Supplement for "SFQRA: Scaled Factor-augmented Quantile Regression with Aggregation in Conditional Mean Forecasting"

The supplement consists of two parts. Section S1 provides the details to determine the optimal weights in quantile aggregation. Section S2 presents some additional numerical analysis.

S1 Additions on weights in quantile aggregation

We would like to point out that weighing different quantiles appropriately is crucial to SFQRA. Let us look at an experiment. Table S1 below records the optimal weights corresponding to seven different quantiles τ_j when using SFQRA or FQRA. We construct FQRA by aggregating information from traditional FQR estimates according to (11). Five different idiosyncratic errors ϵ_{t+h} are used. One may observe that the weights of SFQRA are sensitive to error distributions, while those of FQRA are less susceptible. Moreover, symmetry over quantiles is not even satisfied for SFQRA. We think that it is due to the scaling step in our SFQRA, which captures the target information as well as the idiosyncratic errors. As a consequence, it produces different factors under distinct error distributions. While the same factors independent of the errors are used in FQRA. In addition, for both FQRA and SFQRA, we observe that contrary to intuitive imagination, the weights of extreme quantiles may be weighted more heavily than a more central one. This also reflects the special aspect of using quantile aggregation for mean prediction, where the median, which is close to the mean, does not dominate the aggregation process.

Table S1: Optimal weights corresponding to five different error distributions

Method	Distribution	$\tau_1 = 0.2$	$\tau_2 = 0.3$	$\tau_3 = 0.4$	$\tau_4 = 0.5$	$\tau_5 = 0.6$	$\tau_6 = 0.7$	$\tau_7 = 0.8$
FQRA	$\mathcal{N}(0,1)$	0.2631	0.1289	0.0768	0.0427	0.0844	0.1405	0.2636
	t_2	0.2545	0.1725	0.0861	0.0440	0.0731	0.1142	0.2557
	t_4	0.2566	0.1379	0.0602	0.0364	0.0875	0.1518	0.2696
	Mixture1	0.2687	0.1295	0.1067	0.0352	0.0653	0.1479	0.2467
	Mixture2	0.2564	0.1401	0.0973	0.0307	0.0529	0.1560	0.2665
SFQRA	$\mathcal{N}(0,1)$	7.5372	-0.3923	-1.4769	-12.4376	-0.6677	1.4873	6.9500
	t_2	0.8743	-1.5161	2.8147	-2.3719	0.8409	-1.4347	1.7928
	t_4	4.1411	-2.4581	2.5449	-6.1794	-1.5974	1.0977	3.4512
	Mixture1	4.8853	2.0254	-4.2280	-4.8868	-0.7511	-0.3161	4.2714
	Mixture2	4.4252	-3.1909	0.0769	-3.7294	0.7155	2.3153	0.3873

Notes. The data generation process is the same as in Section **??** (DGP1). Average optimal weights over 5000 repetitions are recorded.

S2 Additional numerical analysis

S2.1 Numerical analysis of the SFQR method

To better illustrate the performance of the SFQR method, we consider the following data generation process:

$$X_{it} = \boldsymbol{\lambda}_i^\top \boldsymbol{f}_t + e_{it} \text{ and } y_{t+1} = \eta_{t+1} g_t + \epsilon_{t+1},$$

where $\boldsymbol{\lambda}_i = (\lambda_{i1}, \lambda_{i2}, \lambda_{i3})^\top$, $\boldsymbol{f}_t = (g_t, h_{t1}, h_{t2})^\top$ and $g_t = 0.8g_{t-1} + \xi_{t1}$, $h_{t1} = 0.5h_{t-1,1} + \xi_{t2}$, $h_{t2} = 0.2h_{t-1,2} + \xi_{t3}$. The variables ξ_{t1}, ξ_{t2}, ξ_{t3}, λ_{i1}, λ_{i2} and λ_{i3} are all independent draws from $\mathcal{N}(0,1)$, $\eta_{t+1} \sim \mathcal{U}[1,2]$, and the quantile-dependent idiosyncratic errors $\epsilon_{t+1} \sim i.i.d.\ t_v$ for $v = 2, 3, 4$, where t_v denotes the Student t-distribution with v degrees of freedom. In this setting, there are N observable covariates X_{it} $(i = 1, 2, \cdots, N)$, that are loaded on both the target-relevant factor g_t and target-irrelevant factors (h_{t1}, h_{t2}), which can be regarded as the strong factor case.

