Statistics and Its Interface Volume 13 (2020) 151–156

Bi-level variable selection in high dimensional Tobit models*

Hailin Huang, Jizi Shangguan, Yuanzhang Li, and Hua Liang†

To study variable selection for high dimensional Tobit models, we formulate Tobit models to single-index models. We hybrid group variable selection procedures for single index models and univariate regression methods for Tobit models to achieve variable selection for Tobit models with group structures taken into consideration. The procedure is computationally efficient and easily implemented. Finite sample experiments show its promising performance. We also illustrate its utility by analyzing a dataset from an HIV/AIDS study.

Keywords and phrases: Group structure, Group Lasso, Single-index models, Tobit models.

1. INTRODUCTION

With advances in high throughput technologies, many medical studies are complemented with information about biomarkers of each patient. Identification of important biomarkers can lead to better understanding of the mechanism behind disease development, and thus facilitates further clinical diagnosis and prognosis activities. Sometimes the number of biomarkers may be larger than the number of observations, which raises high dimensional problems and brings challenges in data analysis. Even more, we may face the situations that the response is fixed censored due to detection limit (Haab, Dunham and Brown, 2001; Van der Pouw Kraan et al., 1995). For instance, in measuring vial load of HIV/ADIS studies, the half maximal inhibitory concentration (IC50) values in blood serum can not be measured when they are below the detected limitation. Conventional Tobit models (Tobin, 1958) for fixed censored responses and associated estimation methods cannot be directly applied. In addition, among the tons of biomarkers/covariates investigated, maybe only a few are associated with the response variable of interest. Thus, variable selection or dimension reduction is always recommended along with the estimation procedure. To explore the relationship between a fixed censored response variable and a set of high dimensional covariates, we propose a new method to accommodate high dimensional data with fixed censored responses.

Among the many variable selection techniques developed, penalized selection methods have attracted extensive attentions. Penalization methods put penalties on the regression coefficients, which reduces model complexity and can lead to better model fitting. In the literature, some of the most popular work on penalization methods includes Lasso (Tibshirani, 1996), MCP (Zhang, 2010), and SCAD (Fan and Li, 2001). These methods and their variants have also been widely used in high dimensional data analysis (Fan and Li, 2002; Huang, Breheny and Ma, 2012; Gui and Li, 2005). The above methods tackle variable selection problems at individual covariate levels. However, some prior knowledge may introduce group structures as well. For example, in biomarker analysis, biomarkers belonging to the same functional group may perform similarly. Enlightened by this, it may be more desirable to take into account the grouping structure in the variable selection procedure. For this purpose, researchers proposed group variable selection methods and bi-level variable selection methods when the covariates could be grouped, where the former type of methods focuses on selecting important groups, and the latter type of methods targets at selecting important groups as well as identifying important members within the groups (Breheny and Huang, 2009). Some representative examples for these two types of methods include group Lasso, group SCAD, group bridge Lasso, and group exponential Lasso (Yuan and Lin, 2006; Huang, Breheny and Ma, 2012; Wang, Chen and Li, 2007; Breheny and Huang, 2009; Breheny, 2015; Huang et al., 2009). Although variable selection methods that takes group structures into consideration have been extensively studied in various parametric, semiparametric and nonparametric models, the efforts for the Tobit models are still needed. Motivated by the group Lasso method of Yuan and Lin (2006), Liu, Wang and Wu (2013) propose a group Lasso for Tobit models. But their method can only work for low dimensional Tobit models and can not separate noisy and significant covariates within a group, i.e. fails to perform bi-level group variable selection. As far as we know, there no methods that perform bi-level group variable selection for high dimensional Tobit models.

In this article, we propose a bi-level variable selection method for high dimensional Tobit (shorten as BHTobit)

*The authors thank a referee for valuable suggestions and comments. Liang's research was partially supported by NSF grant DMS-1620898.
†Corresponding author.

models, where variable selection procedure is performed within and between groups. The method combines the idea in group selection methods for single index models and univariate regression methods for Tobit models. The procedure is easy to implement and computationally efficient. The rest of the article is organized as follows. Section 2 presents the model. Section 3 describes the variable selection method. Section 4 evaluates the finite-sample performance through simulation studies. Section 5 illustrates the method by analyzing an HIV data.

