

Supplementary Material for "Predictive Modeling of Short-Term Recidivism: A Mixture Cure Rate Approach with Diverse Link Functions"

CHAEYEON YOO^{1,*} AND DIPAK K. DEY¹

¹215 Glenbrook Rd, Storrs, University of Connecticut, USA

1 Additional Simulation Results

This supplementary material provides additional simulation results to support the findings presented in the main text. Specifically, it includes results for scenarios where the true latency distribution follows a lognormal or a loglogistic model, complementing the Weibull based simulations discussed in the main manuscript. Each case evaluates the performance of the four link functions, with true value recovery, coverage, and model assessment evaluation.

In addition, we present Cox-Snell residual plots of the Weibull based simulation that further assess model fit. These plots compare the fit of each link function when used to estimate data generated under each of the other link functions, offering insight into goodness-of-fit and the flexibility of the proposed link structures. These figures and tables are organized to match the structure of the main manuscript and provide supporting evidence for the flexibility of the proposed modeling framework.

*Corresponding author. Email: chaeyeon.yoo@uconn.edu.

Table 1: Coverage probabilities for each parameter for lognormal and loglogistic latency across difference link functions, and sample sizes.

N	Distribution	Parameter	Logit	SLogit	RPLogit	FGLogit
500	Lognormal	γ_1	0.939	0.926	0.941	1.000
		γ_2	0.959	0.936	0.938	1.000
		β_0	0.948	0.947	0.944	0.946
		β_1	0.946	0.951	0.952	0.946
		β_2	0.937	0.953	0.952	0.932
		α	—	0.934	—	1.000
		λ	—	—	0.947	1.000
		σ	0.953	0.945	0.954	0.950
	Loglogistic	γ_1	0.932	0.941	0.938	1.000
		γ_2	0.948	0.953	0.952	1.000
		β_0	0.949	0.950	0.952	0.934
		β_1	0.945	0.945	0.943	0.942
		β_2	0.946	0.932	0.945	0.938
		α	—	0.935	—	1.000
		λ	—	—	0.953	1.000
		σ	0.949	0.934	0.945	0.953
1000	Lognormal	γ_1	0.954	0.936	0.947	1.000
		γ_2	0.949	0.941	0.951	1.000
		β_0	0.937	0.935	0.957	0.953
		β_1	0.944	0.940	0.940	0.945
		β_2	0.937	0.933	0.943	0.945
		α	—	0.946	—	1.000
		λ	—	—	0.951	1.000
		σ	0.943	0.952	0.949	0.948
	Loglogistic	γ_1	0.945	0.943	0.936	1.000
		γ_2	0.954	0.951	0.944	1.000
		β_0	0.939	0.940	0.943	0.959
		β_1	0.940	0.953	0.934	0.949
		β_2	0.936	0.946	0.934	0.947
		α	—	0.941	—	1.000
		λ	—	—	0.942	1.000
		σ	0.951	0.941	0.954	0.949

Table 2: Number of times each model was selected under different true models with lognormal and loglogistic latency

	Criterion	True slogit	True rplogit	True fglogit
Lognormal	LOOIC	91	85	78
	WAIC	91	86	78
	DIC	92	89	87
Loglogistic	LOOIC	91	87	79
	WAIC	91	87	78
	DIC	88	81	71

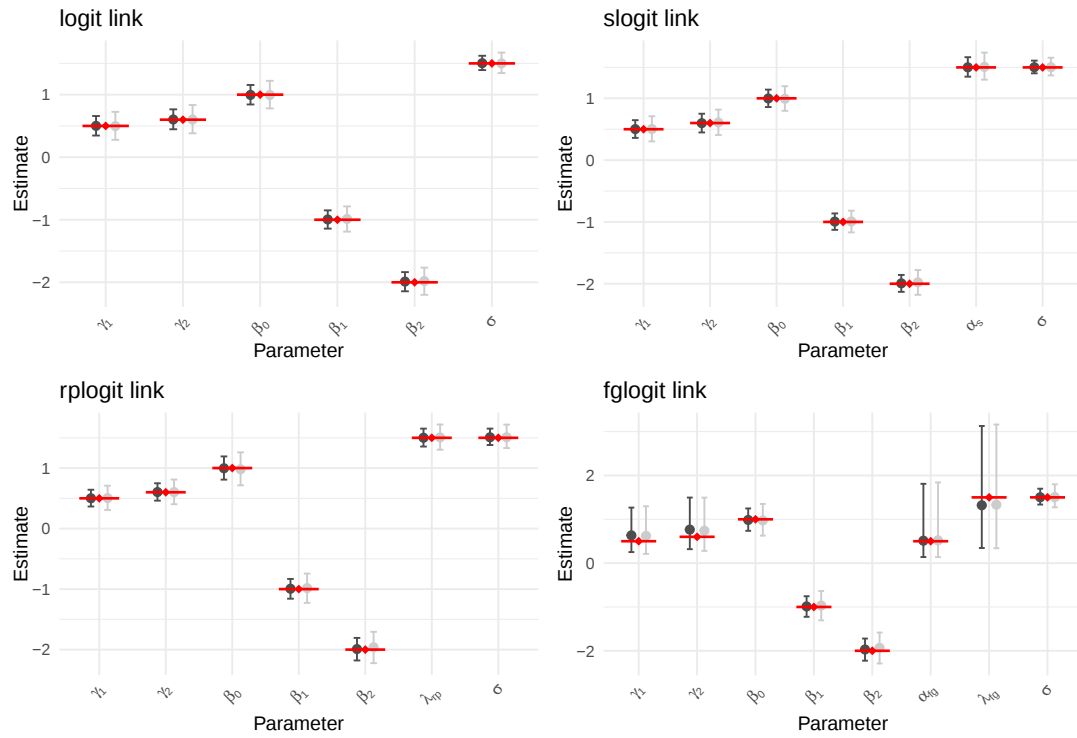


Figure 1: Posterior mean estimates and 95% credible intervals under the Lognormal latency model. Black lines represent $n = 1000$, gray lines represent $n = 500$, and the red horizontal line indicates the true parameter value.