We focus on the mean quantile forecast error (MQFE), defined as follows:

$$\text{MQFE} = \frac{1}{T-h} \sum_{t=1}^{T-h} [\rho_\tau (y_{t+h} - \hat{\beta}_{\tau,0} - \hat{\boldsymbol{\beta}}_\tau^\top \boldsymbol{f}_t)],$$

under different quantile levels, forecasting with only one estimated factor. Tables S2 and S3 represent the in-sample and out-of-sample simulation results of SFQR and FQR methods for $N, T \in \{100, 200\}$ obtained from 5000 replications. Columns 3, 4, and 5 present the average values of MQFE under diverse idiosyncratic error types at the quantile $\tau = 0.2$. Meanwhile, columns 6, 7, and 8 provide corresponding results at the $\tau = 0.5$ level.

Table S2: Simulation for the conditional quantile forecasting performance (in-sample)

		$\tau = 0.2$			$\tau = 0.5$		
	method	d.f.= 2	d.f.= 3	d.f.= 4	d.f.= 2	d.f.= 3	d.f.= 4
N = 100, T = 100	FQR	0.7911	0.6247	0.5779	0.7911	0.6247	0.5779
	SFQR	0.7443	0.5737	0.5252	0.7411	0.5704	0.5209
N = 100, T = 200	FQR	0.7741	0.6076	0.5558	0.7741	0.6076	0.5558
	SFQR	0.7423	0.5717	0.5182	0.7401	0.5692	0.5161
N = 200, T = 100	FQR	0.7680	0.6171	0.5593	0.7680	0.6171	0.5593
	SFQR	0.7298	0.5703	0.5146	0.7258	0.5669	0.5108
N = 200, T = 200	FQR	0.7518	0.5900	0.5382	0.7518	0.5900	0.5382
	SFQR	0.7288	0.5646	0.5123	0.7269	0.5621	0.5111

A cursory examination of these outcomes reveals that FQR's predictive performance is subpar when compared to the SFQR approach across a range of N and T values, as well as at various quantile levels τ. The lower MQFE of SFQR for both in-sample and out-of-sample data indicates its effectiveness. This suggests that the factor estimates derived using this method perform admirably in both model fitting and prediction. Interestingly, an increase in T at the same quantile level τ is

Table S3: Simulation for the conditional quantile forecasting performance (out-of-sample)

	method	$\tau = 0.2$ d.f.= 2	$\tau = 0.2$ d.f.= 3	$\tau = 0.2$ d.f.= 4	$\tau = 0.5$ d.f.= 2	$\tau = 0.5$ d.f.= 3	$\tau = 0.5$ d.f.= 4
N = 100, T = 100	FQR	0.9329	0.7784	0.7161	0.9352	0.7756	0.7374
	SFQR	0.8939	0.7362	0.6671	0.8931	0.7310	0.6853
N = 100, T = 200	FQR	0.8849	0.7180	0.6864	0.8894	0.7236	0.6774
	SFQR	0.8599	0.6893	0.6553	0.8633	0.6915	0.6448
N = 200, T = 100	FQR	0.9399	0.7900	0.7274	0.9423	0.7797	0.7259
	SFQR	0.9092	0.7483	0.6900	0.9079	0.7336	0.6853
N = 200, T = 200	FQR	0.8831	0.7167	0.6908	0.8745	0.7147	0.6752
	SFQR	0.8661	0.6972	0.6706	0.8559	0.6937	0.6544

associated with a decrease in MQFE. This observation is intuitive as larger sample sizes offer more robust parameter estimates, thereby increasing their accuracy. In parallel, an increase in N leads to a smaller in-sample MQFE but a larger out-of-sample MQFE. This is because larger sample sizes introduce higher dimensionality when using the estimated factors for quantile forecasting, thereby including additional noise during the factor extraction process. In summary, the various facets of this numerical analysis strongly demonstrate that our scale-based methodology yields state-of-the-art forecasting performance.