2. MODEL SPECIFICATIONS

Consider the Tobit model with group structures:

$$Y^* = \sum_{j=1}^{J} \mathbf{X}_j \beta_j - \epsilon, \tag{1}$$

where $\mathbf{X}_j$ is the portion of the design matrix $\mathbf{X}$ formed by the predictors in the jth group, and β_j is the associated unknown parameter vector. Let p_j denote the number of members in group j. Then $\mathbf{X}_j$ is an $n \times p_j$ matrix. The total number of explanatory covariates is $p = \sum_j p_j$. Write $\beta = (\beta_1, \cdots, \beta_p)$. For each observation, due to fixed censoring, Y^* may not be fully observed, and we can only observe (Y, δ), where $Y = \max(Y^*, c)$, $\delta = I(Y^* > c)$, c is the known lower detection limit. Without loss of generality, we assume $c = 0$. The realization of ϵ are ϵ_i's, instead of making parametric distributional assumptions such as normality, here we only assume that ϵ_i's are independently and identically distributed (i.i.d.) from an unknown distribution symmetric around 0 with finite variance. To facilitate theoretical derivations, we consider the model with error term "$-\epsilon$" instead of "ϵ". Similar model setting can be found in Huang et al. (2019a,b); Lewbel and Linton (2002).

3. BHTOBIT METHOD

The BHTobit method is inspired by considering a connection between Tobit models and single-index models given by Proposition 1 below. This connection transfers group variable selection of high dimensional Tobit models to that of high dimensional single-index models under mild assumptions, and enables the selection procedure applicable using an existing procedure for high dimensional single-index models. Let $F(\cdot)$ be the distribution function of ϵ.

Proposition 1. *Assume that the latent response Y^* has first $\nu(\geq 3)$ absolute moments, and the density function $f_\epsilon(\cdot)$ of ϵ is continuously differentiable and symmetric around zero. If $\lim_{\epsilon \to -\infty} \epsilon F(\epsilon) = 0$, then*

$$E(Y|\mathbf{X}\beta) = \int_{-\infty}^{\mathbf{X}\beta} F(\epsilon) d\epsilon.$$

Proof. To prove the proposition, we proceed as follows.

$$\begin{aligned} E(Y|\mathbf{X}\beta) &= E\{Y^* I(Y^* > 0)|\mathbf{X}\beta\} \\ &= \int_{-\infty}^{\mathbf{X}\beta} \{\mathbf{X}\beta - \epsilon\} f(\epsilon) d\epsilon \\ &= \mathbf{X}\beta F(\mathbf{X}\beta) - \int_{-\infty}^{\mathbf{X}\beta} \epsilon f(\epsilon) d\epsilon \\ &= \mathbf{X}\beta F(\mathbf{X}\beta) - \left\{ \epsilon F(\epsilon)|_{-\infty}^{\mathbf{X}\beta} - \int_{-\infty}^{\mathbf{X}\beta} F(\epsilon) d\epsilon \right\} \\ &= \int_{-\infty}^{\mathbf{X}\beta} F(\epsilon) d\epsilon. \end{aligned}$$

Proposition 1 indicates that $E(Y|\mathbf{X}\beta)$ can be written as an uncensored single-index model with an index parameter β, and a link function $w(u)$, where $w(u) = \int_{-\infty}^{u} F(\epsilon) d\epsilon$. Therefore, any group variable selection methods for high dimensional single-index models could be applied.

Because of the computational simplicity, we adopt the method proposed by Wang, Xu and Zhu (2012) and further studied by Zeng, Wen and Zhu (2017). They applied a response variable transformation method to deal with the unknown link function of the single index models. No bandwidth selection is needed. To achieve bi-level variable selection, we further replace the group Lasso penalty in Zeng, Wen and Zhu (2017) by group exponential Lasso penalty (Breheny, 2015) to achieve bi-level variable selection. Briefly speaking, the group exponential Lasso penalty has a functional form of $\sum_{j=1}^{J} f(||\beta_j||_1)$, for groups $j = 1, \cdots, J$ and unknown parameter vector β_j associated with group j as specified in equation (1), where $f(\cdot)$ is a concave exponential function and $|| \cdot ||_1$ is L_1 norm. The outer summation of $f(||\beta_j||_1)$ for groups $j = 1, \cdots, J$ considers penalization at group level and the inner L_1 norm imposes the penalty to individual level, which together could yield sparsity at both group level and individual level (Breheny, 2015; Breheny and Huang, 2009).

