Supplementary Materials for "Identification and Estimation under Multiple Versions of Treatment: Mixture-of-Experts Approach"

Kohei Yoshikawa [*†] **Shuichi Kawano** [‡]

January 1, 2026

S1 Additional Results for Simulation Studies

Table S 1: Results of Monte Carlo simulations for $J = 2$ with version structure $(3, 3)$.

p	SNR	n	$(t, v), (t, v')$	Bias	SD	p	SNR	n	$(t, v), (t, v')$	Bias	SD
10	5	500	(0, 0),(0, 1)	–0.551712	0.458374	20	5	500	(0, 0),(0, 1)	–0.707968	0.506136
			(0, 1),(0, 2)	–0.471244	0.593685				(0, 1),(0, 2)	–0.844283	0.527499
			(1, 0),(1, 1)	–0.214661	2.466008				(1, 0),(1, 1)	–0.799751	1.094747
			(1, 1),(1, 2)	–0.815227	2.265468				(1, 1),(1, 2)	–0.743909	1.024243
		1000	(0, 0),(0, 1)	–0.279642	0.391158			1000	(0, 0),(0, 1)	–0.606764	0.363956
			(0, 1),(0, 2)	–0.473747	0.396130				(0, 1),(0, 2)	–0.541640	0.388322
			(1, 0),(1, 1)	–0.396367	0.768891				(1, 0),(1, 1)	–0.705720	0.781761
			(1, 1),(1, 2)	–0.373775	1.290072				(1, 1),(1, 2)	–0.532661	0.812710
		2000	(0, 0),(0, 1)	–0.248394	0.357212			2000	(0, 0),(0, 1)	–0.445068	0.385567
			(0, 1),(0, 2)	–0.394400	0.359251				(0, 1),(0, 2)	–0.364514	0.377402
			(1, 0),(1, 1)	–0.358984	0.582568				(1, 0),(1, 1)	–0.561352	0.619937
			(1, 1),(1, 2)	–0.578707	0.644446				(1, 1),(1, 2)	–0.480460	0.956558
	10	500	(0, 0),(0, 1)	–0.359351	0.610603		10	500	(0, 0),(0, 1)	–0.658402	0.408634
			(0, 1),(0, 2)	–0.377299	0.904225				(0, 1),(0, 2)	–0.780896	0.508393
			(1, 0),(1, 1)	–0.870598	1.050258				(1, 0),(1, 1)	–0.894695	1.945616
			(1, 1),(1, 2)	–0.530266	0.904071				(1, 1),(1, 2)	–0.599389	2.257911
		1000	(0, 0),(0, 1)	–0.223664	0.394502			1000	(0, 0),(0, 1)	–0.528546	0.509852
			(0, 1),(0, 2)	–0.445415	0.381609				(0, 1),(0, 2)	–0.406291	0.508146
			(1, 0),(1, 1)	–0.723602	0.851517				(1, 0),(1, 1)	–1.093205	0.867631
			(1, 1),(1, 2)	–0.333697	1.262244				(1, 1),(1, 2)	–0.477498	0.913059
		2000	(0, 0),(0, 1)	–0.151458	0.343766			2000	(0, 0),(0, 1)	–0.346055	0.458460
			(0, 1),(0, 2)	–0.460660	0.291630				(0, 1),(0, 2)	–0.355370	0.384591
			(1, 0),(1, 1)	–0.438220	0.685977				(1, 0),(1, 1)	–0.843612	0.908934
			(1, 1),(1, 2)	0.094631	4.914994				(1, 1),(1, 2)	–0.272797	1.273020

[*]Graduate School of Mathematics, Kyushu University, 744, Motooka, Nishi-ku, Fukuoka, 819-0395 Japan. e-mail: (yoshikawa.kohei.603 (at) s.kyushu-u.ac.jp)

[†]NTT DATA Mathematical Systems Inc., 1F Shinanomachi Rengakan, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-0016, Japan. e-mail: (yoshikawa (at) msi.co.jp)

