to the estimated values from the first estimation step described in Section 3.1.

A6: Evaluate the auxiliary model score at the m-th element of $\tilde{\boldsymbol{x}}(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_2^0)$, $\tilde{\boldsymbol{x}}_m(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_2^0)$ and at $\hat{\boldsymbol{\beta}}_{2,m}$: $\mathcal{Q}_m(\tilde{\boldsymbol{x}}_m(\hat{\boldsymbol{\theta}}_1, \boldsymbol{\theta}_2^0); \hat{\boldsymbol{\beta}}_{2,m})$.

A7: Update the m-th element of $\boldsymbol{\theta}_2^0$, i.e., $\boldsymbol{\theta}_{2,m}^0$ to $\boldsymbol{\theta}_{2,m}^1$, by keeping the rest of the $m-1$ elements of $\boldsymbol{\theta}_2^0$ at the initial values set at step *A3*. Denote the updated $\boldsymbol{\theta}_2$ vector by $\boldsymbol{\theta}_2^{1_m,0}$ and repeat steps *A4* to *A6* until the weighted squared distance:

$$\begin{aligned}&\left(\mathcal{Q}_m(\tilde{\boldsymbol{x}}_m(\hat{\boldsymbol{\theta}}_1,\boldsymbol{\theta}_2^{1_m,0});\hat{\boldsymbol{\beta}}_{2,m})-\mathcal{Q}_m(\hat{\boldsymbol{x}}_m;\hat{\boldsymbol{\beta}}_{2,m})\right)'\Omega\\&\left(\mathcal{Q}_m(\tilde{\boldsymbol{x}}_m(\hat{\boldsymbol{\theta}}_1\boldsymbol{\theta}_2^{1_m,0});\hat{\boldsymbol{\beta}}_{2,m})-\mathcal{Q}_m(\hat{\boldsymbol{x}}_m;\hat{\boldsymbol{\beta}}_{2,m})\right),\end{aligned}\tag{B.6}$$

where Ω is a positive definite weighting matrix, is minimised. The m-th element of the last updated $\boldsymbol{\theta}_2^{j_m,0}$ that minimises the distance is the EMM estimator of $\boldsymbol{\theta}_{2,m}$, i.e., $\hat{\boldsymbol{\theta}}_{2,m}$.

A8: Set $m=m+1$ and repeat steps *A1*-*A7* until $m=k+N$.[19][20]

In our case, the parameter dimension of the auxiliary model is the same as the one of MFSV, i.e., $\dim\boldsymbol{\theta}=\dim\boldsymbol{\beta}$. Particularly, we have that $\dim\boldsymbol{\theta_{2,m}}=\dim\boldsymbol{\beta_{2,m}}$. Thus, the choice of the weighting matrix Ω at step *A7* can be arbitrary and we set it equal to the identity matrix. In this case of the exact identified model, the minimization of the weighted squared distance boils down to the minimization of the squared distance. In a matter of fact, in our estimation, we find $\hat{\boldsymbol{\theta}}_{2,m}$ that sets $\mathcal{Q}_m(\tilde{\boldsymbol{x}}_m(\hat{\boldsymbol{\theta}}_1,\boldsymbol{\theta}_2^{j_m,0});\hat{\boldsymbol{\beta}}_{2,m})-\mathcal{Q}_m(\hat{\boldsymbol{x}}_m;\hat{\boldsymbol{\beta}}_{2,m})=\mathbf{0}_{\dim\boldsymbol{\beta_{2,m}}\times 1}$. However, sometimes upon convergence, the $\hat{\boldsymbol{\theta}}_{2,m}$ do not fulfil the parameter constraints we define in Section 2 and, therefore, in these cases, we optimise the squared distance to get the EMM estimators.

In the steps *A4* to *A7*, we use Equation (2.4) to compute and update the values of the constant of the ARSV process μ_m^j from the updated m-th element of $\boldsymbol{\theta}_2^j$, i.e., $\boldsymbol{\theta}_{2,m}^j$ and from $\hat{\sigma}_m^2$, that is computed in the first estimation step described in Section 3.1.

[19] In a matter of fact step A.3 is implemented at the beginning before step A0 and not repeated for each m.

[20] We also implement an iterative version of the estimation procedure, where we use the estimated values at step *A8* as new initial values at step *A3*. We then iterate these steps until convergence. However, we find that the estimation results do not change significantly (only up to 3 digits after the comma) from the ones reported in the paper.

C Closed-form Gradient of the Static Factor Model

Consider the static factor model of the first part of the estimation approach:

$$\boldsymbol{y}_t = B^* \boldsymbol{g}_t + \boldsymbol{e}_t \tag{C.1}$$

$$\boldsymbol{g}_t \sim \mathcal{N}\left(\mathbf{0}_{k\times 1}, \Sigma_1^*\right) \tag{C.2}$$

$$\boldsymbol{e}_t \sim \mathcal{N}\left(\mathbf{0}_{N\times 1}, \Sigma_2^*\right), \tag{C.3}$$

where Σ_1^* and Σ_2^* are diagonal matrices and B^* is restricted such that $b_{jj}^* = 1$ and $b_{ij}^* = 0\,\forall j > i, \quad j = 1, \ldots, k, \quad i = j, \ldots, N$. Let $C = B^* \Sigma_1^* B^{*\prime} + \Sigma_2^*$, then the log-likelihood of the model defined above at time t is proportional to:[21]

$$F^{a,1,*}(\boldsymbol{y}_t; \boldsymbol{\beta}_1) = -\frac{1}{N} \ln\left(\det C\right) - \frac{1}{N} \boldsymbol{y}_t' C^{-1} \boldsymbol{y}_t. \tag{C.4}$$

Remember that:

$$\boldsymbol{\beta}_1 = (b_{21}^*, \ldots, b_{Nk}^*, \sigma_1^{*2}, \ldots, \sigma_{k+N}^{*2}). \tag{C.5}$$

The derivative of Equation (C.4) at time t with respect to the p-th element of of $\boldsymbol{\beta}_1$, i.e., $\beta_{1,p}$ is given by

$$\frac{\partial F^{a,1,*}(\boldsymbol{y}_t; \boldsymbol{\beta}_1)}{\partial \beta_{1,p}} = -\operatorname{tr}\left(C^{-1} \frac{\partial C}{\partial \beta_{1,p}}\right) + \operatorname{tr}\left(C^{-1} \frac{\partial C}{\partial \beta_{1,p}} C^{-1} \frac{\boldsymbol{y}_t' \boldsymbol{y}_t}{N}\right). \tag{C.6}$$

Taking the sum over the $t = 1, \ldots, T$ elements of (C.6) yields the score of the static factor model. For each p, $\dfrac{\partial C}{\partial \beta_{1,p}}$ is an $N \times N$ matrix of derivatives that are defined below, depending on what $\beta_{1,p}$ represents: loadings (*Case 1*), unconditional variances of the factors (*Case 2*) or unconditional variances of the idiosyncratic error terms (*Case 3*):

Case 1: $\beta_{1,p} = b_{ij}^*$, $\forall j = 1, \ldots, k$ and $i = j, \ldots, N$.

Denote by B_j^* the j-th column of B^* and let

$$a_j = \sigma_j^* B_j^* \tag{C.7}$$

be a vector of dimension $N \times 1$. Then $\dfrac{\partial C}{\partial b_{ij}^*}$ is an $N \times N$ matrix built as the sum of two $N \times N$ matrixes, in such a way that the first matrix has all elements zeros except for the $i-$th column which is given by the vector a_j and the second matrix

[21] Please note that the static factor in (C.1)- (C.3) differ from the static factor in (A.1)-(A.2) in terms of the identification constraints, as explained in the main text of the paper in Section 3.

has all elements zeros expect for the i-th row that is given by a_j'.

Below we provide an illustration for the matrix $\dfrac{\partial C}{\partial b_{21}^*}$:

$$\frac{\partial C}{\partial b_{21}^*} = \begin{pmatrix} 0 & a_1 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ 0 & a_3 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a_N & 0 & \dots & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ a_1 & a_2 & a_3 & \dots & a_N \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & a_1 & 0 & \dots & 0 \\ a_1 & 2a_2 & a_3 & \dots & a_N \\ 0 & a_3 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & a_N & 0 & \dots & 0 \end{pmatrix} \tag{C.8}$$

Case 2: $\beta_{1,p} = \sigma_j^{*2}$, $\forall j = 1, \dots, k$.

$$\frac{\partial C}{\partial \sigma_j^{*2}} = B^* A_j B^{*\prime}, \tag{C.9}$$

where A_j is a $k \times k$ matrix of zeros with the $j-$th diagonal element equal to 1.

Case 3: $\beta_{1,p} = \sigma_{k+i}^{*2}$, $\forall i = 1, \dots, N$.

$\dfrac{\partial C}{\partial \sigma_{k+i}^{*2}}$ is an $N \times N$ matrix of zeros, except for the i-th diagonal element, which is equal to 1.

D Outliers

Besides the results of Section 4, we would like to make the discussion about some potential outliers that may occur in the second step of the estimation, usually for small samples of observations. The occurrence of such outliers is well documented in the Monte Carlo literature in general, but also in the specific literature on using simulation-based estimation methods, such as the indirect inference, to estimate dynamic models with persistency closely related to overkurtosis or fat-tailedness: i.e., when the values of the parameters approach an edge (which is the case of SV and GARCH models). In this case, one often faces the problem that, within the estimation algorithm, one simulates data within parameter values that jump over these edges, and thus, create outliers in the simulations and in the following parameter values that are optimised. The way of dealing with such issues is quite different in the literature: sometimes the presence of outliers is simply mentioned (e.g., among others, Sahu et al. (2003), page 143), sometimes the number of outliers in the experiments is reported in detail and it is considered very small, so they can be cancelled without invalidating the experiments (e.g., among others, Calzolari et al. (2014), page 165) or sometimes the number of outliers is reported as numbers of estimated values outside the whiskers in the box-plots (e.g., among others, Sentana et al. (2008), pp.17-18).

In our case, we say we have outliers in the replications, when: (1) The estimated autoregressive parameter is at least ten times smaller than the true value, or it is negative or (2) The absolute value of the constant or of the standard deviation parameter is at least ten times larger than its true value.[22]

Tables D.1 to D.4 report the percentage of these kinds of outliers for different number of factors k, number of return series N, sample size T, choice of H and choice of starting values. From analysing the entries of the tables, we see that the value of T is crucial in generating outliers: while large values of T are accompanied by only a few outliers (for $T = 10000$, only up to 0.5% of all simulated series are discarded), working with series of smaller sample sizes is much more affected by the outliers: e.g., for $T = 1000$, large N ($N = 100$), small H (e.g., $H = 10$) and poor starting values, we encounter up to 37% of the replications to have outliers. However, increasing H (to 100) reduces the number of outliers drastically, e.g., only 12%. A further significant reduction is obtained by choosing the QML starting values: for $H = 10$ the percentage of replications with outliers decreases to 5%, while for $H = 100$, it further reduces to 3.6%. Also, for $T = 4000$, we find that increasing H and choosing the QML starting values significantly reduce the

[22]In fact, in this particular case we consider an outlier when the absolute value of the estimated constant of the factors is greater than 9, since the true value is zero.

outliers, especially for large N. Moreover, by setting $H = 10^5/T$ and using the QML starting values, regardless of the choice of T and N, we have only up 3.6% of the replications as outliers, which is much less than by keeping H fixed (e.g., to $H = 10$).

One should notice that when we consider replications with outliers, we refer to the whole set of N series replicated, regardless if the driving estimated parameter characterises the factors or the idiosyncratic errors. However, one should disentangle these effects of parameter misestimations as they affect the validity of the total number of N series per replication differently: if the misspecification regards the parameters of the factors, the number of replications affected by the outliers reduces further: i.e., for the case $T = 1000$, $N = 100$ and $k = 1$, we get only 3 outliers (in 3 out of 1000 replications) in the estimated parameters. These outliers affect all N series replicated in each of those replications, as the factors are common to all N series. So, the number of replications affected by these particular outliers decreases from 3.6% to 0.3%. However, if an outlier occurs in one of the parameters describing one of the idiosyncratic noises, then not all N series in that replication are affected by it, but only one of them. So in this case the number of series (and thus of parameters) affected by outliers is much smaller: i.e., for the case $T = 1000$, $N = 100$ and $k = 1$, only 33 replications out of 1000 exhibit such a problem, which reduces the number of series affected by it to 0.033% outlier with 100000 being the total number of noise series in the 1000 MC replications. Results showing that discarding the 3.6% or the 0.333% number of replications makes almost no difference on the MSE of the estimates can be obtained upon request from the authors. However, one should notice that these types of outliers affect all standard errors of the entire vector of parameter estimates due to the way we compute the variance-covariance matrix of the estimates as described in Section 3. So, if the focus is the standard errors, then we recommend to discard the whole set of replications affected by the outliers. This is what we do, also for simplicity, when presenting the results below.[23] However, one important take away from this discussion is that one could reduce drastically the number of outliers when choosing $H = 10^5/T$ and the QML to get starting values, which we apply in our empirical application presented in Section 5.

[23] Alternatively, one could choose not to discard the outliers and present summary results by means of quantile measures, such as the median and the interquartile ranges. These results are available from the authors upon request.

Table D.1: Percentage of outliers for $H = 10$ and starting values 20% from the true values

	$k = 1$			$k = 2$			$k = 3$		
N \ T	1000	4000	10000	1000	4000	10000	1000	4000	10000
10	3.6	0.0	0.0	5.4	0.0	0.0	6.2	0.2	0.0
20	1.7	0.5	0.0	0.4	0.0	0.0	2.7	0.1	0.0
30	8.6	3.4	0.2	3.7	0.0	0.0	3.5	0.2	0.0
100	37.4	24.2	3.7	22.2	6.3	0.1	23.6	7.9	0.1

Table D.2: Percentage of outliers for $H = 10^5/T$ and starting values 20% from the true values

	$k = 1$			$k = 2$			$k = 3$		
N \ T	1000	4000	10000	1000	4000	10000	1000	4000	10000
10	0.6	0.2	0.0	0.6	0.0	0.0	2.8	0.0	0.0
20	0.3	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.0
30	2.3	1.3	0.2	0.6	0.0	0.0	0.5	0.1	0.0
100	11.9	6.5	3.7	5.7	0.9	0.1	4.4	1.5	0.1

Table D.3: Percentage of outliers for $H = 10$ and QML starting values

	$k = 1$			$k = 2$			$k = 3$		
N \ T	1000	4000	10000	1000	4000	10000	1000	4000	10000
10	1.5	0.4	0.1	1.8	0.5	0.2	4.4	0.6	0.2
20	1.6	0.2	0.1	1.6	1.0	0.3	1.8	0.5	0.4
30	2.8	0.8	0.4	2.1	0.5	0.5	2.9	0.8	0.3
100	5.0	0.9	0.2	6.5	1.6	0.3	7.4	1.4	0.4

Table D.4: Percentage of outliers for $H = 10^5/T$ and QML starting values

	$k = 1$			$k = 2$			$k = 3$		
N \ T	1000	4000	10000	1000	4000	10000	1000	4000	10000
10	1.6	0.3	0.1	0.5	0.4	0.2	2.3	0.3	0.2
20	1.0	0.1	0.1	1.2	0.2	0.3	1.1	0.2	0.4
30	1.5	0.7	0.4	1.7	0.6	0.5	1.6	0.7	0.3
100	3.6	0.6	0.2	3.2	1.2	0.3	3.3	1.4	0.4

E Variance Covariance Estimation Results

Table E.1: Average of the Monte Carlo standard deviation (computed over the R replications) divided by the the r-th replication's asymptotic standard error. The results are displayed for the corresponding MFSV parameter vectors: $H = 10$ and starting values 20% from the true values

		$k=1$			$k=2$			$k=3$		
	Par. \ T	1000	4000	10000	1000	4000	10000	1000	4000	10000
$N=10$	θ	2.1513	0.9355	1.0725	6.0062	0.8292	0.8167	6.7391	0.6474	0.7657
	B	0.8411	0.9671	0.8662	0.5992	0.8424	0.9180	0.5296	0.7431	0.8444
	Σ_1	1.1679	1.1467	0.8212	0.7004	0.9415	0.8350	0.7249	0.9136	0.8770
	Σ_2	0.7783	0.9273	0.8876	0.6447	0.8697	0.8484	0.6965	0.7416	0.8575
	μ_ε	0.4276	0.6386	0.7234	17.4756	0.6150	0.6115	37.4403	0.4332	0.5556
	φ_ε	1.8715	0.9425	1.7522	3.9926	0.7418	0.7595	5.7238	0.5207	0.6043
	$\sigma_{\eta,\varepsilon}$	0.8722	1.1351	1.1940	1.2092	0.9503	0.8735	2.6213	0.6210	0.7763
	μ_f	0.8021	0.9935	0.7481	0.4490	0.8272	0.7240	0.6695	0.5669	0.6897
	φ_f	63.8039	1.2998	0.9270	71.8021	1.1751	0.8414	9.0744	0.6659	0.8207
	$\sigma_{\eta,f}$	1.1783	1.0011	0.9764	0.5446	0.9611	0.8761	1.0777	0.5893	0.9418
$N=20$	θ	0.9898	1.2926	0.8921	0.8708	0.9068	0.8759	0.8328	0.9721	0.8665
	B	0.7944	0.7716	0.8758	0.6684	0.8782	0.9276	0.5338	0.8859	0.9303
	Σ_1	1.2993	0.2066	0.8984	0.7688	0.6848	0.9343	0.5601	0.8020	0.9196
	Σ_2	0.8773	0.9862	0.9335	0.8240	0.9853	0.9207	0.8416	0.9625	0.9059
	μ_ε	0.5179	0.6595	0.6581	0.4337	0.6590	0.6513	0.4903	0.6264	0.6406
	φ_ε	1.7341	1.5392	0.9244	1.6358	1.0486	0.8832	1.9146	1.5984	0.8284
	$\sigma_{\eta,\varepsilon}$	0.9018	1.0721	1.0237	0.7009	0.9560	0.9280	0.7196	0.9569	0.9209
	μ_f	0.7735	0.1955	0.8533	0.4367	0.5416	0.8312	0.3410	0.5242	0.7556
	φ_f	3.4364	32.0932	1.0097	4.1697	1.7602	0.9840	3.0591	1.6175	0.8635
	$\sigma_{\eta,f}$	0.7293	0.8412	1.6873	0.7371	0.9552	1.0052	0.5218	0.7907	0.9101
$N=30$	θ	1.0733	1.5028	1.1615	0.8532	0.9762	0.9067	0.7850	1.1055	0.9117
	B	0.7291	1.0640	1.0403	0.6245	1.0343	1.0036	0.5154	0.9524	0.9848
	Σ_1	0.6169	1.1585	1.0178	0.7487	1.0555	0.9934	0.6689	0.9748	0.9203
	Σ_2	0.8462	0.9898	0.9555	0.8122	0.9389	0.9524	0.7820	0.9315	0.9417
	μ_ε	0.4888	0.6816	0.6595	0.4253	0.6465	0.6508	0.4164	0.6268	0.6412
	φ_ε	2.1191	3.8620	2.0651	1.8637	1.1975	0.8855	1.7185	2.3963	0.8727
	$\sigma_{\eta,\varepsilon}$	0.7154	0.9475	1.0994	0.6988	0.9689	0.9358	0.6680	0.9611	0.9273
	μ_f	0.4555	0.8382	0.9120	0.3843	0.8787	0.8361	0.4380	0.7907	0.7670
	φ_f	16.4430	1.3390	1.1627	2.5636	1.1908	1.0018	4.3175	0.9574	1.1226
	$\sigma_{\eta,f}$	0.4672	1.3060	1.0627	0.4301	1.4221	1.0749	0.8154	1.0456	1.4312
$N=100$	θ	0.8371	1.6526	1.1284	1.1881	1.2442	0.9214	1.3235	1.2429	0.9525
	B	0.6093	0.8841	0.8919	0.7930	0.9628	0.9692	0.7633	0.9711	1.0016
	Σ_1	0.4278	0.8557	0.8723	0.7981	1.0030	0.9592	0.8444	0.9759	0.9808
	Σ_2	0.8419	0.9565	0.9341	0.9191	0.9671	0.9446	0.9397	0.9691	0.9427
	μ_ε	0.4294	0.6649	0.6652	0.5261	0.6746	0.6533	0.5616	0.6887	0.6529
	φ_ε	1.6547	4.7301	2.0985	3.1844	2.8430	1.0058	3.9651	2.9804	1.0800
	$\sigma_{\eta,\varepsilon}$	0.6260	1.0548	1.0572	0.8368	1.0762	0.9844	0.8812	1.1746	0.9886
	μ_f	0.3095	0.7165	0.8178	0.4773	0.7771	0.8247	0.5892	0.7737	0.7911
	φ_f	4.0869	0.7122	1.1902	5.2379	0.8257	1.1485	22.4435	0.9027	1.0947
	$\sigma_{\eta,f}$	0.7388	0.8332	0.8909	1.4696	0.8811	0.9022	1.1960	0.9215	1.0058

