THRESHOLDED LASSO: MONTE CARLOS.

CALLOT, CANER, KOCK, RIQUELME.

Reported statistics. This document contains the Monte Carlo results. Note S and $\hat{S}$ be the active set and the estimated active set. The statistics reported (all averaged across iterations) are:

- MSE: mean square prediction error.
- $|\hat{S} \cap S^c|$: number of false non zero.
- $|S \cap \hat{S}^c|$: number of false zero.
- Perfect Sel. (%): the share of iterations for which we have perfect model selection.
- $\|\hat{\alpha} - \alpha_0\|_1$: ℓ_1 estimation error for the parameters.
- $\|\hat{\alpha} - \alpha_0\|_\infty$: ℓ_∞ estimation error for the parameters.
- $\|\hat{\tau} - \tau_0\|_1$: absolute threshold parameter estimation error.
- C: selected (BIC) thresholding parameter.
- $\hat{\lambda}$: selected (BIC) penalty parameter.

Rows with a white background give results for the Lasso, those with a grey background for the thresholded Lasso.

DGP, Estimation settings. The specifics of each DGP are detailed below the corresponding table. The global settings are described below:

- The regressors are generated by independent draws from a $X_i \sim \mathcal{N}(0, I_m)$.
- The threshold variable is generated by independent draws from a $Q_i \sim \mathcal{U}[0, 1]$.
- The residuals are drawn from a $U_i \sim \mathcal{N}(0, \sigma^2)$.
- An intercept is always estimated.
- The grid we use to search for τ is $[0.15, ..., 0.85]$ by steps of 0.05.
- The grid we use to search for C is $[0.1, 0.2, ..., 5]$. C is selected by BIC.

Experiments. This document contains 5 tables:

(1) A quasi replication of Lee, Seo, and Shin, table 3. Also increasing number of zero parameters
(2) Experiments with no jumps ($\delta = 0$).
(3) Increasing sample size.
(4) Different parameter scale.
(5) Increasing number of non zero parameters.

	τ_0	MSE	$\|J(\hat{\alpha})\cap J(\alpha_0)^c\|$	$\|J(\alpha_0)\cap J(\hat{\alpha})^c\|$	Perfect Sel	$\|\hat{\alpha}-\alpha_0\|_1$	$\|\hat{\alpha}-\alpha_0\|_\infty$	$\|\hat{\tau}-\tau_0\|_1$	C	$\hat{\lambda}$
	0.3	1.20	4.50	0.06	3	3.39	0.85	0.27	-	0.06
		1.22	0.04	0.11	86	3.29	0.85	-	1.46	-
$m=50$	0.4	1.42	5.23	0.07	1	3.68	0.94	0.22	-	0.05
		1.45	0.04	0.16	81	3.55	0.94	-	1.49	-
	0.5	1.55	5.72	0.07	0	3.99	1.00	0.18	-	0.05
		1.58	0.06	0.14	82	3.85	1.01	-	1.45	-
	0.3	1.34	5.59	0.05	1	3.99	0.95	0.25	-	0.07
		1.38	0.04	0.13	85	3.86	0.95	-	1.27	-
$m=100$	0.4	1.56	6.26	0.08	0	4.29	1.03	0.22	-	0.07
		1.60	0.05	0.16	82	4.15	1.03	-	1.25	-
	0.5	1.77	7.27	0.12	0	4.77	1.10	0.19	-	0.07
		1.83	0.07	0.21	78	4.60	1.11	-	1.22	-
	0.3	1.57	7.06	0.10	0	4.65	1.06	0.25	-	0.09
		1.62	0.03	0.19	82	4.49	1.06	-	1.15	-
$m=200$	0.4	1.80	8.10	0.12	0	5.04	1.14	0.22	-	0.09
		1.87	0.03	0.22	79	4.86	1.15	-	1.12	-
	0.5	2.22	9.20	0.26	0	5.82	1.27	0.18	-	0.09
		2.30	0.06	0.40	71	5.60	1.28	-	1.07	-
	0.3	1.73	8.81	0.15	0	5.38	1.16	0.26	-	0.10
		1.81	0.03	0.23	81	5.18	1.17	-	1.04	-
$m=400$	0.4	2.16	9.35	0.33	0	6.17	1.30	0.22	-	0.12
		2.26	0.04	0.47	73	5.94	1.31	-	0.98	-
	0.5	2.84	9.81	0.66	0	7.26	1.46	0.19	-	0.13
		2.96	0.03	0.91	60	7.02	1.47	-	0.90	-

TABLE 1. Lasso (white background) and Thresholded Lasso (grey background). Increasing number of zero parameters and 3 locations of τ_0.