[‡]Faculty of Mathematics, Kyushu University, 744 Motooka, Nishi-ku, Fukuoka, 819-0395 Japan. e-mail: (skawano (at) math.kyushu-u.ac.jp)

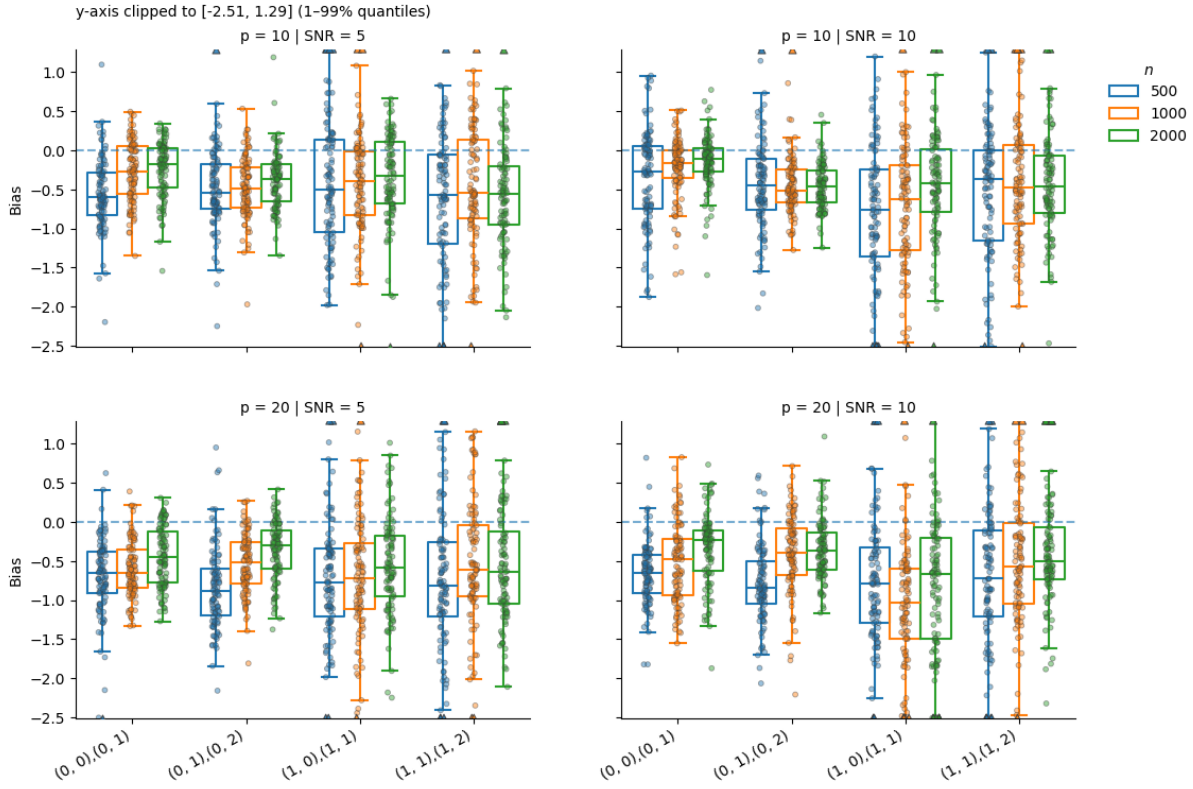

Figure S 1: Distribution of the bias across Monte Carlo replications. For two treatment-version pairs (t, v) and (t, v'), the boxplots summarize the bias stratified by sample size $\{500, 1000, 2000\}$, SNR $\{5, 10\}$, and the covariate dimension $\{10, 20\}$. Points represent individual Monte Carlo replications. The dashed horizontal line indicates zero bias. This plot corresponds to the configuration $J = 2$ with version structure $(3, 3)$.