Note: The ratios for B, Σ_1, Σ_2 are computed by stacking the elements of matrix B and the diagonal of matrices Σ_1 and Σ_2 respectively, in a vector. $\mu_f \equiv (\mu_1, \mu_2, \mu_3)'$, $\mu_\varepsilon \equiv (\mu_4, \ldots, \mu_{3+N})'$, $\varphi_f \equiv (\varphi_1, \varphi_2, \varphi_3)'$, $\varphi_\varepsilon \equiv (\varphi_4, \ldots, \varphi_{3+N})'$, $\sigma_{\eta,f} \equiv (\sigma_{\eta,1}, \sigma_{\eta,2}, \sigma_{\eta,3})'$ and $\sigma_{\eta,\varepsilon} \equiv (\sigma_{\eta,4}, \ldots, \sigma_{\eta,3+N})'$, if $k = 3$.

Table E.2: Average of the Monte Carlo standard deviation (computed over the R replications) divided by the the r-th replication's asymptotic standard error. The results are displayed for the corresponding MFSV parameter vectors: $H = 10^5/T$ and starting values 20% from the true values

		$k=1$			$k=2$			$k=3$		
	Par. \ T	1000	4000	10000	1000	4000	10000	1000	4000	10000
$N=10$	θ	0.9337	1.1374	1.0725	6.0980	0.8433	0.8167	4.1943	0.6642	0.7657
	B	1.0851	0.9975	0.8662	0.6282	0.8525	0.9180	0.6641	0.7653	0.8444
	Σ_1	1.6219	1.1696	0.8212	0.7705	0.9761	0.8350	0.7850	0.9319	0.8770
	Σ_2	0.9692	0.9705	0.8876	0.7305	0.9107	0.8484	0.8841	0.7766	0.8575
	μ_ε	0.5374	0.6847	0.7234	35.7034	0.6207	0.6115	26.1315	0.4604	0.5556
	φ_ε	1.1498	1.8763	1.7522	0.5726	0.7593	0.7595	0.7236	0.5066	0.6043
	$\sigma_{\eta,\varepsilon}$	0.7938	1.1495	1.1940	1.0651	0.9459	0.8735	1.7316	0.6529	0.7763
	μ_f	1.1110	1.0176	0.7481	0.4889	0.8586	0.7240	0.4810	0.5572	0.6897
	φ_f	1.3258	1.3266	0.9270	0.7307	1.1965	0.8414	0.7970	0.6593	0.8207
	$\sigma_{\eta,f}$	1.1586	0.9803	0.9764	0.4964	0.9461	0.8761	0.6433	0.5683	0.9418
$N=20$	θ	0.8592	0.8713	0.8921	0.6955	0.9024	0.8759	0.6056	0.8952	0.8665
	B	0.9167	0.8339	0.8758	0.7100	0.9092	0.9276	0.5803	0.9427	0.9303
	Σ_1	1.3770	0.7840	0.8984	0.9155	0.9133	0.9343	0.6522	0.9560	0.9196
	Σ_2	0.9949	1.0298	0.9335	0.9013	1.0147	0.9207	0.8893	0.9936	0.9059
	μ_ε	0.5807	0.6490	0.6581	0.4478	0.6759	0.6513	0.4350	0.6627	0.6406
	φ_ε	0.8624	0.9102	0.9244	0.5987	0.9493	0.8832	0.5852	0.8775	0.8284
	$\sigma_{\eta,\varepsilon}$	0.9127	0.9444	1.0237	0.6600	0.9624	0.9280	0.6372	0.9613	0.9209
	μ_f	0.8104	0.7312	0.8533	0.5005	0.7722	0.8312	0.3367	0.7151	0.7556
	φ_f	1.0972	0.7096	1.0097	2.0587	0.9021	0.9840	0.5393	0.7732	0.8635
	$\sigma_{\eta,f}$	0.7816	1.0029	1.6873	0.7783	0.9667	1.0052	0.5198	0.8526	0.9101
$N=30$	θ	0.8791	1.2928	1.1615	0.8225	0.9370	0.9067	0.7596	0.9571	0.9117
	B	0.9378	1.1572	1.0403	0.8769	1.0692	1.0036	0.7172	0.9818	0.9848
	Σ_1	1.4443	1.2193	1.0178	1.2526	1.0857	0.9934	0.9959	1.0102	0.9203
	Σ_2	0.9328	0.9745	0.9555	0.9470	0.9559	0.9524	0.8897	0.9539	0.9417
	μ_ε	0.5036	0.6620	0.6595	0.5048	0.6424	0.6508	0.4704	0.6290	0.6412
	φ_ε	1.2310	2.5817	2.0651	0.9262	0.9109	0.8855	1.0612	1.2409	0.8727
	$\sigma_{\eta,\varepsilon}$	0.7968	1.1044	1.0994	0.8251	0.9519	0.9358	0.7342	0.9304	0.9273
	μ_f	0.7171	0.9608	0.9120	0.5136	0.8961	0.8361	0.5464	0.8073	0.7670
	φ_f	0.5715	1.2358	1.1627	0.5258	1.1065	1.0018	0.7208	0.9859	1.1226
	$\sigma_{\eta,f}$	0.6479	1.1423	1.0627	0.7529	1.1940	1.0749	0.7908	1.0753	1.4312
$N=100$	θ	0.9915	1.2097	1.1284	0.9936	0.9962	0.9214	0.9314	0.9861	0.9525
	B	0.7452	0.9491	0.8919	0.9081	1.0035	0.9692	0.8597	1.0035	1.0016
	Σ_1	0.9568	1.0658	0.8723	1.1263	1.0700	0.9592	0.9569	1.0239	0.9808
	Σ_2	1.0272	1.0008	0.9341	1.0449	0.9961	0.9446	1.0571	1.0027	0.9427
	μ_ε	0.5693	0.6730	0.6652	0.6054	0.6843	0.6533	0.6046	0.6518	0.6529
	φ_ε	1.7368	2.3539	2.0985	1.5219	1.2361	1.0058	1.3120	1.2612	1.0800
	$\sigma_{\eta,\varepsilon}$	0.8739	1.0674	1.0572	0.9446	1.0543	0.9844	0.9524	0.9812	0.9886
	μ_f	0.7736	0.9001	0.8178	0.8581	0.8386	0.8247	0.7273	0.8188	0.7911
	φ_f	1.6245	2.1814	1.1902	1.6506	1.1747	1.1485	1.3738	1.0645	1.0947
	$\sigma_{\eta,f}$	0.8655	0.8729	0.8909	1.6505	0.8933	0.9022	1.0031	0.9125	1.0058

Note: The ratios for B, Σ_1, Σ_2 are computed by stacking the elements of matrix B and the diagonal of matrices Σ_1 and Σ_2 respectively, in a vector. $\mu_f \equiv (\mu_1, \mu_2, \mu_3)'$, $\mu_\varepsilon \equiv (\mu_4, \dots, \mu_{3+N})'$, $\varphi_f \equiv (\varphi_1, \varphi_2, \varphi_3)'$, $\varphi_\varepsilon \equiv (\varphi_4, \dots, \varphi_{3+N})'$, $\sigma_{\eta,f} \equiv (\sigma_{\eta,1}, \sigma_{\eta,2}, \sigma_{\eta,3})'$ and $\sigma_{\eta,\varepsilon} \equiv (\sigma_{\eta,4}, \dots, \sigma_{\eta,3+N})'$, if $k = 3$.

Table E.3: Average of the Monte Carlo standard deviation (computed over the R replications) divided by the the r-th replication's asymptotic standard error. The results are displayed for the corresponding MFSV parameter vectors: $H = 10$ and QML starting values

		$k = 1$			$k = 2$			$k = 3$		
	Par. \ T	1000	4000	10000	1000	4000	10000	1000	4000	10000
$N = 10$	θ	1.3749	0.9471	0.8590	1.4831	1.1612	0.8820	3.1548	1.6978	1.3908
	B	1.0291	0.9707	0.8728	0.7218	0.8887	0.9224	0.6858	0.7882	0.8592
	Σ_1	1.8654	1.1609	0.8286	1.0096	1.0035	0.8737	0.8788	1.0421	0.9514
	Σ_2	0.9109	0.9488	0.8870	0.8038	0.9642	0.9118	1.0332	0.9338	0.9420
	μ_ε	0.8265	0.7292	0.6196	2.5623	1.0882	0.6290	13.6900	1.2311	0.8794
	φ_ε	0.9965	0.7545	0.8532	2.0526	1.6976	0.7739	3.4037	6.7315	4.1425
	$\sigma_{\eta,\varepsilon}$	1.0842	0.9093	0.8694	1.2752	1.0557	0.9067	1.6690	1.0430	1.2130
	μ_f	5.1516	4.5366	2.6206	1.4209	2.8884	2.0511	0.8616	0.8751	1.5918
	φ_f	17.2785	1.3311	0.9471	5.4728	1.2421	0.8831	5.7605	0.9408	0.9195
	$\sigma_{\eta,f}$	1.1335	1.0135	0.9835	0.6931	1.0217	0.9081	0.9488	0.7162	0.9752
$N = 20$	θ	1.9195	0.9162	0.9245	1.0828	1.0976	0.9556	0.9350	1.0661	1.0378
	B	0.8952	0.8270	0.8790	0.7200	0.8873	0.9262	0.6355	0.9161	0.9299
	Σ_1	1.5557	0.8831	0.9074	1.0276	0.9530	0.9453	0.7803	0.9987	0.9529
	Σ_2	0.9463	1.0090	0.9422	0.8691	1.0022	0.9388	0.9498	1.0099	0.9347
	μ_ε	3.9256	0.6247	0.9061	1.7596	1.1644	1.0377	0.8158	1.0169	0.8186
	φ_ε	0.8747	0.8867	0.9191	0.9855	0.9343	0.8862	1.9313	1.6393	1.1914
	$\sigma_{\eta,\varepsilon}$	2.7269	0.8765	0.9775	1.1589	0.9483	0.9354	0.7472	1.0599	1.3556
	μ_f	1.0383	1.8380	0.8506	0.9927	1.5864	1.7124	0.5357	1.2865	2.3548
	φ_f	7.8915	7.2637	0.8258	3.8337	8.3027	0.9824	2.5686	0.8041	0.8734
	$\sigma_{\eta,f}$	0.7395	0.7383	1.0352	0.7676	0.8531	0.9716	0.5541	0.8006	0.9198
$N = 30$	θ	1.9572	1.1509	1.1235	0.9811	0.9592	1.0535	0.9154	1.0835	1.0699
	B	0.8718	1.1213	1.0418	0.8139	1.0385	1.0031	0.6179	0.9547	0.9832
	Σ_1	1.2207	1.2409	1.0143	1.1167	1.0727	0.9949	0.8900	0.9997	0.9445
	Σ_2	0.8927	0.9423	0.9572	0.9058	0.9380	0.9619	0.8393	0.9454	0.9587
	μ_ε	4.5632	1.0887	1.1293	1.3525	0.9237	1.4885	1.5625	1.1028	1.5190
	φ_ε	0.9471	1.0455	0.9122	0.9390	0.8829	0.8867	0.9784	0.9431	0.8884
	$\sigma_{\eta,\varepsilon}$	2.2872	1.2637	0.9543	0.9813	0.9158	0.9371	0.9891	0.9338	1.2188
	μ_f	0.7389	5.5825	2.7688	0.5249	1.0431	1.6736	0.5848	2.3266	0.7815
	φ_f	10.8654	5.3555	5.8856	2.6040	1.0434	1.0050	2.6545	3.4451	0.9684
	$\sigma_{\eta,f}$	0.6400	1.1795	13.4307	0.6278	1.0692	1.0738	0.7812	5.2584	0.9565
$N = 100$	θ	1.5642	1.0718	0.9442	1.5514	1.3289	0.9539	1.6064	1.0137	1.1217
	B	0.7070	0.9107	0.9059	0.8479	0.9676	0.9696	0.7975	0.9729	1.0021
	Σ_1	0.8993	0.8881	0.8739	1.0984	1.0070	0.9588	0.9577	0.9774	0.9825
	Σ_2	0.9936	0.9669	0.9432	0.9769	0.9662	0.9458	0.9901	0.9711	0.9442
	μ_ε	2.8160	0.9816	0.7609	3.8957	2.1873	0.7686	4.2378	1.2081	1.4513
	φ_ε	1.0824	1.0984	1.0061	1.0540	1.1635	1.0213	1.1157	1.0474	1.0306
	$\sigma_{\eta,\varepsilon}$	2.1406	1.3884	1.1045	1.5581	1.7456	1.0486	2.3688	0.9624	1.4306
	μ_f	0.7778	2.8918	0.8252	0.8280	0.7968	0.8245	0.7543	0.7833	0.7922
	φ_f	10.9135	0.7490	1.1927	8.3497	0.8468	1.1525	6.6584	0.9046	1.0960
	$\sigma_{\eta,f}$	0.9792	0.8886	0.8663	1.2750	0.8953	0.9046	1.0966	0.9134	1.0055

Note: The ratios for B, Σ_1, Σ_2 are computed by stacking the elements of matrix B and the diagonal of matrices Σ_1 and Σ_2 respectively, in a vector. $\mu_f \equiv (\mu_1, \mu_2, \mu_3)'$, $\mu_\varepsilon \equiv (\mu_4, \ldots, \mu_{3+N})'$, $\varphi_f \equiv (\varphi_1, \varphi_2, \varphi_3)'$, $\varphi_\varepsilon \equiv (\varphi_4, \ldots, \varphi_{3+N})'$, $\sigma_{\eta,f} \equiv (\sigma_{\eta,1}, \sigma_{\eta,2}, \sigma_{\eta,3})'$ and $\sigma_{\eta,\varepsilon} \equiv (\sigma_{\eta,4}, \ldots, \sigma_{\eta,3+N})'$, if $k = 3$.

