Supplementary Material to "Combining variable screening methods for model averaging in high-dimensional data analysis"

Zhihao Zhao · Yuhong Yang · Li Wen

In the Supplementary Material, we present the proof of Theorem 1 in Section A.1. Proofs of Proposition 1 and Proposition 2 are detailed in Sections A.2 and A.3, respectively, and the Corollary 1 is provided in Section A.4. In addition, Sections A.5 provides a simplified example illustrating the advantages of RTMS. Sections A.6 and A.7 include some additional simulation results. Section A.8 presents the variable rankings for two real data sets using different methods discussed in the paper. Finally, Section A.9 provides a description of the R package *RTMS*, which is available at https://github.com/zhzhao07/RTMS.

A.1 Proof of Theorem 1

Recall that $\nabla_L(x_j)$ is the set of ranking procedures with the largest L ranks for the j-th variable in set M_*, S_0 represents the best set of $K-L$ methods in terms of the largeness of $A_n^{(k)}$, and $\bar{A}_n = \sum_{k\in S_0} A_n^{(k)}/(K-L)$. The value of RTMS for the j-th variable is

$$\text{RTMS}(x_j) = \frac{\sum_{k\notin\nabla_L(x_j)} R_{k,j}}{K-L}.$$

Then, we have

$$\text{RTMS}(x_j) \le \frac{\sum_{k\in S_0} R_{k,j}}{K-L} \le \frac{\sum_{k\in S_0}(1+c_k A_n^{(k)})\varrho_n^*}{K-L} \le (1+c_0\bar{A}_n)\varrho_n^*$$

Zhihao Zhao
School of Statistics, Capital University of Economics and Business, Beijing 100070, China

Yuhong Yang
Yau Mathematical Sciences Center, Tsinghua University, Beijing 100084, China

Li Wen
Center for Evidence-Based and Translational Medicine, Zhongnan Hospital of Wuhan University, Wuhan 430071, China
E-mail: wlwendy1008@163.com

with probability going to 1, where $c_0 = \max_{k \in S_0} c_k$.

As $\bar{R}_j$ is the rank of RTMS(x_j) and let $J_n = (1 + c_0 \bar{A}_n)\varrho_n^*$. We next show that there exists a constant $\tilde{c} > 0$ such that

$$\Pr\left(\max_{j \in M_*} \bar{R}_j \leq (1 + \tilde{c}\bar{A}_n)\varrho_n^*\right) \to 1.$$

If there exists a certain $j \notin M_*$ for which RTMS$(x_j) \leq J_n$ is true, then it follows that at least one ranking satisfies $R_{k,j} \leq J_n$, as determined by the computation of RTMS. Therefore, there are at most $(K-L)\lceil J_n \rceil - \varrho_n^*$ many variables with RTMS values no larger than J_n, and then

$$\begin{aligned}
\max_{j \in M_*} \bar{R}_j &\leq (K-L)\lceil J_n \rceil - \varrho_n^* + \varrho_n^* = (K-L)\lceil (1 + c_0\bar{A}_n)\varrho_n^* \rceil \\
&\leq (K-L)((1 + c_0\bar{A}_n)\varrho_n^* + 1) \\
&\leq (1 + \tilde{c}\bar{A}_n)\varrho_n^*
\end{aligned}$$

in probability, where $\tilde{c}$ needs to satisfy

$$\tilde{c} > \frac{K-L}{\bar{A}_n} + \frac{K-L}{\bar{A}_n \varrho_n^*} + (K-L)c_0 - \frac{1}{\bar{A}_n}.$$

Note that $\tilde{c}$, when $\bar{A}_n$ is of a constant order or tends to infinity at a very slow rate, $\tilde{c}$ exists as required and it is large than a linear function of c_0. $\square$

A.2 Proof of Proposition 1

For any $x_j \in M_*$, with probability going to 1, its RTMS score satisfies

$$\text{RTMS}(x_j) = \frac{\sum_{k \notin \nabla_L(x_j)} R_{k,j}}{K-L} \leq \max_{k \notin \nabla_L(x_j)} R_{k,j} \leq \varrho_n^* \max_{1 \leq k \leq K} \alpha_n^{(k)}.$$

Moreover, the worst average rank of variables in M_* under RTMS satisfies

$$\max_{x_j \in M_*} \bar{R}_j \leq K \max_{1 \leq k \leq K} q_n^{(k)} + \varrho_n^*$$

with probability approaching 1. If $K \max_{1 \leq k \leq K} q_n^{(k)} / \varrho_n^* \to 0$, then RTMS ranking is approximately consistent. $\square$

Remark 1 As shown in Table S10, the situation described above appears to be reflected in our results on the supermarket dataset. In this case, the variable X2830 was ranked very poorly by the random forest. However, due to our proposed method, this variable was ultimately ranked highly.

A.3 Proof of Proposition 2

With probability approaching 1, we bound the worst average rank by considering two cases: the well-ranked variables ($x_j \in M_* \setminus \mathcal{B}_k$) satisfy $\bar{R}_j \leq \varrho_n^* + \max_{1\leq k\leq K} K q_n^{(k)}$, while poorly ranked ones ($x_j \in \mathcal{B}_k$) have ranks at most $\varrho_n^* + K\alpha_n^o \varrho_n^*$. Then, the worst average rank of variables in M_* under RTMS satisfies

$$\max_{x_j \in M_*} \bar{R}_j \leq \varrho_n^* + \max\{ \sum_{1\leq k\leq K} q_n^{(k)},\ K\alpha_n^o \varrho_n^* \}$$

with probability approaching 1. Thus, if $\max_{1\leq k\leq K} q_n^{(k)} = O(\alpha_n^o \varrho_n^*)$, then RTMS ranking is α_n^o-consistent. □

A.4 Proof of Corollary 1

Note that $|M_{\nu_n}^c \cap M_*| = \sum_{j\in M_*} I(\bar{R}_j > \nu_n)$. For RTMS, there exists the $\tilde{c}$, the $\max_{j\in M_*} \bar{R}_j \leq (1 + \tilde{c}\bar{A}_n)\varrho_n^*$ holds with probability 1. Given any positive constant c, we have the following results:

(i) When the RTMS has consistency in ranking, then for any $c > 0$, the $|M_{\nu_n}^c \cap M_*| = o(1)$ holds.

(ii) When the RTMS has approximately consistency in ranking, suppose $\bar{A}_n = o(\varrho_n^*)$ and ϱ_n^* tends to infinity at a certain rate, then $|M_{\nu_n}^c \cap M_*| = o(\varrho_n^*)$ holds.

(iii) When the RTMS has A_n consistency in ranking and $\bar{A}_n = O(1)$, as long as c is sufficiently large and greater than or equal to $\tilde{c}$, then all variables in M_* survive. □

A.5 A toy illustration of the advantages of RTMS

Consider $n = 100$ and $p = 1000$, and we assume the true variables $x_1, x_2, \ldots, x_5$ have non-zero coefficients. Suppose the variable screening method δ_k is $A_n^{(k)}$-consistent for $k = 1, 2, 3$ with $\{\mathcal{Q}(\delta_1), \mathcal{Q}(\delta_2), \mathcal{Q}(\delta_3)\} = \{13/5, 18/5, 1000/5\}$ and $\{A_n^{(1)}, A_n^{(2)}, A_n^{(3)}\} = \{\log n, n^{1/3}, p\}$. Next, we will compare the performance of the RTMS method with the method that ranks directly based on average without trimming, referred to as RMS.

The rankings of these 5 variables under the three methods, as well as the RTMS and RMS, are shown in Table S1. Here, each of the p variables has rankings generated by different ranking methods. For simplicity, we display the rankings of the true variables among the 1000 total variables and select the ranking of a false variable x_{10} for comparison.