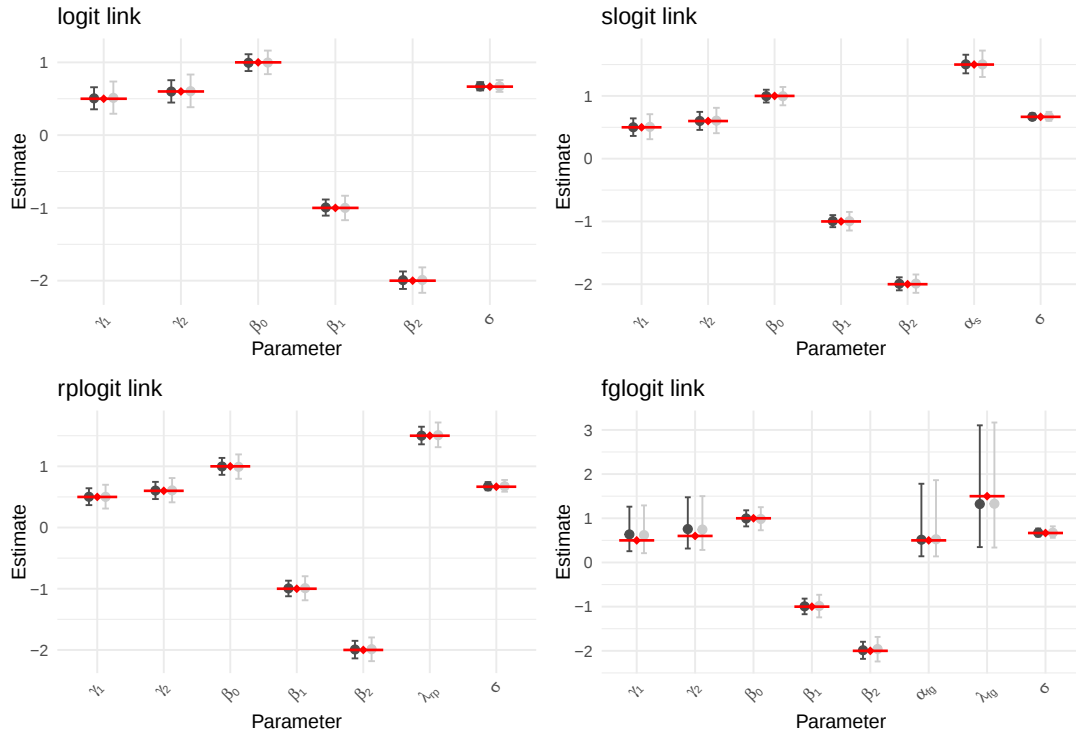


Figure 2: Posterior mean estimates and 95% credible intervals under the Loglogistic latency model. Black lines represent $n = 1000$, gray lines represent $n = 500$, and the red horizontal line indicates the true parameter value.