S2.2 Additional numerical results of Section 4.2

In Section 4.2 of the main document, we only show the out-of-sample performance results for the eight methods, and here in Table S4 we supplement the in-sample fitting results. The performance of the DGP2 case is similar to that of the DGP1 case. While OLS is prone to over-fitting, the MSEs of the other seven methods are significantly better. Among these, sPCA has the worst out-of-sample performance, likely due to asymmetrically distributed target-related factors and temporal and cross-sectional correlation in the errors e_{it}. Scaling based solely on linear regression coefficients may negatively impact factor estimates. The poor performance of FR and FQRA compared to SFQRA is likely due to their inability to incorporate information about the target variable during factor extraction. In contrast, both SFQRA and t-SFQRA consistently outperform the other methods, with t-SFQRA slightly improving forecasting accuracy over SFQRA.

S2.3 Numerical analysis of weak relevant factors

Different from DGP2 in Section 4.2, in this subsection, we consider a weak relevant factors framework, where each factor is correlated with only a subset of covariates, and only the relevant factors are embedded in the covariates. Specifically, the data generation process (DGP3) is:

$$\mathbf{X}_t = \mathbf{\Lambda} \boldsymbol{f}_t + \boldsymbol{e}_t \text{ and } y_{t+1} = f_{t1} + \epsilon_{t+1},$$

where $\mathbf{\Lambda} = (\boldsymbol{\lambda}_1, \cdots, \boldsymbol{\lambda}_N)^\top$, $\boldsymbol{\lambda}_i = (\lambda_{i1}, \lambda_{i2})^\top$ and $\boldsymbol{f}_t = (f_{t1}, f_{t2})^\top$. The matrix $\mathbf{\Lambda}$ has only the first $d_N = \lceil T/\log(T) \rceil$ nonzero rows and satisfies that $\lambda_{i1}, \lambda_{i2} \sim i.i.d.\ \mathcal{U}(-1, 1)$ for $i = 1, 2, \cdots, d_N$. In addition, $\boldsymbol{e}_t \sim i.i.d.\ \mathcal{N}(\mathbf{0}, \mathbf{I}_N)$, $f_{t1}, f_{t2} \sim i.i.d.\ \mathcal{N}(0, 1)$ and the same six types of idiosyncratic errors ϵ_{t+1} as in DGP1 are considered. In DGP3, there are d_N observable covariates X_{it} $(i = 1, \cdots, d_N)$ loaded on the factor $\boldsymbol{f}_t$, and the remaining $(N - d_N)$ observable covariates X_{it} $(i = d_N + 1, d_N + 2, \cdots, N)$ contain only the error term e_{it}. This can be considered as the weak factor case. We are still interested in both the in-sample and out-of-sample average MSEs for the eight methods.

The simulated results are reported in Tables S5 and S6. The phenomena they exhibit are typically similar to those discussed in Section 4.2. Comparing DGP2 with DGP3, one may notice that the relationship between y_{t+1} and the factor f_{t1} in DGP3 remains unchanged at each quantile level, while the one in DGP2 is different at each quantile level as η_t is a random variable instead of a constant. Thus it makes intuitive sense to assign a more complex forecasting model such as the quantile aggregation method to DGP2 and a general linear regression method to DGP3. However, our simulated results indicate that SFQRA and t-SFQRA not only achieve the most outstanding forecasting performance in DGP2, but also outperforms linear regression methods in DGP3, which implies that SFQRA would be an efficient forecasting method in general cases, no matter how the data generation progress would be.