Due to the identifiability issue in single index models, the estimation method for single index models can only identify β up to a scale constant, say $\beta_S = \beta/\gamma$, similar discussion can be found in the literature (Ichimura, 1993; Liang et al., 2010; Härdle, Hall and Ichimura, 1993; Carroll et al., 1997; Wang, Xu and Zhu, 2012; Zeng, Wen and Zhu, 2017). Thus, a second step estimation is needed to recover the scale constant γ. Given the initial estimator $\hat{\beta}_S$, for the second step estimation, any univariate regression method for Tobit models can be used to recover the scale constant γ. We state the BHTobit method as follows.

BHTobit Method:

Step 1. Find $\hat{\beta}_S$ by minimizing the objective function:

$$\frac{1}{2n} ||F_n(Y) - \mathbf{X}\beta||^2 + \sum_{j=1}^{J} f(||\beta_j||_1 | \lambda, \tau),$$

Table 1. Estimation accuracy: estimated values and associated MSE based on the three methods for the four settings with different censoring proportions. The true values are $(1.5, 1.5, -1, 1, 1)$

Censor	method	Estimated values					MSE
scenario (i)							
15%	BHTobit	1.566	1.559	−0.958	1.015	0.89	0.386
	ELasso	0.806	0.758	−0.383	0.416	0.429	2.34
	GLasso	1.018	0.998	−0.564	0.592	0.501	1.253
30%	BHTobit	1.477	1.492	−0.919	0.873	0.806	0.645
	ELasso	0.571	0.577	−0.285	0.205	0.243	3.561
	GLasso	0.792	0.776	−0.417	0.373	0.338	2.361
scenario (ii)							
Censor	method	Estimated values					MSE
15%	BHTobit	1.511	1.489	−0.928	1.000	0.752	0.384
	ELasso	0.775	0.762	−0.368	0.395	0.363	2.453
	GLasso	1.021	0.988	−0.565	0.586	0.470	1.375
30%	BHTobit	1.536	1.584	−0.910	0.943	0.791	0.537
	ELasso	0.547	0.573	−0.263	0.301	0.241	3.588
	GLasso	0.752	0.798	−0.441	0.436	0.309	2.304
scenario (iii)							
Censor	method	Estimated values					MSE
15%	BHTobit	1.580	1.565	−0.934	0.824	0.790	0.561
	ELasso	0.694	0.724	−0.348	0.342	0.390	2.601
	GLasso	0.933	0.942	−0.512	0.536	0.484	1.570
30%	BHTobit	1.609	1.586	−0.95	0.933	0.867	0.555
	ELasso	0.531	0.545	−0.259	0.244	0.263	3.733
	GLasso	0.777	0.778	−0.431	0.395	0.367	2.303
scenario (iv)							
Censor	method	Estimated values					MSE
15%	BHTobit	1.61	1.558	−0.977	0.949	0.85	0.412
	ELasso	0.713	0.677	−0.333	0.323	0.329	2.807
	GLasso	0.938	0.955	−0.532	0.53	0.444	1.456
30%	BHTobit	1.519	1.527	−0.841	0.892	0.707	0.518
	ELasso	0.487	0.511	−0.204	0.213	0.205	3.922
	GLasso	0.74	0.749	−0.379	0.4	0.293	2.538

where $F_n(\cdot)$ is defined as $F_n(t) = n^{-1}\sum_{i=1}^n 1_{Y_i \le t}$, and $f(\cdot|\lambda,\tau)$ is defined as
$f(\theta|\lambda,\tau) = \frac{\lambda^2}{\tau}\left\{1 - \exp(-\frac{\tau\theta}{\lambda})\right\}$, λ, τ are two positive tuning parameters.

Step 2. Replacing β_S by $\hat{\beta}_S$, we have

$$Y_i^* \approx \gamma \cdot X_i^\top \hat{\beta}_S + \epsilon_i. \quad (2)$$

Then, we apply least absolute deviation method (Powell, 1984) to recover the scale parameter γ, on account of the robustness of the method. That is, we obtain $\hat{\gamma}$ by minimizing the objective function:

$$\hat{L}_n(\gamma) = \frac{1}{n}\sum_{i=1}^n \left| Y_i - \max(0, \gamma \cdot X_i^\top \hat{\beta}_S) \right|.$$

Numerically, the algorithm is easy to implement. For step one, the group descent algorithm for group exponential Lasso can be directly applied to obtain estimate $\hat{\beta}_S$, with the tuning parameters being selected by cross-validation (Breheny, 2015). And $\hat{\gamma}$ can be then obtained by applying the Fitzenberger algorithm (Fitzenberger, 1997) for least absolute regression.

