SUPPLEMENTARY MATERIAL FOR

Subgroup analysis and correlated covariates detection for high-dimensional longitudinal data

This supplementary material includes three parts. Specifically, Section B1 presents additional simulation results with initial values generated by the second method (3.6). Section B2 provides the MATLAB code for two different initial values. Section B3 provides the detected countries in the empirical application.

B1.Additional Simulation Results

Table S1. The accuracy of group structure recovery for the proposed, MH and WZ's methods in the low-dimensional scenario.

			Proposed			MH			WZ		
m_i	n	p	$ARI_{\boldsymbol{\alpha}}$	$\hat{K}$	F_1	$ARI_{\boldsymbol{\alpha}}$	$\hat{K}$	F_1	$ARI_{\boldsymbol{\alpha}}$	$\hat{K}$	F_1
5	30	50	0.684	3.110	0.100	0.700	3.350	0.095	0.601	2.760	0.075
	30	100	0.662	2.935	0.060	0.689	3.265	0.065	0.538	2.520	0.035
	50	50	0.727	3.660	0.020	0.728	3.740	0.020	0.701	3.480	0.020
	50	100	0.692	3.720	0.015	0.700	3.915	0.015	0.657	3.490	0.020
	100	50	0.737	4.790	0	0.736	4.820	0	0.736	4.790	0
	100	100	0.753	4.560	0.005	0.751	4.655	0.005	0.752	4.525	0.005
10	30	50	0.824	2.970	0.320	0.834	3.050	0.325	0.754	2.755	0.290
	30	100	0.822	2.865	0.355	0.833	2.960	0.355	0.757	2.705	0.325
	50	50	0.849	3.290	0.110	0.857	3.385	0.110	0.844	3.270	0.110
	50	100	0.865	3.255	0.140	0.867	3.295	0.135	0.851	3.210	0.125
	100	50	0.886	4.045	0.035	0.885	4.045	0.035	0.881	3.975	0.035
	100	100	0.895	3.810	0.060	0.893	3.825	0.060	0.891	3.745	0.055
20	30	50	0.973	3	0.825	0.974	3.010	0.820	0.957	2.960	0.800
	30	100	0.974	2.980	0.850	0.974	2.980	0.850	0.943	2.910	0.790
	50	50	0.992	3.030	0.875	0.992	3.03	0.875	0.992	3.035	0.870
	50	100	0.977	3.055	0.765	0.977	3.055	0.765	0.976	3.045	0.765
	100	50	0.988	3.110	0.650	0.988	3.110	0.650	0.989	3.110	0.655
	100	100	0.989	3.100	0.695	0.989	3.100	0.695	0.989	3.100	0.690
30	30	50	0.992	2.985	0.970	0.994	2.990	0.975	0.987	2.975	0.960
	30	100	0.992	2.990	0.955	0.992	2.990	0.955	0.979	2.960	0.925
	50	50	0.997	3.005	0.945	0.997	3.005	0.945	0.995	3	0.945
	50	100	0.996	3.005	0.940	0.996	3.005	0.940	0.996	3.005	0.940
	100	50	0.997	3.040	0.910	0.997	3.040	0.910	0.997	3.040	0.910
	100	100	0.998	3.035	0.925	0.998	3.035	0.925	0.997	3.035	0.920

Table S2. The accuracy of covariates structure recovery for the proposed and WZ's methods in the low-dimensional scenario.

			Proposed					WZ				
m_i	n	p	$ARI_{\lvert\boldsymbol{\beta}\rvert}$	F_2	F_3	FN	FP	$ARI_{\lvert\boldsymbol{\beta}\rvert}$	F_2	F_3	FN	FP
5	30	50	0.896	0.135	0.095	1.435	0	0.668	0	0	2.865	0.005
	30	100	0.920	0.020	0.015	2.135	0	0.842	0	0	3.590	0.010
	50	50	0.935	0.345	0.300	0.945	0	0.703	0	0	2.040	0
	50	100	0.958	0.265	0.240	1.145	0	0.889	0	0	2.155	0
	100	50	0.988	0.870	0.830	0.185	0	0.712	0	0	1.825	0
	100	100	0.990	0.800	0.775	0.280	0	0.899	0	0	1.840	0
10	30	50	0.930	0.390	0.345	1.035	0	0.698	0	0	2.155	0
	30	100	0.961	0.325	0.285	1.050	0	0.892	0	0	2.060	0
	50	50	0.950	0.555	0.525	0.760	0	0.710	0	0	1.870	0
	50	100	0.976	0.590	0.570	0.660	0	0.901	0	0	1.785	0
	100	50	0.991	0.910	0.910	0.140	0	0.716	0	0	1.735	0
	100	100	0.997	0.940	0.940	0.095	0	0.900	0	0	1.790	0
20	30	50	0.960	0.635	0.605	0.610	0	0.713	0	0	1.790	0
	30	100	0.973	0.575	0.560	0.750	0	0.899	0	0	1.835	0
	50	50	0.977	0.810	0.810	0.355	0	0.717	0	0	1.710	0
	50	100	0.984	0.770	0.770	0.430	0	0.900	0	0	1.805	0
	100	50	0.999	0.995	0.995	0.010	0	0.717	0	0	1.700	0
	100	100	0.998	0.975	0.975	0.050	0	0.903	0	0	1.720	0
30	30	50	0.976	0.810	0.805	0.365	0	0.713	0	0	1.790	0
	30	100	0.981	0.715	0.715	0.525	0	0.903	0	0	1.700	0
	50	50	0.984	0.880	0.880	0.240	0	0.715	0	0	1.740	0
	50	100	0.987	0.820	0.820	0.355	0	0.904	0	0	1.680	0
	100	50	1	1	1	0	0	0.719	0	0	1.655	0
	100	100	0.999	0.985	0.985	0.030	0	0.903	0	0	1.710	0

Table S3. Ablation study for the proposed method in the low-dimensional scenario.