Table E.4: Average of the Monte Carlo standard deviation (computed over the R replications) divided by the the r-th replication's asymptotic standard error. The results are displayed for the corresponding MFSV parameter vectors: $H = 10^5/T$ and QML starting values

		$k=1$			$k=2$			$k=3$		
	Par. \ T	1000	4000	10000	1000	4000	10000	1000	4000	10000
	θ	1.2578	0.9388	0.8590	0.7725	1.0282	0.8820	0.9781	1.7295	1.3908
$N=10$	B	1.1281	0.9993	0.8728	0.7271	0.8986	0.9224	0.6799	0.8612	0.8592
	Σ_1	1.6586	1.1736	0.8286	0.8765	1.0314	0.8737	0.8218	1.0658	0.9514
	Σ_2	1.0139	0.9831	0.8870	0.8569	0.9921	0.9118	1.0883	0.9696	0.9420
	μ_ε	0.9711	0.7480	0.6196	0.4272	1.0827	0.6290	0.5496	1.2389	0.8794
	φ_ε	0.7643	0.7729	0.8532	0.8589	0.9460	0.7739	1.8176	6.5759	4.1425
	$\sigma_{\eta,\varepsilon}$	0.9090	0.9261	0.8694	0.5876	1.0499	0.9067	0.6477	1.0433	1.2130
	μ_f	1.1461	2.9471	2.6206	0.5662	2.1374	2.0511	0.6408	1.3444	1.5918
	φ_f	15.8503	1.3537	0.9471	3.2585	1.2590	0.8831	3.4412	1.0027	0.9195
	$\sigma_{\eta,f}$	1.2724	0.9856	0.9835	0.5719	0.9983	0.9081	0.7569	0.7518	0.9752
	θ	1.3023	0.8816	0.9245	1.2495	0.9862	0.9556	0.7424	1.0306	1.0378
$N=20$	B	0.9348	0.8370	0.8790	0.7362	0.9056	0.9262	0.6274	0.9448	0.9299
	Σ_1	1.4510	0.7965	0.9074	0.9448	0.9257	0.9453	0.7233	1.0000	0.9529
	Σ_2	1.0072	1.0395	0.9422	0.9302	1.0331	0.9388	0.9673	1.0330	0.9347
	μ_ε	2.0487	0.7625	0.9061	2.9314	1.1221	1.0377	0.5626	0.9932	0.8186
	φ_ε	0.8698	0.9039	0.9191	0.7266	0.9496	0.8862	1.0243	1.3240	1.1914
	$\sigma_{\eta,\varepsilon}$	1.5498	0.8971	0.9775	1.3605	0.9644	0.9354	0.7423	1.0471	1.3556
	μ_f	0.8555	0.7307	0.8506	0.8642	1.5096	1.7124	0.4057	0.9062	2.3548
	φ_f	3.8362	0.5969	0.8258	2.4100	0.8752	0.9824	1.1914	1.0544	0.8734
	$\sigma_{\eta,f}$	0.7254	0.7186	1.0352	0.7701	0.8841	0.9716	0.5414	0.8711	0.9198
	θ	1.1265	1.3129	1.1235	1.3024	0.9743	1.0535	0.8533	1.0411	1.0699
$N=30$	B	0.9415	1.1673	1.0418	0.8826	1.0706	1.0031	0.7690	0.9823	0.9832
	Σ_1	1.4658	1.2213	1.0143	1.2581	1.0915	0.9949	1.0562	1.0301	0.9445
	Σ_2	0.9618	0.9727	0.9572	0.9675	0.9668	0.9619	0.9342	0.9750	0.9587
	μ_ε	1.6630	1.1965	1.1293	2.6515	0.9045	1.4885	0.9345	1.2504	1.5190
	φ_ε	0.8332	1.1391	0.9122	0.9650	0.8948	0.8867	0.9562	0.9935	0.8884
	$\sigma_{\eta,\varepsilon}$	1.1342	1.7808	0.9543	1.4877	0.9288	0.9371	0.7863	1.0720	1.2188
	μ_f	1.2922	3.6338	2.7688	0.7971	0.9003	1.6736	0.5783	1.5822	0.7815
	φ_f	3.8863	2.9184	5.8856	1.4282	1.0892	1.0050	1.3498	0.9580	0.9684
	$\sigma_{\eta,f}$	0.6450	6.5806	13.4307	0.7565	1.1095	1.0738	0.8133	0.9771	0.9565
	θ	2.2432	1.0013	0.9442	1.2580	1.4200	0.9539	1.1634	1.5041	1.1217
$N=100$	B	0.7609	0.9563	0.9059	0.9064	1.0065	0.9696	0.8597	1.0073	1.0021
	Σ_1	0.9820	1.0510	0.8739	1.1215	1.0656	0.9588	0.9560	1.0266	0.9825
	Σ_2	1.0565	1.0013	0.9432	1.0473	0.9979	0.9458	1.0589	1.0048	0.9442
	μ_ε	5.4528	0.8827	0.7609	2.0884	2.5818	0.7686	1.9302	3.3306	1.4513
	φ_ε	1.1516	1.0336	1.0061	1.0910	1.0682	1.0213	1.0494	1.1542	1.0306
	$\sigma_{\eta,\varepsilon}$	2.7661	1.1128	1.1045	1.4284	1.8831	1.0486	1.4742	2.0580	1.4306
	μ_f	0.7974	3.3959	0.8252	0.8516	0.8347	0.8245	0.7320	0.8192	0.7922
	φ_f	7.6621	0.6616	1.1927	4.8341	1.0982	1.1525	3.0783	0.9136	1.0960
	$\sigma_{\eta,f}$	0.8778	0.8211	0.8663	1.6922	0.8693	0.9046	1.0241	0.8989	1.0055

Note: The ratios for B, Σ_1, Σ_2 are computed by stacking the elements of matrix B and the diagonal of matrices Σ_1 and Σ_2 respectively, in a vector. $\mu_f \equiv (\mu_1, \mu_2, \mu_3)'$, $\mu_\varepsilon \equiv (\mu_4, \dots, \mu_{3+N})'$, $\varphi_f \equiv (\varphi_1, \varphi_2, \varphi_3)'$, $\varphi_\varepsilon \equiv (\varphi_4, \dots, \varphi_{3+N})'$, $\sigma_{\eta,f} \equiv (\sigma_{\eta,1}, \sigma_{\eta,2}, \sigma_{\eta,3})'$ and $\sigma_{\eta,\varepsilon} \equiv (\sigma_{\eta,4}, \dots, \sigma_{\eta,3+N})'$, if $k = 3$.

F MSE Results

Table F.1: Average MSE of the MFSV parameter vectors: $H = 10$ and starting values 20% from the true values

	T	$k=1$			$k=2$			$k=3$		
	Par.	1000	4000	10000	1000	4000	10000	1000	4000	10000
$N=10$	θ	0.0447	0.0107	0.0054	0.0985	0.0146	0.0075	0.2523	0.0357	0.0157
	B	0.0005	0.0001	0.0000	0.0009	0.0002	0.0001	0.0018	0.0004	0.0001
	Σ_1	0.1559	0.0505	0.0194	0.0919	0.0295	0.0116	0.0688	0.0231	0.0087
	Σ_2	0.0209	0.0057	0.0024	0.0204	0.0057	0.0025	0.0217	0.0059	0.0025
	μ_ε	0.1674	0.0389	0.0215	0.5480	0.0735	0.0400	1.7592	0.2413	0.1070
	φ_ε	0.0010	0.0002	0.0002	0.0011	0.0002	0.0001	0.0023	0.0002	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0196	0.0038	0.0017	0.0366	0.0052	0.0020	0.0873	0.0106	0.0045
	μ_f	0.1005	0.0269	0.0109	0.0606	0.0176	0.0071	0.0585	0.0153	0.0068
	φ_f	0.0104	0.0000	0.0000	0.0050	0.0001	0.0000	0.0030	0.0002	0.0001
	$\sigma_{\eta,f}$	0.0063	0.0002	0.0001	0.0068	0.0009	0.0003	0.0112	0.0019	0.0009
$N=20$	θ	0.0432	0.0079	0.0029	0.0396	0.0071	0.0027	0.0549	0.0073	0.0036
	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0010	0.0003	0.0001
	Σ_1	0.1622	0.0436	0.0193	0.0970	0.0272	0.0113	0.0780	0.0203	0.0086
	Σ_2	0.0201	0.0053	0.0021	0.0201	0.0054	0.0021	0.0199	0.0053	0.0022
	μ_ε	0.1708	0.0275	0.0103	0.1898	0.0296	0.0116	0.3201	0.0368	0.0198
	φ_ε	0.0008	0.0002	0.0001	0.0008	0.0002	0.0001	0.0014	0.0002	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0164	0.0029	0.0011	0.0182	0.0032	0.0011	0.0339	0.0036	0.0017
	μ_f	0.1030	0.0419	0.0111	0.0642	0.0251	0.0066	0.0557	0.0197	0.0052
	φ_f	0.0043	0.0001	0.0000	0.0019	0.0001	0.0000	0.0044	0.0002	0.0001
	$\sigma_{\eta,f}$	0.0035	0.0078	0.0002	0.0075	0.0039	0.0003	0.0121	0.0037	0.0005
$N=30$	θ	0.0326	0.0079	0.0028	0.0258	0.0060	0.0024	0.0329	0.0059	0.0024
	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0010	0.0002	0.0001
	Σ_1	0.1664	0.0462	0.0194	0.0993	0.0265	0.0117	0.0783	0.0208	0.0086
	Σ_2	0.0186	0.0050	0.0020	0.0186	0.0050	0.0020	0.0187	0.0050	0.0021
	μ_ε	0.1210	0.0287	0.0101	0.1128	0.0257	0.0102	0.1833	0.0300	0.0120
	φ_ε	0.0016	0.0006	0.0002	0.0012	0.0002	0.0001	0.0012	0.0002	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0144	0.0033	0.0011	0.0136	0.0029	0.0011	0.0166	0.0033	0.0012
	μ_f	0.1284	0.0259	0.0118	0.0729	0.0156	0.0070	0.0546	0.0127	0.0055
	φ_f	0.0011	0.0000	0.0000	0.0010	0.0001	0.0000	0.0024	0.0002	0.0001
	$\sigma_{\eta,f}$	0.0059	0.0006	0.0001	0.0053	0.0008	0.0003	0.0089	0.0013	0.0005
$N=100$	θ	0.0308	0.0072	0.0027	0.0246	0.0056	0.0021	0.0213	0.0048	0.0018
	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0008	0.0002	0.0001
	Σ_1	0.1588	0.0460	0.0190	0.1013	0.0289	0.0116	0.0761	0.0217	0.0086
	Σ_2	0.0183	0.0048	0.0019	0.0180	0.0048	0.0019	0.0182	0.0048	0.0019
	μ_ε	0.1180	0.0261	0.0100	0.1116	0.0242	0.0090	0.1114	0.0239	0.0091
	φ_ε	0.0018	0.0015	0.0002	0.0010	0.0007	0.0001	0.0010	0.0006	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0137	0.0031	0.0011	0.0132	0.0027	0.0010	0.0133	0.0027	0.0010
	μ_f	0.1051	0.0313	0.0110	0.0629	0.0188	0.0067	0.0488	0.0143	0.0051
	φ_f	0.0004	0.0000	0.0000	0.0013	0.0001	0.0000	0.0026	0.0003	0.0001
	$\sigma_{\eta,f}$	0.0070	0.0002	0.0001	0.0074	0.0007	0.0002	0.0104	0.0015	0.0004

Note: The MSEs for B, Σ_1, Σ_2 are computed by stacking the elements of matrix B and the diagonal of matrices Σ_1 and Σ_2 respectively, in a vector. $\mu_f \equiv (\mu_1, \mu_2, \mu_3)'$, $\mu_\varepsilon \equiv (\mu_4, \ldots, \mu_{3+N})'$, $\varphi_f \equiv (\varphi_1, \varphi_2, \varphi_3)'$, $\varphi_\varepsilon \equiv (\varphi_4, \ldots, \varphi_{3+N})'$, $\sigma_{\eta,f} \equiv (\sigma_{\eta,1}, \sigma_{\eta,2}, \sigma_{\eta,3})'$ and $\sigma_{\eta,\varepsilon} \equiv (\sigma_{\eta,4}, \ldots, \sigma_{\eta,3+N})'$ if $k = 3$.

Table F.2: Average MSE of the MFSV parameter vectors: $H = 10^5/T$ and starting values 20% from the true values

	Par. \ T	$k=1$			$k=2$			$k=3$		
		1000	4000	10000	1000	4000	10000	1000	4000	10000
	θ	0.0372	0.0108	0.0054	0.0438	0.0131	0.0075	0.1778	0.0310	0.0157
$N=10$	B	0.0005	0.0001	0.0000	0.0009	0.0002	0.0001	0.0018	0.0004	0.0001
	Σ_1	0.1588	0.0505	0.0194	0.0943	0.0295	0.0116	0.0716	0.0231	0.0087
	Σ_2	0.0210	0.0057	0.0024	0.0209	0.0057	0.0025	0.0218	0.0059	0.0025
	μ_ε	0.1332	0.0397	0.0215	0.2097	0.0652	0.0400	1.2268	0.2080	0.1070
	φ_ε	0.0010	0.0003	0.0002	0.0011	0.0002	0.0001	0.0020	0.0002	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0154	0.0034	0.0017	0.0189	0.0042	0.0020	0.0557	0.0086	0.0045
	μ_f	0.1019	0.0272	0.0109	0.0639	0.0174	0.0071	0.0532	0.0157	0.0068
	φ_f	0.0004	0.0000	0.0000	0.0010	0.0001	0.0000	0.0018	0.0002	0.0001
	$\sigma_{\eta,f}$	0.0015	0.0002	0.0001	0.0044	0.0007	0.0003	0.0091	0.0019	0.0009
	θ	0.0282	0.0071	0.0029	0.0266	0.0064	0.0027	0.0361	0.0073	0.0036
$N=20$	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0010	0.0003	0.0001
	Σ_1	0.1631	0.0437	0.0193	0.0970	0.0272	0.0113	0.0781	0.0203	0.0086
	Σ_2	0.0200	0.0053	0.0021	0.0201	0.0054	0.0021	0.0201	0.0054	0.0022
	μ_ε	0.0990	0.0250	0.0103	0.1136	0.0268	0.0116	0.1971	0.0385	0.0198
	φ_ε	0.0008	0.0001	0.0001	0.0009	0.0001	0.0001	0.0011	0.0002	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0119	0.0026	0.0011	0.0141	0.0027	0.0011	0.0213	0.0036	0.0017
	μ_f	0.0978	0.0256	0.0111	0.0605	0.0162	0.0066	0.0518	0.0124	0.0052
	φ_f	0.0004	0.0000	0.0000	0.0014	0.0001	0.0000	0.0021	0.0002	0.0001
	$\sigma_{\eta,f}$	0.0019	0.0004	0.0002	0.0052	0.0007	0.0003	0.0093	0.0012	0.0005
	θ	0.0260	0.0070	0.0028	0.0222	0.0057	0.0024	0.0223	0.0055	0.0024
$N=30$	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0010	0.0002	0.0001
	Σ_1	0.1707	0.0458	0.0194	0.1021	0.0265	0.0117	0.0779	0.0208	0.0086
	Σ_2	0.0189	0.0050	0.0020	0.0188	0.0050	0.0020	0.0189	0.0050	0.0021
	μ_ε	0.0917	0.0249	0.0101	0.0935	0.0243	0.0102	0.1108	0.0276	0.0120
	φ_ε	0.0010	0.0004	0.0002	0.0008	0.0001	0.0001	0.0010	0.0002	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0116	0.0028	0.0011	0.0116	0.0026	0.0011	0.0141	0.0029	0.0012
	μ_f	0.0978	0.0246	0.0118	0.0595	0.0153	0.0070	0.0480	0.0123	0.0055
	φ_f	0.0002	0.0000	0.0000	0.0009	0.0001	0.0000	0.0017	0.0002	0.0001
	$\sigma_{\eta,f}$	0.0013	0.0002	0.0001	0.0033	0.0007	0.0003	0.0063	0.0011	0.0005
	θ	0.0235	0.0063	0.0027	0.0195	0.0051	0.0021	0.0176	0.0045	0.0018
$N=100$	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0008	0.0002	0.0001
	Σ_1	0.1592	0.0476	0.0190	0.1001	0.0287	0.0116	0.0773	0.0218	0.0086
	Σ_2	0.0181	0.0048	0.0019	0.0181	0.0048	0.0019	0.0181	0.0048	0.0019
	μ_ε	0.0858	0.0232	0.0100	0.0841	0.0221	0.0090	0.0877	0.0224	0.0091
	φ_ε	0.0008	0.0003	0.0002	0.0009	0.0002	0.0001	0.0009	0.0002	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0104	0.0025	0.0011	0.0101	0.0023	0.0010	0.0109	0.0024	0.0010
	μ_f	0.0910	0.0284	0.0110	0.0585	0.0169	0.0067	0.0496	0.0131	0.0051
	φ_f	0.0007	0.0001	0.0000	0.0009	0.0001	0.0000	0.0017	0.0002	0.0001
	$\sigma_{\eta,f}$	0.0013	0.0002	0.0001	0.0026	0.0005	0.0002	0.0061	0.0010	0.0004

Note: The MSEs for B, Σ_1, Σ_2 are computed by stacking the elements of matrix B and the diagonal of matrices Σ_1 and Σ_2 respectively, in a vector. $\mu_f \equiv (\mu_1, \mu_2, \mu_3)'$, $\mu_\varepsilon \equiv (\mu_4, \ldots, \mu_{3+N})'$, $\varphi_f \equiv (\varphi_1, \varphi_2, \varphi_3)'$, $\varphi_\varepsilon \equiv (\varphi_4, \ldots, \varphi_{3+N})'$, $\sigma_{\eta,f} \equiv (\sigma_{\eta,1}, \sigma_{\eta,2}, \sigma_{\eta,3})'$ and $\sigma_{\eta,\varepsilon} \equiv (\sigma_{\eta,4}, \ldots, \sigma_{\eta,3+N})'$ if $k = 3$.

Table F.3: Average MSE of the MFSV parameter vectors: $H = 10$ and QML starting values

		$k=1$			$k=2$			$k=3$		
	Par. \ T	1000	4000	10000	1000	4000	10000	1000	4000	10000
$N=10$	θ	0.0448	0.0087	0.0034	0.0879	0.0153	0.0039	0.1154	0.0219	0.0184
	B	0.0005	0.0001	0.0000	0.0009	0.0002	0.0001	0.0018	0.0004	0.0001
	Σ_1	0.1589	0.0505	0.0194	0.0942	0.0296	0.0116	0.0727	0.0228	0.0087
	Σ_2	0.0210	0.0057	0.0024	0.0209	0.0057	0.0025	0.0221	0.0059	0.0025
	μ_ε	0.1617	0.0289	0.0115	0.4576	0.0762	0.0171	0.7123	0.1158	0.1155
	φ_ε	0.0011	0.0002	0.0001	0.0018	0.0003	0.0001	0.0045	0.0012	0.0007
	$\sigma_{\eta,\varepsilon}$	0.0207	0.0033	0.0012	0.0393	0.0066	0.0022	0.0670	0.0284	0.0158
	μ_f	0.1598	0.0265	0.0106	0.1472	0.0209	0.0066	0.1322	0.0214	0.0059
	φ_f	0.0008	0.0000	0.0000	0.0030	0.0001	0.0000	0.0047	0.0005	0.0002
	$\sigma_{\eta,f}$	0.0035	0.0002	0.0001	0.0069	0.0008	0.0003	0.0144	0.0020	0.0010
$N=20$	θ	0.0354	0.0071	0.0028	0.0351	0.0085	0.0031	0.0504	0.0102	0.0039
	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0010	0.0003	0.0001
	Σ_1	0.1625	0.0436	0.0193	0.0971	0.0272	0.0114	0.0773	0.0203	0.0086
	Σ_2	0.0200	0.0053	0.0021	0.0202	0.0054	0.0021	0.0200	0.0054	0.0022
	μ_ε	0.1318	0.0248	0.0096	0.1625	0.0288	0.0101	0.2803	0.0500	0.0212
	φ_ε	0.0009	0.0001	0.0001	0.0009	0.0002	0.0001	0.0016	0.0003	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0151	0.0027	0.0010	0.0170	0.0032	0.0011	0.0353	0.0076	0.0028
	μ_f	0.1156	0.0264	0.0108	0.0705	0.1254	0.0489	0.0853	0.0461	0.0052
	φ_f	0.0012	0.0000	0.0000	0.0024	0.0002	0.0000	0.0061	0.0002	0.0001
	$\sigma_{\eta,f}$	0.0017	0.0002	0.0001	0.0051	0.0006	0.0003	0.0147	0.0012	0.0005
$N=30$	θ	0.0319	0.0075	0.0026	0.0277	0.0059	0.0027	0.0357	0.0070	0.0024
	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0010	0.0002	0.0001
	Σ_1	0.1693	0.0456	0.0195	0.1020	0.0265	0.0117	0.0785	0.0209	0.0086
	Σ_2	0.0189	0.0050	0.0020	0.0188	0.0050	0.0020	0.0189	0.0050	0.0021
	μ_ε	0.1182	0.0264	0.0092	0.1235	0.0248	0.0095	0.1972	0.0365	0.0122
	φ_ε	0.0008	0.0001	0.0000	0.0009	0.0001	0.0001	0.0011	0.0002	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0132	0.0029	0.0010	0.0131	0.0027	0.0010	0.0213	0.0039	0.0013
	μ_f	0.1556	0.0649	0.0119	0.0904	0.0242	0.0472	0.0709	0.0157	0.0053
	φ_f	0.0028	0.0002	0.0000	0.0015	0.0001	0.0000	0.0026	0.0003	0.0001
	$\sigma_{\eta,f}$	0.0011	0.0002	0.0001	0.0029	0.0006	0.0003	0.0079	0.0013	0.0004
$N=100$	θ	0.0273	0.0062	0.0024	0.0236	0.0052	0.0020	0.0208	0.0051	0.0018
	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0008	0.0002	0.0001
	Σ_1	0.1598	0.0479	0.0190	0.1002	0.0288	0.0116	0.0780	0.0217	0.0086
	Σ_2	0.0183	0.0048	0.0019	0.0182	0.0048	0.0019	0.0182	0.0048	0.0019
	μ_ε	0.1038	0.0229	0.0088	0.1069	0.0226	0.0089	0.1093	0.0270	0.0090
	φ_ε	0.0008	0.0001	0.0001	0.0008	0.0001	0.0001	0.0008	0.0001	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0115	0.0023	0.0009	0.0121	0.0024	0.0009	0.0121	0.0024	0.0010
	μ_f	0.1041	0.0320	0.0111	0.0693	0.0188	0.0067	0.0517	0.0145	0.0050
	φ_f	0.0005	0.0000	0.0000	0.0012	0.0001	0.0000	0.0025	0.0003	0.0001
	$\sigma_{\eta,f}$	0.0013	0.0002	0.0001	0.0032	0.0007	0.0002	0.0080	0.0013	0.0004