4. SIMULATION STUDIES

We conducted simulation studies to evaluate the finite sample performance of the proposed method and compare it with the existing ones like group Lasso (GLasso; Yuan and Lin, 2006) and Elastic Lasso (ELasso; Zou and Hastie, 2005). We generated Y^*, from the following model,

$$Y^* = \mathbf{X}^\top \beta - \epsilon,$$

where components of $\mathbf{X}$ are independent standard normal, with β, n and p being specified below. We considered two different families of the error: standard normal and standard Laplace. The observed response Y_i are set as $Y_i = \max(Y_i^*, c)$ with c chosen to lead to censoring

Table 2. Selection accuracy: the average number of the true zero groups/individuals coefficients that were correctly set to zero (C), and the average number of the truly nonzero groups/individuals incorrectly set to zero (I) based on the three methods for the four settings with different censoring proportions

censor	Method	C (group)	I (group)	C (individual)	I (individual)	Inclusion of β_6
scenario (i)						
15%	BHTobit	3	3.5	5	4.33	0.28
	ELasso	2.98	5	4.95	5.4	0.02
	GLasso	3	1.15	5	6.7	1
30%	BHTobit	2.98	3.83	4.95	4.58	0.22
	ELasso	2.9	3.92	4.72	4.2	0
	GLasso	2.98	1.27	4.97	7.3	1
scenario (ii)						
15%	BHTobit	3	3.73	5	4.6	0.25
	ELasso	3	4.62	4.95	4.88	0
	GLasso	2.98	1.3	4.97	7.45	1
30%	BHTobit	2.98	4.3	4.97	5.65	0.18
	ELasso	2.88	3.73	4.7	4.2	0
	GLasso	2.92	1.43	4.9	8.05	1
scenario (iii)						
15%	BHTobit	3	6.08	5	7.3	0.22
	ELasso	2.98	8.82	4.92	9.6	0.08
	GLasso	2.98	3.35	4.97	17.7	1
30%	BHTobit	3	3.9	4.97	4.35	0.1
	ELasso	2.92	5.5	4.78	6	0
	GLasso	3	1.68	5	9.35	1
scenario (iv)						
15%	BHTobit	3	4	5	4.5	0.18
	ELasso	2.98	5.65	4.9	5.9	0.05
	GLasso	3	1.98	5	10.85	1
30%	BHTobit	2.98	4.08	4.97	4.5	0.15
	ELasso	2.85	4.9	4.67	5.1	0.05
	GLasso	2.88	0.88	4.88	5.22	1

proportion (Cens): Cen = 15% or Cen = 30%. We consider the following scenarios, which extends similar scenarios in Liu, Wang and Wu (2013) to high dimensional settings:

(i) $p=200$, $\beta=(\underbrace{1.5,1.5}_{g_1},\underbrace{1,-1}_{g_2},\underbrace{1,0}_{g_3},\underbrace{0,0,0,0}_{g_4},\underbrace{0,0,0,0,0}_{g_5},\cdots,\underbrace{0,0,0,0,0}_{g_{42}})$, where g denotes the predefined group structures, from g_5 to g_{42}, each group has 5 elements, $n = 150$, ϵ follows standard normal distribution.

(ii) The same setting as (i) except that ϵ is standard Laplace distribution.

(iii) The same setting as (i) except that $p = 400$ and $\beta = (\underbrace{1.5,1.5}_{g_1},\underbrace{1,-1}_{g_2},\underbrace{1,0}_{g_3},\underbrace{0,0,0,0}_{g_4},\underbrace{0,0,0,0,0}_{g_5},\cdots,\underbrace{0,0,0,0,0}_{g_{82}})$.

(iv) The same setting as (ii) except that $p = 400$ and $\beta = (\underbrace{1.5,1.5}_{g_1},\underbrace{1,-1}_{g_2},\underbrace{1,0}_{g_3},\underbrace{0,0,0,0}_{g_4},\underbrace{0,0,0,0,0}_{g_5},\cdots,\underbrace{0,0,0,0,0}_{g_{82}})$.

As a comparison, the results of GLasso and ELasso (Yuan and Lin, 2006; Zou and Hastie, 2005) are also reported. To compare the performance of different methods, we shall use two criteria; selection accuracy, which demonstrated by number of correctly selected and incorrectly selected covariates; and mean squared errors (MSE) between $\hat{\beta}$ and β. Tables 1 and 2 summarizes the results.