Increasing number of zeros.

- Sample size: $n = 200$.
- $\beta = [2, 2, 2, 2, 2, 0, ..., 0]$
- $\delta = [2, -2, 2, -2, 2, 0, ..., 0]$
- The length β and δ is $m = 50, 100, 200, 400$ so that each model has $2m + 1$ parameters.
- The intercept is set to 0.
- The results are based on 1000 replications.
- $\sigma^2 = 0.25$.
- We experiment with $\tau = 0.3, 0.4, 0.5$.

	MSE	$\|J(\hat{\alpha}) \cap J(\alpha_0)^c\|$	$\|J(\alpha_0) \cap J(\hat{\alpha})^c\|$	Perfect Sel	$\|\hat{\alpha} - \alpha_0\|_1$	$\|\hat{\alpha} - \alpha_0\|_\infty$	C	$\hat{\lambda}$
$m = 50$	0.29	1.56	0.00	23	0.60	0.16	-	0.07
	0.29	0.21	0.00	81	0.56	0.16	0.73	-
$m = 100$	0.30	1.56	0.00	23	0.65	0.17	-	0.08
	0.31	0.18	0.00	83	0.61	0.17	0.61	-
$m = 200$	0.31	1.45	0.00	27	0.70	0.18	-	0.09
	0.32	0.15	0.00	86	0.66	0.18	0.53	-
$m = 400$	0.32	1.44	0.00	27	0.74	0.19	-	0.10
	0.33	0.12	0.00	89	0.71	0.19	0.46	-

TABLE 2. Lasso (white background) and Thresholded Lasso (grey background). No threshold effect ($\delta = 0$), $n = 200$, 4 different length of the parameter vector.

No Jumps.

- Sample size: $n = 200$.
- $\beta = [2, 2, 2, 2, 2, 0, ..., 0]$
- $\delta = [0, ..., 0]$
- The length β and δ is $m = 50, 100, 200, 400$ so that each model has $2m + 1$ parameters.
- The intercept is set to 0.
- The results are based on 1000 replications.
- $\tau_0 = 0.5$.
- $\sigma^2 = 0.25$.

		MSE	$\|J(\hat{\alpha})\cap J(\alpha_0)^c\|$	$\|J(\alpha_0)\cap J(\hat{\alpha})^c\|$	Perfect Sel	$\|\hat{\alpha}-\alpha_0\|_1$	$\|\hat{\alpha}-\alpha_0\|_\infty$	$\|\hat{\tau}-\tau_0\|_1$	C	$\hat{\lambda}$
$\tau_0 = 0.3$	$n = 50$	10.04	1.83	4.92	0	14.72	1.99	0.30	-	0.58
		10.64	0.29	5.51	0	14.66	1.99	-	0.67	-
	$n = 100$	3.34	7.22	1.09	0	7.92	1.51	0.27	-	0.15
		3.53	0.12	1.38	45	7.63	1.51	-	1.32	-
	$n = 200$	1.46	5.56	0.08	1	4.07	1.00	0.25	-	0.07
		1.50	0.04	0.16	82	3.95	1.00	-	1.25	-
	$n = 500$	0.76	3.31	0.01	6	2.27	0.64	0.17	-	0.04
		0.76	0.01	0.02	97	2.23	0.64	-	0.95	-
	$n = 1000$	0.50	2.62	0.00	10	1.51	0.45	0.06	-	0.03
		0.50	0.00	0.01	98	1.49	0.45	-	0.81	-
$\tau_0 = 0.5$	$n = 50$	8.98	1.81	4.84	0	14.56	2.00	0.21	-	0.48
		9.52	0.24	5.43	0	14.48	2.00	-	0.62	-
	$n = 100$	4.73	5.41	2.15	0	10.05	1.75	0.20	-	0.21
		4.94	0.12	2.62	23	9.84	1.75	-	1.00	-
	$n = 200$	1.83	7.41	0.12	0	4.83	1.14	0.18	-	0.07
		1.89	0.06	0.21	78	4.66	1.14	-	1.22	-
	$n = 500$	0.86	4.32	0.01	2	2.53	0.69	0.18	-	0.04
		0.87	0.01	0.04	96	2.48	0.69	-	0.96	-
	$n = 1000$	0.55	3.27	0.00	8	1.70	0.49	0.08	-	0.03
		0.55	0.01	0.01	98	1.67	0.49	-	0.80	-

