Table1: Test accuracy (%)) on heterophilous graphs with flip noise and standard semi-supervised learning setting.

Datasets	p	GCN	Coteaching	RTGNN	GNN Cleaner	ERASE	GOAL-V1	GOAL-V2
Chameleon	0.2	47.14	47.63	46.59	47.81	45.37	**55.69**	54.32
	0.4	40.55	41.53	40.39	41.15	41.72	**54.50**	53.49
	0.6	36.94	39.43	37.61	38.44	35.92	50.83	**52.01**
Actor	0.2	27.05	27.10	26.83	25.47	25.96	**29.75**	28.04
	0.4	26.39	26.95	25.89	25.14	24.36	**29.19**	27.95
	0.6	25.99	25.03	24.27	23.60	24.28	27.74	**27.88**
Squirrel	0.2	28.14	29.60	27.49	28.71	26.35	**34.99**	34.78
	0.4	27.83	29.36	26.22	27.18	26.03	**34.79**	33.94
	0.6	26.54	28.89	25.96	27.09	24.98	33.09	**33.59**

Table2: Test accuracy (%) with uniform noise and standard semi-supervised learning setting.

Datasets	p	GCN	Coteaching	RTGNN	GNN Cleaner	ERASE	GOAL-V1	GOAL-V2
Cora	0.2	76.20	74.27	74.58	78.15	79.74	**80.10**	79.97
	0.4	71.51	72.59	73.17	74.25	77.06	**78.97**	77.62
	0.6	59.43	62.61	63.52	64.81	68.74	71.83	**73.71**
Citeseer	0.2	65.64	66.10	67.50	68.16	68.95	**70.35**	70.10
	0.4	60.85	63.35	63.72	65.43	66.19	**68.68**	68.02
	0.6	50.37	55.79	54.37	54.18	56.25	61.53	**64.78**
Pubmed	0.2	74.67	75.61	74.28	75.83	76.07	**78.43**	77.48
	0.4	70.32	71.84	70.25	71.64	72.31	**76.93**	75.97
	0.6	67.85	68.13	69.15	70.28	69.37	**76.71**	75.38
Chameleon	0.2	49.00	47.16	46.58	48.31	47.75	**53.21**	53.03
	0.4	45.87	45.96	42.19	44.62	44.37	**52.75**	51.83
	0.6	40.01	39.82	39.86	41.29	40.83	50.46	**51.47**
Actor	0.2	27.29	27.69	26.90	27.39	26.15	**29.83**	29.57
	0.4	26.93	27.43	26.73	26.81	25.49	**29.45**	29.38
	0.6	26.60	27.40	25.18	26.04	24.33	29.12	**29.14**
Squirrel	0.2	31.02	30.15	30.56	29.28	28.37	**35.28**	34.93
	0.4	30.58	30.22	29.75	28.03	28.29	**34.30**	33.95
	0.6	27.75	27.56	26.80	27.49	25.14	32.61	**33.28**

From the above tables, we can see that GOAL achieves the best performance in all cases. This further shows that it can consistently provide superior results under the semi-supervised setting.