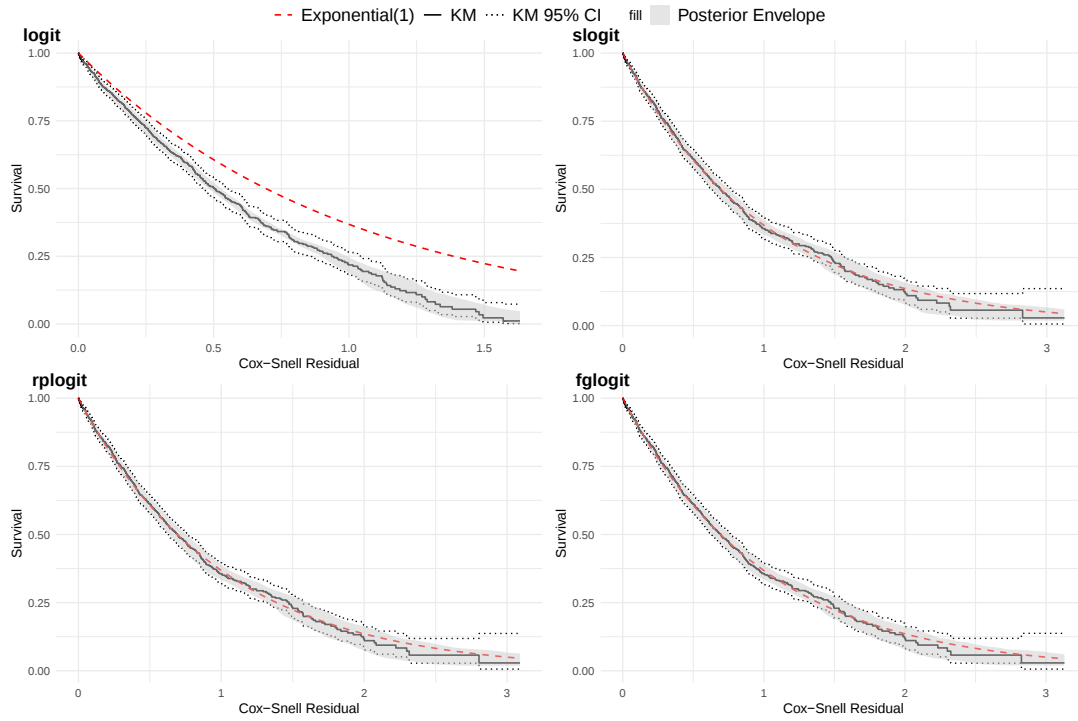


Figure 3: Cox-Snell residual plot with Weibull slogit link being true model

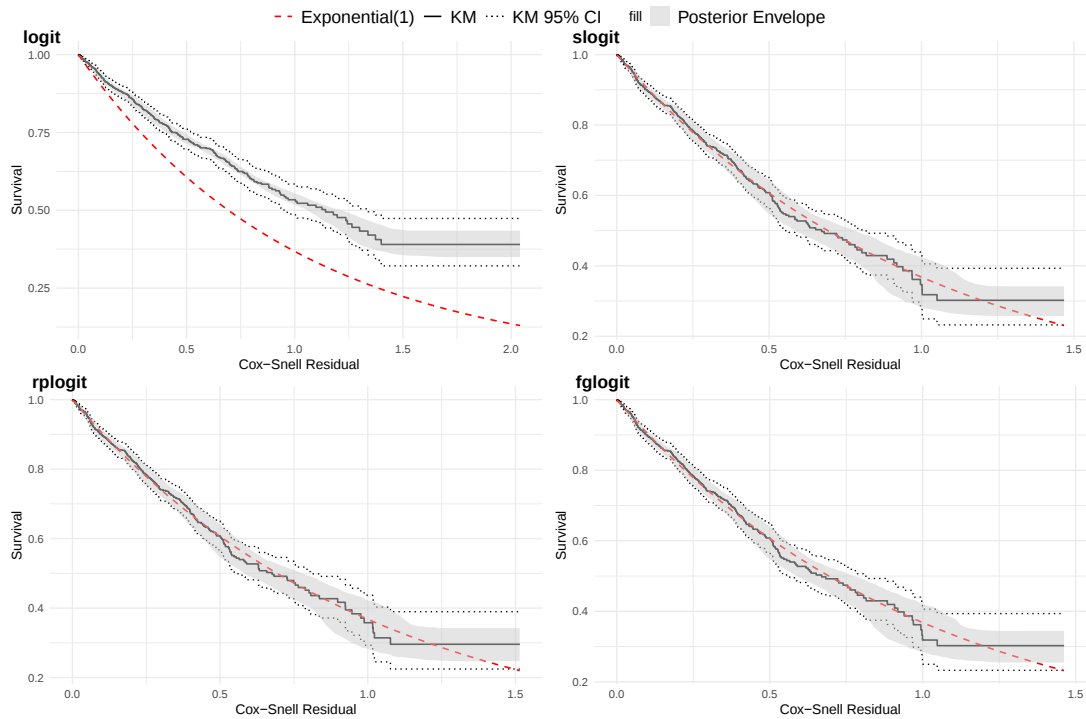


Figure 4: Cox-Snell residual plot with Weibull rplogit link being true model

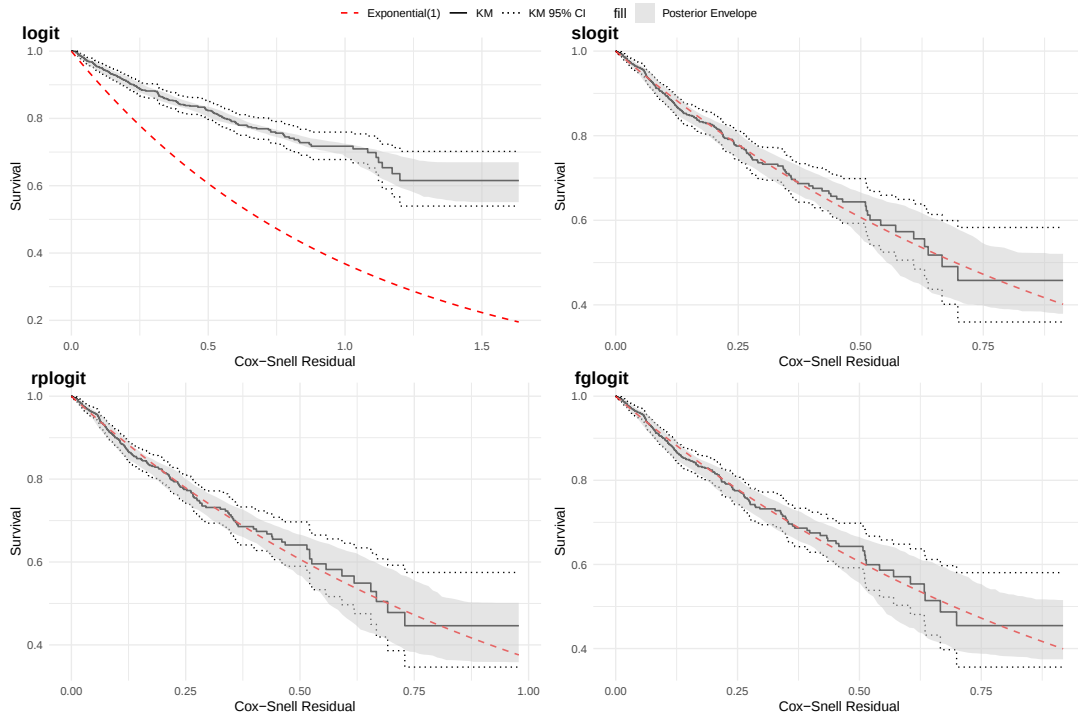