Table S 2: Results of Monte Carlo simulations for $J = 3$ with version structure $(2, 2, 2)$.

p	SNR	n	$(t, v), (t, v')$	Bias	SD	p	SNR	n	$(t, v), (t, v')$	Bias	SD
10	5	500	(0, 0),(0, 1)	−0.794535	0.620591	20	5	500	(0, 0),(0, 1)	−0.899272	0.322530
			(1, 0),(1, 1)	−0.822343	0.876196				(1, 0),(1, 1)	−0.844009	0.781225
			(2, 0),(2, 1)	0.128274	11.122714				(2, 0),(2, 1)	−0.898328	0.960484
		1000	(0, 0),(0, 1)	−0.667052	0.402377			1000	(0, 0),(0, 1)	−0.904489	0.271696
			(1, 0),(1, 1)	−0.714652	0.605393				(1, 0),(1, 1)	−0.926727	0.738980
			(2, 0),(2, 1)	0.069483	5.965427				(2, 0),(2, 1)	−0.881478	0.975845
		2000	(0, 0),(0, 1)	−0.400396	0.399054			2000	(0, 0),(0, 1)	−0.741543	0.350526
			(1, 0),(1, 1)	−0.556791	0.630337				(1, 0),(1, 1)	−0.610226	0.767296
			(2, 0),(2, 1)	−0.239225	0.751491				(2, 0),(2, 1)	−0.503652	0.637189
	10	500	(0, 0),(0, 1)	−0.629530	0.588353		10	500	(0, 0),(0, 1)	−0.937347	0.400495
			(1, 0),(1, 1)	−0.468826	0.763242				(1, 0),(1, 1)	−0.821011	0.937891
			(2, 0),(2, 1)	0.347013	5.927685				(2, 0),(2, 1)	−0.858508	1.227051
		1000	(0, 0),(0, 1)	−0.224204	0.360983			1000	(0, 0),(0, 1)	−0.623687	0.434492
			(1, 0),(1, 1)	−0.356690	1.188595				(1, 0),(1, 1)	−0.592114	1.034627
			(2, 0),(2, 1)	−0.218996	2.112210				(2, 0),(2, 1)	−0.352578	1.157178
		2000	(0, 0),(0, 1)	−0.119485	0.233503			2000	(0, 0),(0, 1)	−0.321010	0.362165
			(1, 0),(1, 1)	−0.236888	0.527463				(1, 0),(1, 1)	−0.262908	0.757589
			(2, 0),(2, 1)	−0.121069	1.418014				(2, 0),(2, 1)	−0.159216	0.861098

