

Variable Selection For Discrete Competing Risks Models

Quanfeng Yang

Ecology

Institute of Zoology, CAS

yqf1983626931@163.com

June 29, 2016

Application

Example: Congressional careers in the USA.

R1: *retirement*, retire

R2: *ambition*, might be ambitious and seek an alternative office

R3: *primary*, might lose a primary election

R4: *general*, might lose a general election

This research aim is to seek for covariates probably influence the transition process of a Congressman from his first election up to one of the competing events *general*, *primary*, *retirement or ambition* and fit a appropriate penalized competing risk model matching the example.

Variable	Description
Duration	Time (in terms served) the incumbent has spent in Congress prior to the election cycle
Age	Incumbent's age (in years) at each election cycle, centered around 51
Republican	Member of the Republican party 0: no, 1: yes
PriorMargin	The incumbent's margin of victory in his or her previous election, centered around 35
Leadership	Prestige position 0: otherwise, 1: member is in the House leadership and/or is a chair of a standing House committee
OpenGub	Open gubernatorial seat available in the incumbent's state 0: no, 1: yes
OpenSen	Open Senatorial seat available in the incumbent's state 0: no, 1: yes
Scandal	Incumbent was involved in an ethical or sexual misconduct scandal or when the incumbent was under criminal investigation 0: no, 1: yes
Redistricting	The incumbent's district was substantially redistricted 0: no, 1: yes

TABLE 1: *Description of the variables of the Congressional career data.*

Age
Republican
PriorMargin
Leadership
Open Gub.
Open Sen.
Scandal
Redistricting

Age:Republican
Age:PriorMargin
Age:Leadership
Age:OpenGub.
Age:OpenSen.
Age:Scandal
Age:Redistricting
Republican:PriorMargin
Republican:OpenGub.
Republican:OpenSen.
Republican:Scandal
Republican:Redistricting
PriorMargin:Leadership
PriorMargin:OpenGub.
PriorMargin:OpenSen.
PriorMargin:Scandal
PriorMargin:Redistricting
Leadership:OpenGub.
Leadership:OpenSen.
Leadership:Scandal
OpenGub.:OpenSen.
OpenGub.:Redistricting
OpenSen.:Scandal
OpenSen.:Redistricting

Tuning parameter: ζ_1, ζ_2

5-fold cross validation

$\zeta_1=6.0$ and $\zeta_2=2.64$

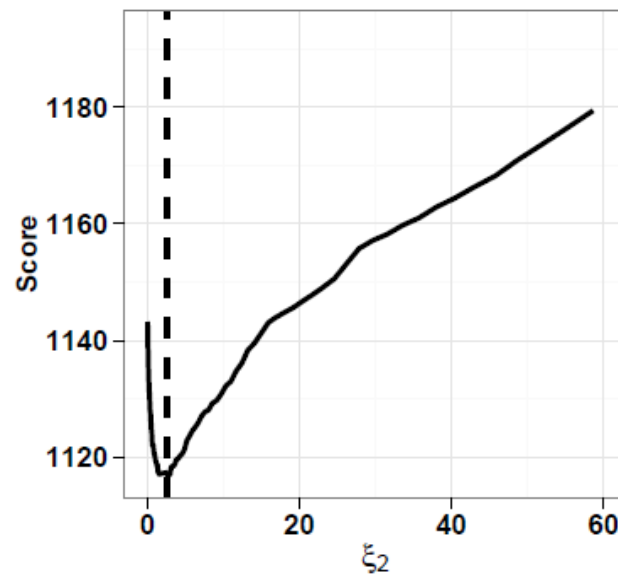


FIGURE 1: *Cross validation score subject to penalty parameter ζ_2 for $\zeta_1 = 6.0$ for the Congressional career data.*