Based on the results in Table S1, it can be observed that, by comparing $\mathcal{Q}(\delta)$, the RTMS method inherits the strength of the top ranking method. For δ_3, with $A_n^{(3)} = p$, ranking x_5 as 1000 results in the worst performance

Table S1: Comparison of the performances between RTMS and RMS

Variable	**Rank**				
	δ_1	δ_2	δ_3	**RMS (Score)**	**RTMS (Score)**
x_1	8	1	5	1 (4.67)	2 (3.00)
x_2	1	8	2	3 (6.00)	1 (1.50)
x_3	10	7	4	5 (6.33)	8 (5.50)
x_4	11	12	3	8 (8.00)	9 (7.00)
x_5	13	18	1000	179 (343.67)	13 (15.50)
x_{10}	3	15	7	9 (8.33)	7 (5.0)
$\mathcal{Q}(\delta)$	13/5	18/5	1000/5	179/5	13/5

among all methods ($\mathcal{Q}(\delta_3) = 1000/5$). RMS, which does not use the trimmed approach to remove this poor screening method, results in poor performance.

A.6 Some additional results for variable screening

This section will present the simulation results for Cases 2-5 in Section **??** of the main text.

Table S2: The values of $\mathcal{Q}(\delta)$ in Case 2

	σ	Measure	DC-SIS	SIRS	SIS	SOIL-ARM	SOIL-AIC	SOIL-BIC	RF-S	RMedian	RTMS-3S	RTMS-4I	RTMS-7M	RTMS-7-4M
The first twelve are non-zero														
$\rho = 0.5$		Mean	32.581	31.778	62.203	1.000	1.000	1.000	57.513	3.830	20.610	1.000	3.781	1.704
	$\sigma = 1$	Median	25.792	24.458	67.708	1.000	1.000	1.000	66.375	1.292	14.375	1.000	2.500	1.000
		SD	23.673	23.565	18.611	0.000	0.000	0.000	24.144	7.937	17.744	0.006	3.824	1.789
		Mean	33.396	31.650	61.768	3.124	14.554	11.423	62.918	5.905	21.554	5.931	5.344	2.931
	$\sigma = 3$	Median	27.500	26.708	67.583	1.333	11.542	4.708	70.750	2.833	15.875	3.417	3.750	2.000
		SD	23.702	22.876	17.582	8.429	16.110	16.509	21.044	9.000	18.072	7.652	4.388	2.900
		Mean	41.230	39.985	65.348	36.898	44.883	45.880	65.559	21.619	29.331	35.178	17.763	13.234
	$\sigma = 5$	Median	38.417	38.958	69.208	36.083	44.917	46.000	73.625	12.125	24.583	27.250	11.708	7.208
		SD	24.372	24.538	15.134	26.489	22.210	24.535	18.636	21.939	21.233	24.384	16.202	15.923
$\rho = 0.8$		Mean	45.687	46.265	51.856	16.098	16.169	15.123	51.215	5.371	45.118	9.067	5.809	2.020
	$\sigma = 1$	Median	49.000	48.500	56.500	6.625	6.375	6.458	58.375	1.833	47.583	4.458	4.292	1.167
		SD	24.186	23.176	22.133	20.848	21.948	20.428	28.313	9.881	24.551	11.922	4.991	2.138
		Mean	45.207	45.993	51.333	56.567	53.908	52.993	56.395	11.583	45.185	34.118	9.655	4.480
	$\sigma = 3$	Median	46.250	45.667	53.917	60.208	59.083	56.958	64.583	4.917	46.542	32.625	7.625	2.833
		SD	23.839	23.478	21.510	19.373	20.256	20.314	25.596	16.538	24.093	18.009	7.059	4.463
		Mean	44.615	42.682	54.450	64.141	62.890	64.032	59.848	22.859	43.247	46.585	17.339	12.482
	$\sigma = 5$	Median	44.833	41.750	58.250	66.625	66.125	68.333	67.417	14.208	44.750	45.750	12.750	7.167
		SD	23.761	23.888	20.486	15.294	15.222	15.453	22.226	20.668	23.198	18.567	13.886	13.888
The random position of twelve are non-zero														
$\rho = 0.5$		Mean	14.955	13.263	76.504	1.000	1.000	1.000	56.442	4.397	14.731	1.003	3.619	1.993
	$\sigma = 1$	Median	9.083	8.500	78.042	1.000	1.000	1.000	69.875	2.458	8.708	1.000	3.042	1.417
		SD	15.620	13.404	6.144	0.000	0.000	0.000	27.422	5.285	15.646	0.015	2.538	1.459
		Mean	19.693	16.947	77.093	1.003	14.813	3.703	63.792	7.987	19.035	3.327	6.311	4.171
	$\sigma = 3$	Median	14.250	10.250	78.292	1.000	15.083	1.000	71.792	4.292	11.042	2.917	4.917	2.750
		SD	18.127	17.406	5.365	0.014	1.448	5.597	21.250	11.526	18.961	1.804	5.866	3.835
		Mean	27.319	25.442	76.663	4.358	17.691	13.161	68.182	13.372	27.277	8.203	10.903	9.304
	$\sigma = 5$	Median	22.625	20.250	78.000	1.458	15.667	14.208	76.417	7.792	21.833	5.208	7.208	6.250
		SD	19.634	19.445	6.524	11.362	9.682	13.339	18.177	15.201	19.763	11.746	10.740	9.684
$\rho = 0.8$		Mean	26.270	25.545	77.932	1.000	1.000	1.000	60.878	9.972	26.629	1.012	5.911	3.554
	$\sigma = 1$	Median	21.667	19.250	79.917	1.000	1.000	1.000	68.250	5.708	19.875	1.000	4.375	2.542
		SD	17.648	18.255	5.196	0.000	0.000	0.000	20.605	11.799	19.234	0.029	4.711	2.653
		Mean	28.829	27.461	76.615	1.409	11.302	1.935	64.314	10.849	28.795	2.669	7.773	4.986
	$\sigma = 3$	Median	23.542	20.958	78.375	1.000	14.417	1.000	71.667	5.958	22.500	2.667	5.792	3.750
		SD	19.519	20.148	6.315	3.187	6.048	3.141	19.577	13.041	20.438	1.486	6.056	3.887
		Mean	39.511	36.465	77.062	20.648	28.944	24.603	70.945	23.214	38.705	19.621	18.454	16.908
	$\sigma = 5$	Median	35.542	32.083	78.583	6.500	17.500	16.583	76.833	14.417	34.000	9.542	11.042	10.583
		SD	20.436	19.919	6.118	25.413	20.727	23.610	14.286	20.971	20.666	22.153	17.911	17.315

Table S3: The values of $\mathcal{Q}(\delta)$ in Case 3

	σ	Measure	DC-SIS	SIRS	SIS	SOIL-ARM	SOIL-AIC	SOIL-BIC	RF-S	Rmedian	RTMS-3S	RTMS-4I	RTMS-7M	RTMS-7-4M
$\rho = 0.5$		Mean	23.735	21.719	63.593	1.000	1.000	1.000	48.690	5.831	21.071	1.002	3.649	2.161
	$\sigma = 1$	Median	14.400	12.000	68.000	1.000	1.000	1.000	50.950	1.550	10.850	1.000	1.800	1.150
		SD	24.980	24.489	27.303	0.000	0.000	0.000	34.717	11.782	24.794	0.014	5.065	2.503
		Mean	30.427	25.865	66.416	11.457	12.610	10.960	54.301	13.913	25.751	10.664	10.737	8.666
	$\sigma = 3$	Median	19.200	17.200	74.050	1.500	1.200	1.200	56.850	4.000	15.300	1.300	3.500	2.200
		SD	28.283	25.545	26.728	22.759	24.901	21.503	31.370	21.532	26.676	21.351	17.822	16.354
		Mean	34.033	31.960	67.832	44.062	48.497	41.955	61.850	30.193	31.565	38.358	27.549	26.412
	$\sigma = 5$	Median	23.500	22.850	72.250	43.800	47.950	37.000	70.200	21.550	22.400	30.350	18.900	15.950
		SD	29.416	29.350	23.893	33.045	31.191	32.130	30.785	27.570	28.261	30.400	25.996	26.117
$\rho = 0.8$		Mean	1.037	1.023	2.688	1.001	1.001	1.001	2.597	1.003	1.022	1.001	1.003	1.002
	$\sigma = 1$	Median	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		SD	0.157	0.102	10.683	0.007	0.007	0.007	9.985	0.016	0.100	0.007	0.023	0.014
		Mean	1.055	1.041	2.375	16.700	21.479	20.312	5.494	1.495	1.032	11.072	2.306	1.148
	$\sigma = 3$	Median	1.000	1.000	1.000	1.800	2.050	2.150	1.000	1.000	1.000	1.650	1.000	1.000
		SD	0.256	0.203	9.762	26.791	30.865	30.274	16.587	3.488	0.169	17.931	3.009	0.926
		Mean	1.057	1.069	4.497	41.718	44.387	44.108	10.726	3.390	1.045	25.171	4.980	1.878
	$\sigma = 5$	Median	1.000	1.000	1.000	41.800	45.650	42.600	1.200	1.000	1.000	18.800	3.300	1.000
		SD	0.219	0.289	14.694	33.173	32.622	33.494	23.152	10.186	0.188	22.313	5.902	3.839