TABLE 3. Lasso (white background) and Thresholded Lasso (grey background). Increasing sample size with $m = 100$ and 2 locations of τ_0.

Increasing sample size.

- Sample size: $n = 50, 100, 200, 500, 1000$.
- $\beta = [2, 2, 2, 2, 2, 0, ..., 0]$
- $\delta = [2, -2, 2, -2, 2, 0, ..., 0]$
- The length β and δ is $m = 100$ so that each model has $2m + 1$ parameters.
- The intercept is set to 0.
- The results are based on 1000 replications.
- $\sigma^2 = 0.25$.

		MSE	$|J(\hat{\alpha}) \cap J(\alpha_0)^c|$	$|J(\alpha_0) \cap J(\hat{\alpha})^c|$	Perfect Sel	$\|\hat{\alpha} - \alpha_0\|_1$	$\|\hat{\alpha} - \alpha_0\|_\infty$	$\|\hat{\tau} - \tau_0\|_1$	C	$\hat{\lambda}$
	$n = 100$	0.75	0.57	4.49	0	3.43	0.50	0.25	-	0.15
		0.78	0.03	5.15	0	3.45	0.50	-	0.47	-
$a = 0.5$	$n = 200$	0.57	0.50	3.21	0	2.92	0.50	0.27	-	0.10
		0.58	0.01	3.95	0	2.93	0.50	-	0.48	-
	$n = 1000$	0.31	2.75	0.04	9	1.37	0.32	0.10	-	0.03
		0.31	0.00	0.06	94	1.35	0.32	-	0.75	-
	$n = 100$	0.34	0.01	9.73	0	1.03	0.10	0.30	-	0.18
		0.34	0.00	9.78	0	1.03	0.10	-	0.12	-
$a = 0.1$	$n = 200$	0.32	0.04	8.60	0	0.94	0.10	0.34	-	0.13
		0.32	0.00	8.97	0	0.95	0.10	-	0.23	-
	$n = 1000$	0.27	0.15	4.20	0	0.64	0.10	0.10	-	0.04
		0.27	0.00	4.80	0	0.65	0.10	-	0.39	-
	$n = 100$	0.50	0.41	5.50	0	2.27	0.30	0.28	-	0.15
		0.52	0.02	6.22	0	2.30	0.30	-	0.46	-
$a = 0.3$	$n = 200$	0.38	0.29	4.00	0	1.89	0.30	0.32	-	0.10
		0.39	0.01	4.56	0	1.91	0.30	-	0.44	-
	$n = 1000$	0.31	0.60	1.74	1	1.38	0.30	0.10	-	0.04
		0.31	0.00	2.21	5	1.38	0.30	-	0.51	-
	$n = 100$	1.87	1.12	3.52	0	6.31	1.00	0.22	-	0.18
		1.94	0.05	4.21	0	6.31	1.00	-	0.56	-
$a = 1$	$n = 200$	1.09	3.95	1.16	0	4.46	0.86	0.21	-	0.09
		1.12	0.04	1.54	39	4.39	0.86	-	0.88	-
	$n = 1000$	0.34	2.98	0.00	9	1.43	0.35	0.08	-	0.03
		0.34	0.00	0.01	99	1.41	0.35	-	0.83	-
	$n = 100$	4.68	5.32	2.12	0	10.01	1.76	0.20	-	0.21
		4.89	0.10	2.61	21	9.80	1.76	-	1.02	-
$a = 2$	$n = 200$	1.81	7.44	0.11	0	4.74	1.12	0.18	-	0.07
		1.87	0.05	0.21	78	4.57	1.12	-	1.23	-
	$n = 1000$	0.56	3.18	0.00	7	1.70	0.49	0.07	-	0.03
		0.56	0.01	0.01	98	1.68	0.49	-	0.79	-

TABLE 4. Lasso (white background) and Thresholded Lasso (grey background). Increasing parameter scale, 3 sample sizes, $\tau_0 = 0.5$.