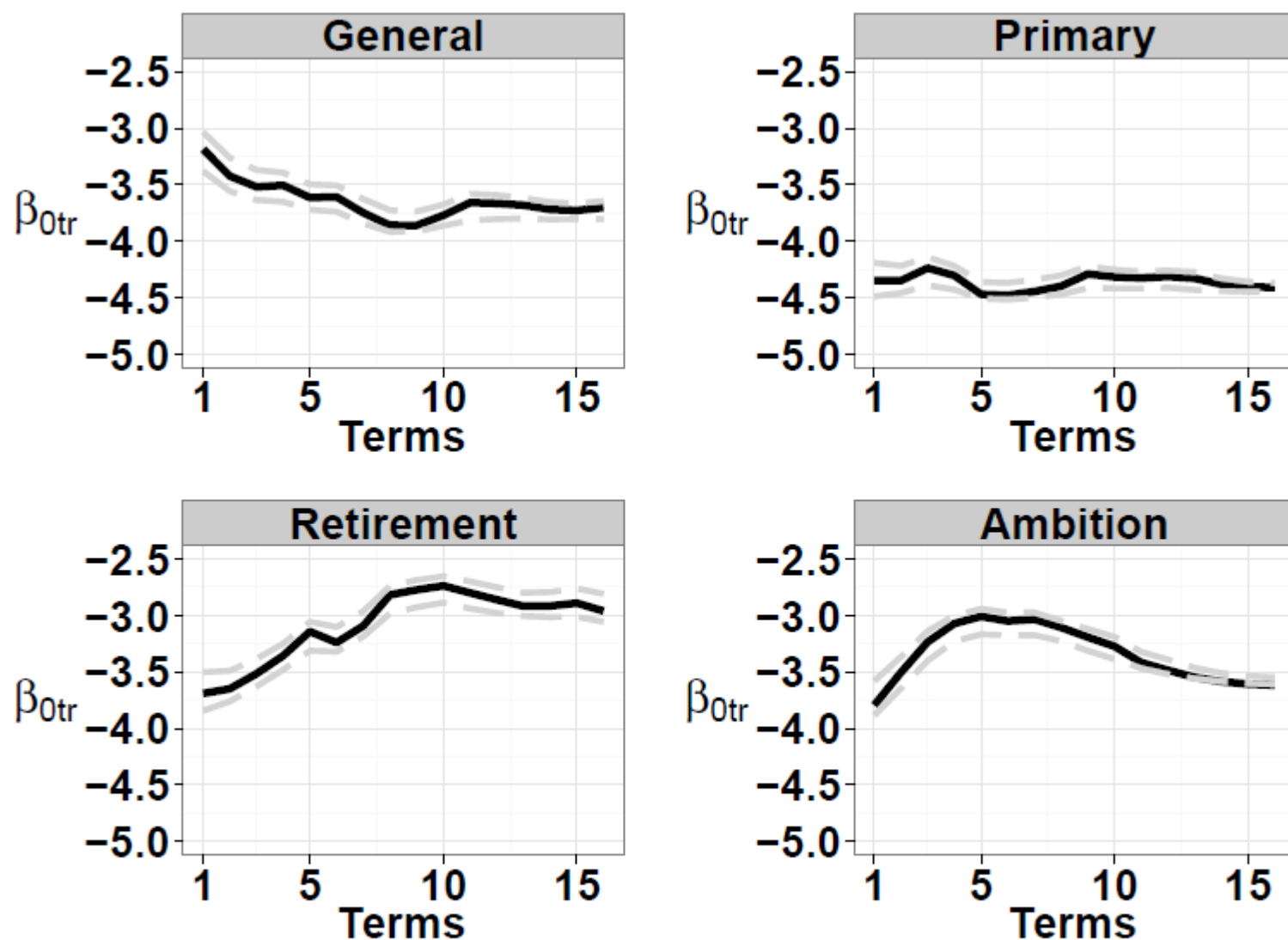


FIGURE 2: *Parameter estimates of the cause-specific time-varying baseline effects for the Congressional careers data. Dashed lines represent the 95% pointwise bootstrap interval.*