Table S4: The values of $\mathcal{Q}(\delta)$ in Case 4 at the 0.05 significance level

		σ	F-test for ϱ_n^{G*}	AIC for ϱ_n^{G*}	Measure	DC-SIS	SIRS	SIS	SOIL-ARM	SOIL-AIC	SOIL-BIC	RF-S	RTMS-3S	RTMS-4I	RTMS-7M	RTMS-7-4M
$\alpha = 1$	$\rho = 0$	$\sigma = 1$	17	14	Mean	52.681	51.993	54.222	28.078	24.864	23.394	55.222	52.082	25.027	30.019	27.204
					Median	54.324	53.618	55.618	30.618	22.324	19.235	56.735	54.147	19.676	26.147	23.088
					SD	5.317	5.943	4.513	18.773	19.301	18.579	3.873	6.274	19.092	15.291	17.150
		$\sigma = 3$	7	6	Mean	68.294	67.707	107.864	16.539	28.109	18.024	103.918	65.063	17.308	26.149	20.942
					Median	66.571	64.714	116.143	2.929	22.714	3.643	115.214	59.929	5.857	17.429	10.786
					SD	38.786	40.186	31.368	32.577	27.264	34.080	35.345	40.877	29.744	27.968	28.140
		$\sigma = 5$	4	3	Mean	39.588	33.674	114.736	10.351	45.153	17.974	109.401	30.780	13.948	14.790	12.303
					Median	15.750	14.000	119.125	1.500	41.250	2.000	96.375	11.375	5.250	5.000	3.750
					SD	52.682	45.241	80.663	33.311	30.111	39.264	90.826	45.518	33.320	28.205	26.113
	$\rho = 0.5$	$\sigma = 1$	20	16	Mean	37.932	37.115	44.296	21.220	18.194	17.249	44.504	36.302	16.346	18.387	16.314
					Median	40.475	38.950	45.975	21.825	9.525	8.650	46.750	38.550	10.950	13.925	10.275
					SD	9.741	10.353	5.118	17.246	16.652	15.508	5.921	10.493	14.453	13.141	13.891
		$\sigma = 3$	10	8	Mean	19.219	16.382	55.465	19.958	19.791	19.190	50.569	14.803	15.528	9.310	7.653
					Median	8.350	6.900	64.550	4.250	7.200	3.500	48.250	5.900	3.650	3.550	2.500
					SD	23.095	20.481	31.878	28.179	27.689	28.003	37.803	19.354	23.941	14.653	13.196
		$\sigma = 5$	7	5	Mean	6.415	4.868	35.214	27.813	36.201	23.655	35.574	4.178	20.354	6.911	4.217
					Median	2.000	1.643	2.571	3.286	25.000	2.857	7.857	1.571	3.643	1.929	1.429
					SD	13.695	9.217	44.746	42.758	34.988	37.562	47.983	8.491	33.432	10.994	7.765
$\alpha = 1.5$	$\rho = 0$	$\sigma = 1$	8	7	Mean	99.859	95.866	106.938	15.089	15.062	15.203	92.346	96.292	15.428	35.154	24.550
					Median	106.313	100.250	112.750	2.563	2.688	2.688	98.000	100.125	2.688	27.750	15.063
					SD	20.683	21.598	16.584	29.350	29.681	30.766	22.702	22.093	29.591	24.802	26.542
		$\sigma = 3$	3	3	Mean	47.478	40.375	146.577	1.198	9.262	1.262	38.260	40.368	1.738	7.765	4.533
					Median	11.833	11.833	149.667	1.000	1.000	1.000	11.167	9.667	1.000	2.000	1.333
					SD	74.527	61.419	106.908	0.815	17.524	0.760	63.537	65.783	1.978	13.315	7.611
		$\sigma = 5$	2	2	Mean	8.370	7.403	89.025	1.128	59.013	3.933	6.930	6.943	1.673	2.355	2.043
					Median	1.500	1.500	1.000	1.000	62.750	1.000	1.500	1.500	1.000	1.000	1.000
					SD	31.118	26.587	143.924	0.840	24.630	13.033	17.828	27.194	1.642	5.446	4.273
	$\rho = 0.5$	$\sigma = 1$	11	10	Mean	62.388	60.385	75.086	17.762	19.074	18.811	59.043	59.408	17.262	24.145	19.238
					Median	68.045	66.091	78.864	2.727	2.773	2.727	62.455	61.318	2.727	17.545	10.045
					SD	23.071	23.998	13.888	25.712	26.820	27.019	21.216	24.065	25.139	21.879	21.749
		$\sigma = 3$	5	4	Mean	5.254	4.366	42.795	10.910	14.800	12.099	5.148	4.461	7.447	3.669	2.620
					Median	1.200	1.200	1.200	1.400	1.200	1.200	1.600	1.200	1.200	1.000	1.000
					SD	12.966	8.449	60.879	29.849	40.646	35.575	10.293	11.453	19.906	8.692	5.955
		$\sigma = 5$	3	2	Mean	1.015	1.020	2.542	3.900	28.375	4.558	1.038	1.010	2.688	1.375	1.032
					Median	1.000	1.000	1.000	1.333	25.667	1.000	1.000	1.000	1.000	1.000	1.000
					SD	0.084	0.155	21.731	18.974	35.196	25.427	0.270	0.066	11.655	2.455	0.152

In Table S4, we conducted two searches for M_G^* based on the F-test and AIC, and we have listed the sizes ϱ_n^{G*} of the sets of important variables. Overall, as σ increases, the size of the important variable set that needs to be determined decreases. Furthermore, we present the values of $\mathcal{Q}(\delta)$ under the large ϱ_n^{G*} for eleven methods in the case of two decaying coefficients. Findings are as follows:

1. When the size of the model ϱ_n^{G*} is relatively large, the $\mathcal{Q}(\delta)$ values provided by individual screening methods are generally large, especially for the three screening methods. In contrast, the performances of all four RTMS methods are superior to those of individual methods, with RTMS-4I and RTMS-7-4M showing particularly better performance.
2. When the value of ϱ_n^{G*} is small, different methods tend to have smaller estimates for $\mathcal{Q}(\delta)$. However, RTMS still outperforms individual methods.

Table S5 displays the values of $\mathcal{Q}(\delta)$ for the nonlinear Case 5. According to the setup, screening out x_{12} is very challenging. Therefore, individual methods