S2.4 Additional real data results

In Figure S1, we incorporate the t-SFQRA approach of the four different targets 'GDPC1', 'INDPRO', 'HWIURATIOx' and 'DFSARG3Q086SBEA' for the FRED-QD dataset when the number of factors is changed. It can be seen that the prediction performance of t-SFQRA and SFQRA is relatively close and performs better than the other three methods. It is noteworthy that t-SFQRA uses only some of the macroeconomic variables that have a large aggregated quantile correlation with the target, with fewer covariates, but is less time efficient. This is because t-SFQRA has to do an extra step of variable screening, which results in additional computational complexity but does not lead to additional forecasting benefits. Therefore, we do not recommend t-SFQRA for the FRED-QD dataset.

Furthermore, based on Figure 5 in the main document, we speculate that SFQRA is an appropriate choice for analyzing the FRED-QD dataset, and the t-SFQRA method would be more suitable for the FRED-MD dataset. To illustrate, we present in Figure S2 the magnitude of the aggregated quantile correlations of 'DPCERA3M086SBEA' and 'HWI'[1] with the remaining 119 variables in the FRED-MD dataset. The upper panel shows the full values, while the lower panel is truncated and hides those values that are less than 0.11. It can be found that they all have a high correlation mainly with the 'Consumption, Orders, and Inventories', 'Output and Income' and 'Labor Market' categories of macroeconomic indicators.

Finally, a selection of four variables from the block, which showcases a notably high correlation at

[1]'DPCERA3M086SBEA' denotes Real Personal Consumption Expenditures; 'HWI' denotes Help-Wanted Index for United States.