	General			Primary			Retirement			Ambition		
	ML	pen.	sd	ML	pen.	sd	ML	pen.	sd	ML	pen.	sd
Age	0.069	0.046	0.008	0.071	0.046	0.011	0.070	0.068	0.008	-0.034	-0.037	0.007
Republican	0.255	0	0.005	-0.188	0	0.002	-0.201	0	0.009	0.343	0	0.018
PriorMargin	-0.078	-0.060	0.005	0.006	0.001	0.005	-0.007	-0.005	0.003	-0.010	-0.004	0.002
Leadership	-0.272	0	0.087	-2.779	0	0.081	-0.393	0	0.065	0.033	0	0.080
Open Gub.	0.815	0.205	0.116	0.598	0.181	0.097	0.227	0.109	0.077	0.528	0.208	0.121
Open Sen.	-0.638	-0.243	0.125	-0.215	-0.193	0.134	-0.086	0.062	0.125	1.136	0.878	0.134
Scandal	3.750	2.689	0.370	3.215	3.272	0.428	1.921	1.611	0.441	-3.118	-1.532	0.073
Redistricting	2.548	1.617	0.447	1.465	1.149	0.499	-0.563	0.431	0.251	0.574	0.801	0.309
Age:Republican	0.007	0.011	0.007	-0.045	-0.010	0.007	0.041	0.030	0.009	-0.038	-0.029	0.009
Age:PriorMargin	0.001	0.000	0.000	-0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.0000
Age:Leadership	0.014	0	0.002	-0.117	0	0.002	0.018	0	0.002	-0.269	0	0.001
Age:OpenGub.	-0.006	0	0	0.034	0	0	-0.016	0	0	-0.011	0	0
Age:OpenSen.	-0.005	0	0.001	-0.074	0	0.001	-0.039	0	0.004	-0.015	0	0.002
Age:Scandal	-0.106	0	0	0.022	0	0	0.090	0	0	0.009	0	0
Age:Redistricting	-0.001	0.007	0.016	-0.066	-0.039	0.018	0.174	0.097	0.031	0.037	0.018	0.016
Republican:PriorMargin	0.016	0.005	0.004	-0.041	-0.016	0.005	-0.008	-0.004	0.004	0.015	0.012	0.004
Republican:OpenGub.	-0.532	-0.342	0.200	-4.282	-1.337	0.147	-0.147	-0.233	0.201	-0.063	0.294	0.184
Republican:OpenSen.	0.323	0	0.001	-0.092	0	0.002	0.802	0	0.010	-0.260	0	0.011
Republican:Scandal	0.007	0	0.021	2.121	0	0.054	0.182	0	0.005	-1.418	0	0.001
Republican:Redistricting	-1.833	0	0.076	0.447	0	0.059	1.247	0	0.050	-0.276	0	0.051
PriorMargin:Leadership	0.025	0	0	-0.009	0	0	-0.008	0	0.001	0.057	0	0
PriorMargin:OpenGub.	0.020	0	0	-0.001	0	0.001	0.008	0	0.001	0.009	0	0.001
PriorMargin:OpenSen.	-0.016	0	0.001	-0.019	0	0.002	0.013	0	0.002	0.011	0	0.004
PriorMargin:Scandal	0.006	0.007	0.005	-0.017	-0.010	0.004	-0.071	-0.019	0.006	-0.028	-0.001	0
PriorMargin:Redistricting	0.066	0.037	0.019	0.000	-0.002	0.003	0.030	0.010	0.006	-0.013	-0.009	0.007
Leadership:OpenGub.	-5.168	0	0.117	-1.693	0	0.087	1.054	0	0.359	-5.402	0	0.116
Leadership:OpenSen.	-4.513	0	0	-0.941	0	0	1.001	0	0	-6.053	0	0
Leadership:Scandal	-0.213	-0.029	0.594	-4.212	-1.803	0.733	-8.621	-1.925	0.756	-0.897	-0.108	0.047
OpenGub.:OpenSen.	-0.436	0	0	0.124	0	0	-0.280	0	0	-0.429	0	0
OpenGub.:Redistricting	-0.175	0.172	0.663	-4.274	-0.415	0.125	-5.297	-0.666	0.237	2.751	2.126	0.932
OpenSen.:Scandal	-2.277	0	0.307	-1.482	0	0.206	-8.270	0	0.266	-3.311	0	0.058
OpenSen.:Redistricting	0.914	0	0.052	-4.560	0	0.006	-0.522	0	0.031	1.771	0	0.147