Table S5: The values of $\mathcal{Q}(\delta)$ in Case 5

	ϱ_n^*	σ	Measure	DC-SIS	SIRS	SIS	SOIL-ARM	SOIL-AIC	SOIL-BIC	RF-S	RTMS-3S	RTMS-4I	RTMS-7M	RTMS-7-4M
$\rho = 0.5$	4	$\sigma = 1$	Mean	106.410	168.228	168.323	173.285	160.123	166.413	108.160	104.000	123.815	97.370	88.018
			Median	103.000	180.500	184.625	190.625	174.250	175.125	98.250	90.250	115.875	77.750	71.750
			SD	73.759	57.905	60.887	61.814	65.761	66.837	73.697	76.103	64.269	68.244	70.905
		$\sigma = 3$	Mean	118.135	186.060	161.623	162.435	167.600	168.278	133.860	112.823	132.570	110.938	103.755
			Median	115.750	199.875	173.000	178.500	181.750	183.250	132.125	91.250	128.375	104.000	94.750
			SD	70.619	53.262	63.729	63.605	64.670	65.113	80.870	72.298	67.134	65.675	68.363
		$\sigma = 5$	Mean	104.735	176.428	158.355	172.690	172.013	168.068	151.878	94.530	150.373	112.245	109.063
			Median	91.375	190.125	165.125	177.750	188.750	179.500	163.625	84.250	158.625	106.125	98.250
			SD	68.541	53.919	66.464	58.127	60.359	63.409	77.905	63.524	69.428	64.949	67.882
	3	$\sigma = 1$	Mean	17.770	188.533	159.213	193.583	182.680	187.550	23.687	12.577	109.833	47.480	29.530
			Median	8.667	214.333	159.333	214.000	181.667	196.667	1.667	8.000	95.833	40.500	21.167
			SD	25.896	94.991	102.302	99.753	97.654	100.968	68.406	14.188	73.670	36.989	27.989
		$\sigma = 3$	Mean	25.400	203.323	167.547	180.040	186.157	182.797	46.347	16.923	116.543	57.143	36.520
			Median	12.500	229.833	169.333	208.667	210.500	210.000	2.667	10.000	110.833	50.000	22.667
			SD	35.473	93.102	94.850	100.789	97.094	100.931	101.118	20.763	82.117	46.752	34.203
		$\sigma = 5$	Mean	36.083	203.407	172.030	200.290	201.037	197.383	86.300	22.797	141.500	70.173	52.053
			Median	26.167	211.333	168.333	212.000	225.000	205.667	9.500	17.167	127.333	55.333	36.667
			SD	38.109	86.704	100.303	93.236	88.460	96.621	117.442	22.846	89.509	56.810	50.335
$\rho = 0.8$	4	$\sigma = 1$	Mean	94.689	162.746	132.296	177.448	174.675	173.220	110.571	70.211	144.550	94.501	85.361
			Median	77.125	170.750	133.000	198.875	182.250	181.375	97.750	45.750	148.125	81.125	67.125
			SD	76.540	64.826	76.169	60.168	51.731	54.383	72.158	64.576	56.819	58.564	61.493
		$\sigma = 3$	Mean	93.009	159.043	136.199	180.306	173.815	177.254	127.919	73.570	145.675	100.890	92.846
			Median	73.250	172.000	151.500	188.750	187.125	187.125	123.625	48.250	148.750	85.875	74.125
			SD	73.736	68.312	79.960	51.270	53.464	53.505	73.762	67.170	59.630	64.319	67.234
		$\sigma = 5$	Mean	91.500	167.090	124.868	179.208	177.184	178.856	117.435	72.831	139.960	93.165	84.756
			Median	76.625	184.500	128.000	188.500	188.875	186.500	114.500	54.000	144.875	79.500	68.625
			SD	69.553	65.121	73.869	53.845	53.497	52.968	70.566	62.906	57.370	56.294	59.648
	3	$\sigma = 1$	Mean	4.927	185.115	147.248	181.503	178.543	178.567	1.810	6.133	98.307	42.375	21.673
			Median	4.333	187.000	151.833	194.000	190.167	195.000	1.667	5.667	88.833	37.167	18.333
			SD	3.203	99.615	103.085	104.005	99.198	99.804	0.374	1.644	64.902	31.686	16.377
		$\sigma = 3$	Mean	5.030	181.400	148.570	181.345	184.250	186.040	1.890	6.163	100.683	44.792	23.628
			Median	4.333	190.667	146.833	190.167	198.333	210.333	1.667	6.000	94.333	40.667	19.333
			SD	2.720	106.443	105.182	98.556	97.114	100.501	0.481	1.444	65.346	33.174	19.057
		$\sigma = 5$	Mean	6.250	190.428	141.723	181.322	192.265	193.515	1.992	6.845	105.185	48.812	25.080
			Median	4.333	203.500	145.333	189.833	209.833	212.333	2.000	6.000	96.667	44.667	19.667
			SD	6.083	101.324	100.608	100.214	93.376	97.954	0.634	3.234	64.832	35.912	20.617

may show their best performance by selecting only the other three variables. Here, we present the cases where $\varrho_n^* = 4$ and $\varrho_n^* = 3$. The findings of the simulation results are as follows.

1. When $\varrho_n^* = 4$, it is challenging for most methods to screen all 4 variables. However, at this juncture, the RTMS method still exhibits some improvement compared to individual methods.
2. When $\varrho_n^* = 3$, the nonlinear variable x_{12} does not need to be screened, and we can observe that DC-SIS outperforms all individual methods. However, RTMS-3S performs similarly to DC-SIS. Moreover, when the variables exhibit higher correlations, RF performs the best.

Table S6: The probability of selecting true variables under different thresholds in Case 1

	σ	ν_n	DC-SIS	SIRS	SIS	SOIL-ARM	SOIL-AIC	SOIL-BIC	RF-S	RMedian	RTMS-3S	RTMS-4I	RTMS-7M	RTMS-7-4M
$\rho = 0.5$	$\sigma = 1$	2n	0.888	0.897	0.823	1.000	1.000	1.000	0.758	0.980	0.928	1.000	0.974	0.990
		3n/logn	0.938	0.943	0.839	1.000	1.000	1.000	0.802	0.992	0.960	1.000	0.992	0.998
		n/logn	0.888	0.897	0.823	1.000	1.000	1.000	0.758	0.980	0.928	1.000	0.974	0.990
	$\sigma = 3$	2n	0.937	0.948	0.828	1.000	0.984	0.985	0.821	0.992	0.959	0.996	1.000	1.000
		3n/logn	0.884	0.881	0.748	0.997	0.568	0.848	0.736	0.968	0.909	0.983	0.961	0.983
		n/logn	0.814	0.825	0.726	0.993	0.196	0.711	0.678	0.896	0.856	0.884	0.888	0.937
	$\sigma = 5$	2n	0.928	0.938	0.768	0.932	0.838	0.832	0.736	0.952	0.950	0.896	0.972	0.972
		3n/logn	0.822	0.844	0.662	0.836	0.446	0.614	0.636	0.856	0.874	0.760	0.846	0.886
		n/logn	0.746	0.754	0.644	0.762	0.160	0.418	0.578	0.736	0.796	0.612	0.690	0.782
$\rho = 0.8$	$\sigma = 1$	2n	1.000	1.000	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000
		3n/logn	1.000	1.000	1.000	1.000	0.988	0.988	1.000	1.000	1.000	1.000	1.000	1.000
		n/logn	1.000	1.000	1.000	1.000	0.798	0.790	1.000	1.000	1.000	0.988	1.000	1.000
	$\sigma = 3$	2n	1.000	1.000	1.000	0.820	0.712	0.732	1.000	1.000	1.000	0.928	1.000	1.000
		3n/logn	1.000	1.000	1.000	0.718	0.584	0.602	1.000	1.000	1.000	0.664	0.980	1.000
		n/logn	1.000	1.000	1.000	0.646	0.510	0.582	1.000	1.000	1.000	0.614	0.798	1.000
	$\sigma = 5$	2n	1.000	1.000	1.000	0.692	0.694	0.684	0.992	0.998	1.000	0.874	1.000	1.000
		3n/logn	1.000	1.000	1.000	0.562	0.312	0.526	0.984	0.994	1.000	0.592	0.930	0.998
		n/logn	1.000	1.000	1.000	0.536	0.106	0.492	0.982	0.990	1.000	0.558	0.684	0.994

Table S7: The probability of selecting true variables under different thresholds in Case 2