			AS.Subtraction				AS.Addition			
m_i	n	p	$ARI_{\|\boldsymbol{\beta}\|}$	F_2	FN	FP	$ARI_{\|\boldsymbol{\beta}\|}$	F_2	FN	FP
5	30	50	0.866	0.005	1.980	0	0.800	0.120	1.605	0
	30	100	0.908	0.005	2.465	0	0.898	0.125	2.295	0
	50	50	0.869	0.005	1.965	0	0.827	0.340	1.130	0
	50	100	0.927	0.010	1.980	0	0.929	0.450	1.380	0
	100	50	0.875	0.010	1.960	0	0.866	0.850	0.550	0
	100	100	0.929	0.010	1.960	0	0.957	0.950	0.595	0
10	30	50	0.867	0	2.025	0	0.819	0.260	1.305	0
	30	100	0.926	0	2.015	0	0.930	0.390	1.355	0
	50	50	0.869	0	1.990	0	0.849	0.700	0.805	0
	50	100	0.927	0	2.000	0	0.948	0.830	0.865	0
	100	50	0.874	0	1.925	0	0.881	0.955	0.295	0
	100	100	0.929	0	1.930	0	0.965	0.995	0.370	0
20	30	50	0.870	0	1.975	0	0.846	0.515	0.855	0
	30	100	0.926	0	2.020	0	0.940	0.735	1.100	0
	50	50	0.870	0	1.990	0	0.882	0.980	0.265	0
	50	100	0.927	0	1.990	0	0.963	0.995	0.420	0
	100	50	0.880	0	1.830	0	0.894	1	0.050	0
	100	100	0.932	0	1.850	0	0.971	1	0.200	0
30	30	50	0.873	0	1.945	0	0.866	0.735	0.575	0
	30	100	0.928	0	1.980	0	0.947	0.875	0.910	0
	50	50	0.874	0	1.935	0	0.888	0.985	0.165	0
	50	100	0.929	0	1.965	0	0.966	0.995	0.325	0
	100	50	0.883	0	1.780	0	0.895	1	0.035	0
	100	100	0.934	0	1.815	0	0.976	1	0.060	0

Table S4. The accuracy of parameter estimation for the proposed and WZ's methods in the high-dimensional scenario.

			Proposed		WZ	
m_i	n	p	$\text{mRMSE}_{\boldsymbol{\alpha}} * 10^2$ ($\text{sRMSE}_{\boldsymbol{\alpha}} * 10^2$)	$\text{mRMSE}_{\boldsymbol{\beta}} * 10^3$ ($\text{sRMSE}_{\boldsymbol{\beta}} * 10^3$)	$\text{mRMSE}_{\boldsymbol{\alpha}} * 10^2$ ($\text{sRMSE}_{\boldsymbol{\alpha}} * 10^2$)	$\text{mRMSE}_{\boldsymbol{\beta}} * 10^3$ ($\text{sRMSE}_{\boldsymbol{\beta}} * 10^3$)
30	200	500	9.771	2.273	9.836	62.011
			(5.866)	(1.157)	(5.873)	(14.192)
	200	1000	10.184	1.517	10.218	43.558
			(6.474)	(0.762)	(6.453)	(10.553)
	500	500	7.289	1.421	7.321	56.554
			(4.733)	(0.785)	(4.715)	(14.470)
	500	1000	7.795	0.928	7.814	40.098
			(4.813)	(0.529)	(4.816)	(10.237)
50	200	500	5.906	1.872	5.940	58.439
			(2.494)	(0.971)	(2.489)	(14.306)
	200	1000	5.774	1.265	5.826	42.119
			(1.530)	(0.679)	(1.525)	(10.015)
	500	500	3.705	1.095	3.726	55.371
			(1.317)	(0.557)	(1.319)	(14.532)
	500	1000	3.701	0.781	3.739	38.828
			(1.473)	(0.440)	(1.493)	(10.231)
100	200	500	4.283	1.168	4.307	54.003
			(1.304)	(0.682)	(1.315)	(14.371)
	200	1000	4.138	0.836	4.159	39.110
			(1.224)	(0.422)	(1.242)	(10.275)
	500	500	2.683	0.743	2.707	54.231
			(0.777)	(0.406)	(0.782)	(14.515)
	500	1000	2.619	0.547	2.625	38.962
			(0.756)	(0.324)	(0.773)	(10.323)

Remark .1. *Obviously, the first method (3.1) can give relatively better initial values because it takes into account the correlation.*

B2.The MATLAB Code

The corresponding MATLAB code is available at the following site:MatlabCode

B3.The detected countries in empirical application

The corresponding detected countries are available at the following site:DetectedCountries