Note: The MSEs for B, Σ_1, Σ_2 are computed by stacking the elements of matrix B and the diagonal of matrices Σ_1 and Σ_2 respectively, in a vector. $\mu_f \equiv (\mu_1, \mu_2, \mu_3)'$, $\mu_\varepsilon \equiv (\mu_4, \ldots, \mu_{3+N})'$, $\varphi_f \equiv (\varphi_1, \varphi_2, \varphi_3)'$, $\varphi_\varepsilon \equiv (\varphi_4, \ldots, \varphi_{3+N})'$, $\sigma_{\eta,f} \equiv (\sigma_{\eta,1}, \sigma_{\eta,2}, \sigma_{\eta,3})'$ and $\sigma_{\eta,\varepsilon} \equiv (\sigma_{\eta,4}, \ldots, \sigma_{\eta,3+N})'$ if $k = 3$.

Table F.4: Average MSE of the MFSV parameter vectors: $H = 10^5/T$ and QML starting values

	Par. \ T	$k=1$ 1000	$k=1$ 4000	$k=1$ 10000	$k=2$ 1000	$k=2$ 4000	$k=2$ 10000	$k=3$ 1000	$k=3$ 4000	$k=3$ 10000
$N=10$	θ	0.0384	0.0083	0.0034	0.0552	0.0150	0.0039	0.1241	0.0308	0.0184
	B	0.0005	0.0001	0.0000	0.0009	0.0002	0.0001	0.0018	0.0004	0.0001
	Σ_1	0.1579	0.0506	0.0194	0.0938	0.0296	0.0116	0.0726	0.0231	0.0087
	Σ_2	0.0210	0.0057	0.0024	0.0209	0.0057	0.0025	0.0221	0.0059	0.0025
	μ_ε	0.1390	0.0271	0.0115	0.2727	0.0752	0.0171	0.7937	0.1881	0.1155
	φ_ε	0.0011	0.0002	0.0001	0.0020	0.0003	0.0001	0.0045	0.0014	0.0007
	$\sigma_{\eta,\varepsilon}$	0.0161	0.0030	0.0012	0.0285	0.0066	0.0022	0.0639	0.0259	0.0158
	μ_f	0.0991	0.0267	0.0106	0.0626	0.0165	0.0066	0.0914	0.0139	0.0059
	φ_f	0.0004	0.0000	0.0000	0.0021	0.0001	0.0000	0.0049	0.0005	0.0002
	$\sigma_{\eta,f}$	0.0014	0.0002	0.0001	0.0053	0.0007	0.0003	0.0115	0.0020	0.0010
$N=20$	θ	0.0275	0.0068	0.0028	0.0293	0.0081	0.0031	0.0358	0.0127	0.0039
	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0010	0.0003	0.0001
	Σ_1	0.1632	0.0437	0.0193	0.0970	0.0272	0.0114	0.0783	0.0203	0.0086
	Σ_2	0.0201	0.0053	0.0021	0.0201	0.0054	0.0021	0.0201	0.0054	0.0022
	μ_ε	0.0955	0.0237	0.0096	0.1301	0.0264	0.0101	0.1927	0.0722	0.0212
	φ_ε	0.0008	0.0001	0.0001	0.0009	0.0002	0.0001	0.0015	0.0003	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0117	0.0025	0.0010	0.0138	0.0028	0.0011	0.0221	0.0068	0.0028
	μ_f	0.0948	0.0247	0.0108	0.0636	0.1313	0.0489	0.0562	0.0258	0.0052
	φ_f	0.0008	0.0000	0.0000	0.0013	0.0001	0.0000	0.0028	0.0003	0.0001
	$\sigma_{\eta,f}$	0.0015	0.0002	0.0001	0.0048	0.0006	0.0003	0.0094	0.0013	0.0005
$N=30$	θ	0.0283	0.0064	0.0026	0.0235	0.0055	0.0027	0.0256	0.0062	0.0024
	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0010	0.0002	0.0001
	Σ_1	0.1686	0.0457	0.0195	0.1021	0.0266	0.0117	0.0783	0.0208	0.0086
	Σ_2	0.0189	0.0050	0.0020	0.0188	0.0050	0.0020	0.0188	0.0050	0.0021
	μ_ε	0.1042	0.0225	0.0092	0.1006	0.0231	0.0095	0.1329	0.0318	0.0122
	φ_ε	0.0007	0.0001	0.0000	0.0008	0.0001	0.0001	0.0011	0.0002	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0112	0.0024	0.0010	0.0113	0.0026	0.0010	0.0154	0.0037	0.0013
	μ_f	0.0962	0.0247	0.0119	0.0705	0.0153	0.0472	0.0458	0.0120	0.0053
	φ_f	0.0002	0.0000	0.0000	0.0010	0.0001	0.0000	0.0019	0.0002	0.0001
	$\sigma_{\eta,f}$	0.0011	0.0002	0.0001	0.0034	0.0006	0.0003	0.0062	0.0011	0.0004
$N=100$	θ	0.0239	0.0059	0.0024	0.0198	0.0053	0.0020	0.0180	0.0047	0.0018
	B	0.0005	0.0001	0.0000	0.0007	0.0002	0.0001	0.0008	0.0002	0.0001
	Σ_1	0.1610	0.0479	0.0190	0.0997	0.0289	0.0116	0.0775	0.0218	0.0086
	Σ_2	0.0184	0.0048	0.0019	0.0182	0.0048	0.0019	0.0182	0.0048	0.0019
	μ_ε	0.0882	0.0214	0.0088	0.0862	0.0234	0.0089	0.0912	0.0244	0.0090
	φ_ε	0.0007	0.0001	0.0001	0.0007	0.0001	0.0001	0.0007	0.0001	0.0001
	$\sigma_{\eta,\varepsilon}$	0.0097	0.0022	0.0009	0.0100	0.0023	0.0009	0.0106	0.0023	0.0010
	μ_f	0.0913	0.0283	0.0111	0.0582	0.0168	0.0067	0.0479	0.0129	0.0050
	φ_f	0.0026	0.0000	0.0000	0.0030	0.0001	0.0000	0.0019	0.0002	0.0001
	$\sigma_{\eta,f}$	0.0011	0.0001	0.0001	0.0026	0.0005	0.0002	0.0062	0.0010	0.0004

Note: The MSEs for B, Σ_1, Σ_2 are computed by stacking the elements of matrix B and the diagonal of matrices Σ_1 and Σ_2 respectively, in a vector. $\mu_f \equiv (\mu_1, \mu_2, \mu_3)'$, $\mu_\varepsilon \equiv (\mu_4, \ldots, \mu_{3+N})'$, $\varphi_f \equiv (\varphi_1, \varphi_2, \varphi_3)'$, $\varphi_\varepsilon \equiv (\varphi_4, \ldots, \varphi_{3+N})'$, $\sigma_{\eta,f} \equiv (\sigma_{\eta,1}, \sigma_{\eta,2}, \sigma_{\eta,3})'$ and $\sigma_{\eta,\varepsilon} \equiv (\sigma_{\eta,4}, \ldots, \sigma_{\eta,3+N})'$ if $k = 3$.

G Heat-Maps of the Average MSE

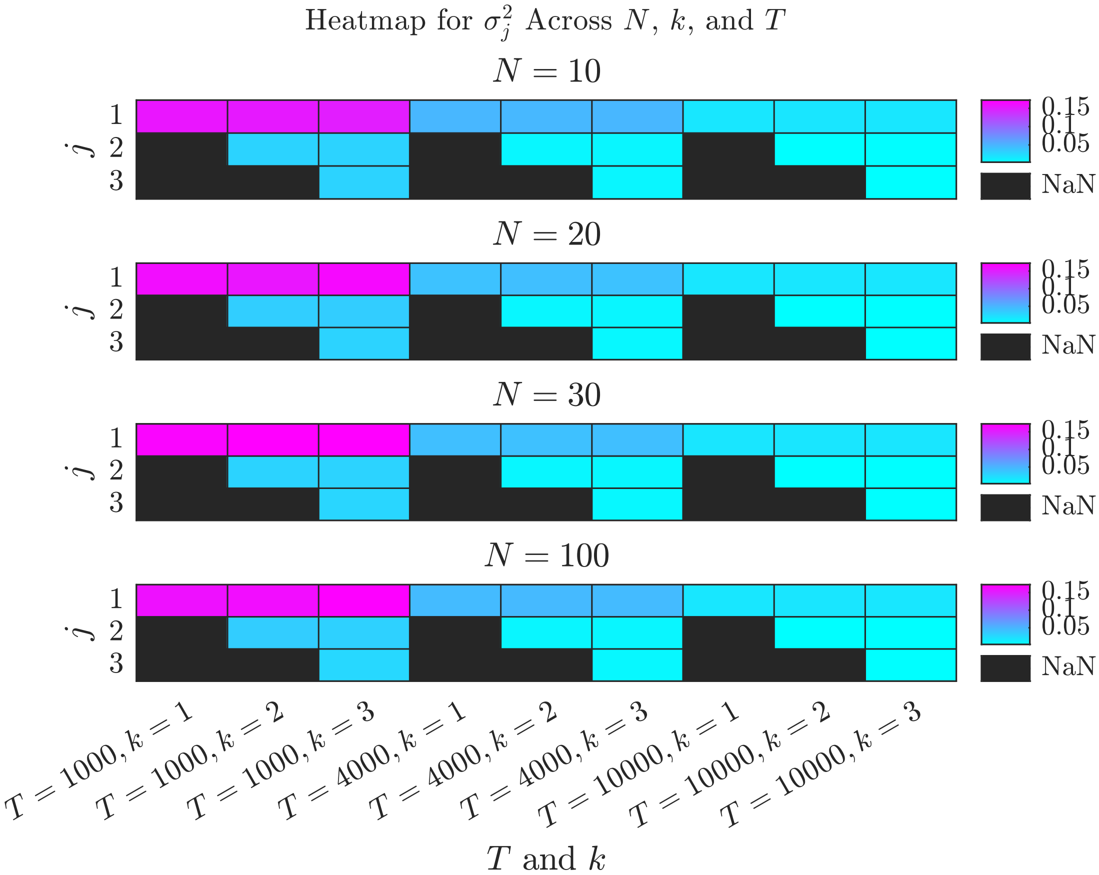

Figure G.1: Heat-map of average MSE of the factor variances: QML starting values, $H = 10^5/T$. $\gamma_j \equiv \sigma_j$, where $j = 1, 2, 3$. 'NaN' stands for non-existent elements, such as σ_j^2 for $j > k$.

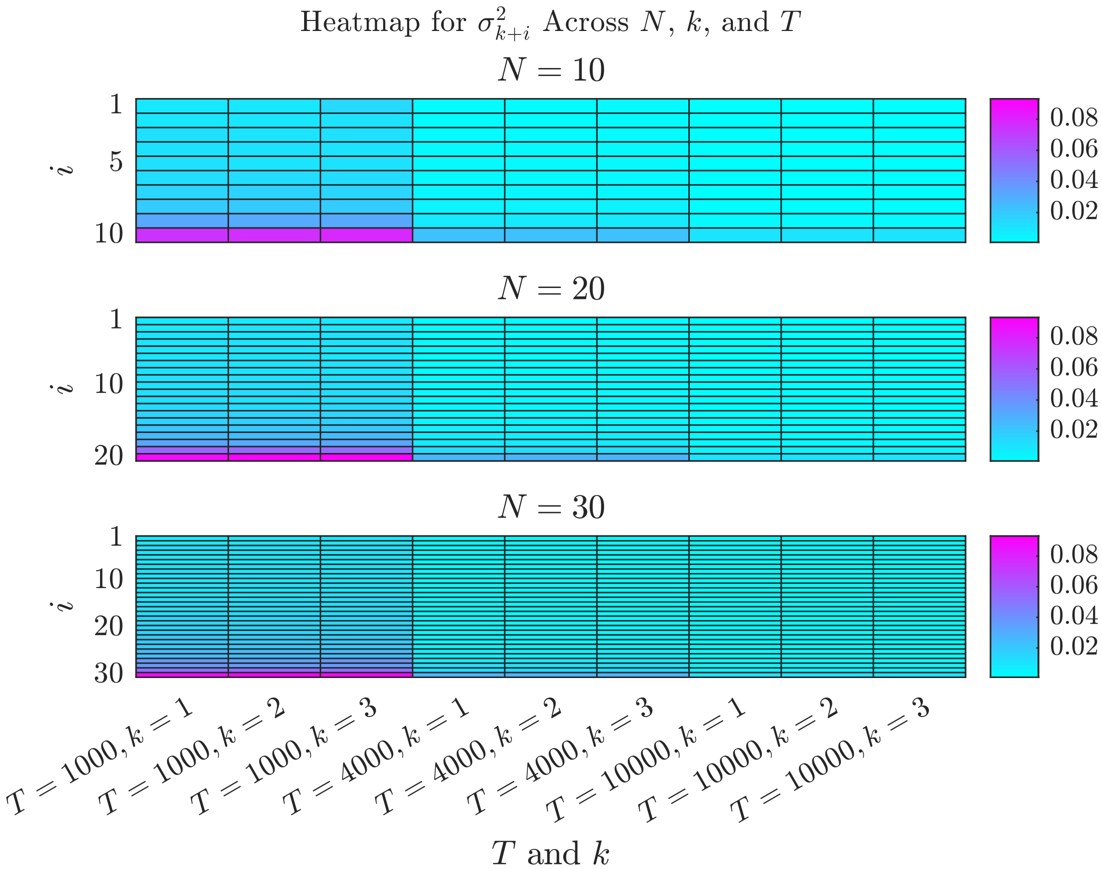

Figure G.2: Heat-map of average MSE of the idiosyncratic error variances: $N = 10, 20, 30$, QML starting values, $H = 10^5/T$.

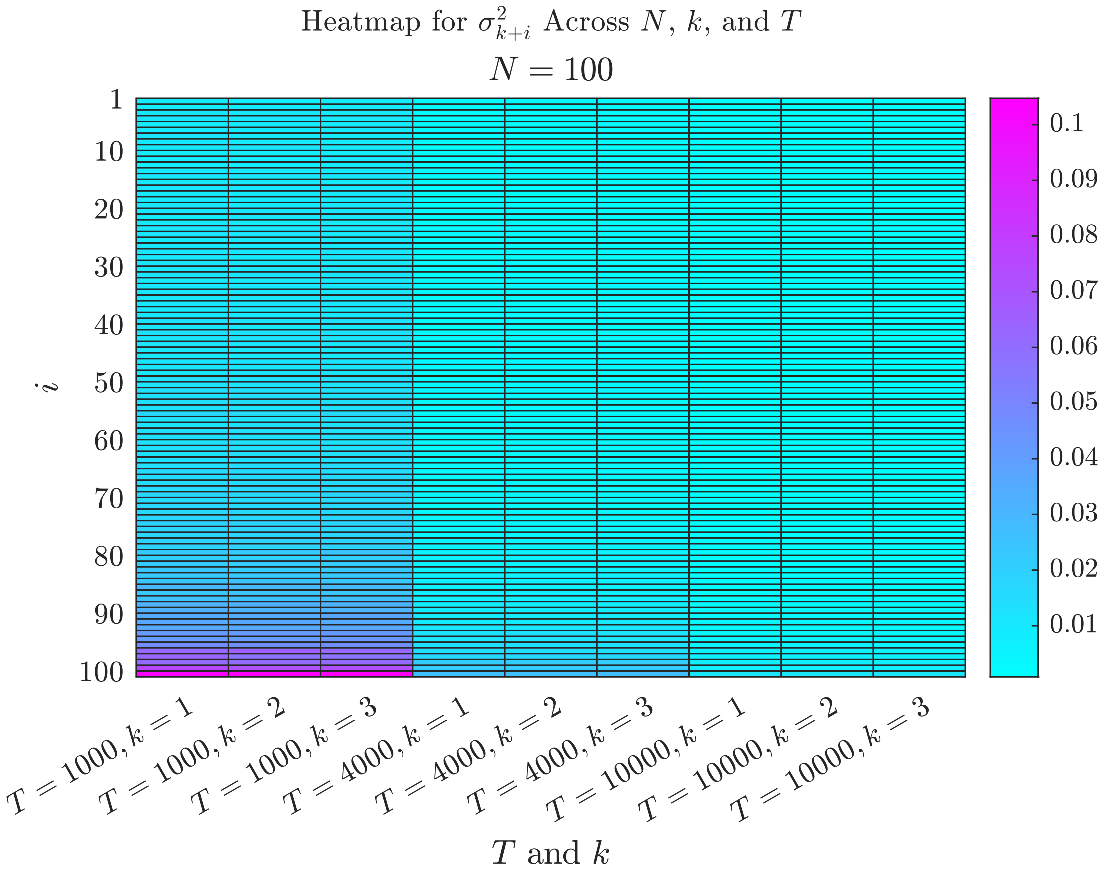

Figure G.3: Heat-map of average MSE of the idiosyncratic error variances: $N = 100$, QML starting values, $H = 10^5/T$.