Figure 5: Cox-Snell residual plot with Weibull fglogit link being true model

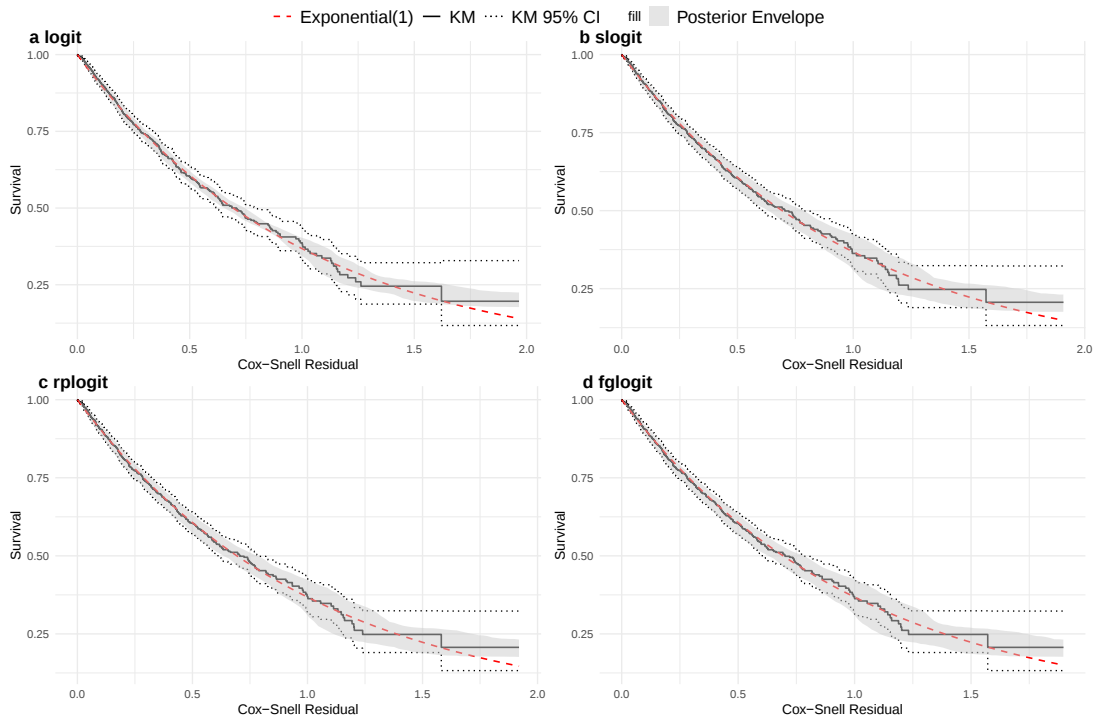


Figure 6: Cox-Snell residual plot with Lognormal logit link being true model

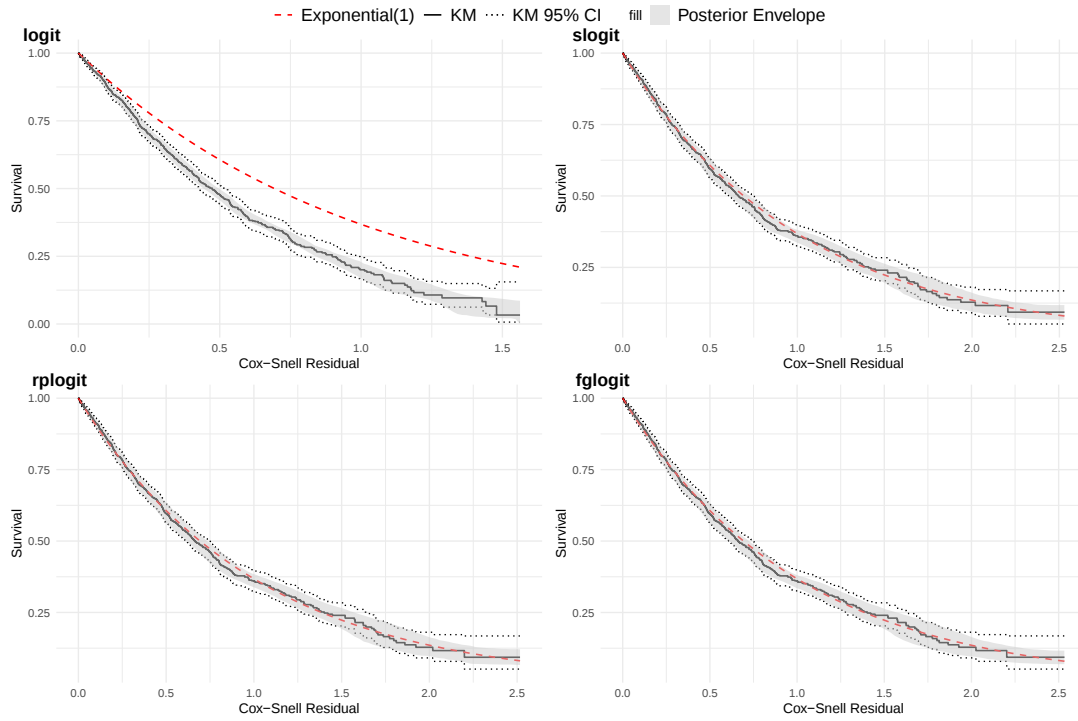