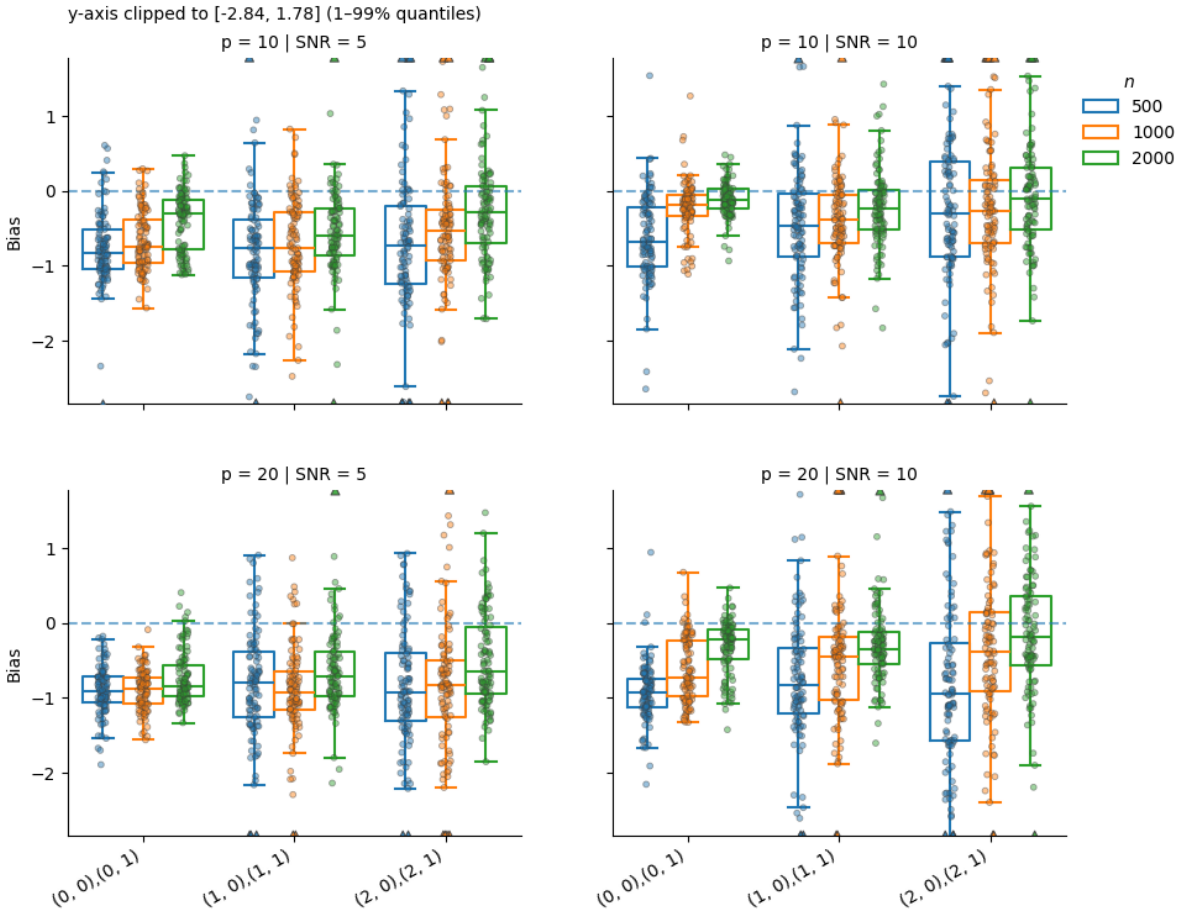

Figure S 2: Distribution of the bias across Monte Carlo replications. For two treatment-version pairs (t, v) and (t, v'), the boxplots summarize the bias stratified by sample size $\{500, 1000, 2000\}$, SNR $\{5, 10\}$, and the covariate dimension $\{10, 20\}$. Points represent individual Monte Carlo replications. The dashed horizontal line indicates zero bias. This plot corresponds to the configuration $J = 3$ with version structure $(2, 2, 2)$.

Table S 3: Results of Monte Carlo simulations for $J = 3$ with version structure $(3, 3, 3)$.