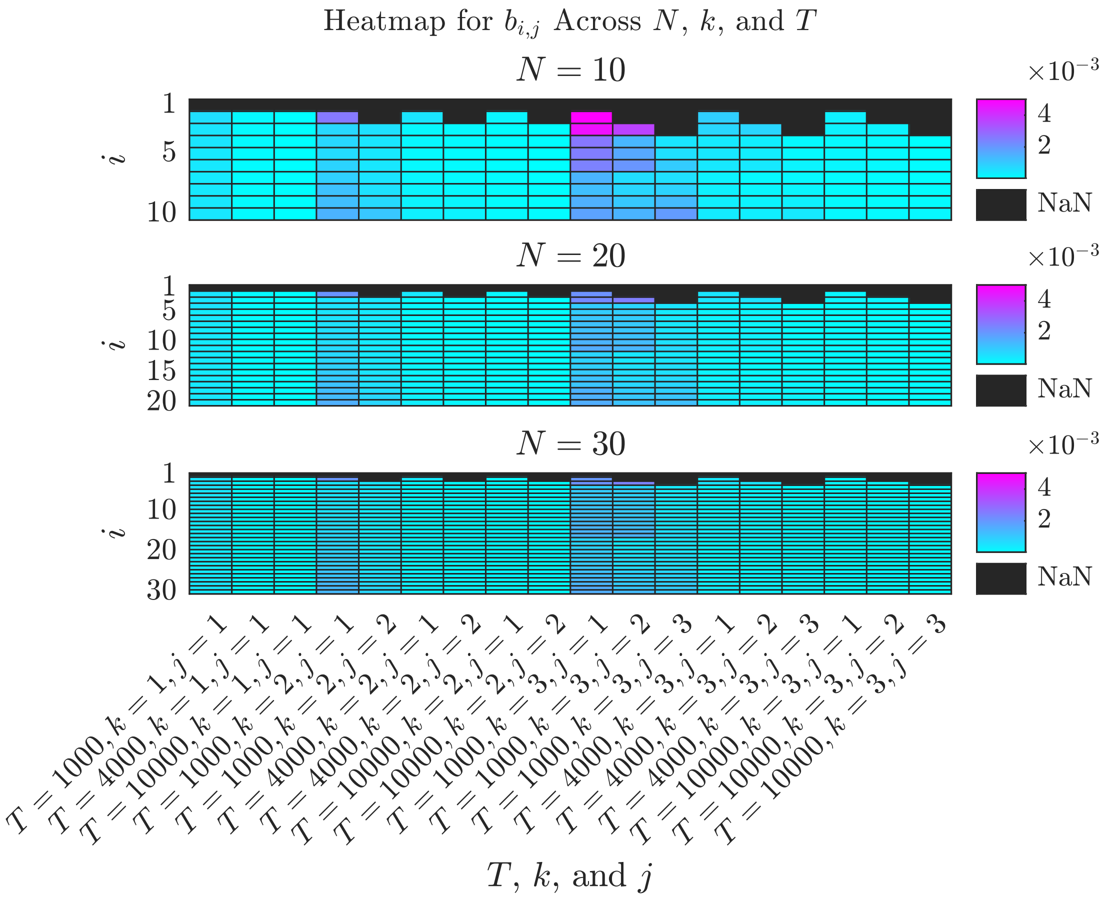

Figure G.4: Heat-map of average MSE of the loadings B: $N = 10, 20, 30$, QML starting values, $H = 10^5/T$. 'NaN' stands for non-existent elements, such as the diagonal and upper triangle elements of the loading matrix that are restricted to 1 and 0, respectively.

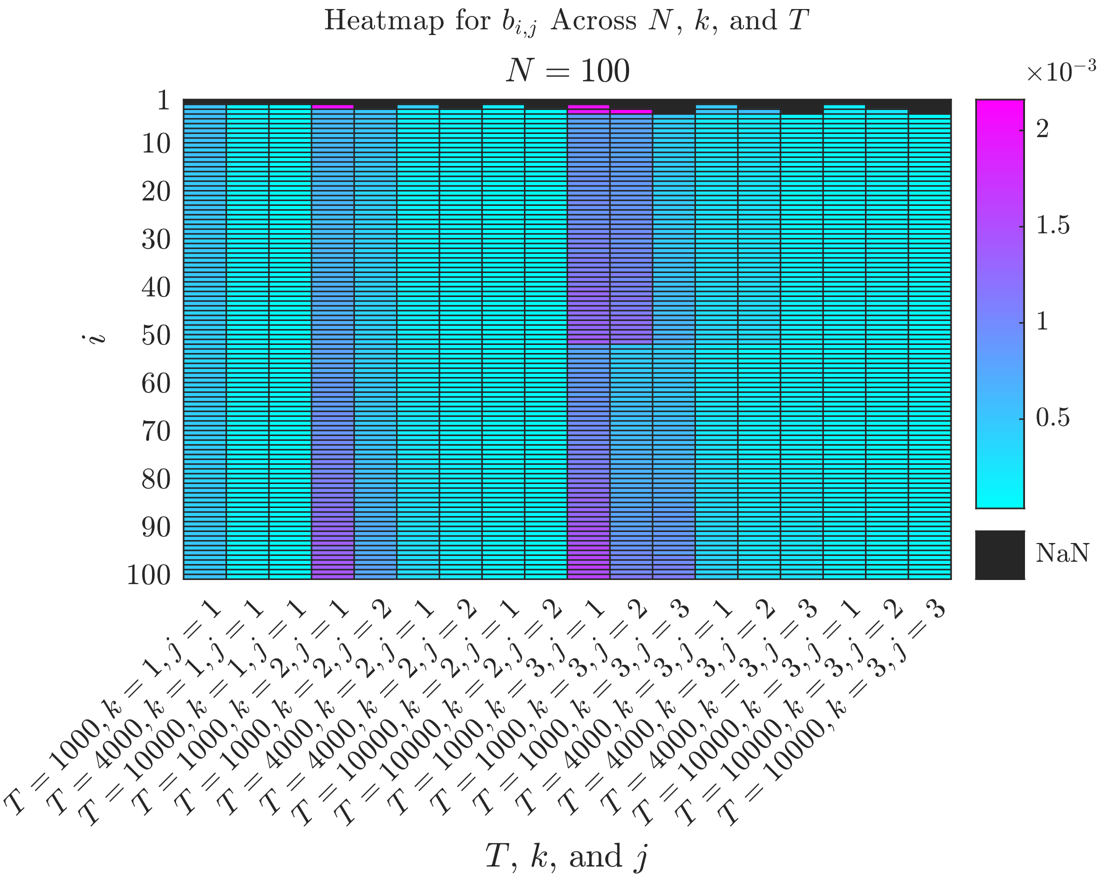

Figure G.5: Heat-map of average MSE of the loadings B: $N = 100$, QML starting values, $H = 10^5/T$. 'NaN' stands for non-existent elements, such as the diagonal and upper triangle elements of the loading matrix that are restricted to 1 and 0, respectively.

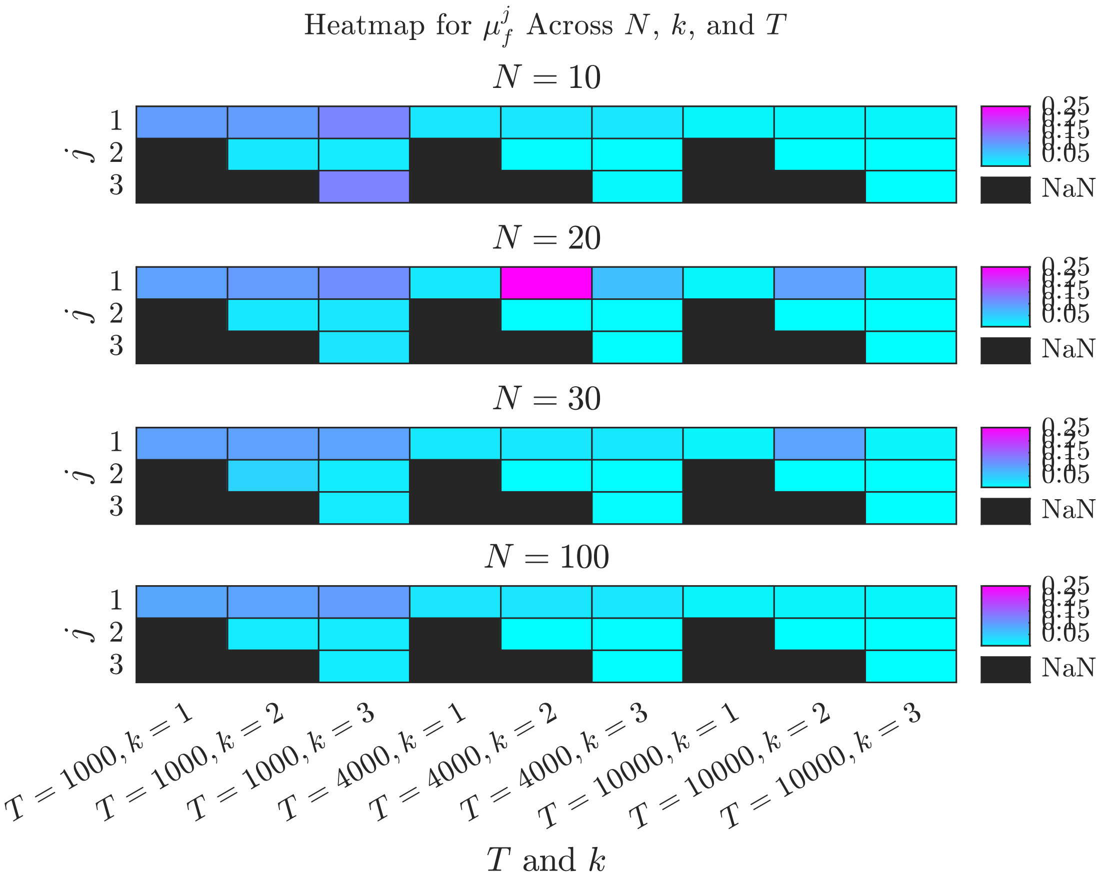

Figure G.6: Heat-map of average MSE of the μ parameters of the factors ARSV equations: QML starting values, $H = 10^5/T$. 'NaN' stands for non-existent elements, such as μ_f^j for $j > k$.

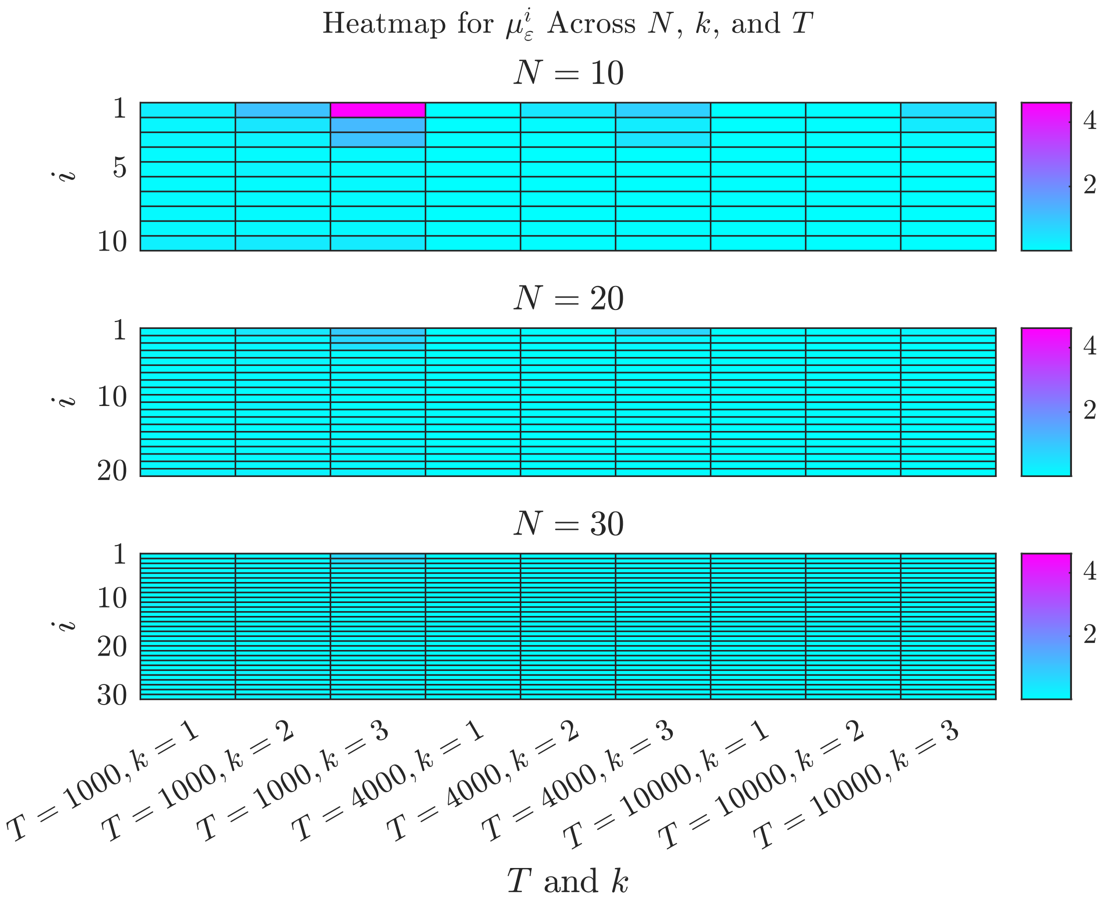

Figure G.7: Heat-map of average MSE of the μ parameters of the idiosyncratic noise ARSV equations: $N = 10, 20, 30$, QML starting values, $H = 10^5/T$.

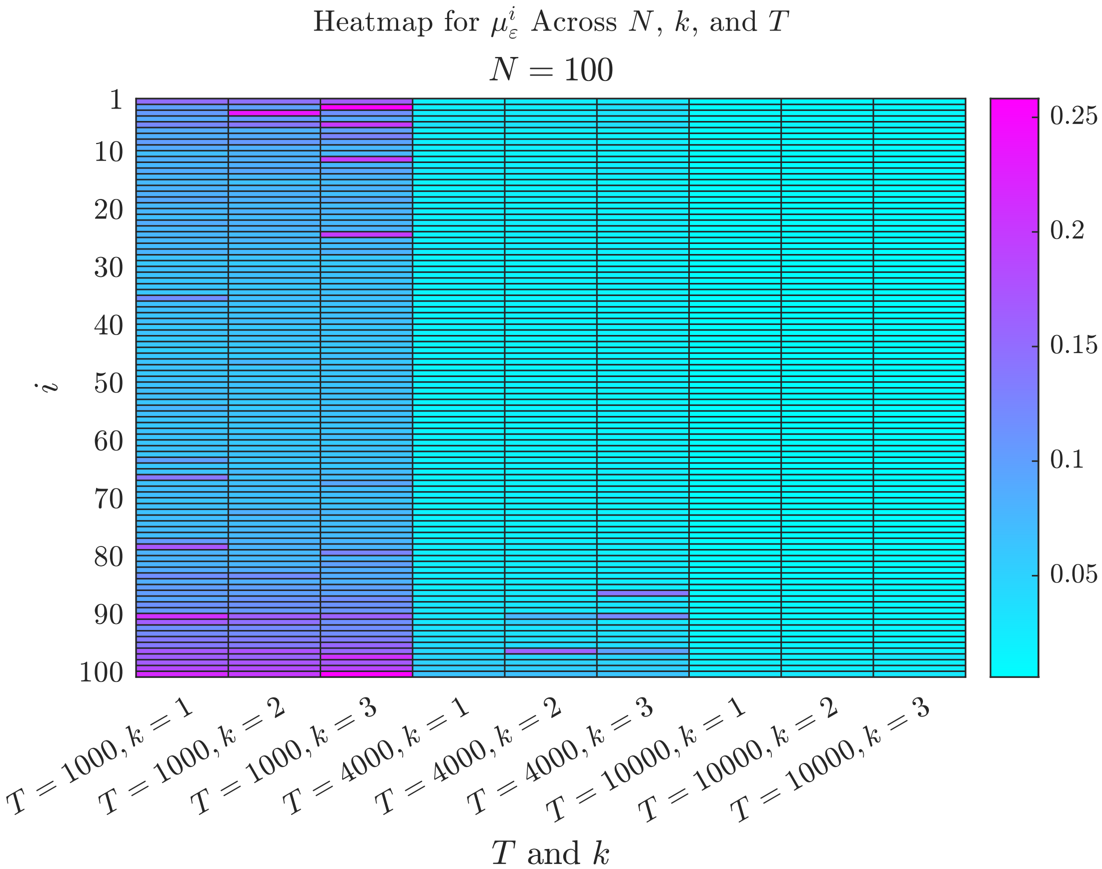

Figure G.8: Heat-map of average MSE of the μ parameters of the idiosyncratic noise ARSV equations: $N = 100$, QML starting values, $H = 10^5/T$.