Figure 7: Cox-Snell residual plot with Lognormal slogit link being true model

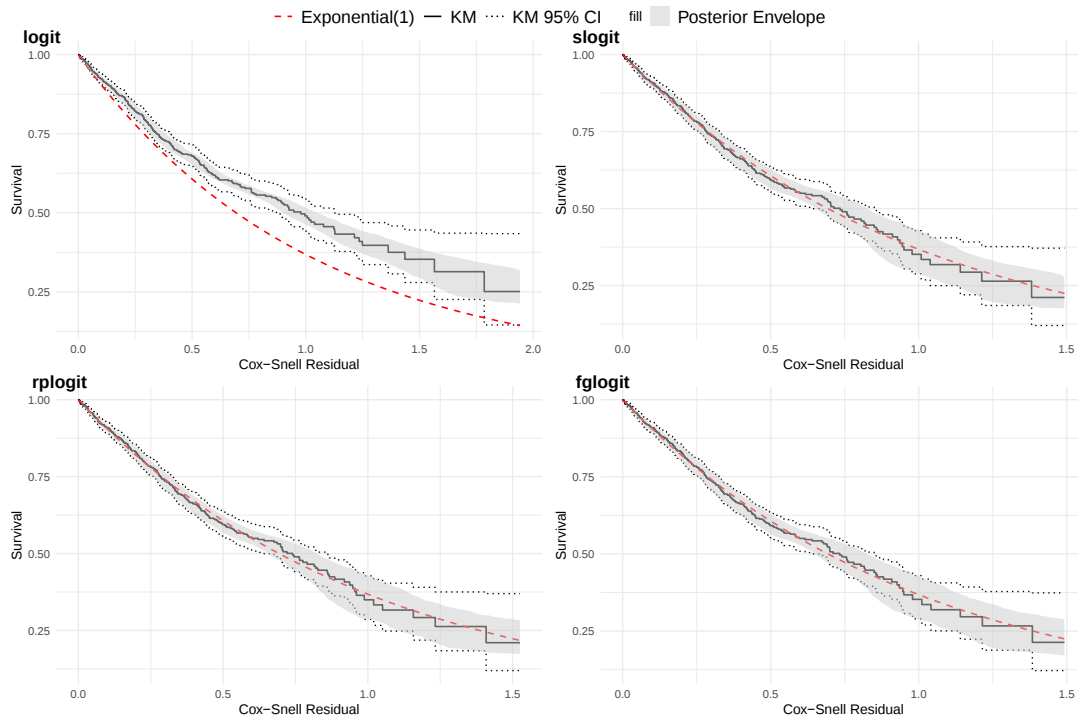


Figure 8: Cox-Snell residual plot with Lognormal rplogit link being true model

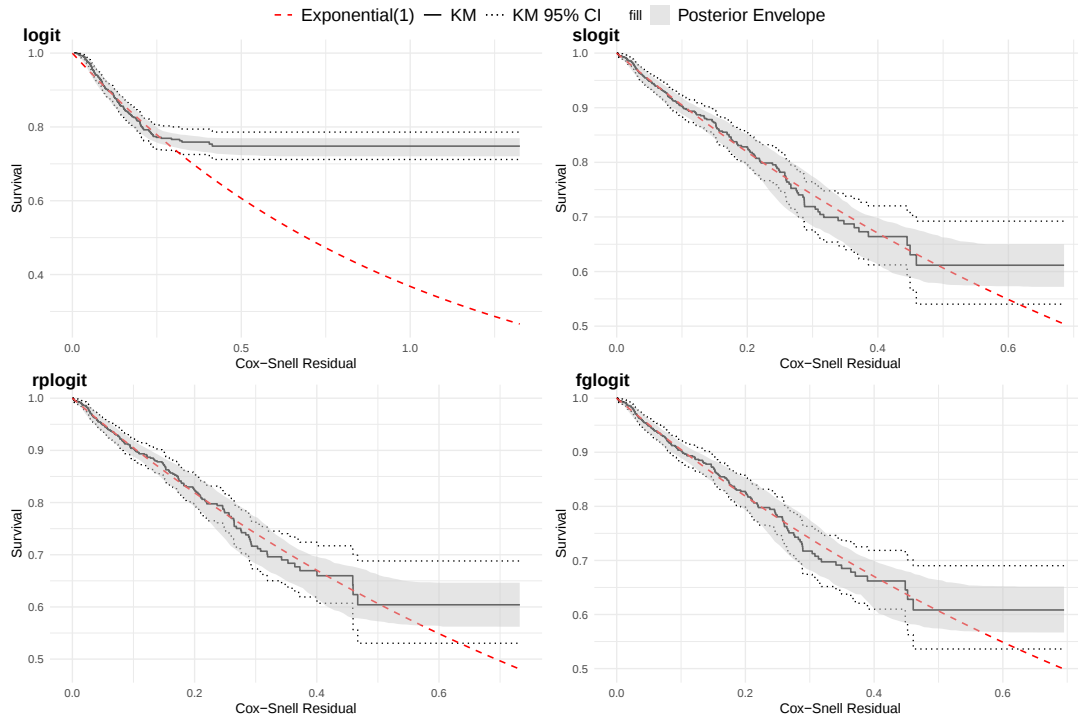


Figure 9: Cox-Snell residual plot with Lognormal fglogit link being true model