With respect to the estimation accuracy, it can be seen that BHTobit method always has much smaller MSE than ELasso and GLasso. More specifically, for the five nonzero components, the estimated coefficients based on BHTobit method are close to the true values, while the estimated coefficients based on ELasso and GLasso significantly deviate from the true values. For selection accuracy, BHTobit method selects more correctly identified covariates and less falsely identified covariates compared with ELasso and GLasso at the individual level. At the group level, BHTobit method selects similar correctly identified groups and slightly more falsely identified groups compared with GLasso. It is also worthy to mention that, for β_6, compared with GLasso, BHTobit method tends to not include it in the final model, which yields sparsity within a group and performs bi-level variable selection, while GLasso always includes it.

Table 3. The results for the HIV study

Selected variables	Estimated coefficients	Selected groups
Bcell8	−3.82	Bcells
Bcell24	−7.12	Bcells
Bcell29	−0.02	Bcells
Bcell43	−2.02	Bcells
Bcell49	−22.32	Bcells
Antibody2	7.3	Antibody
Antibody3	6.7	Antibody

With censoring proportion increases, the estimation accuracy and selection accuracy decrease for all methods. This is not surprising as higher censoring indicates more information loss. However, the magnitude of decreasing is not remarkable for BHTobit method. Tables 1 and 2 also indicate that the results corresponding to normal and Laplace errors share a similar pattern, which may demonstrate the robustness of BHTobit method to normal and non-normal error distributions.

5. ANALYSIS OF AN HIV/AIDS STUDY

A main target to develop vaccine for HIV infection is to generate protective humoral response. It has been observed that some patients with HIV infection can produce potent serum antibodies, which can neutralize HIV isolates. To study the mechanisms of the generation of potent neutralizing antibodies, an efficient strategy is to investigate the behaviors of the B cells and Antibodies in such HIV-infected patients, and the relationship between B cells and Antibodies with serum antibody neutralization activity (determined by IC50). In an AIDS clinical study, 42 HIV-infected patients' observations were obtained and one third of their IC50 values are left censored at 20, while there are 55 biomarkers/covariates. The 55 covariates are formed into two groups: B cells group and neutralizing serum antibody group.

Our goal is to find the covariates and associated groups which have significant effects on IC50. To achieve this goal, we apply our proposed BHTobit method with the 55 covariates and consider the group structures within them, with IC50 as the response variable. All covariates were standardized.

Both B cells and antibody groups are both selected. The tuning parameters selected by 3-fold cross-validation are $\lambda = 0.049$ and $\tau = 0.334$. The selected covariates and associated groups, with their estimated coefficients are listed in Table 3. Our proposed method identified 5 covariates in the B cells group and 2 covariates in the antibody group. In particular, it can be seen that there is a negative correlation between BCELL8 (CD19-CD20+, percentage of total B cells), BCELL49 (IgD-B220-CD27-, percentage of IgDB cells) and IC50, and there is a positive relationship between Antibody2 (anti-dsDNA) and IC50. These suggests that B cells exhaustion may have a negative impact on the HIV neutralizing activity, while HIV-specific antibodies can facilitate the development of HIV neutralizing, similar findings were also reported in the literature (Gray et al., 2009; Petrovas et al., 1999; Haynes et al., 2005; McGranahan et al., 2016). It is interesting that both GLasso and ELasso identify no important biomarkers.

Received 27 April 2019

REFERENCES

Breheny, P. (2015). The group exponential lasso for bi-level variable selection. *Biometrics* **71** 731–740. MR3402609

Breheny, P. and Huang, J. (2009). Penalized methods for bi-level variable selection. *Statistics and Its Interface* **2** 369–380. MR2540094

Carroll, R. J., Fan, J., Gijbels, I. and Wand, M. P. (1997). Generalized partially linear single-index models. *Journal of the American Statistical Association* **92** 477–489. MR1467842

Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association* **96** 1348–1360. MR1946581

Fan, J. and Li, R. (2002). Variable selection for Cox's proportional hazards model and frailty model. *Annals of Statistics* **30** 74–99. MR1892656

Fitzenberger, B. (1997). Computational aspects of censored quantile regression. *Lecture Notes-Monograph Series* 171–186.

Gray, E. S., Taylor, N., Wycuff, D., Moore, P. L., Tomaras, G. D., Wibmer, C. K., Puren, A., DeCamp, A., Gilbert, P. B., Wood, B. et al. (2009). Antibody specificities associated with neutralization breadth in plasma from human immunodeficiency virus type 1 subtype C-infected blood donors. *Journal of Virology* **83** 8925–8937.