	σ	ν_n	DC-SIS	SIRS	SIS	SOIL-ARM	SOIL-AIC	SOIL-BIC	RF-S	RMedian	RTMS-3S	RTMS-4I	RTMS-7M	RTMS-7-4M
$\rho = 0.5$	$\sigma = 1$	2n	0.960	0.960	0.844	1.000	1.000	1.000	0.884	0.999	0.980	1.000	1.000	1.000
		3n/logn	0.897	0.898	0.769	1.000	1.000	1.000	0.812	0.993	0.928	1.000	0.993	0.999
		n/logn	0.818	0.825	0.749	1.000	1.000	1.000	0.751	0.978	0.854	1.000	0.965	0.991
	$\sigma = 3$	2n	0.958	0.961	0.829	0.998	0.992	0.992	0.861	0.998	0.978	0.999	1.000	1.000
		3n/logn	0.883	0.889	0.740	0.998	0.869	0.977	0.764	0.989	0.919	0.987	0.988	0.997
		n/logn	0.790	0.800	0.720	0.988	0.720	0.932	0.704	0.950	0.835	0.946	0.928	0.980
	$\sigma = 5$	2n	0.940	0.945	0.808	0.949	0.938	0.940	0.825	0.978	0.963	0.957	0.986	0.991
		3n/logn	0.861	0.873	0.712	0.904	0.534	0.840	0.718	0.938	0.893	0.898	0.937	0.957
		n/logn	0.761	0.765	0.687	0.855	0.173	0.774	0.653	0.852	0.805	0.811	0.818	0.888
$\rho = 0.8$	$\sigma = 1$	2n	0.933	0.933	0.898	0.985	0.988	0.990	0.920	0.997	0.941	0.995	1.000	1.000
		3n/logn	0.888	0.886	0.850	0.983	0.985	0.985	0.866	0.991	0.890	0.986	0.984	0.998
		n/logn	0.859	0.860	0.842	0.911	0.912	0.911	0.820	0.969	0.861	0.928	0.935	0.990
	$\sigma = 3$	2n	0.933	0.933	0.892	0.879	0.913	0.905	0.893	0.991	0.933	0.949	0.997	1.000
		3n/logn	0.875	0.878	0.845	0.843	0.844	0.844	0.828	0.965	0.878	0.869	0.958	0.989
		n/logn	0.853	0.854	0.833	0.739	0.731	0.732	0.775	0.924	0.858	0.756	0.829	0.946
	$\sigma = 5$	2n	0.934	0.938	0.900	0.820	0.839	0.837	0.871	0.978	0.943	0.912	0.991	0.993
		3n/logn	0.884	0.885	0.837	0.737	0.688	0.728	0.793	0.926	0.886	0.762	0.928	0.963
		n/logn	0.856	0.858	0.825	0.635	0.607	0.611	0.748	0.871	0.863	0.652	0.741	0.898

Table S8: The probability of selecting true variables under different thresholds in Case 3

	σ	ν_n	DC-SIS	SIRS	SIS	SOIL-ARM	SOIL-AIC	SOIL-BIC	RF-S	RMedian	RTMS-3S	RTMS-4I	RTMS-7M	RTMS-7-4M
$\rho = 0.5$	$\sigma = 1$	2n	0.956	0.958	0.834	0.936	0.932	0.939	0.889	0.964	0.958	0.946	0.972	0.970
		3n/logn	0.877	0.889	0.752	0.892	0.743	0.894	0.806	0.904	0.892	0.901	0.915	0.923
		n/logn	0.810	0.818	0.732	0.859	0.580	0.862	0.751	0.863	0.833	0.847	0.861	0.879
	$\sigma = 3$	2n	0.961	0.970	0.850	0.989	0.988	0.989	0.911	0.988	0.970	0.990	0.993	0.994
		3n/logn	0.900	0.904	0.783	0.982	0.978	0.979	0.844	0.966	0.910	0.980	0.976	0.981
		n/logn	0.838	0.840	0.764	0.968	0.968	0.970	0.801	0.931	0.852	0.968	0.935	0.957
	$\sigma = 5$	2n	0.956	0.958	0.834	0.936	0.932	0.939	0.889	0.964	0.958	0.946	0.972	0.970
		3n/logn	0.877	0.889	0.752	0.892	0.743	0.894	0.806	0.904	0.892	0.901	0.915	0.923
		n/logn	0.810	0.818	0.732	0.859	0.580	0.862	0.751	0.863	0.833	0.847	0.861	0.879
$\rho = 0.8$	$\sigma = 1$	2n	1.000	1.000	0.999	1.000	1.000	1.000	0.998	1.000	1.000	1.000	1.000	1.000
		3n/logn	1.000	1.000	0.996	1.000	1.000	1.000	0.997	1.000	1.000	1.000	1.000	1.000
		n/logn	1.000	1.000	0.996	1.000	1.000	1.000	0.988	1.000	1.000	1.000	1.000	1.000
	$\sigma = 3$	2n	1.000	1.000	0.999	0.984	0.972	0.975	0.994	0.999	1.000	0.989	1.000	1.000
		3n/logn	1.000	1.000	0.999	0.971	0.964	0.964	0.989	0.999	1.000	0.969	0.995	1.000
		n/logn	1.000	1.000	0.999	0.964	0.955	0.955	0.983	0.996	1.000	0.957	0.981	0.998
	$\sigma = 5$	2n	1.000	1.000	0.996	0.940	0.931	0.933	0.988	0.997	1.000	0.975	1.000	1.000
		3n/logn	1.000	1.000	0.991	0.911	0.900	0.899	0.980	0.993	1.000	0.915	0.990	0.997
		n/logn	1.000	1.000	0.990	0.887	0.876	0.876	0.971	0.988	1.000	0.884	0.942	0.992

A.7 Some additional results for MA

Here, we will provide additional results for the simulations presented in Section **??** of the main text, focusing on Settings 1-6, Setting 8, and the simulations for L_2B and Bagging.

Setting 1: Decaying coefficients

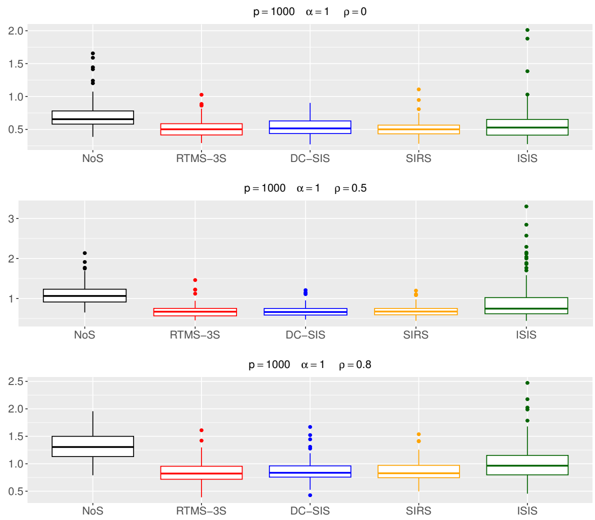

Fig. S1: Boxplots of estimation errors by L_1-ARM: Comparison between no variable screening (NoS) and variable screening using different ranking methods: RTMS-3S, DC-SIS, SIRS, and ISIS.

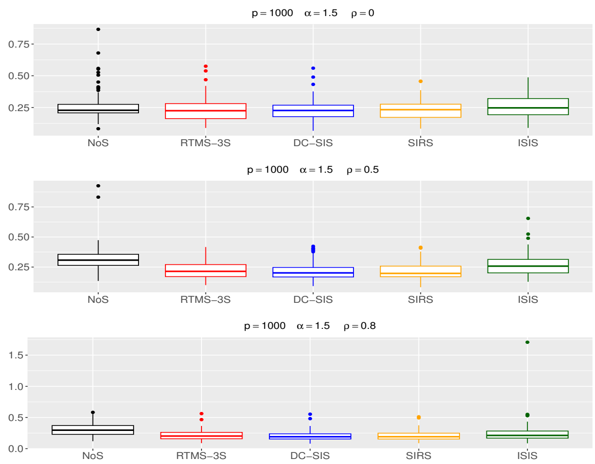

Fig. S2: Boxplots of estimation errors by L_1-ARM: Comparison between no variable screening (NoS) and variable screening using different ranking methods: RTMS-3S, DC-SIS, SIRS, and ISIS.

Setting 2: Evenly spaced coefficients

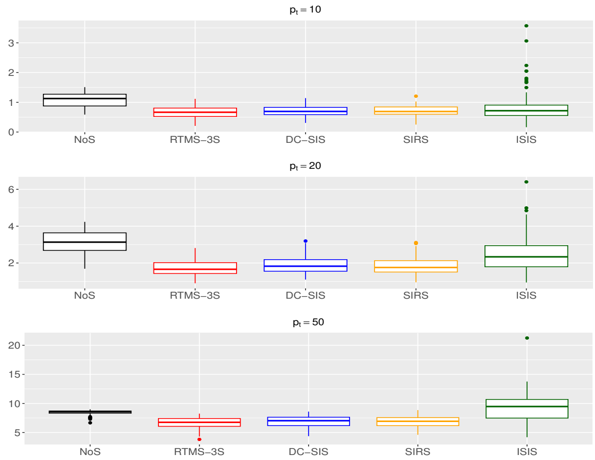

Fig. S3: Boxplots of estimation errors by L_1-ARM: Comparison between no variable screening (NoS) and variable screening using different ranking methods: RTMS-3S, DC-SIS, SIRS, and ISIS.