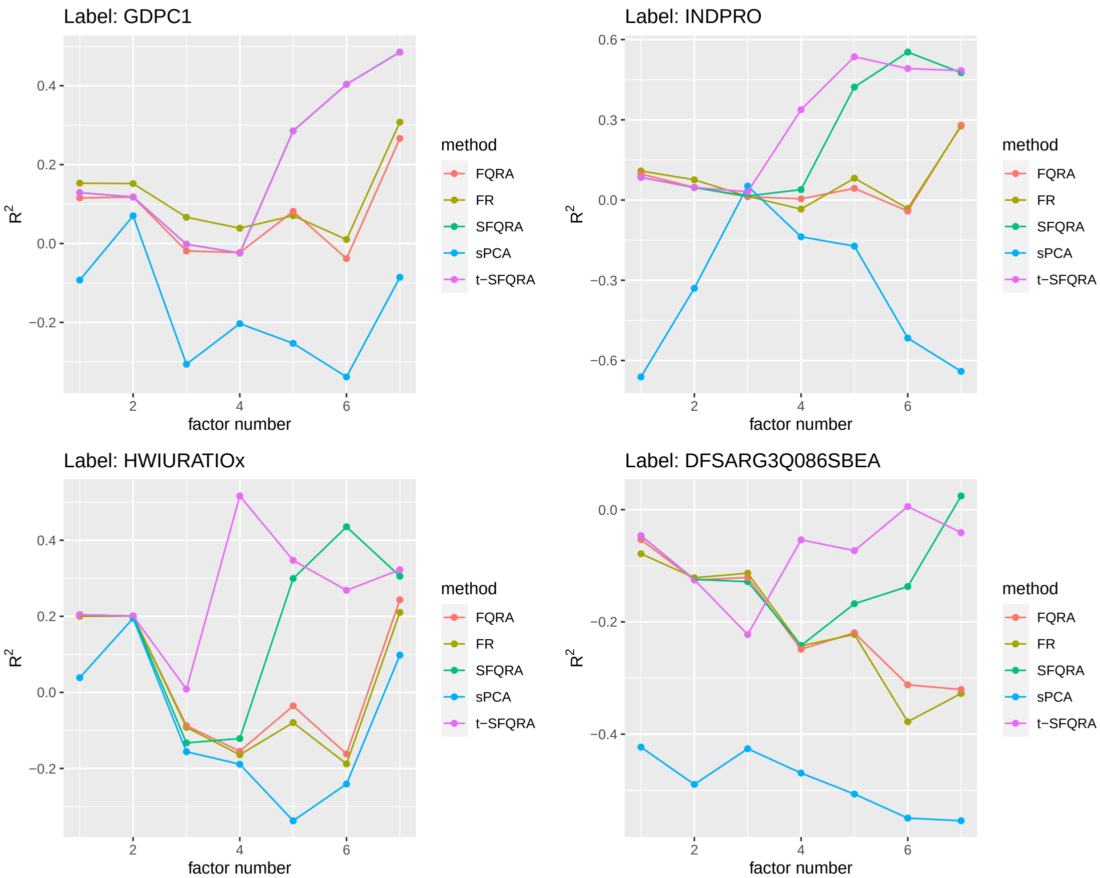

Figure S1: Out-of-sample forecasting performance with different factor numbers for FRED-QD database. (In the top left plot, the two curves representing methods SFQRA and t-SFQRA almost overlap.)

the center of Figure 5's left panel in main document, has been further evaluated. The out-of-sample performance of these variables is demonstrated in Figure S3. Upon reviewing these results, we find a striking similarity with our previous findings, thereby strengthening the recommendation for method t-SFQRA. Therefore, before adopting methods SFQRA or t-SFQRA, we may perform a simple analysis of the data to decide which method to use.

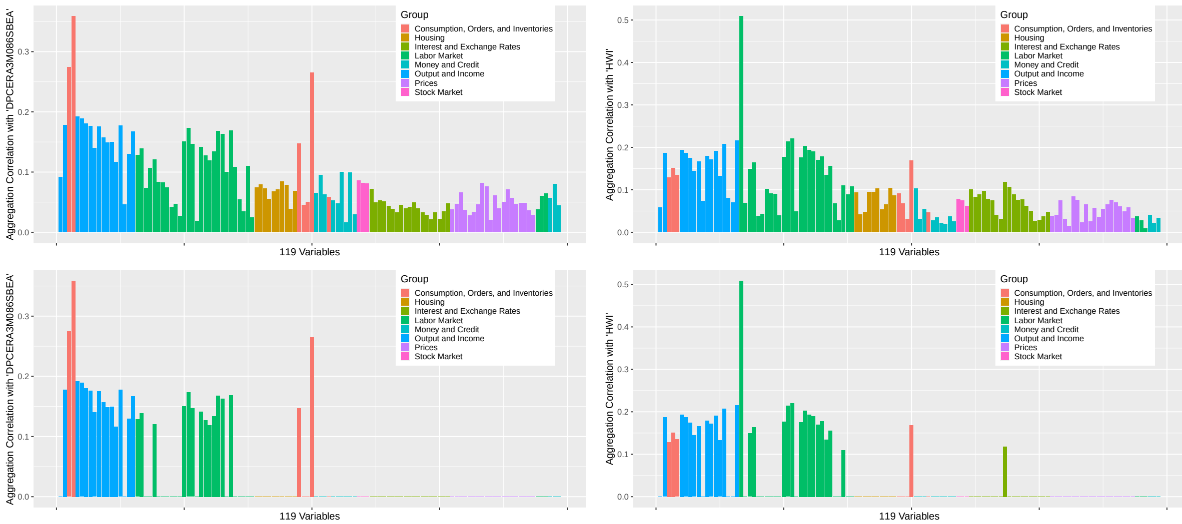

Figure S2: Aggregated Quantile Correlation for FRED-MD database. Upper panel: all values; lower panel: truncated with small values hidden.