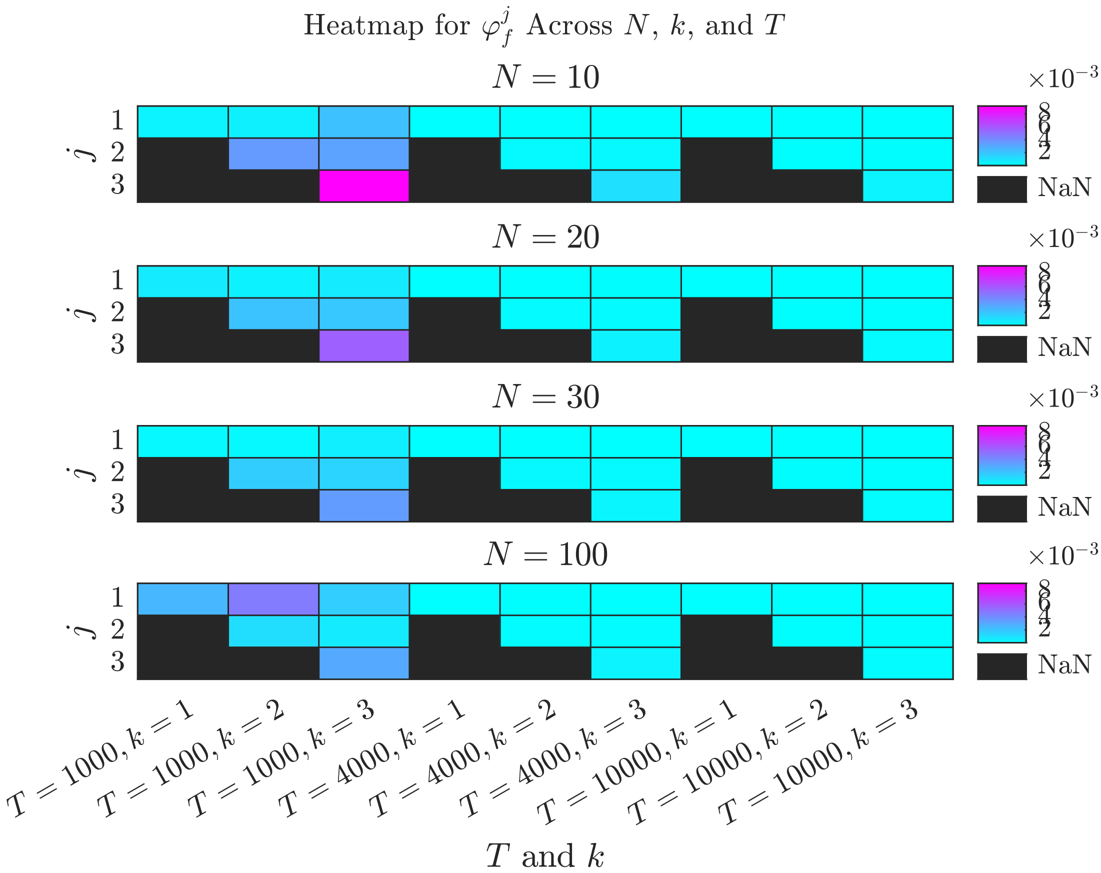

Figure G.9: Heat-map of average MSE of the AR parameters, φ of the factors ARSV equations: QML starting values, $H = 10^5/T$. 'NaN' stands for non-existent elements, such as φ_f^j for $j > k$.

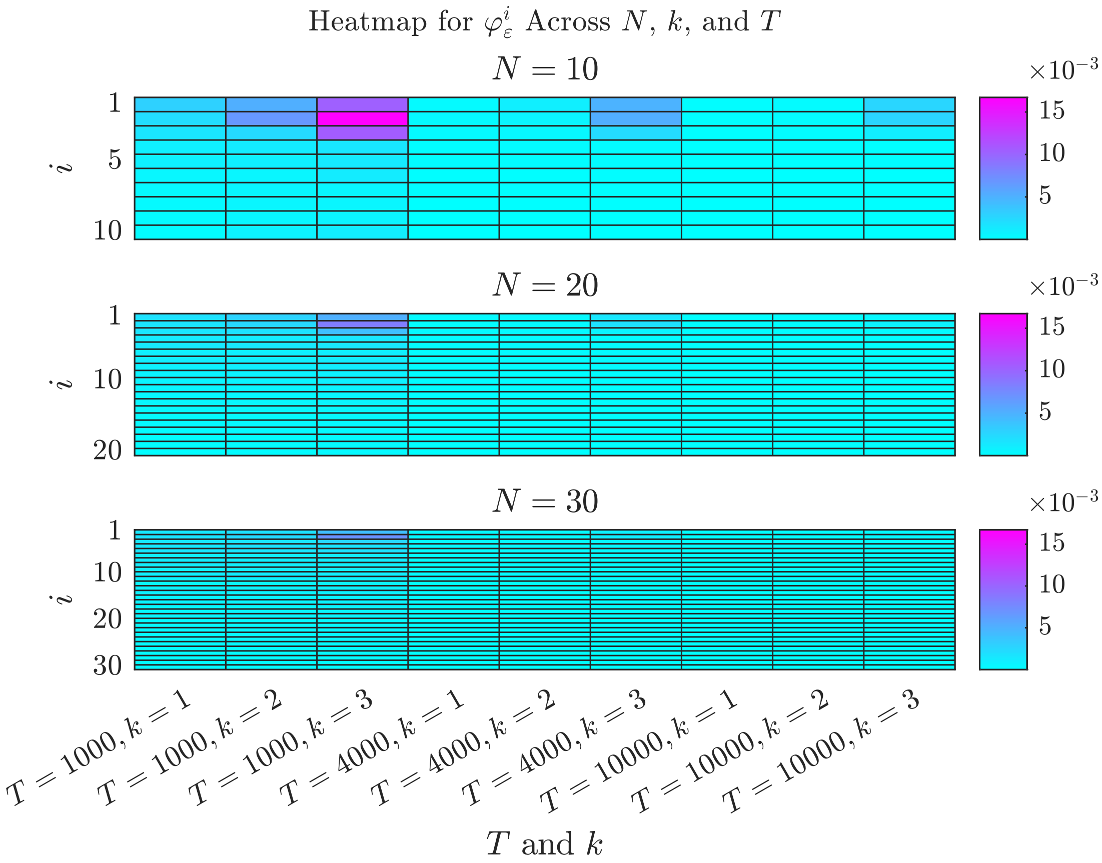

Figure G.10: Heat-map of average MSE of the AR parameters, φ of the idiosyncratic error ARSV equations: $N = 10, 20, 30$, QML starting values, $H = 10^5/T$.

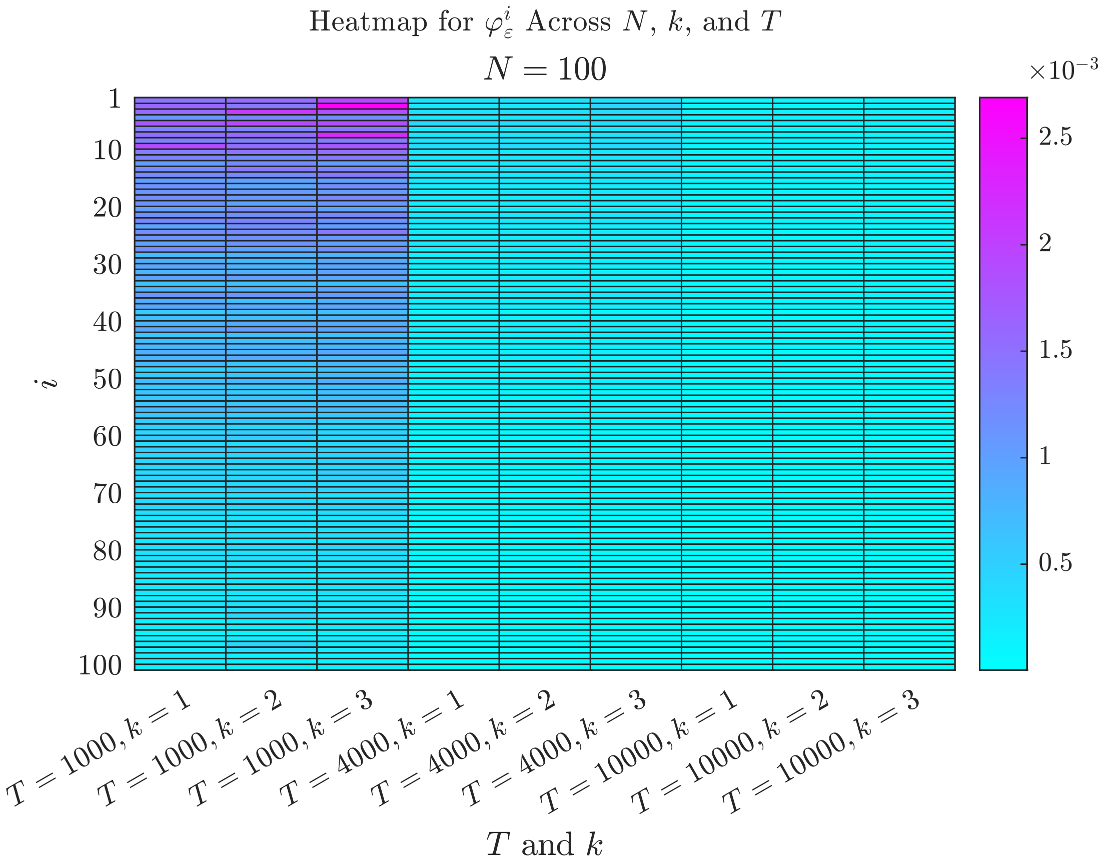

Figure G.11: Heat-map of average MSE of the AR parameters, φ of the idiosyncratic error ARSV equations: $N = 100$, QML starting values, $H = 10^5/T$.

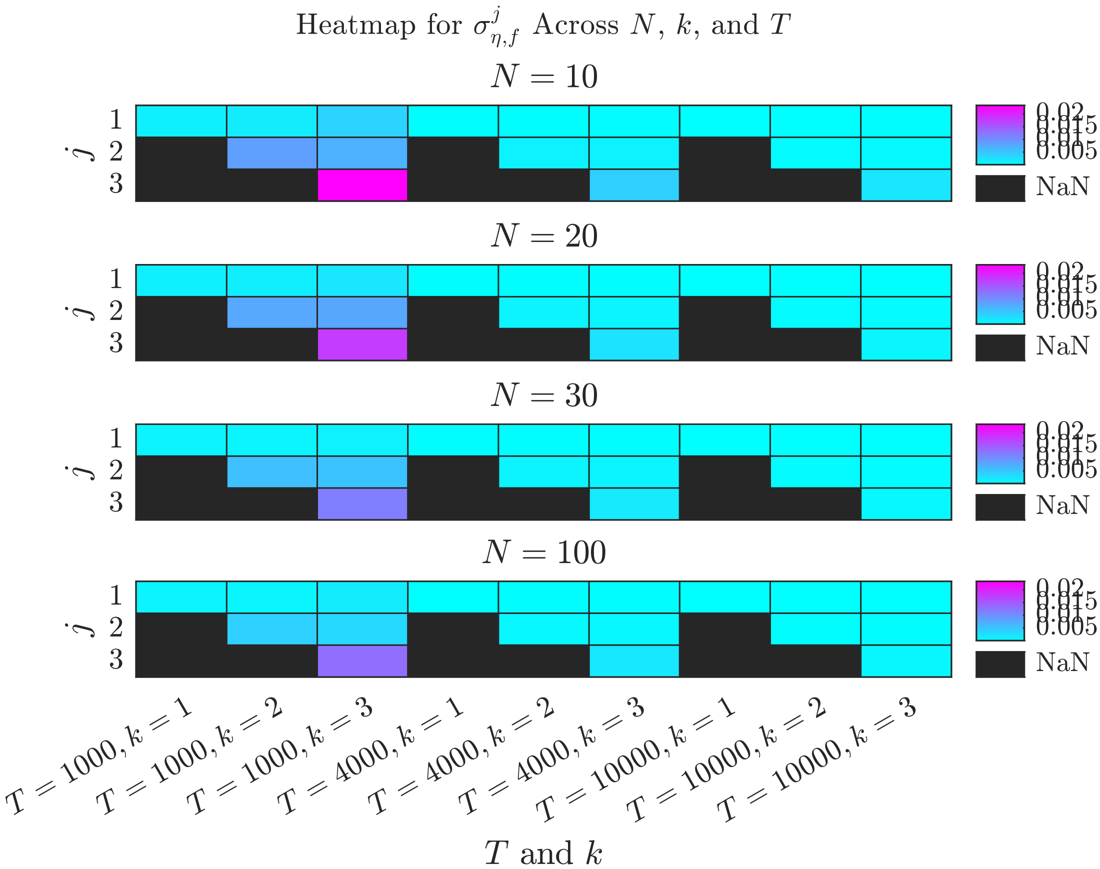

Figure G.12: Heat-map of average MSE of the noise standard deviations of the factors ARSV equations: QML starting values, $H = 10^5/T$. 'NaN' stands for non-existent elements, such as $\sigma^j_{\eta,f}$ for $j > k$.

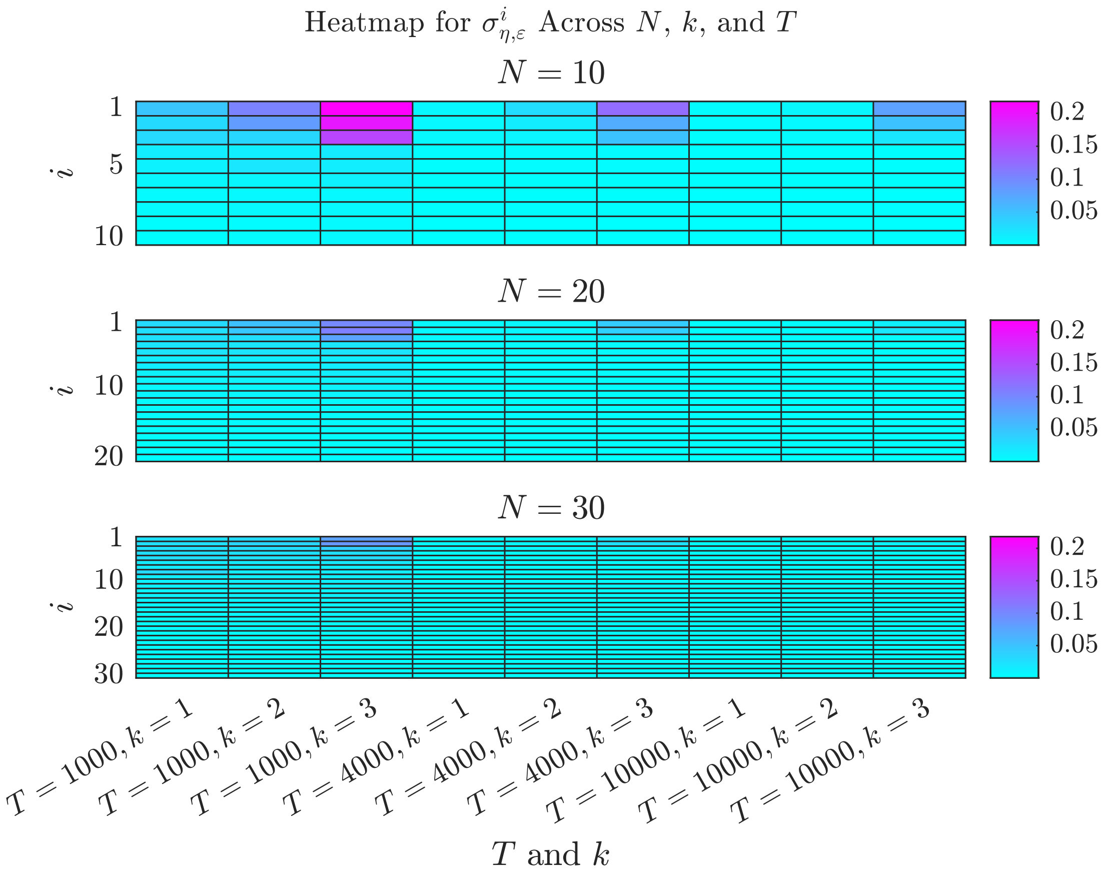

Figure G.13: Heat-map of average MSE of the noise standard deviations of the idiosyncratic error ARSV equations: $N = 10, 20, 30$ QML starting values, $H = 10^5/T$.

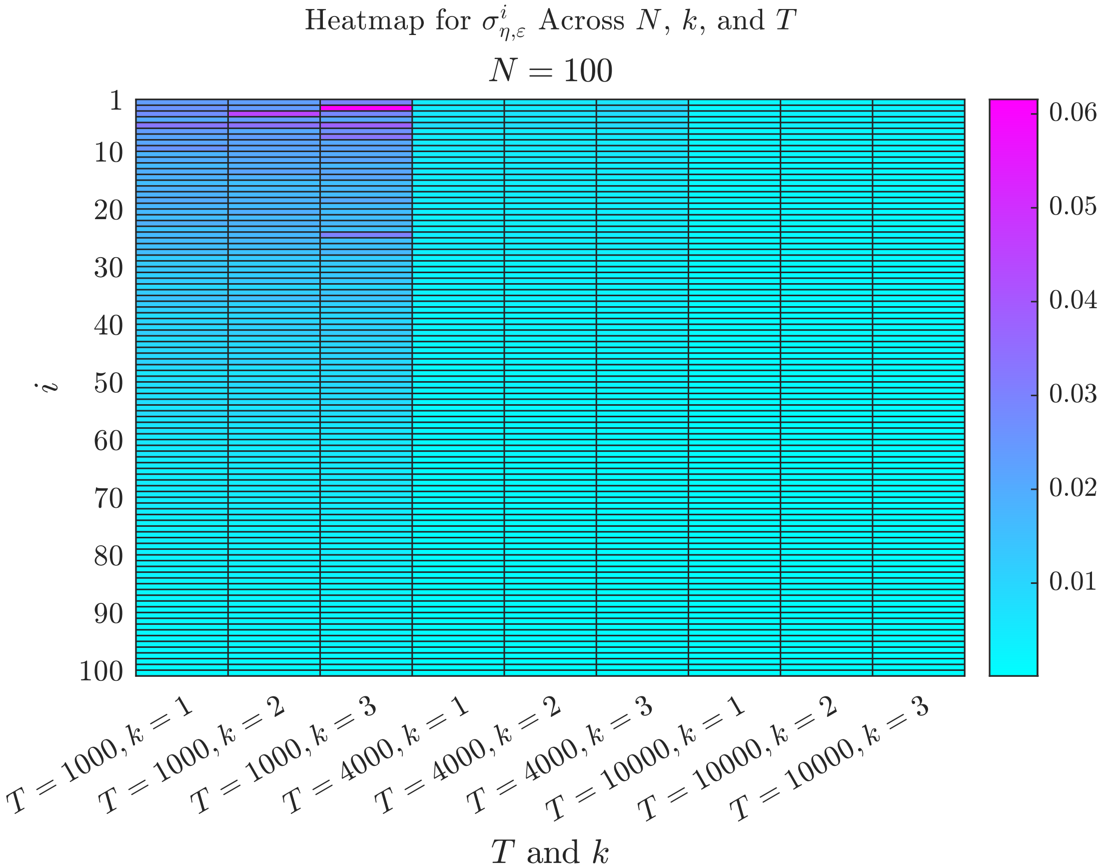

Figure G.14: Heat-map of average MSE of the noise standard deviations of the idiosyncratic error ARSV equations: $N = 100$ QML starting values, $H = 10^5/T$.

H List of Ticker Symbols and Companies

Table H.1: List of the ticker symbols and names of the companies, ordered by market capitalization. The ordering is done column wise (the columns are continued on the next page) from top to bottom and from left to right. Thus JNJ has the largest market capitalization, XOM the second largest, and JWN the lowest.