From Figure S3, as the sparsity decreases, the performances of ISIS gradually improve, while the performances of RTMS-3S remain consistently better than individual screening methods.

Setting 3: Randomly positioned coefficients

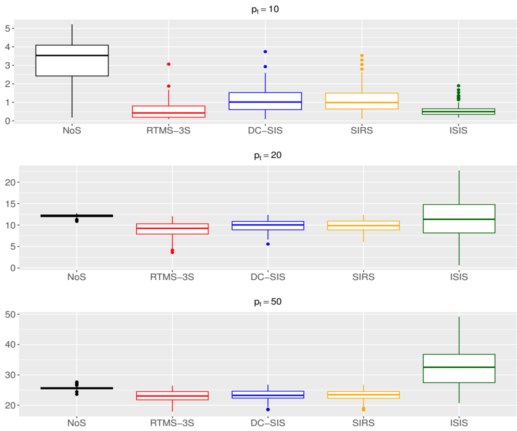

Fig. S4: Boxplots of estimation errors by L_1-ARM: Comparison between no variable screening (NoS) and variable screening using different ranking methods: RTMS-3S, DC-SIS, SIRS, and ISIS.

Setting 4: Example 1 in Zhu et al. (2011)

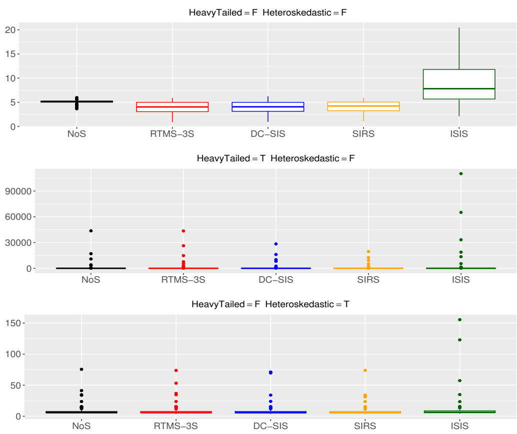

Fig. S5: Boxplots of estimation errors by L_1-ARM: Comparison between no variable screening (NoS) and variable screening using different ranking methods: RTMS-3S, DC-SIS, SIRS, and ISIS.

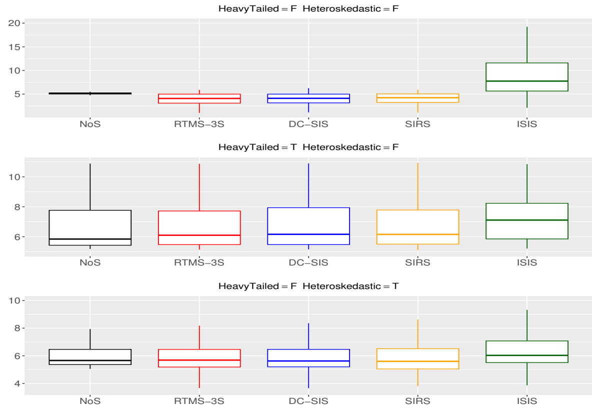

Fig. S6: Boxplots of estimation errors by L_1-ARM without outliners: Comparison between no variable screening (NoS) and variable screening using different ranking methods: RTMS-3S, DC-SIS, SIRS, and ISIS.

Setting 5: Example 1.d in Li et al. (2012)

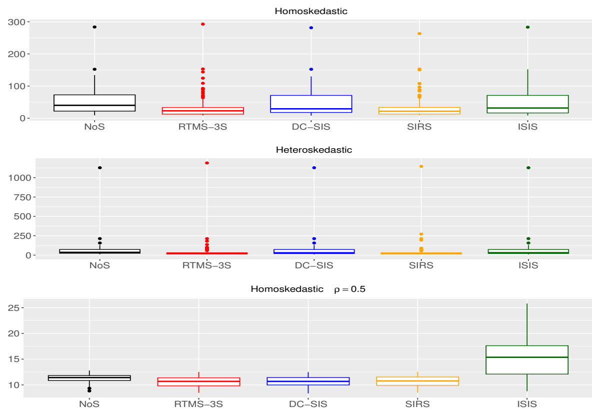

Fig. S7: Boxplots of estimation errors by L_1-ARM: Comparison between no variable screening (NoS) and variable screening using different ranking methods: RTMS-3S, DC-SIS, SIRS, and ISIS.

Setting 6: Simulated example II in Fan and Lv (2008)

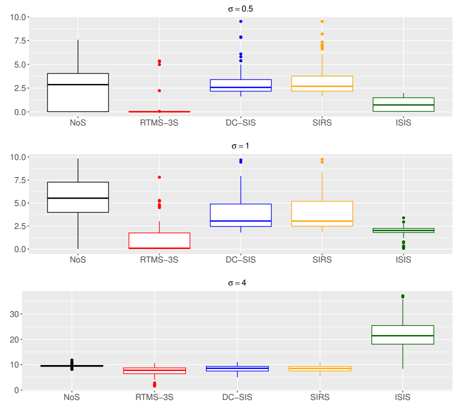

Fig. S8: Boxplots of estimation errors by L_1-ARM: No variable screeing (NoS) versus retaining different numbers of variables: $2n$, $3n/\log n$, $2n/\log n$ and $n/\log n$ under variable ranking methods DC-SIS and RTMS-3S.

From Figure S8, when the value of σ is small, ISIS performs the best among the individual screening methods because it is the only method that can identify the important variable x_4. However, the combined screening method still outperforms ISIS because the ranking of x_4 obtained from the combined screening is even higher.

Setting 7:

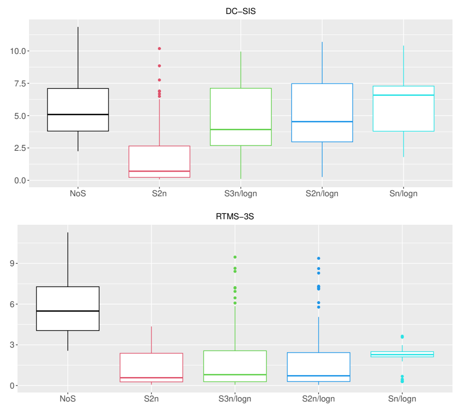

Fig. S9: Boxplots of MA estimation errors based on L_1-ARM: No variable screening (NoS) versus retaining different numbers of variables: $2n$, $3n/\log n$, $2n/\log n$ and $n/\log n$ under variable ranking methods DC-SIS (upper) and RTMS-3S (lower).

Setting 8: Mixed coefficient sizes

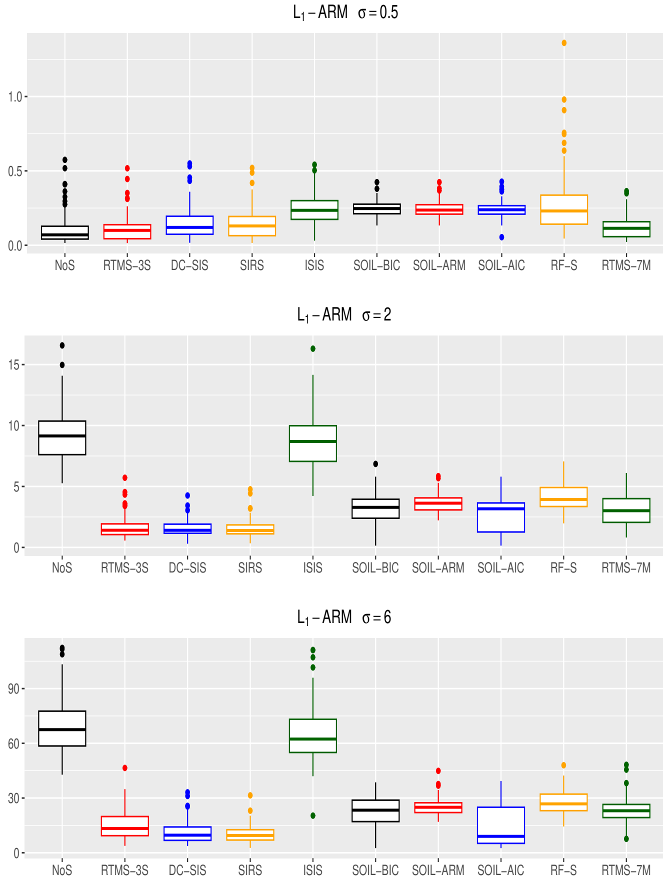

Fig. S10: Boxplots of estimation errors by L_1-ARM: Comparison between no variable screening (NoS) and variable screening using different ranking methods: RTMS-3S, RTMS-7M, DC-SIS, SIRS, ISIS, SOIL-ARM, SOIL-AIC, SOIL-BIC and RF-S.