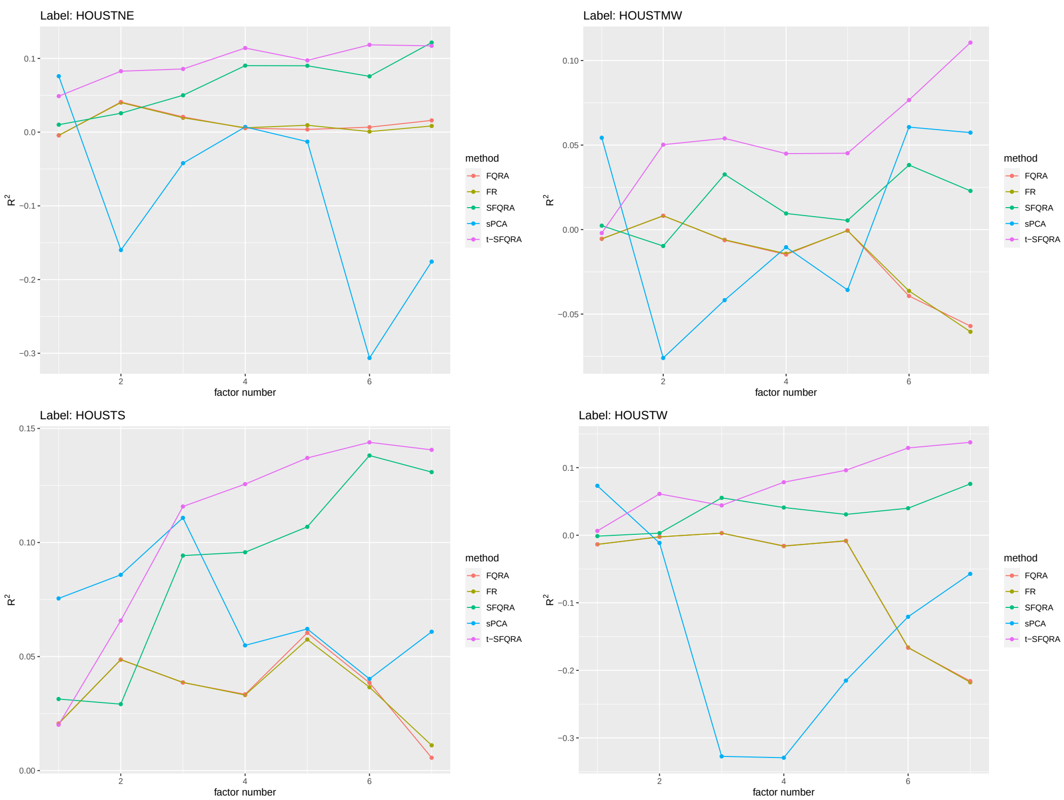

Figure S3: Out-of-sample forecasting performance of the other four targets on housing starts with different factor numbers for FRED-MD database. (In the bottom right plot, the two curves representing methods FQRA and FR almost overlap.)

Table S4: Simulation for the fitting performance in DGP2 framework (in-sample)