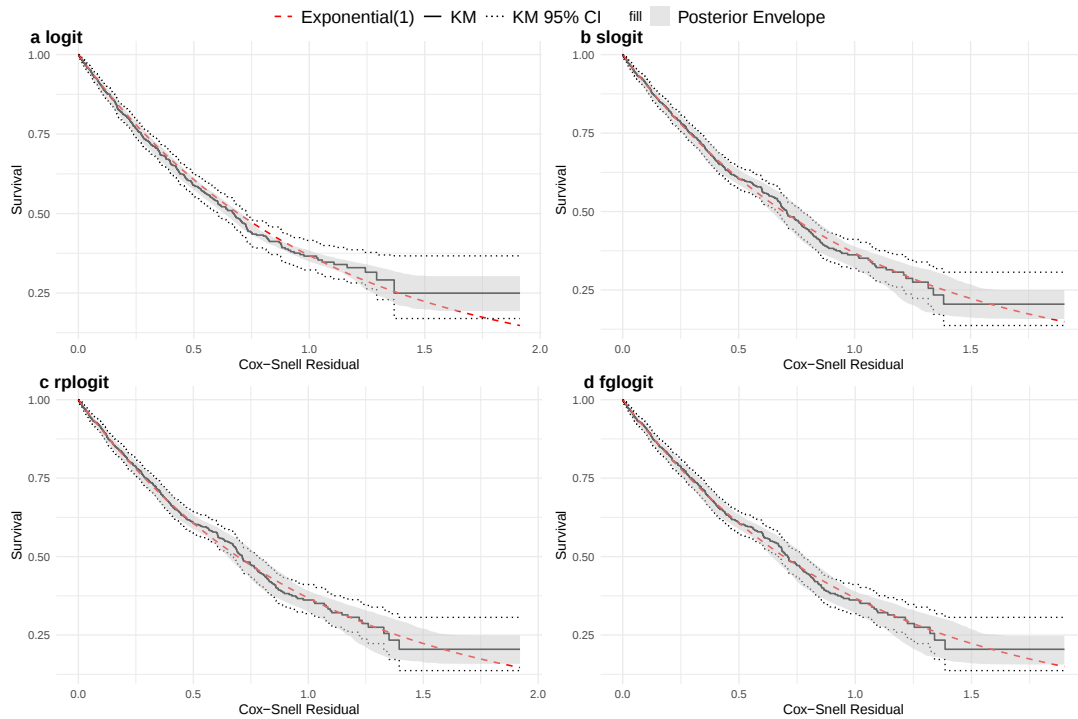


Figure 10: Cox-Snell residual plot with Loglogistic logit link being true model

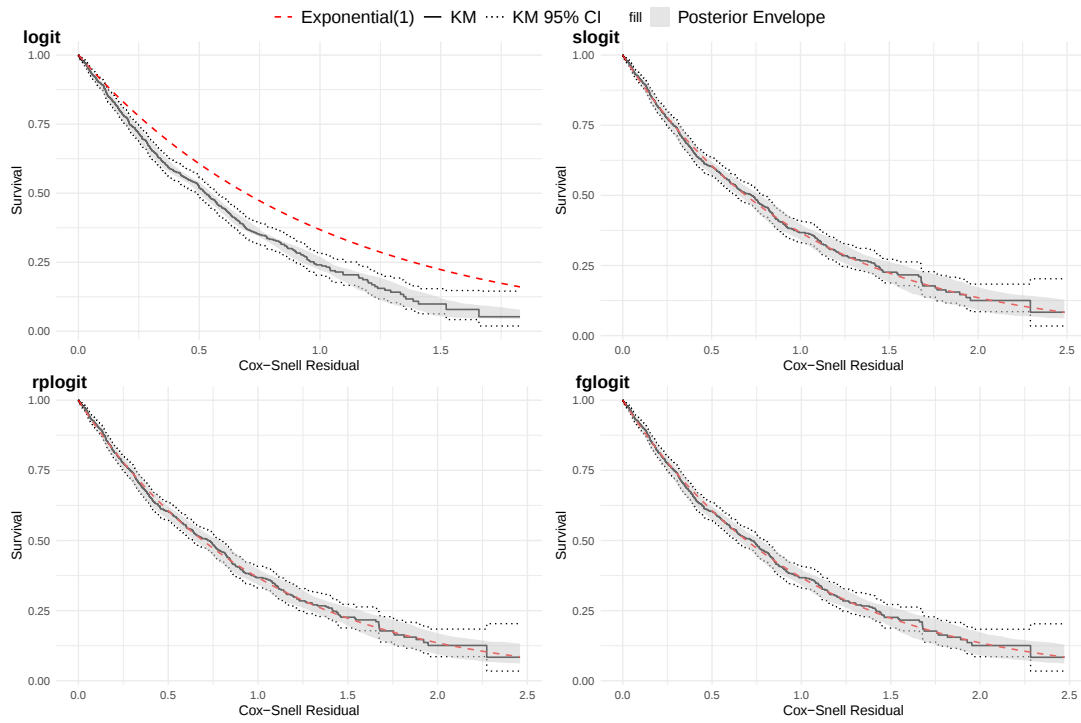


Figure 11: Cox-Snell residual plot with Loglogistic slogit link being true model