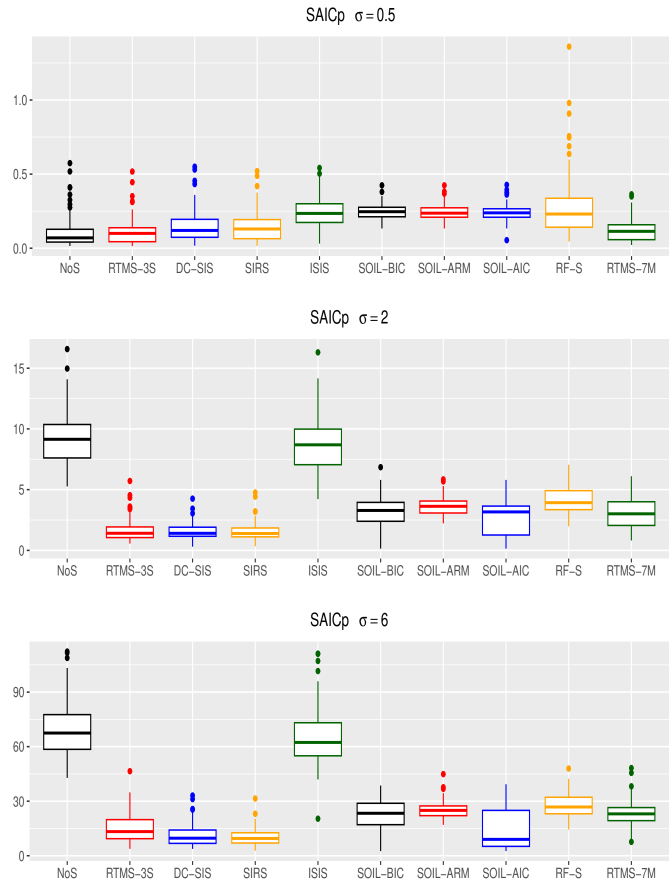

Fig. S11: Boxplots of estimation errors by SAICp: Comparison between no variable screening (NoS) and variable screening using different ranking methods: RTMS-3S, RTMS-7M, DC-SIS, SIRS, ISIS, SOIL-ARM, SOIL-AIC, SOIL-BIC and RF-S.

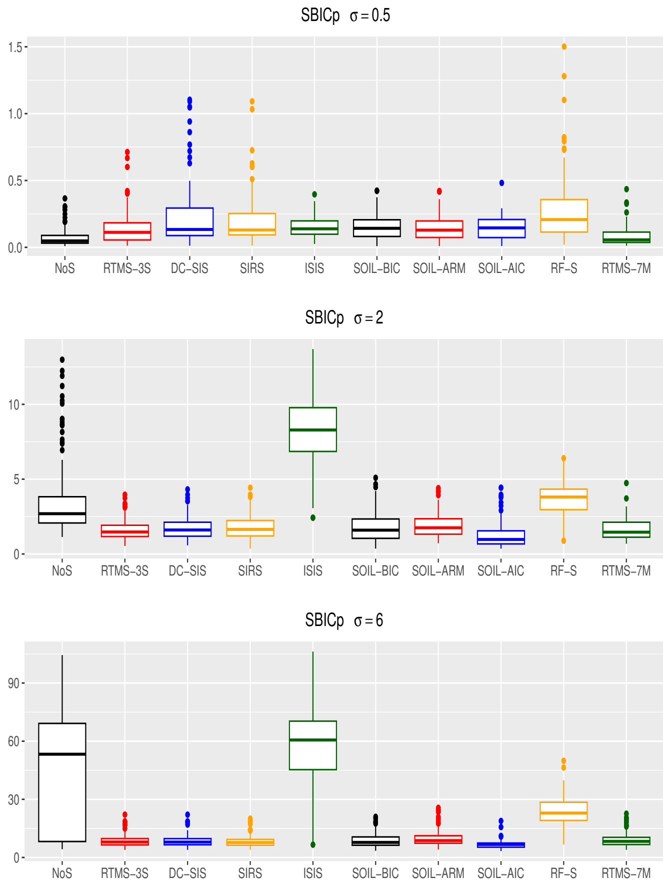

Fig. S12: Boxplots of estimation errors by SBICp: Comparison between no variable screening (NoS) and variable screening using different ranking methods: RTMS-3S, RTMS-7M, DC-SIS, SIRS, ISIS, SOIL-ARM, SOIL-AIC, SOIL-BIC and RF-S.

Simulations for L_2B and Bagging

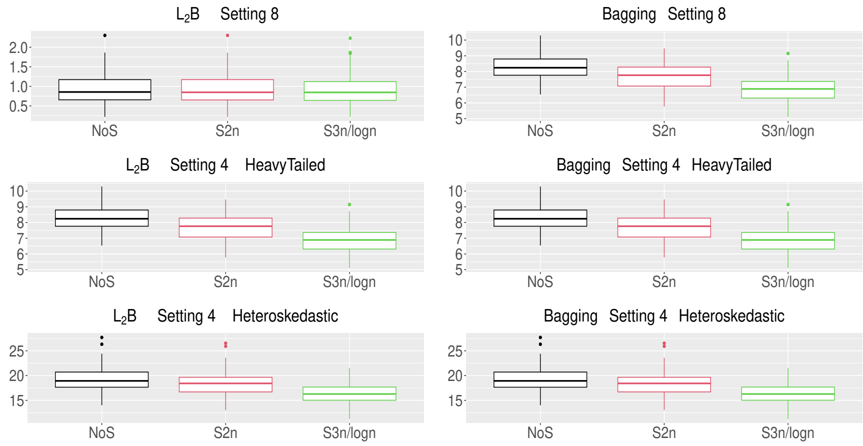

Fig. S13: Boxplots of estimation errors by L$_2$B (left) and Bagging (right): Comparison between no variable screeing (NoS) and retaining different numbers of variables: $2n$ and $3n/\log n$ under RTMS-3S.

In Setting 8, as demonstrated in Figures S10-S12, we observed that the RTMS-7M approach, which integrates seven screening methods, typically surpasses the performance of each screening method when used individually. However, RTMS-3S, which merges three variable screening techniques, consistently shows superior performance.

As depicted in Figure S13, among non-parametric methods, the L_2B method is not sensitive to variable screening, but Bagging shows a significant improvement in estimation performance after variable screening.

A.8 Additional tables for the real data examples

Table S9: Top 14 genes for different screening measures for Bardet data

RANK	**DC-SIS**	**SIRS**	**ISIS**	**SOIL-ARM**	**SOIL-AIC**	**SOIL-BIC**	**RF-S**	**RTMS-3S**	**RTMS-4I**	**RTMS-7M**
1	X153	X96	X153	X153	X153	X153	X96	X153	X153	X153
2	X96	X153	X55	X87	X87	X87	X153	X96	X87	X55
3	X55	X37	X99	X185	X185	X185	X37	X55	**X185**	X99
4	X37	X180	X199	X180	X180	X180	X180	X37	X96	X87
5	X180	X157	X87	X200	X200	X200	X11	X99	X37	X96
6	X5	X24	X60	X62	X62	X62	X151	X180	X180	X37
7	X199	X134	X85	X99	X187	X55	X38	X199	**X200**	X185
8	X24	X140	X177	X55	X55	X99	X99	X24	X62	X180
9	X157	X5	X146	X187	X99	X187	X62	X157	X99	X199
10	X11	X38	X5	X136	X136	X136	X156	X5	X55	X42
11	X60	X131	X42	X54	X54	X42	X199	X60	X187	X177
12	X52	X67	X180	X140	X140	X85	X140	X134	X136	X140
13	X131	X143	X11	X42	X42	X177	X131	X11	X140	X11
14	X134	X168	X4	X85	X85	X109	X5	X131	X42	X134