JNJ	Johnson & Johnson	VFC	VF Corp
XOM	Exxon Mobil Corp	XEL	Xcel Energy Inc
JPM	JPMorgan Chase & Co	STI	SunTrust Banks Inc
PG	Procter & Gamble	GLW	Corning Inc
BAC	Bank of Amarica Corp	PPG	PPG Industries Inc
INTC	Intel Corp	ED	Consolidated Edison Inc
CVX	Chevron Corp	LUV	Southwest Airlines Co
PFE	Pfizer	AMD	Advanced Micro Devices Inc
DIS	Walt Disney Co	PH	Parker-Hannifin Corp
WFC	Wells Fargo & Co	MSI	Motorola Solutions Inc
BA	Boeing Co	PPL	PPL Corp
CMCSA	Comcast Corp	ES	Eversource Energy
KO	Coca-Cola Co	ROK	Rockwell Automation Inc
PEP	PepsiCo Inc	SWK	Stanley Black & Decker Inc
MCD	McDonald's Corp	EIX	Edison International
WMT	Walmart Inc	WY	Weyerhaeuser Co
ABT	Abbott Laboratories	HRS	Harris Corp
IBM	International Business Machines Corp	BLL	Ball Corporation
MMM	Minnesota Mining and Manufacturing Co	KR	Kroger Co
UNP	Union Pacific Corp	IP	International Paper Co
MDT	Medtronic PLC	NUE	Nucor Corp
HON	Honeywell International Inc	ETR	Entergy Corp
TMO	Thermo Fisher Scientific Inc	OMC	Omnicom Group Inc
LLY	Eli Lilly & Co	HES	Hess Corp
UTX	United Technologies Corporation	HSY	Hershey Co
TXN	Texas Instruments Inc	CMS	CMS Energy Corp
MO	Altria Group Inc	CNP	CenterPoint Energy Inc
LOW	Lowe's Cos Inc	WDC	Western Digital Corp
NEE	NextEra Energy Inc	LEN	Lennar Corp
DHR	Danaher Corp	K	Kellogg Co
GE	General Electric Co	MYL	Mylan
CAT	Caterpillar Inc	HST	Host Hotels & Resorts Inc
AXP	American Express Co	MRO	Marathon Oil Corp
BMY	Bristol-Myers Squibb Co	GWW	WW Grainger Inc
COP	ConocoPhillips	DOV	Dover Corp
USB	US Bancorp	CAG	Conarga Brands Inc
CVS	CVS Health Corp	IFF	International Flavors & Fragrances Inc
ADP	Automatic Data Processing Inc	L	Loews Corp
BDX	Becton Dickinson and Co	CINF	Cincinnati Financial Corp
TJX	TJX Cos Inc	LNC	Lincoln National Corp
DUK	Duke Energy Corp	APA	APA Corp
D	Dominion Energy Inc	NBL	Noble Energy Inc
SYK	Stryker Corp	UDR	UDR Inc
CL	Colgate-Palmolive Co	CMA	Comerica Inc
PNC	PNC Financial Services Group	TXT	Textron Inc
SO	Southern Co	KSU	Kansas City Southern
SPGI	S&P Global Inc	CTL	Lumen Technologies Inc
DE	Deere & Co	HRL	Hormel Foods Corp
RTN	Raytheon Co	TAP	Molson Coors Beverage Co

OXY	Occidental Petroleum Corp	LNT	Alliant Energy Corp
EXC	Exelon Corp	PKI	PerkinElmer Inc
MMC	Marsh & McLennan Cos Inc	NI	NiSource Inc
NOC	Northrop Grumman Corp	FRT	Federal Realty Investment Trust
GD	General Dynamics Corp	PNW	Pinnacle West Capital Corp
FDX	FedEx Corp	HAS	Hasbro Inc
ECL	Ecolab Inc	PVH	PVH Corp
WBA	Walgreens Boots Alliance Inc	AVY	Avery Dennison Corp
TGT	Target Corp	ZION	Zions Bancorp NA
APD	Air Products and Chemicals Inc	SNA	Snap-on Inc
PGR	Progressive Corp	IPG	Interpublic Group of Cos Inc
KMB	Kimberly-Clark Corp	SEE	Sealed Air Corp
ADI	Analog Devices Inc	CPB	Campbell Soup Co
AEP	Americal Electric Power Co Inc	ALK	Alaska Air Group Inc
AMAT	Applied Materials Inc	FL	Foot Locker Inc
AIG	American International Group Inc	PNR	Pentair PLC
BAX	Baxter International Inc	XRX	Xerox Holdings Corp
HUM	Humana Inc	HP	Helmerich & Payne Inc
AFL	Aflac Inc	ROL	Rollins Inc
F	Ford Motor Co	NWL	Newell Brands Inc
ETN	Eaton Corp PLC	GPS	Gap Inc
WMB	Williams Cos Inc	LEG	Leggett & Platt Inc
SYY	Sysco Corp	HRB	H&R Block Inc
HPQ	HP Inc	JEF	Jefferies Financial Group Inc
PEG	Public Service Enterprise Group Inc	JWN	Nordstrom Inc

I Principal Component of the Real Data

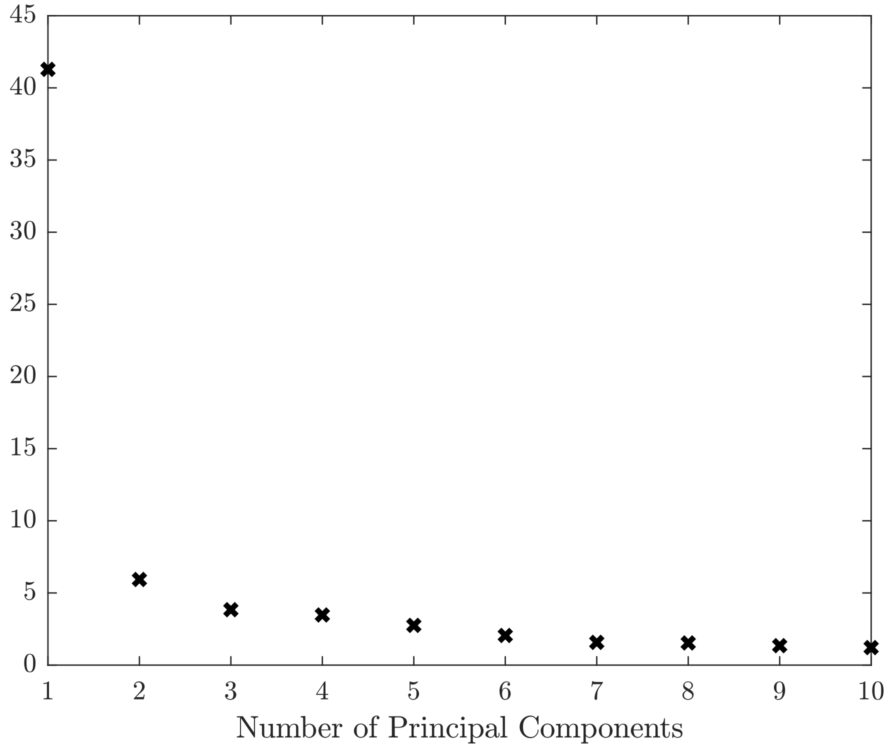

Figure I.1: Plot of the largest 10 sorted eigenvalues of the sample covariance matrix of the standardized returns of our empirical application.

J Empirical Results for $k = 1, 2, 3$

Although as suggested by the tests of Onatski (2010) and Ahn and Horenstein (2013) and PCA, we only need one factor to capture the commonality in the stochastic volatility dynamics of the series, we present here, for completeness, the results from estimating the MFSV model for more factors. We choose $k = 3$ to be the maximum, as it has also been a common choice in the existing Bayesian estimation literature (Han, 2005; Nardari and Scruggs, 2007). We report the estimation results of our procedure below namely in Table J.1 for estimating the SV parameters of the factors and in figures J.1 to J.3, for estimating the SV parameters of the idiosyncratic error series.

The first finding of applying our estimation procedure to MFSV models with more factors than the "true" ones is that it detects this mismatch by delivering not significant estimates or anomalies in the estimates, as reported in Table J.1: while the estimates of the SV model for the idiosyncratic errors are similar in values to the ones of $k = 1$ and the AR parameters of the SV model of the factors are all very close to 1, which indicate high persistency in the volatility dynamics of the series, similar to the case $k = 1$, most of the SV parameters of the second and third factor are estimated to be not significant and some of them have unusual high standard errors. This might be due to the fact that the accurate ML estimation of the static factor model in the first step requires some degree of variation in the factors (Bai and Li, 2012). This is not the case when estimating more factors than necessary, which in fact, should have variation zero. In that case, the rank condition on the factor loading matrix is violated (Anderson and Rubin, 1956), and the distribution of the static factors is not identified (Cox, 2017). In other words, estimating factors with almost zero variance may indicate that the model is not correctly specified and that there are fewer factors describing the true structure. Therefore, the empirical results for the case of $k = 2$ and $k = 3$ should be handled with care as they refer to misspecified MFSV models.

To bring further evidence in this direction, we run a small simulation study with $N = 10$ and $N = 100$ return series, $T = 4000$ observations and $k = 2$ factors. Figures J.4 and J.5 below show the first 10 principal components of the $R = 1000$ simulated data sets for $N = 10$ and $N = 100$ respectively. The plots clearly indicate the number of factors to be equal to 2, which is precisely the number of factors we simulate. On these simulated series, we apply our estimation procedure by choosing $k = 1, 2, 3$ factors. We find that the Fisher information matrix of the auxiliary model for $k = 3$ is not of full rank with serious effects on the standard errors and the inference based on them. In fact for $k = 1$ and $k = 2$ the estimation works smoothly, while for $k = 3$, the first step estimation reveals zero variance of the third factor, as depicted in Figure J.8 for $N = 10$ and in Figure J.9 for $N = 100$.

The fact that for $k = 2$ and $k = 3$ we deal with misspecified MFSV models for our real dataset is also detectable when running the Bayesian procedure of Kastner et al. (2017), which in fact sets the constant in the factor ARSV model to 0 and instead specifies a scaling parameter for the variance of the factor.[24] Figures J.6 and J.7 show that, for $k = 3$, the scaling of the second factor is ill-behaved. Specifically, we see in Figure J.6 that the Markov Chain does not converge for the scaling factor. Figure J.7 shows that the estimated marginal density of the scaling factor is bi-modal, with a lot of probability mass around zero. Thus, the Bayesian procedure also faces "problems" when the number of factors becomes larger than the true one.

Table J.1: Estimates of the ARSV parameters of the k factors. Standard errors in parentheses.

	$k=1$	$k=2$	$k=3$		$k=1$	$k=2$	$k=3$
σ_1^2	0.2918	0.3042	0.3466	φ_1	0.9752	0.9739	0.9738
	(0.0303)	(0.3533)	(0.0513)		(0.0043)	(0.5210)	(0.0083)
σ_2^2		0.0000	0.0096	φ_2		0.9876	0.9913
		(0.0000)	(0.0036)			(0.5847)	(0.0029)
σ_3^2			0.1328	φ_3			0.9854
			(0.0793)				(0.0068)
μ_1	-1.5866	-1.5301	-1.3825	$\sigma_{\eta,1}$	0.1863	0.1871	0.1827
	(0.1156)	(10.6368)	(0.6025)		(0.0164)	(3.4561)	(0.1696)
μ_2		-15.5277	-5.2052	$\sigma_{\eta,2}$		0.1672	0.1392
		(140.0762)	(0.3620)			(9.4228)	(0.0215)
μ_3			-2.6855	$\sigma_{\eta,3}$			0.1964
			(3.1305)				(0.4754)

[24]The variance of the factor is multiplied by this scaling parameter (Broto and Ruiz, 2004). In the specific context of our MFSV model, the scaling parameters are multiplied then with the respective column of the loadings matrix, allowing, thus, for diagonal elements that are not restricted to 1.

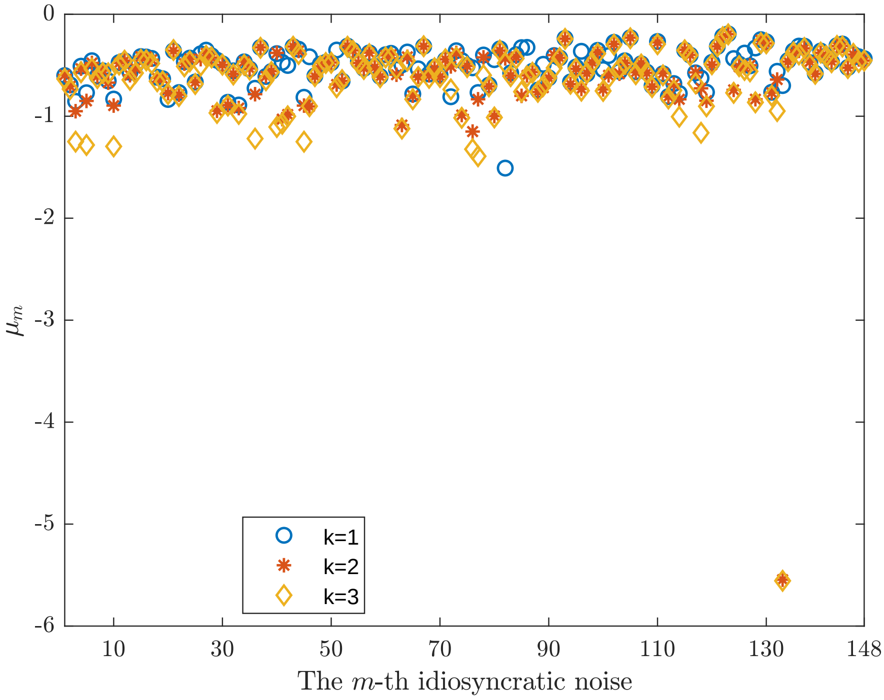

Figure J.1: Estimated constant of ARSV or the m-th idiosyncratic error, $m = 1, \ldots, 148$ stocks

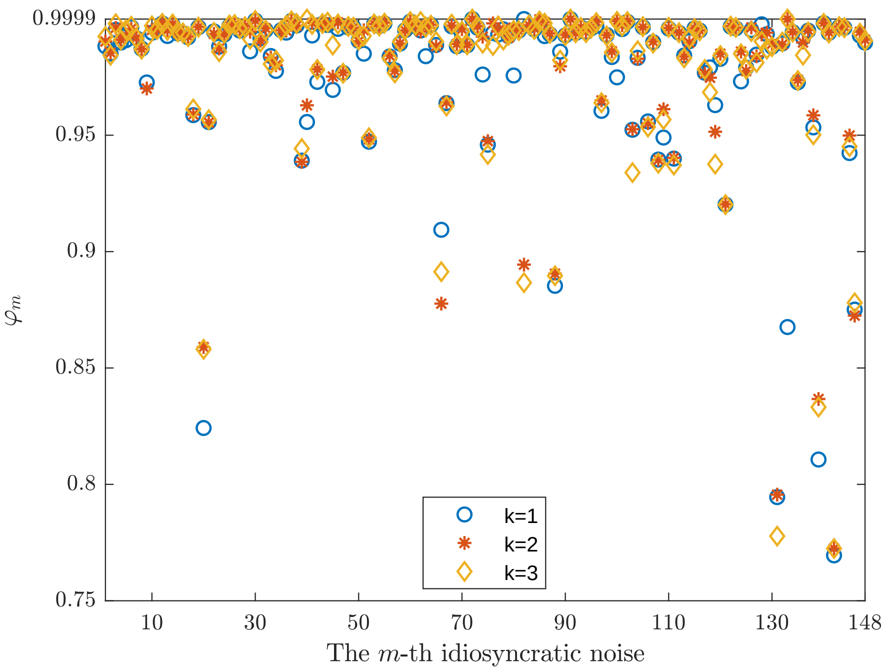

Figure J.2: Estimated autoregressive parameter of ARSV for the m-th idiosyncratic error, $m = 1, \ldots, 148$ stocks

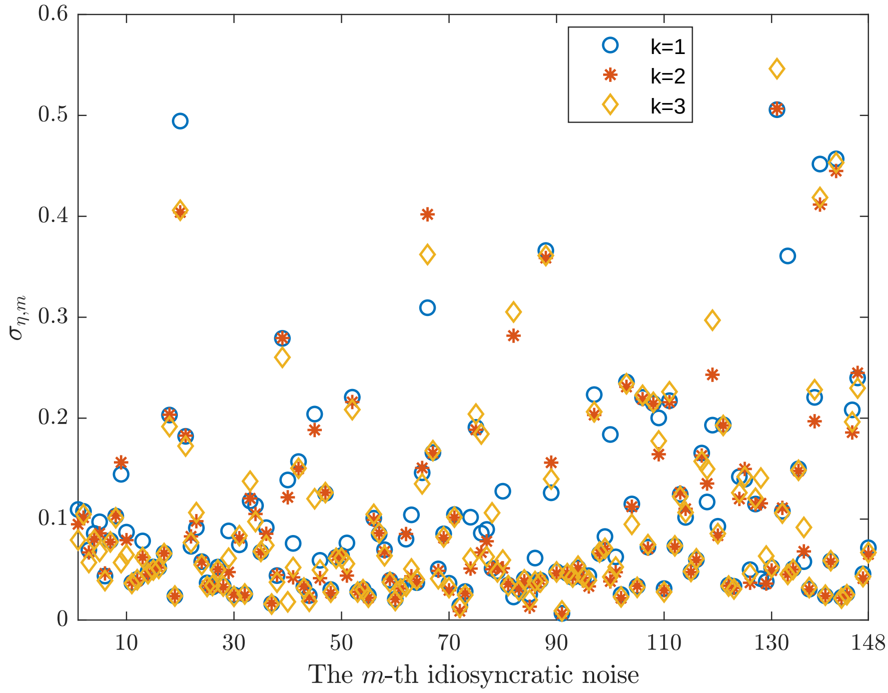

Figure J.3: Estimated standard deviation parameter of ARSV or the m-th idiosyncratic error, $m = 1, \ldots, 148$ stocks

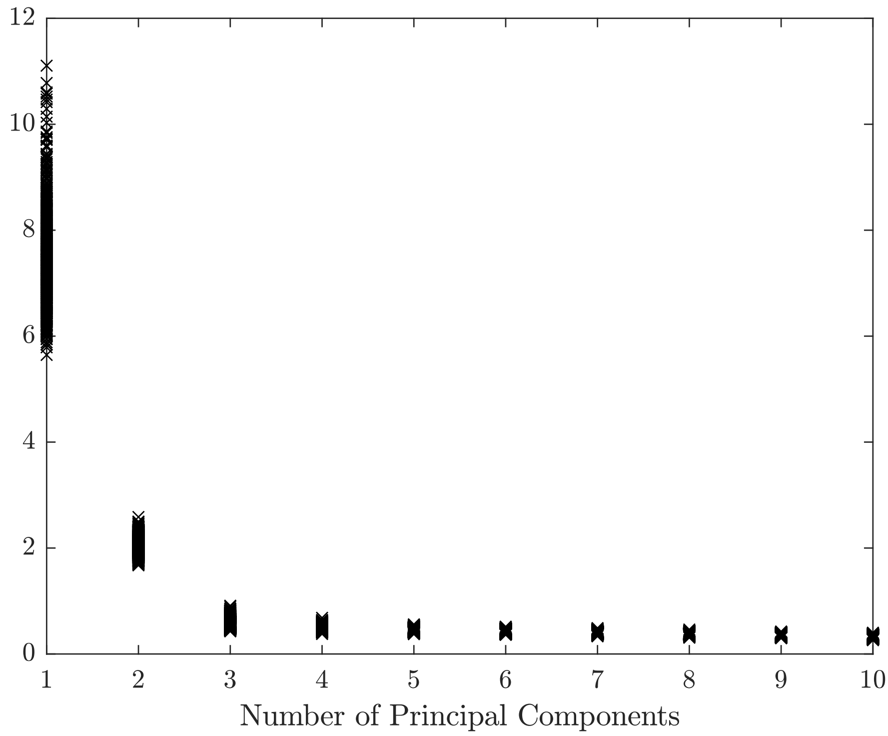

Figure J.4: Plot of the sorted eigenvalues of the sample covariance matrix of the 1000 simulated standardised return vectors for $N = 10$.