Table 2: Parameter estimates for the Congressional careers data. Ordinary maximum likelihood estimates are denoted by “ML”, the penalized estimates are denoted by “pen.”. Estimated standard errors for the penalized model obtained by a bootstrap approach are given in the columns denoted by “sd”.

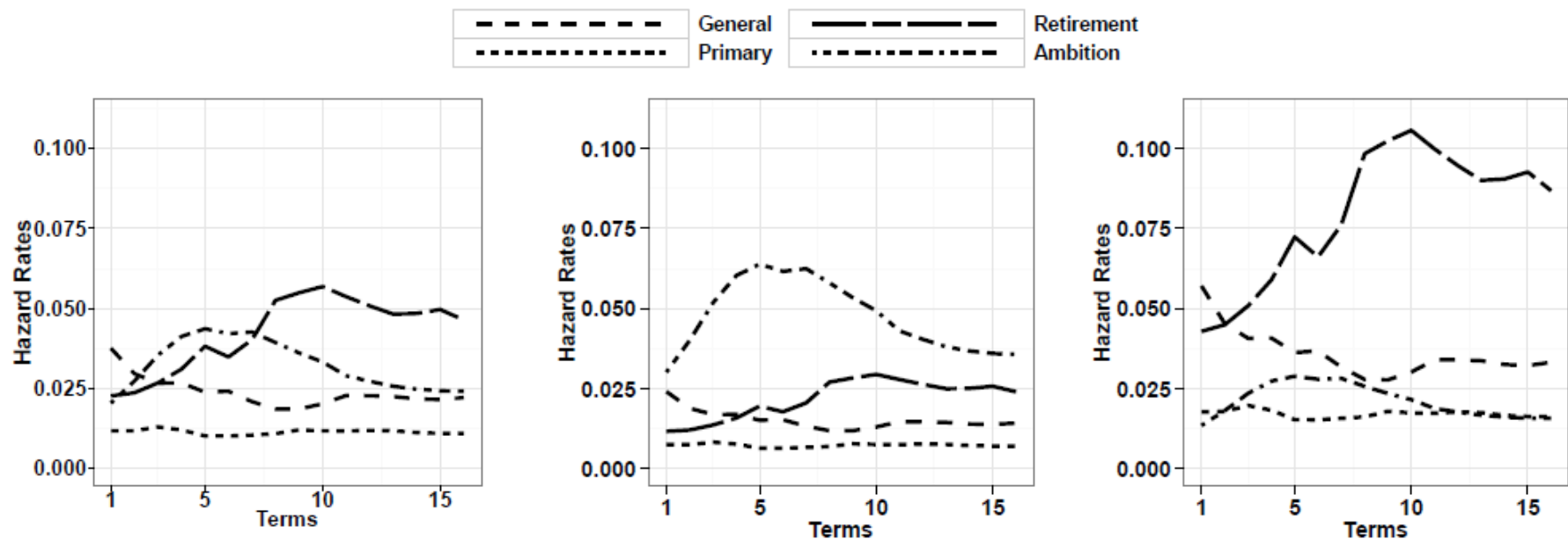


FIGURE 3: *Estimated cause-specific hazard rates over time for the Congressional careers data.*