p	SNR	n	$(t, v), (t, v')$	Bias	SD	p	SNR	n	$(t, v), (t, v')$	Bias	SD
10	5	500	(0, 0),(0, 1)	–0.640001	0.561198	20	5	500	(0, 0),(0, 1)	–0.914625	0.685699
			(0, 1),(0, 2)	–0.813596	0.703809				(0, 1),(0, 2)	–0.923367	0.519712
			(1, 0),(1, 1)	–0.612394	1.117086				(1, 0),(1, 1)	–1.051582	1.546998
			(1, 1),(1, 2)	–0.811394	1.231529				(1, 1),(1, 2)	–0.567335	1.896159
			(2, 0),(2, 1)	–0.515518	3.047348				(2, 0),(2, 1)	–0.894565	1.610217
			(2, 1),(2, 2)	–0.301902	5.465927				(2, 1),(2, 2)	–0.673712	2.104935
		1000	(0, 0),(0, 1)	–0.438777	0.582824			1000	(0, 0),(0, 1)	–0.754259	0.408918
			(0, 1),(0, 2)	–0.616872	0.650947				(0, 1),(0, 2)	–0.336225	4.639719
			(1, 0),(1, 1)	–0.573595	1.267191				(1, 0),(1, 1)	–0.792802	1.367510
			(1, 1),(1, 2)	–0.516553	0.869275				(1, 1),(1, 2)	–0.855256	1.146068
			(2, 0),(2, 1)	–0.625169	1.141048				(2, 0),(2, 1)	–0.362423	1.345928
			(2, 1),(2, 2)	0.397583	3.907518				(2, 1),(2, 2)	–0.934545	1.403665
		2000	(0, 0),(0, 1)	–0.359673	0.413675			2000	(0, 0),(0, 1)	–0.639999	0.332662
			(0, 1),(0, 2)	–0.510531	0.521412				(0, 1),(0, 2)	–0.523703	0.490996
			(1, 0),(1, 1)	–0.502398	0.982730				(1, 0),(1, 1)	–0.505319	0.866375
			(1, 1),(1, 2)	–0.364993	0.883738				(1, 1),(1, 2)	–0.622937	1.256386
			(2, 0),(2, 1)	–0.329380	1.551912				(2, 0),(2, 1)	–0.611052	1.128800
			(2, 1),(2, 2)	–0.300967	2.461647				(2, 1),(2, 2)	–0.427441	1.264493
	10	500	(0, 0),(0, 1)	–0.728693	0.425431		10	500	(0, 0),(0, 1)	–0.874807	0.448140
			(0, 1),(0, 2)	–0.441111	0.638482				(0, 1),(0, 2)	–0.761149	0.580205
			(1, 0),(1, 1)	–0.505315	1.260564				(1, 0),(1, 1)	–0.935901	1.285237
			(1, 1),(1, 2)	–0.667623	1.482289				(1, 1),(1, 2)	–0.408446	1.960015
			(2, 0),(2, 1)	0.341506	7.438462				(2, 0),(2, 1)	–0.817912	1.986666
			(2, 1),(2, 2)	–1.052885	4.190656				(2, 1),(2, 2)	–0.525056	1.985713
		1000	(0, 0),(0, 1)	–0.415863	0.476534			1000	(0, 0),(0, 1)	–0.788573	0.426185
			(0, 1),(0, 2)	–0.426336	0.485171				(0, 1),(0, 2)	–0.580748	0.537924
			(1, 0),(1, 1)	–0.843505	1.741912				(1, 0),(1, 1)	–0.772932	1.619990
			(1, 1),(1, 2)	–0.534846	1.383282				(1, 1),(1, 2)	–0.534977	1.526662
			(2, 0),(2, 1)	–0.517098	1.620798				(2, 0),(2, 1)	–0.821309	1.747896
			(2, 1),(2, 2)	0.151595	4.443936				(2, 1),(2, 2)	–0.121784	1.893568
		2000	(0, 0),(0, 1)	–0.306081	0.714981			2000	(0, 0),(0, 1)	–0.551456	0.495435
			(0, 1),(0, 2)	–0.432849	0.498216				(0, 1),(0, 2)	–0.306021	0.649342
			(1, 0),(1, 1)	–0.356723	1.123591				(1, 0),(1, 1)	–0.699923	2.215479
			(1, 1),(1, 2)	–0.625488	1.133405				(1, 1),(1, 2)	–0.588227	1.791089
			(2, 0),(2, 1)	–0.470448	2.209943				(2, 0),(2, 1)	–0.773832	1.687227
			(2, 1),(2, 2)	–0.493539	2.050104				(2, 1),(2, 2)	–0.234762	1.300314