Gui, J. and Li, H. (2005). Penalized Cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data. *Bioinformatics* **21** 3001–3008.

Haab, B. B., Dunham, M. J. and Brown, P. O. (2001). Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biology* **2** research0004–1.

Härdle, W., Hall, P. and Ichimura, H. (1993). Optimal smoothing in single-index models. *The Annals of Statistics* **21** 157–178. MR1212171

Haynes, B. F., Fleming, J., Clair, E. W. S., Katinger, H., Stiegler, G., Kunert, R., Robinson, J., Scearce, R. M., Plonk, K., Staats, H. F. et al. (2005). Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. *Science* **308** 1906–1908.

Huang, J., Breheny, P. and Ma, S. (2012). A selective review of group selection in high-dimensional models. *Statistical Science* **27** 481–499. MR3025130

Huang, J., Ma, S., Xie, H. and Zhang, C.-H. (2009). A group bridge approach for variable selection. *Biometrika* **96** 339–355. MR2507147

Huang, H., Tang, Y., Li, Y. and Liang, H. (2019a). Estimation in additive models with fixed censored responses. *Journal of Nonparametric Statistics* **31** 131–143. MR3902146

Huang, H., Li, Y., Liang, H. and Tang, Y. (2019b). Estimation of single-index models with fixed censored responses. *Statistica Sinica*, to appear. MR3902146

Ichimura, H. (1993). Semiparametric least squares (SLS) and weighted SLS estimation of single-index models. *Journal of Econometrics* **58** 71–120. MR1230981

Lewbel, A. and Linton, O. (2002). Nonparametric censored and truncated regression. *Econometrica* **70** 765–779. MR1913830

Liang, H., Liu, X., Li, R. and Tsai, C. L. (2010). Estimation and testing for partially linear single-index models. *The Annals of Statistics* **38** 3811–3836. MR2766869

Liu, X., Wang, Z. and Wu, Y. (2013). Group variable selection and estimation in the Tobit censored response model. *Computational Statistics & Data Analysis* **60** 80–89. MR3007020

McGranahan, N., Furness, A. J., Rosenthal, R., Ramskov, S., Lyngaa, R., Saini, S. K., Jamal-Hanjani, M., Wilson, G. A., Birkbak, N. J., Hiley, C. T. et al. (2016). Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351** 1463–1469.

Petrovas, C., Vlachoyiannopoulos, P., Kordossis, T. and Moutsopoulos, H. (1999). Anti-phospholipid antibodies in HIV infection and SLE with or without anti-phospholipid syndrome: comparisons of phospholipid specificity, avidity and reactivity with β2-GPI. *Journal of Autoimmunity* **13** 347–355.

Powell, J. L. (1984). Least absolute deviations estimation for the censored regression model. *Journal of Econometrics* **25** 303–325. MR0752444

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B* **58** 267–288. MR1379242

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica* **26** 24–36. MR0090462

Van der Pouw Kraan, T., Boeije, L., Smeenk, R., Wijdenes, J. and Aarden, L. A. (1995). Prostaglandin-E2 is a potent inhibitor of human interleukin 12 production. *Journal of Experimental Medicine* **181** 775–779.

Wang, L., Chen, G. and Li, H. (2007). Group SCAD regression analysis for microarray time course gene expression data. *Bioinformatics* **23** 1486–1494.

Wang, T., Xu, P.-R. and Zhu, L.-X. (2012). Non-convex penalized estimation in high-dimensional models with single-index structure. *Journal of Multivariate Analysis* **109** 221–235. MR2922865

Yuan, M. and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society. Series B* **68** 49–67. MR2212574

Zeng, B., Wen, X. M. and Zhu, L. (2017). A link-free sparse group variable selection method for single-index model. *Journal of Applied Statistics* **44** 2388–2400. MR3683050

Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics* **38** 894–942. MR2604701

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society. Series B* **67** 301–320. MR2137327

Hailin Huang
Department of Statistics
George Washington University
Washington, D.C. 20052
USA
E-mail address: hhl1988@gwmail.gwu.edu

Jizi Shangguan
Department of Statistics
George Washington University
Washington, D.C. 20052
USA
E-mail address: shangguanjizi@gwmail.gwu.edu

Yuanzhang Li
Department of Statistics
George Washington University
Washington, D.C. 20052
USA
E-mail address: yli201301@email.gwu.edu

Hua Liang
Department of Statistics
George Washington University
Washington, D.C. 20052
USA
E-mail address: hliang@gwu.edu