Table S10: Top 25 variables for different screening measures for the supermarket data

RANK	**DC-SIS**	**SIRS**	**ISIS**	**SOIL-ARM**	**SOIL-AIC**	**SOIL-BIC**	**RF-S**	**RTMS-3S**	**RTMS-4I**	**RTMS-7M**
1	X5	X5	X5	X11	X5	X5	X5	X5	X11	X5
2	X11	X10	X3	X5	X6	X6	X11	X11	X5	X11
3	X10	X3	X10	X139	X7	X10	X10	X10	X139	X10
4	X139	X11	X11	X62	X10	X11	X139	X3	X10	X139
5	X3	X139	X139	X6	X11	X56	X3	X139	X6	X6
6	X41	X39	X6	X10	X21	X62	X42	X39	X7	X3
7	X61	X41	X39	X56	X56	X139	X61	X41	X62	X7
8	X39	X53	X62	X410	X62	X272	X53	X6	X56	X62
9	X6	X73	X53	X2830	X139	X410	X39	X53	X410	X39
10	X44	X6	X176	X21	X410	X2830	X41	X62	**X2830**	X53
11	X53	X33	X30	X272	X2830	X21	X71	X73	X21	X56
12	X73	X44	X113	X7	X48	X48	X6	X44	X272	X42
13	X28	X62	X42	X48	X272	X7	X129	X61	X48	X176
14	X129	X30	X207	X107	X107	X107	X64	X30	X107	X113
15	X62	X16	X2830	X24	X301	X24	X162	X33	X24	X30
16	X33	X61	X301	X301	X24	X145	X86	X176	X301	X48
17	X30	X129	X410	X4981	X142	X301	X28	X28	**X4981**	X21
18	X176	X28	X1213	X3	X4981	X4981	X328	X129	X71	X71
19	X75	X75	X1582	X39	X521	X142	X207	X42	X3	X129
20	X38	X51	X1452	X145	X145	X113	X21	X16	X145	X167
21	X50	X176	X272	X142	X113	X2	X44	X75	X39	X207
22	X16	X50	X4981	X176	X2	X71	X167	X113	X142	X145
23	X42	X7	X4679	X42	X71	X176	X113	X50	X42	X2
24	X64	X167	X1893	X207	X167	X167	X400	X38	X113	X142
25	X76	X38	X4303	X53	X52	X491	X176	X7	X53	X24

A.9 Software

We have developed the *RTMS* (Recommendation and Trimmed Mean Scoring) package to facilitate the implementation of our RTMS method. This package integrates seven variable ranking methods mentioned in the paper: three variable screening methods (ISIS, SIRS, and DC-SIS) and four variable importance ranking methods (SOIL-AIC, SOIL-BIC, SOIL-ARM, and RF). With this package, users can select at least $K \geq 2$ of these methods and then choose

Table S11: Median (standard error) of prediction errors with different variable screening methods for the two datasets

	Bardet data			Supermarket data		
	L_1-ARM	**SAICp**	**SBICp**	**L_1-ARM**	**SAICp**	**SBICp**
NoS	0.0109 (0.00032)	0.0104 (0.00031)	0.0115 (0.00033)	0.1477 (0.00225)	0.1444 (0.00231)	0.1443 (0.00238)
RTMS-3S	0.0099 (0.00031)	0.0098 (0.00025)	0.0108 (0.00029)	0.1247 (0.00151)	0.1168 (0.00112)	0.1267 (0.00131)
RTMS-4I	**0.0096 (0.00028)**	**0.0090 (0.00026)**	0.0101 (0.00024)	0.0993 (0.00112)	**0.1029 (0.00100)**	**0.1081 (0.00122)**
RTMS-7M	0.0099 (0.00029)	0.0097 (0.00027)	0.0106 (0.00025)	0.1111 (0.00101)	0.1065 (0.00124)	0.1154 (0.00133)
DC-SIS	0.0102 (0.00032)	0.0100 (0.00027)	0.0110 (0.00028)	0.1283 (0.00133)	0.1178 (0.00105)	0.1271 (0.00119)
SIRS	0.0102 (0.00033)	0.0101 (0.00028)	0.0111 (0.00029)	0.1283 (0.00146)	0.1184 (0.00104)	0.1284 (0.00131)
ISIS	0.0102 (0.00031)	0.0097 (0.00026)	0.0106 (0.00027)	0.1120 (0.00171)	0.1179 (0.00179)	0.1185 (0.00153)
SOIL-ARM	0.0097 (0.00027)	0.0090 (0.00027)	**0.0100 (0.00022)**	0.1007 (0.00114)	0.1030 (0.00107)	0.1086 (0.00115)
SOIL-AIC	0.0099 (0.00041)	0.0100 (0.00037)	0.0112 (0.00036)	0.0995 (0.00108)	0.1039 (0.00103)	0.1085 (0.00122)
SOIL-BIC	0.0098 (0.00030)	0.0093 (0.00029)	0.0100 (0.00027)	**0.0990 (0.00107)**	0.1035 (0.00106)	0.1107 (0.00122)
RF-S	0.0099 (0.00031)	0.0096 (0.00026)	0.0104 (0.00028)	0.1274 (0.00146)	0.1194 (0.00125)	0.1287 (0.00133)

Table S12: Median (standard error) of prediction errors using 'New', 'New*' and 'Old' versions for the two datasets

		RTMS-3S	**DC-SIS**	**SIRS**	**ISIS**	**RF-S**
Bardet data (SBICp)	**Old**	0.0108 (0.00029)	0.0110 (0.00028)	0.0111 (0.00029)	0.0106 (0.00027)	0.0104 (0.00028)
	New	0.0098 (0.00314)	0.0100 (0.00341)	0.0096 (0.00340)	0.0095 (0.00300)	0.0101 (0.00332)
	New*	0.0109 (0.00288)	0.0109 (0.00307)	0.0111 (0.00330)	0.0102 (0.00305)	0.0107 (0.00298)
Supermarket data (SAICp)	**Old**	0.1168 (0.00112)	0.1178 (0.00105)	0.1184 (0.00104)	0.1179 (0.00179)	0.1194 (0.00125)
	New	0.0994 (0.01139)	0.0998 (0.01270)	0.0998 (0.01162)	0.1135 (0.01615)	0.1042 (0.01258)
	New*	0.1165 (0.01347)	0.1198 (0.01309)	0.1200 (0.01222)	0.1155 (0.01656)	0.1196 (0.01386)

Table S13: Evaluating predictive performance differences between 'New' and 'Old' versions with t-test and DM-test

	Paired t-test-Test Mean Difference (p-value)		**Diebold-Mariano Value (p-value)**	
	Bardet data	Supermarket data	Bardet data	Supermarket data
RTMS-3S	0.0008 (0.0000)	0.0181 (0.0000)	6.8762 (0.0000)	11.3600 (0.0000)
DC-SIS	0.0009 (0.0000)	0.0183 (0.0000)	6.4579 (0.0000)	11.5321 (0.0000)
SIRS	0.0012 (0.0000)	0.0177 (0.0000)	7.8707 (0.0000)	11.7175 (0.0000)
ISIS	0.0006 (0.0000)	0.0050 (0.0449)	5.2451 (0.0000)	2.0308 (0.0423)
RF-S	0.0007 (0.0000)	0.0149 (0.0000)	5.4632 (0.0004)	8.6645 (0.0000)

the top r (default is 3) variables from each ranking procedure as strongly recommended and protected variables. The largest L $(1 \leq L < K)$ ranks are removed in line with the goal of the final variable ranking, and a trimmed average is then calculated for the remaining variables. This approach results in a more informative variable combination.

References

FAN, J. and LV, J. (2008). Sure independence screening for ultrahigh dimensional feature space. *Journal of the Royal Statistical Society: Series B* **70**, 849–911.

LI, R., ZHONG, W. and ZHU, L. (2012). Feature screening via distance correlation learning. *Journal of the American Statistical Association* **107**, 1129–1139.

ZHU, L. P., LI, L., LI, R. and ZHU, L. X. (2011). Model-free feature screening for ultrahigh-dimensional data. *Journal of the American Statistical Association* **106**, 1464–1475.