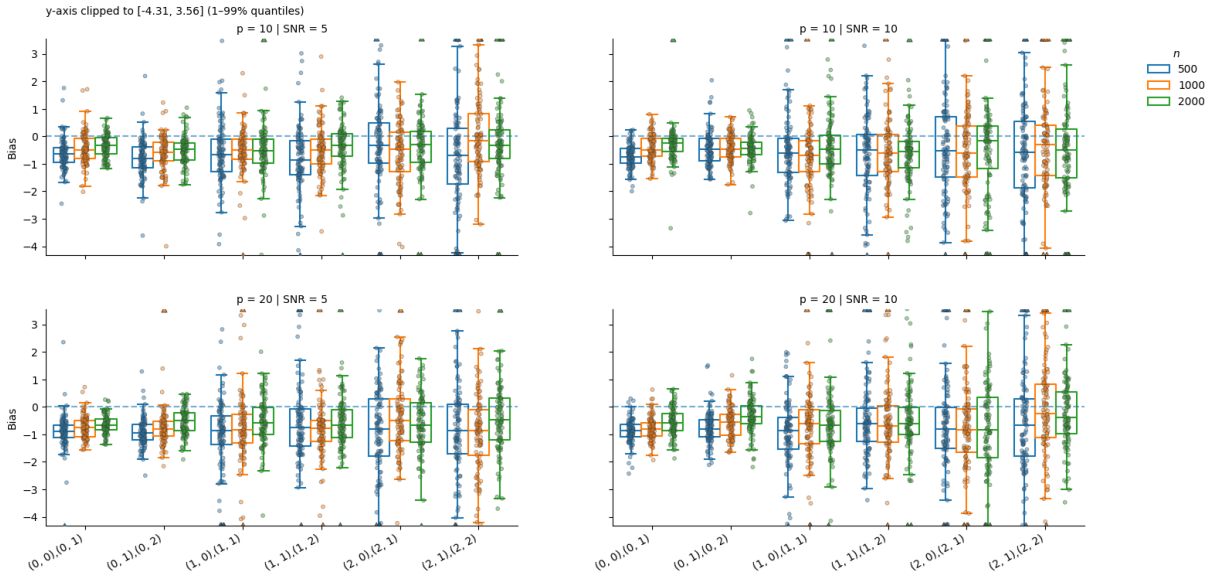

Figure S 3: Distribution of the bias across Monte Carlo replications. For two treatment-version pairs (t, v) and (t, v'), the boxplots summarize the bias stratified by sample size $\{500, 1000, 2000\}$, SNR $\{5, 10\}$, and the covariate dimension $\{10, 20\}$. Points represent individual Monte Carlo replications. The dashed horizontal line indicates zero bias. This plot corresponds to the configuration $J = 3$ with version structure $(3, 3, 3)$.

S2 Additional Results for Empirical Analysis

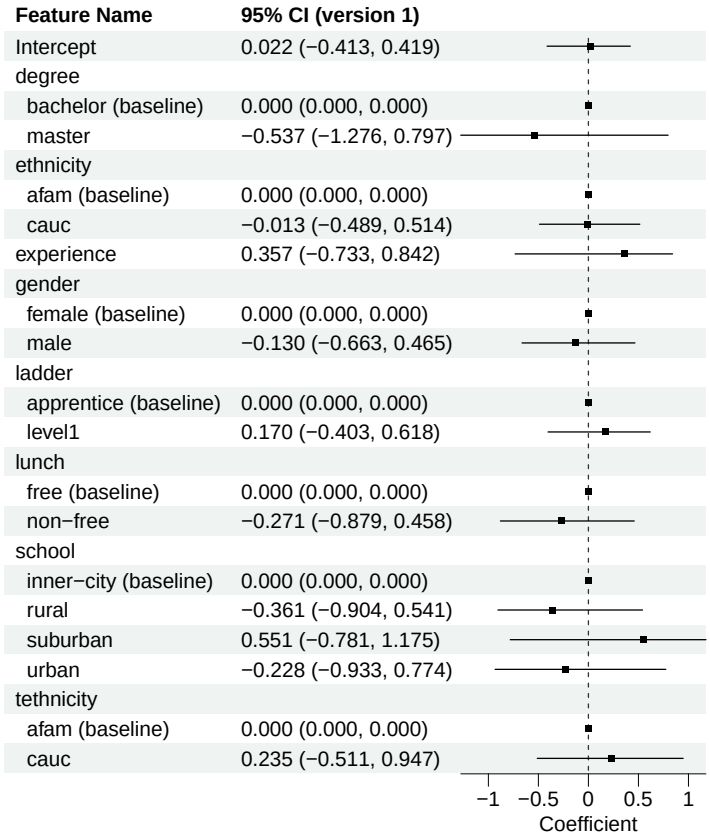

Figure S4: Forest plot of the estimated coefficients of the gating model for the latent version 1 in the `regular` class. Points denote mean estimates, and horizontal bars represent bootstrap 95% confidence intervals based on 100 bootstrap resamples.