Varying parameter scale.

- Sample size: $n = 100, 200$.
- $\beta = a[1, 1, 1, 1, 1, 0, ..., 0]$
- $\delta = a[1, -1, 1, -1, 1, 0, ..., 0]$
- $a = 0.5, 0.1, 0.3, 1, 2$ is the scale of the non zero parameters.
- The length β and δ is $m = 100$ so that each model has $2m + 1$ parameters.
- The intercept is set to 0.
- The results are based on 1000 replications.
- $\tau_0 = 0.5$.
- $\sigma^2 = 0.25$.

		MSE	$|J(\alpha)\cap J(\alpha_0)^c|$	$|J(\alpha_0)\cap J(\hat\alpha)^c|$	Perfect Sel	$\|\hat\alpha-\alpha_0\|_1$	$\|\hat\alpha-\alpha_0\|_\infty$	$\|\hat\tau-\tau_0\|_1$	C	$\hat\lambda$
	$\tau_0=0.3$	0.34	0.10	0.00	90	0.36	0.25	0.25	-	0.09
		0.34	0.01	0.00	99	0.35	0.25	-	0.23	-
$m_1=1$	$\tau_0=0.4$	0.35	0.09	0.00	91	0.36	0.26	0.22	-	0.09
		0.35	0.01	0.00	99	0.36	0.26	-	0.20	-
	$\tau_0=0.5$	0.36	0.10	0.00	90	0.37	0.27	0.23	-	0.09
		0.36	0.01	0.00	99	0.37	0.27	-	0.19	-
	$\tau_0=0.3$	1.84	1.11	0.19	31	2.59	0.90	0.19	-	0.08
		1.86	0.01	0.28	78	2.57	0.91	-	0.68	-
$m_1=5$	$\tau_0=0.4$	2.03	1.11	0.20	32	2.67	0.93	0.18	-	0.08
		2.05	0.02	0.30	76	2.64	0.93	-	0.63	-
	$\tau_0=0.5$	2.05	1.01	0.16	35	2.57	0.91	0.17	-	0.08
		2.06	0.02	0.27	79	2.55	0.91	-	0.61	-
	$\tau_0=0.3$	5.08	2.85	0.81	5	6.54	1.36	0.19	-	0.08
		5.12	0.06	1.06	51	6.48	1.36	-	1.09	-
$m_1=10$	$\tau_0=0.4$	4.84	2.68	0.66	7	6.17	1.28	0.18	-	0.08
		4.88	0.05	0.89	57	6.11	1.28	-	1.01	-
	$\tau_0=0.5$	5.05	2.56	0.65	7	6.10	1.25	0.18	-	0.08
		5.09	0.04	0.90	59	6.05	1.25	-	0.95	-
	$\tau_0=0.3$	19.93	9.92	4.35	0	23.72	1.87	0.20	-	0.07
		20.27	0.31	5.56	10	23.45	1.88	-	2.45	-
$m_1=25$	$\tau_0=0.4$	19.13	10.32	3.56	0	22.76	1.88	0.20	-	0.07
		19.48	0.30	4.63	11	22.46	1.88	-	2.31	-
	$\tau_0=0.5$	18.32	9.90	3.05	0	21.53	1.75	0.23	-	0.07
		18.62	0.30	3.94	19	21.25	1.75	-	2.09	-

TABLE 5. Lasso (white background) and Thresholded Lasso (grey background). Increasing number of non zero parameters (m_1), fixed number of zeros ($m_0 = 100$), and 3 locations of τ_0.

Varying number of non zero parameters.

- Sample size: $n = 200$.
- $\beta = [2, ..., 2, 0, ..., 0]$
- $\delta = [2, ..., 2, 0, ..., 0]$
- β and δ contain both $m_0 = 100$ parameters equal to zero.
- β and δ contain both $m_1 = 1, 5, 10, 50$ parameters equal to one.
- The length β and δ is $m = m_0 + m_1$ so that each model has $2m + 1$ parameters.
- The intercept is set to 0.
- The results are based on 1000 replications.
- $\sigma^2 = 0.25$.