	distribution	$\mathcal{N}(0,1)$	t_3	t_4	t_5	Mixture1	Mixture2
	OLS	0	0	0	0	0	0
	FR	9.4537	11.4713	10.3705	10.0253	10.5991	8.8644
	sPCA	12.9510	14.7591	14.0027	13.5302	14.2276	12.5002
N = 100	t-PCA	11.3200	13.1225	12.2829	11.8490	12.5883	10.8138
T = 100	FQRA	9.4537	11.4713	10.3705	10.0253	10.5991	8.8644
	t-qcPCA	11.2645	13.0900	13.1925	11.7889	12.4000	10.7123
	SFQRA	3.5222	5.4988	4.5647	4.2312	4.8565	3.2109
	t-SFQRA	3.5499	5.6481	4.5933	4.3874	4.9876	3.4258
	OLS	1.2397	2.1693	1.7095	1.5482	1.8741	0.9874
	FR	9.6810	11.6804	10.6900	10.3606	11.0883	9.3808
	sPCA	13.2368	15.0748	14.1499	13.8811	14.5743	12.8171
N = 100	t-PCA	12.2139	14.0749	13.1250	12.8614	13.5585	11.7738
T = 200	FQRA	9.6810	11.6804	10.6900	10.3606	11.0883	9.3808
	t-qcPCA	11.6002	13.4950	12.5445	12.2961	12.9503	11.3098
	SFQRA	4.1434	5.8371	4.9909	4.7827	5.3227	3.6394
	t-SFQRA	4.2618	5.0719	4.0439	4.7273	5.4921	3.7519
	OLS	0	0	0	0	0	0
	FR	9.0586	11.2018	10.2360	9.7768	10.5683	8.8938
	sPCA	13.0747	14.7582	13.8306	13.3860	14.1794	12.7602
N = 200	t-PCA	10.7386	12.5203	11.5319	11.1809	11.8720	10.3916
T = 100	FQRA	9.0586	11.2018	10.2360	9.7768	10.5683	8.8938
	t-qcPCA	11.2672	12.9021	12.0059	11.5562	12.4006	10.9751
	SFQRA	3.1367	5.0809	3.9979	3.8029	4.4088	2.7662
	t-SFQRA	3.2613	5.1457	3.0742	3.83242	4.5611	2.1614
	OLS	0	0	0	0	0	0
	FR	9.6322	11.3898	10.6949	10.5112	10.9319	9.1216
	sPCA	13.2372	15.0227	14.3568	14.2288	14.5697	12.7361
N = 200	t-PCA	11.2469	13.0058	12.3282	12.1510	12.5175	10.7075
T = 200	FQRA	9.6322	11.3898	10.6949	10.5112	10.9319	9.1216
	t-qcPCA	11.5598	13.3384	12.6736	12.5793	12.8617	10.0613
	SFQRA	3.6758	5.2438	4.3873	4.2092	4.6897	3.0603
	t-SFQRA	3.6539	5.1012	4.4566	4.3548	4.7299	3.0590

Table S5: Simulation for the fitting performance in the weak relevant factor framework (in-sample)

	distribution	$\mathcal{N}(0,1)$	t_3	t_4	t_5	Mixture1	Mixture2
N = 100 T = 100	OLS	0	0	0	0	0	0
	FR	1.4426	3.3286	2.3949	2.0561	2.7168	0.9515
	sPCA	1.0118	2.8657	1.9417	1.6237	2.2564	0.5298
	t-PCA	1.5937	3.2107	2.0034	1.8615	2.4432	0.7809
	FQRA	1.4426	3.3285	2.3948	2.0561	2.7167	0.9515
	t-qcPCA	1.0102	2.8639	1.9351	1.6136	2.2474	0.5155
	SFQRA	0.9507	2.7692	1.8649	1.5489	2.1649	0.4748
	t-SFQRA	0.9456	2.7633	1.8612	1.5463	2.1620	0.4718
N = 100 T = 200	OLS	0.4970	1.5177	0.9816	0.8249	1.1250	0.2489
	FR	1.4689	3.4973	2.4443	2.1239	2.7289	0.9821
	sPCA	1.0160	3.0173	1.9741	1.6658	2.2732	0.5210
	t-PCA	1.1515	3.5109	2.6673	2.2388	2.3092	0.5314
	FQRA	1.4688	3.4973	2.4442	2.1238	2.7288	0.9820
	t-qcPCA	0.9948	3.0002	1.9619	1.6557	2.2589	0.5106
	SFQRA	0.9802	2.9669	1.9304	1.6229	2.2203	0.4907
	t-SFQRA	0.9778	2.9591	1.9240	1.6196	2.2179	0.4865
N = 200 T = 100	OLS	0	0	0	0	0	0
	FR	1.4519	3.3291	2.4480	2.1182	2.6901	0.9661
	sPCA	1.0093	2.8585	1.9788	1.6729	2.2284	0.5319
	t-PCA	1.2765	3.3383	2.2276	1.8743	2.3875	0.6682
	FQRA	1.4519	3.3290	2.4480	2.1180	2.6900	0.9661
	t-qcPCA	0.9912	2.8471	1.9663	1.6599	2.2134	0.5194
	SFQRA	0.9529	2.7795	1.9084	1.6117	2.1465	0.4784
	t-SFQRA	0.9496	2.7744	1.8989	1.6073	2.1402	0.4773
N = 200 T = 200	OLS	0	0	0	0	0	0
	FR	1.4724	3.6746	2.4816	2.1295	2.7146	0.9768
	sPCA	1.0158	3.1921	2.0136	1.6771	2.2571	0.5201
	t-PCA	1.1227	3.2199	2.1324	1.8743	2.3577	0.6614
	FQRA	1.4723	3.6745	2.4816	2.1294	2.7145	0.9768
	t-qcPCA	1.0095	3.1787	2.0003	1.6612	2.2431	0.5006
	SFQRA	0.9808	3.1486	1.9729	1.6344	2.2089	0.4899
	t-SFQRA	0.9777	3.1453	1.9687	1.6298	2.2023	0.4845