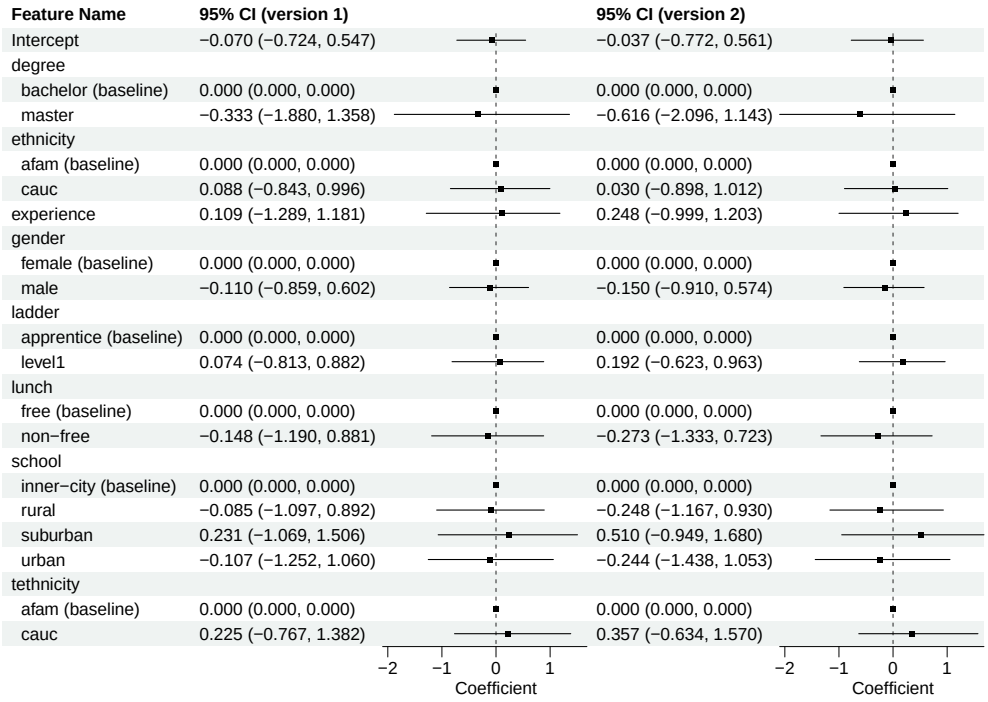

Figure S 5: Forest plot of the estimated coefficients of the gating model for the `regular` class. The plot is compared with latent versions 1 and 2. Points denote mean estimates, and horizontal bars represent bootstrap 95% confidence intervals based on 100 bootstrap resamples.