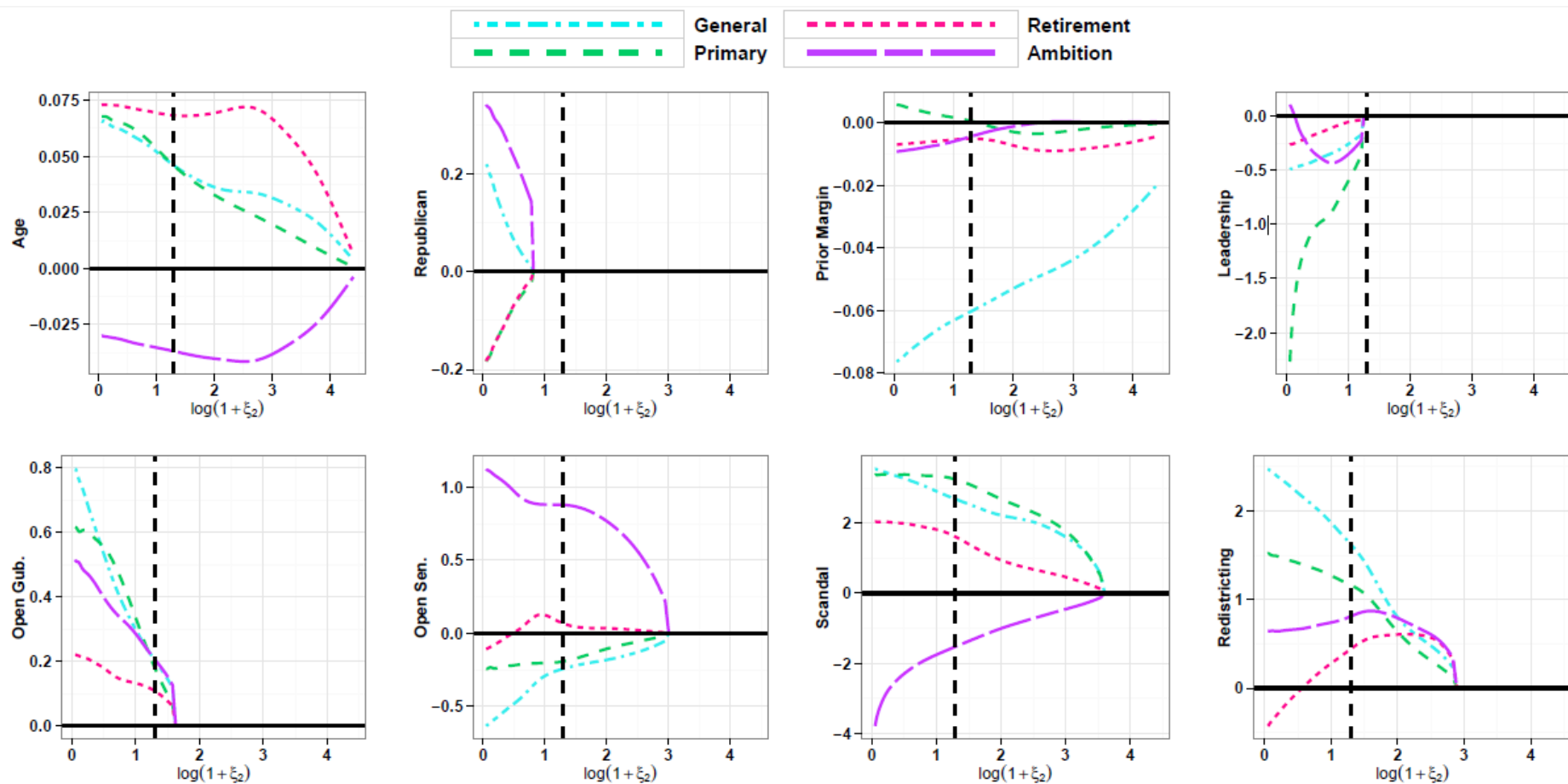


Figure 4 Coefficients paths of the main effects for the Congressional career data.

Conclusion

A penalized discrete competing risk model:

the cause-specific hazard rates

penalization method: to decrease parameters;

variable selection instead of parameter selection;

an additional penalty term: to avoid adjacent parameters of the baseline effects have completely different values.

Thank you for your listening!