Table S6: Simulation for the forecasting performance in the weak relevant factor framework (out-of-sample)

	distribution	$\mathcal{N}(0,1)$	t_3	t_4	t_5	Mixture1	Mixture2
	OLS	∞	∞	∞	∞	∞	∞
	FR	1.5453	3.5834	2.5584	2.2096	2.8671	1.0233
	sPCA	1.1076	3.1798	2.1204	1.7897	2.4354	0.5849
N = 100	t-PCA	1.1084	3.1771	2.1117	1.7756	2.4213	0.5708
T = 100	FQRA	1.5452	3.5834	2.5583	2.2095	2.8670	1.0232
	t-qcPCA	1.1123	3.2198	2.1045	1.7961	2.5120	0.5844
	SFQRA	1.0767	3.1908	2.1062	1.7743	2.4314	0.5438
	t-SFQRA	1.0721	3.1910	2.1009	1.7723	2.4287	0.5401
	OLS	2.0323	6.1163	4.0312	3.4102	4.6320	1.0135
	FR	1.5291	3.5484	2.5212	2.1908	2.8007	1.0244
	sPCA	1.0640	3.0939	2.0640	1.7358	2.3604	0.5515
N = 100	t-PCA	1.0598	3.0882	2.0529	1.7277	2.3530	0.5396
T = 200	FQRA	1.5290	3.5483	2.5212	2.1907	2.8006	1.0243
	t-qcPCA	1.0652	3.0928	2.0627	1.7366	2.3619	0.5512
	SFQRA	1.0361	3.0864	2.0466	1.7131	2.3502	0.5252
	t-SFQRA	1.0322	3.0869	2.0433	1.7095	2.3486	0.5231
	OLS	∞	∞	∞	∞	∞	∞
	FR	1.5450	3.6851	2.5755	2.2109	2.8411	1.0357
	sPCA	1.1001	3.2884	2.1420	1.7912	2.4229	0.5899
N = 200	t-PCA	1.0954	3.2912	2.1376	1.7884	2.4231	0.5651
T = 100	FQRA	1.5449	3.6850	2.5755	2.2108	2.8411	1.0356
	t-qcPCA	1.0945	3.2842	2.1397	1.7825	2.4233	0.5801
	SFQRA	1.0688	3.2923	2.1254	1.7732	2.4227	0.5494
	t-SFQRA	1.0655	3.2930	2.1202	1.7698	2.4186	0.5433
	OLS	∞	∞	∞	∞	∞	∞
	FR	1.5165	3.7482	2.5834	2.1904	2.8070	1.0150
	sPCA	1.0593	3.3083	2.1095	1.7351	2.3727	0.5495
N = 200	t-PCA	1.0597	3.3109	2.1032	1.7293	2.3686	0.5443
T = 200	FQRA	1.5164	3.7482	2.5834	2.1903	2.8069	1.0150
	t-qcPCA	1.0508	3.3006	2.1011	1.7234	2.3664	0.5417
	SFQRA	1.0353	3.2931	2.0918	1.7103	2.3593	0.5226
	t-SFQRA	1.0311	3.2929	2.0883	1.7077	2.3564	0.5201

Notes. The symbol '∞' indicates that the value exceeds 10000, rendering the MSE uninformative.