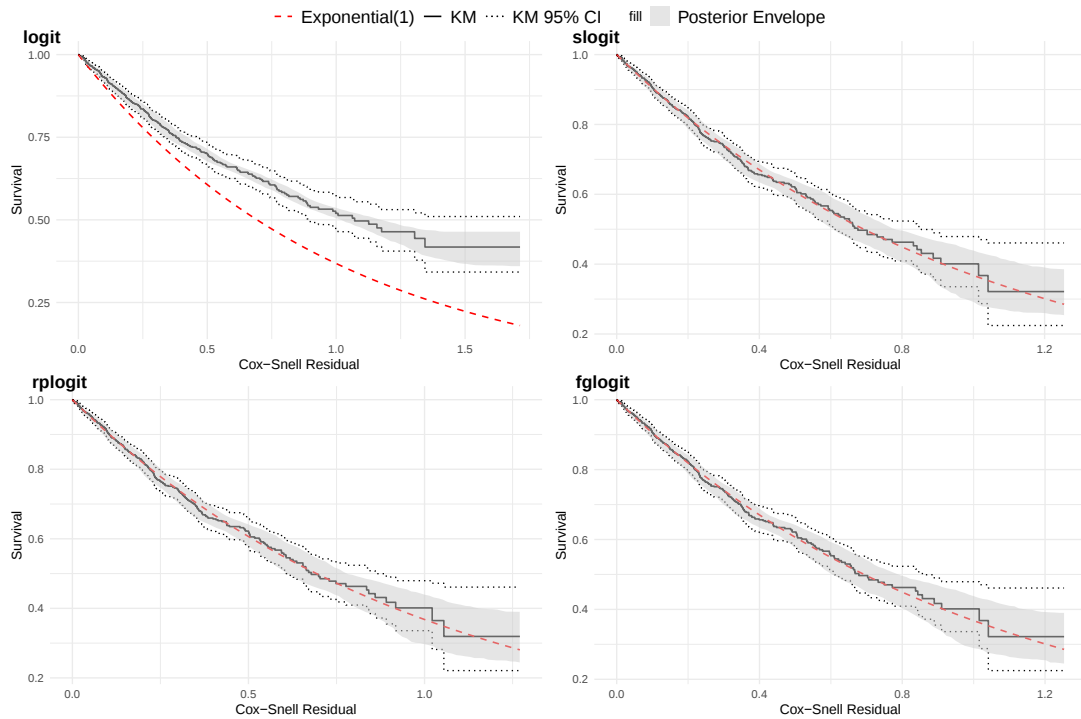


Figure 12: Cox-Snell residual plot with Loglogistic rplogit link being true model

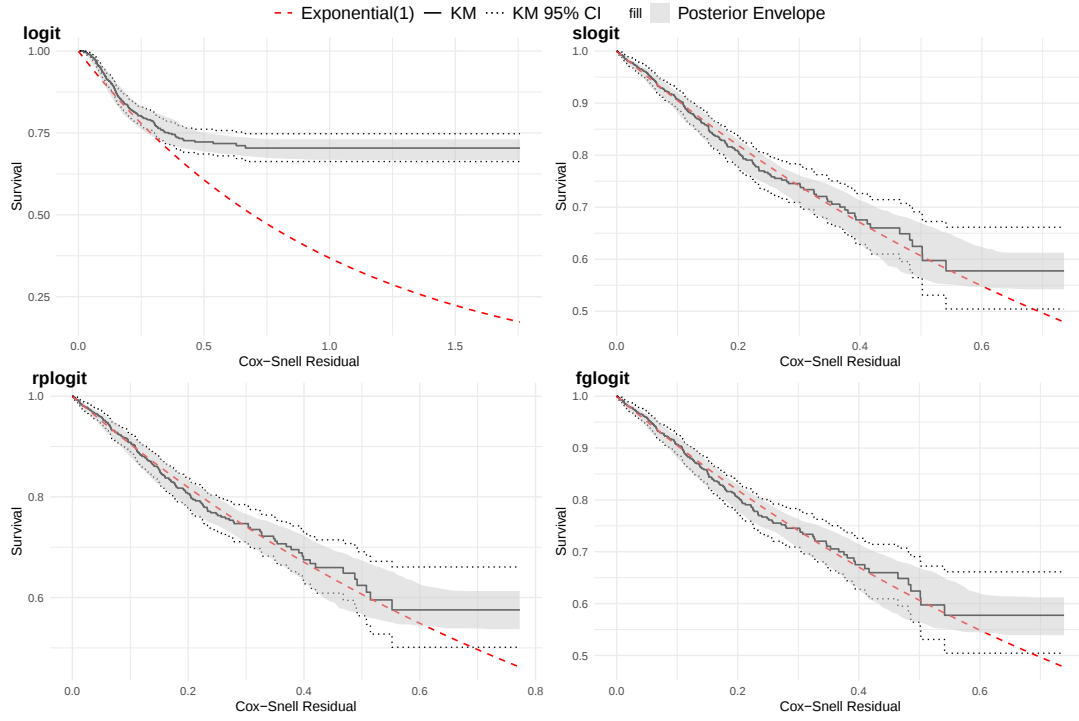


Figure 13: Cox-Snell residual plot with Loglogistic fglogit link being true model

