## **Confidence Scores Evaluation for NL3**

	Freq				Beta			
ECE Score	kr-vs-kp	2.73	3.08	2.40	2.29	2.10	3.27	- 3.2
	mushroom	2.19	3.09	2.39	2.25	1.96	3.27	- 3.0
	iris	3.26	3.34	2.48	3.23	2.15	2.40	- 2.8
	spambase	2.34	3.12	2.43	2.08	2.14	3.31	
	tic-tac-toe	2.50	3.10	2.38	2.41	2.06	3.25	- 2.6
	sick	1.35	3.10	2.38	1.17	2.30	3.16	- 2.4
	waveform	2.37	2.98	2.29	2.16	1.98	3.17	- 2.2
	car	2.51	3.05	2.39	2.36	2.73	3.27	- 2.0
	vote	2.64	3.37	2.59	2.80	2.63	2.53	- 1.8
	ionosphere	3.08	2.96	2.12	3.30	1.84	2.05	- 1.6
Brier Score Loss		0.70	0.45	0.40	0.00	0.45	0.44	0.75
	kr-vs-kp	0.78	0.45	0.43	0.68	0.45	0.41	
	mushroom	0.95	0.45	0.43	0.72	0.48	0.42	- 0.70
	iris	0.62	0.47	0.46	0.60	0.47	0.43	- 0.65
	spambase	0.83	0.45	0.44	0.70	0.50	0.42	
	tic-tac-toe	0.58	0.46	0.44	0.61	0.51	0.42	- 0.60
	sick	0.94	0.45	0.43	0.79	0.56	0.42	- 0.55
	waveform	0.84	0.45	0.43	0.65	0.47	0.42	
	car	0.75	0.46	0.44	0.70	0.52	0.42	- 0.50
	vote	0.87	0.45	0.43	0.73	0.51	0.42	- 0.45
	ionosphere	0.73	0.43	0.41	0.48	0.39	0.40	- 0.40
		Crowd-Certain	Tao	Sheng	Crowd-Certain	Tao	Sheng	0.70