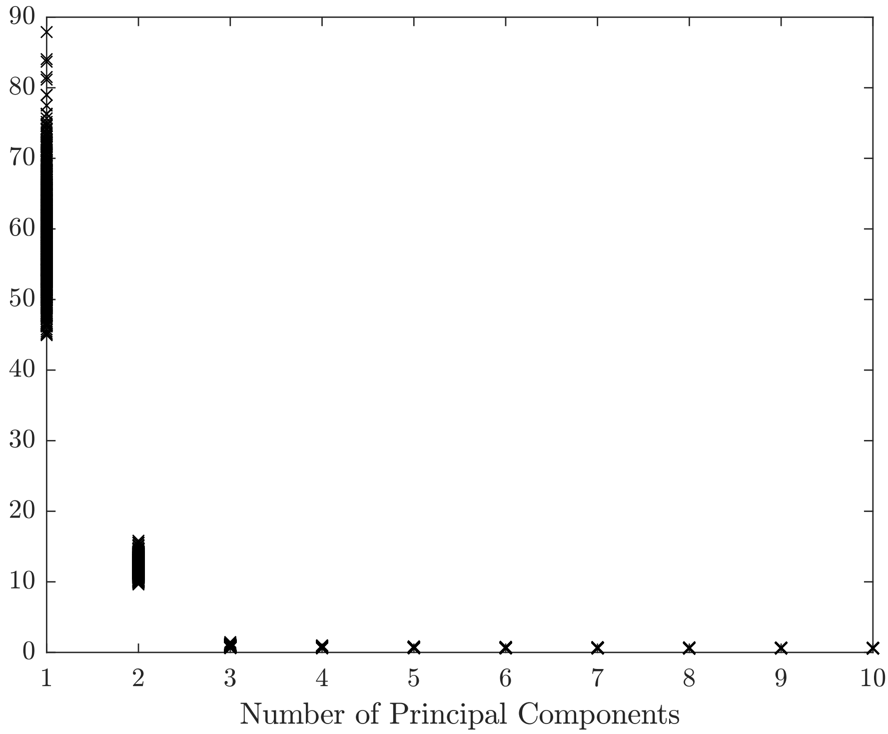

Figure J.5: Plot of the largest ten sorted eigenvalues of the sample covariance matrix of the 1000 simulated standardised return vectors for $N = 100$.

Table J.2: Computational time of our approach and the Bayesian alternative in hours

	# CPUs	$\mathbf{k=1}$	$\mathbf{k=2}$	$\mathbf{k=3}$
Kastner et al. (2017)	1	58.9626	47.0225	51.2809
Our approach	1	3.5516	3.1101	4.3578
	40	0.1558	0.1830	0.2327

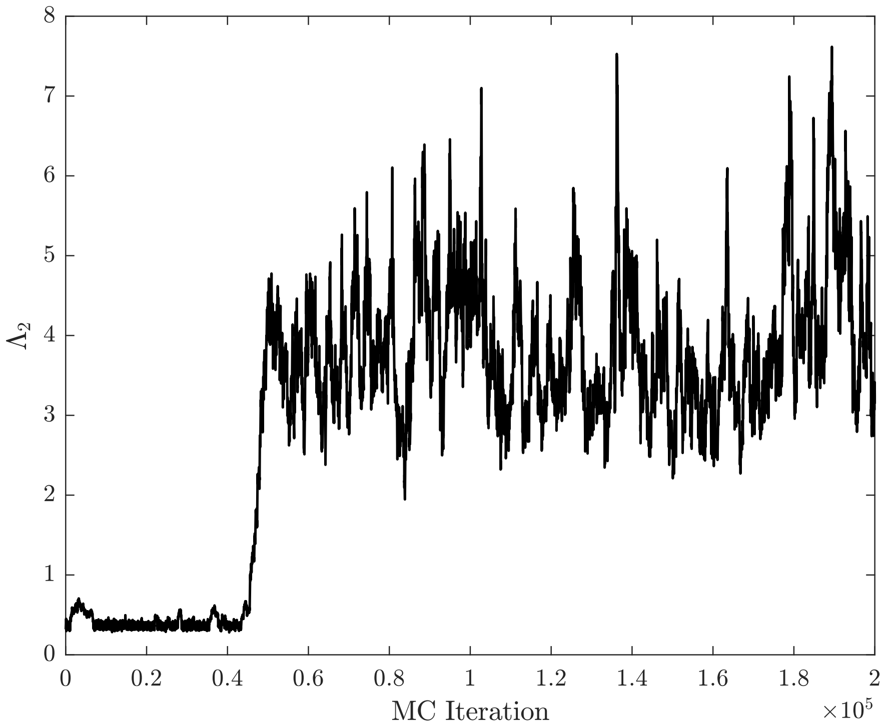

Figure J.6: Trace plot of the MCMC draws for the scaling parameter of the second factor for $k = 3$.

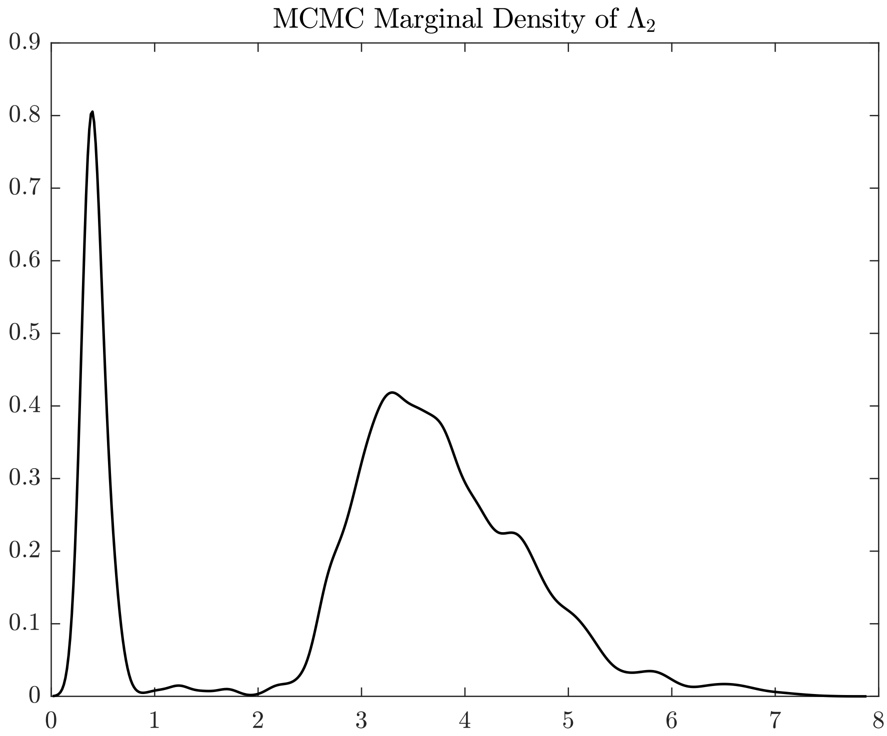

Figure J.7: Estimated MCMC density of the scaling parameter of the second factor for $k = 3$.

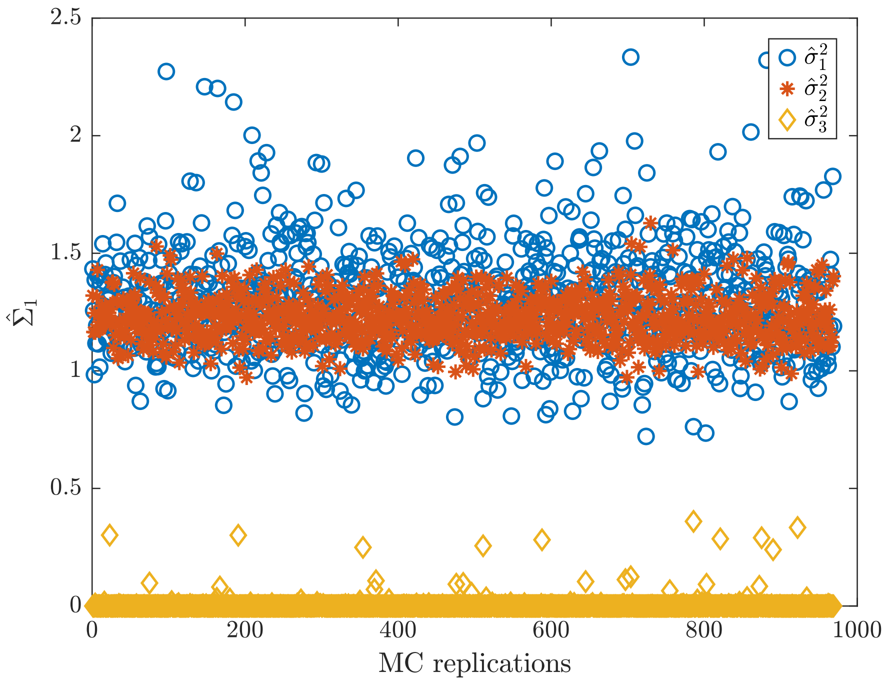

Figure J.8: Plot of the estimated unconditional factor variances for $N = 10$ for the number of replications from 1 to 1000.

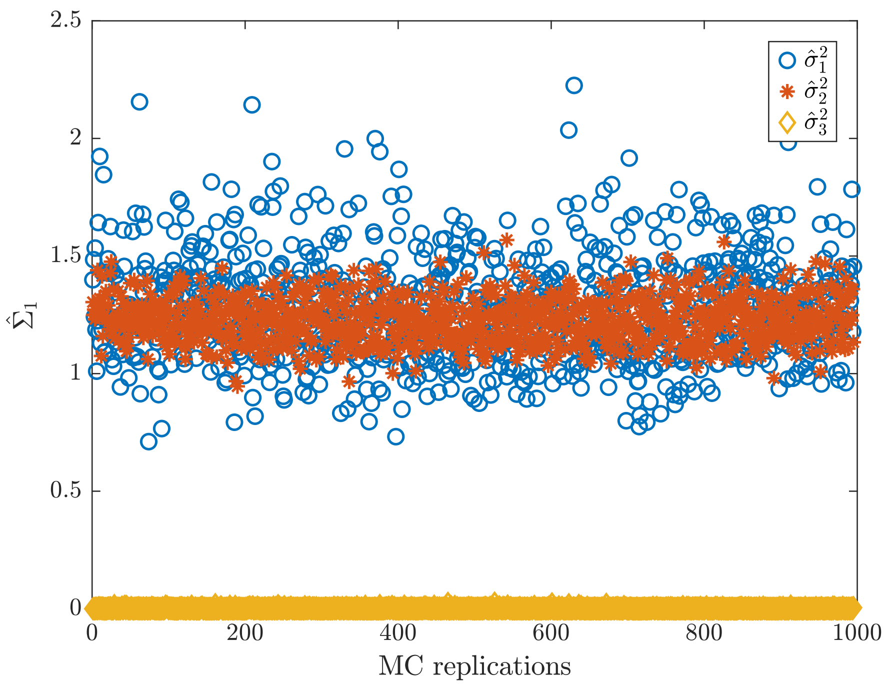

Figure J.9: Plot of the estimated unconditional factor variances for $N = 100$ for the number of replications from 1 to 1000

K Empirical Results for $k = 1$

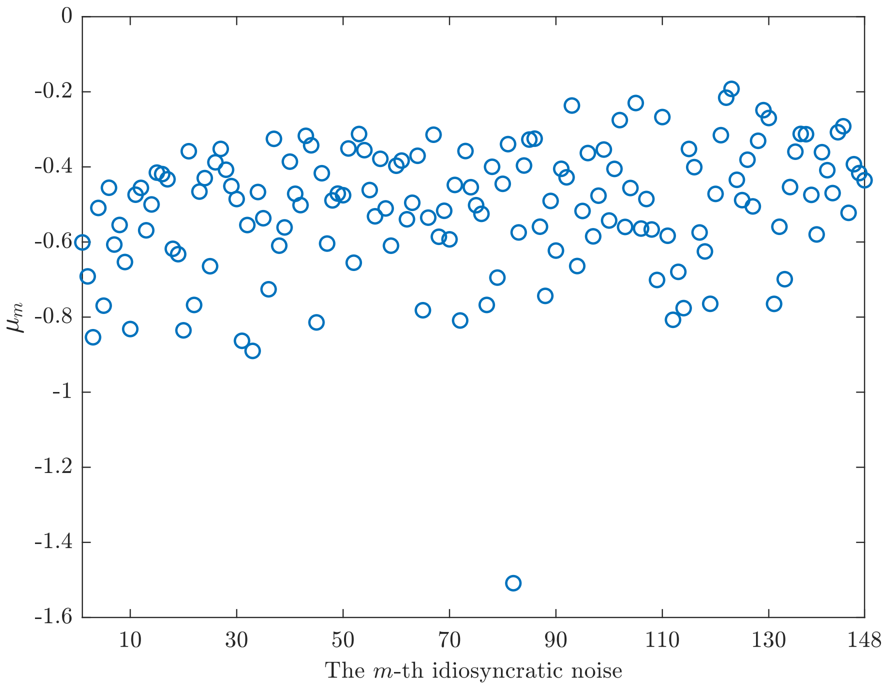

Figure K.1: Estimated constant of ARSV for each m-th idiosyncratic error, where $m = 1, \ldots, 148$ stocks

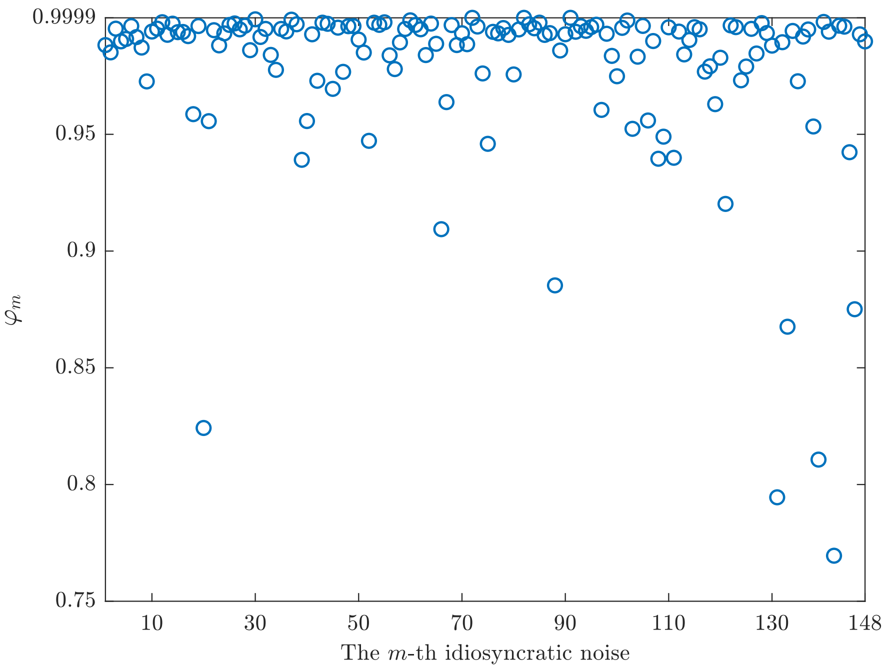

Figure K.2: Estimated autoregressive parameter of ARSV for each m-th idiosyncratic error, where $m = 1, \ldots, 148$ stocks

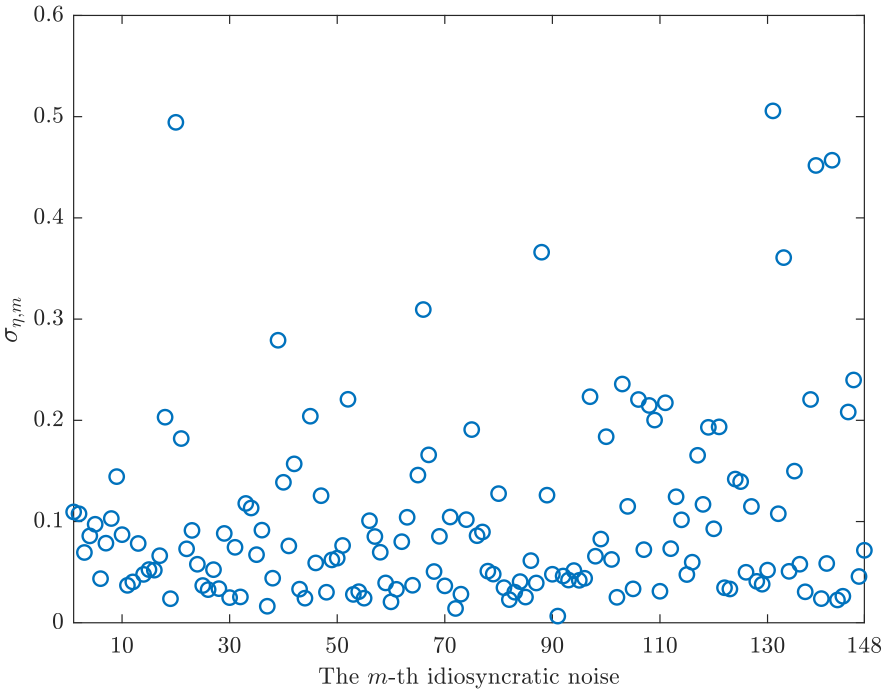

Figure K.3: Estimated standard deviation parameter of ARSV for each m-th idiosyncratic error, where $m = 1, \ldots, 148$ stocks