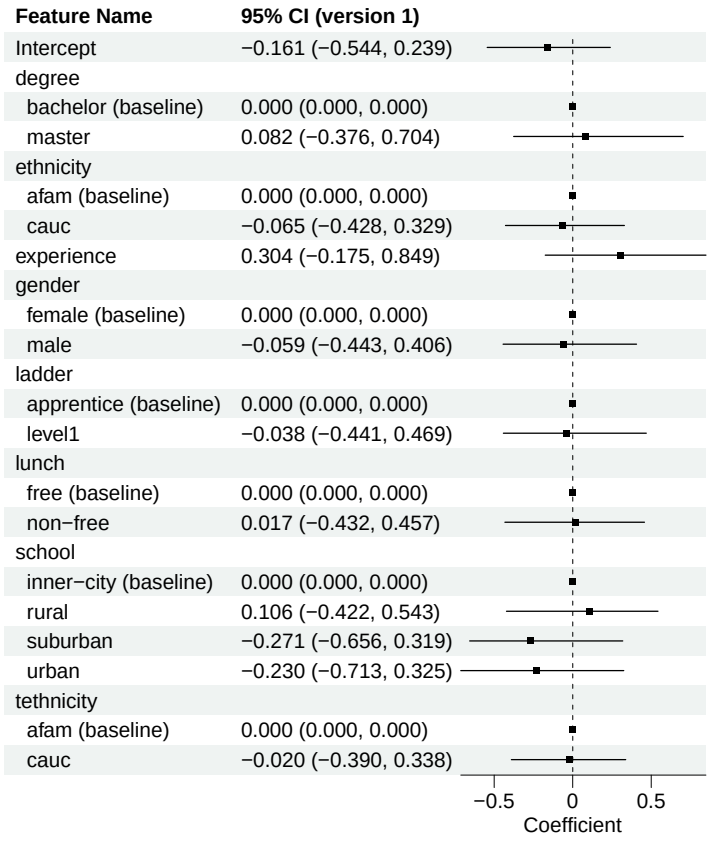

Figure S 6: Forest plot of the estimated coefficients of the gating model for the latent version 1 in the `regular+aide` class. Points denote mean estimates, and horizontal bars represent bootstrap 95% confidence intervals based on 100 bootstrap resamples.

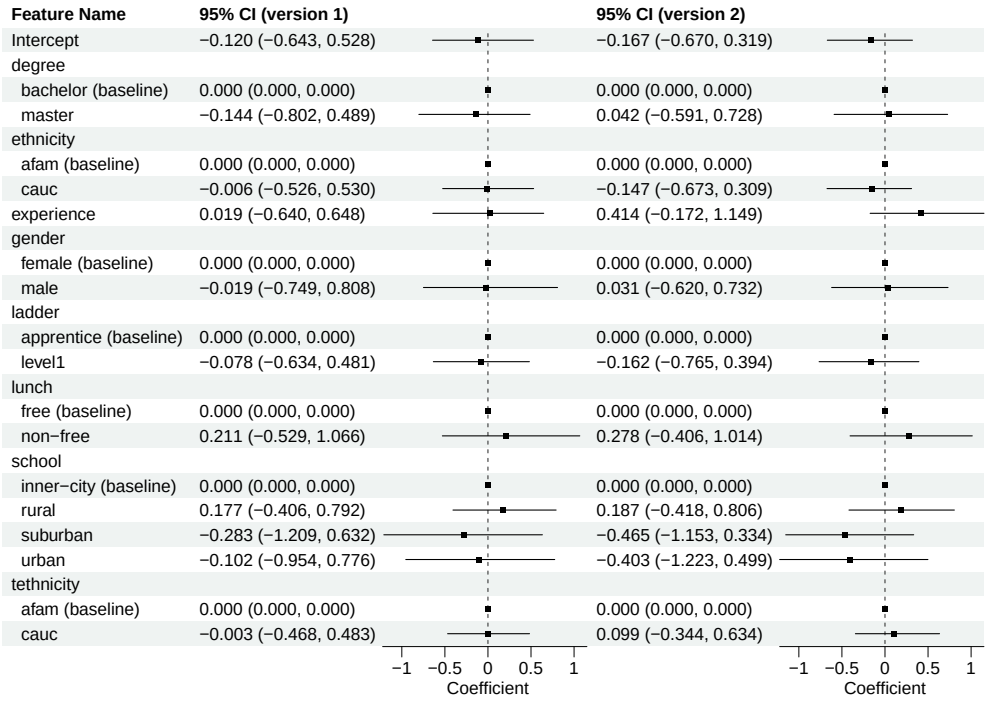

Figure S 7: Forest plot of the estimated coefficients of the gating model for the `regular+aide` class. The plot is compared with latent versions 1 and 2. Points denote mean estimates, and horizontal bars represent bootstrap 95% confidence intervals based on 100 bootstrap resamples.