2 Additional Real Data Analysis results

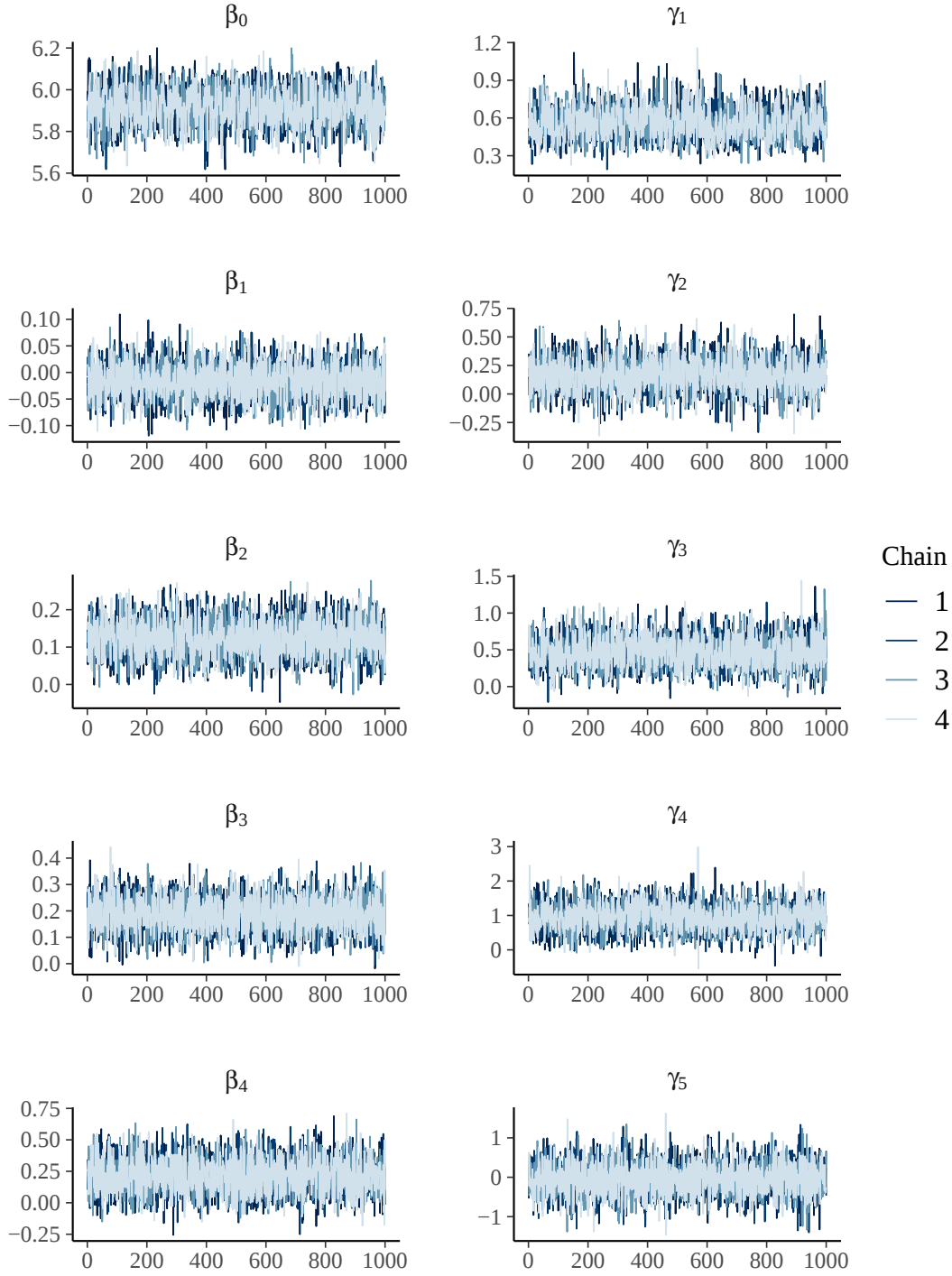


Figure 14: Trace plots of posterior samples for the latency and incidence parameters γ_j and β_j for real data analysis.

Table 3: LLGFG model: Full posterior means and 95% HPD intervals for latency variables.

Parameter	Mean	HPD 95% Interval
Latency Variables		
Age	-0.016	[-0.077, 0.047]
Month Supervised	0.122	[0.032, 0.217]
Race (ref: White)		
Black	0.187	[0.063, 0.31]
Hispanic	0.214	[-0.043, 0.482]
Other	-0.337	[-0.631, -0.052]
Sex Female	-0.353	[-1.459, 0.819]
Supervising Unit (ref: Fort Dodge Correctional Facility)		
Clarinda.Correctional.Facility	-0.141	[-0.348, 0.072]
Iowa.Correctional.Institution.for.Women	0.368	[-0.799, 1.487]
Iowa.Medical...Classification.Center	0.049	[-0.12, 0.21]
Mount.Pleasant.Correctional.Facility	-0.048	[-0.221, 0.132]
Newton.Correctional.Facility	-0.160	[-0.362, 0.035]
North.Central.Correctional.Facility	-0.102	[-0.324, 0.113]
Other	-0.330	[-0.609, -0.037]
Supervision Type (ref: Work Prison)		
Work Release	0.079	[-0.041, 0.197]
Supervision End Reason (ref: Parole Granted)		
Paroled.w.Immediate.Discharge	0.916	[0.549, 1.304]
Released.to.Special.Sentence	-0.425	[-0.629, -0.213]
Discharged.Expiration.of.Sentence	0.713	[0.553, 0.884]
Paroled.to.Detainer	0.401	[-0.093, 0.971]
Offense Type (ref: Drug)		
Other	0.076	[-0.091, 0.255]
Property	-0.140	[-0.268, -0.01]
Public.Order	-0.062	[-0.247, 0.132]
Violent	-0.089	[-0.236, 0.079]
Risk Ranking (ref: High)		
Low	0.243	[0.122, 0.363]
Moderate	0.135	[0.001, 0.27]
σ_{llg}	1.845	[1.748, 1.946]

Table 4: LLGFG model: Full posterior means and 95% HPD intervals for incidence variables.

Parameter	Mean	HPD 95% Interval
Incidence Variables		
Age	0.553	[0.342, 0.82]
Month Supervised	0.166	[-0.101, 0.443]
Race (ref: White)		
Black	0.466	[0.094, 0.9]
Hispanic	0.953	[0.215, 1.731]
Other	-0.040	[-0.823, 0.738]
Sex Female	0.612	[-0.705, 1.967]
Supervising Unit (ref: Fort Dodge Correctional Facility)		
Clarinda.Correctional.Facility	0.569	[0.045, 1.163]
Iowa.Correctional.Institution.for.Women	0.505	[-0.841, 1.893]
Iowa.Medical...Classification.Center	-2.767	[-3.751, -1.9]
Mount.Pleasant.Correctional.Facility	0.001	[-0.456, 0.463]
Newton.Correctional.Facility	1.129	[0.597, 1.785]
North.Central.Correctional.Facility	0.728	[0.191, 1.339]
Other	1.212	[0.523, 2.045]
Supervision Type (ref: Work Prison)		
Work Release	-0.216	[-0.597, 0.124]
Supervision End Reason (ref: Parole Granted)		
Paroled.w.Immediate.Discharge	0.839	[-0.264, 1.96]
Released.to.Special.Sentence	-0.543	[-1.171, 0.044]
Discharged.Expiration.of.Sentence	0.939	[0.381, 1.552]
Paroled.to.Detainer	1.997	[0.56, 3.218]
Offense Type (ref: Drug)		
Other	-0.487	[-1.109, 0.043]
Property	0.703	[0.303, 1.172]
Public.Order	1.386	[0.785, 2.117]
Violent	1.223	[0.748, 1.791]
α_{fg}	0.634	[0.436, 0.910]
λ_{fg}	0.334	[0.182, 0.568]