1 Supplementary Materials

WEIGHTING	MODEL	α	β	γ	AUROC	ECE	Epsilon at 50% accuracy
Published	Cross entropy	1.0	-	-	0.562±0.05	0.094±0.01	0.274±0.06
Published	DROID	0.56	0.44	-	**0.730±0.01**	0.103±0.01	0.465±0.17
Published	NGEBM	0.5	0.5	-	0.689±0.02	**0.086±0.01**	3.400±1.23
Equal	(Naive) REBIS	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	0.720±0.02	0.088±0.01	1.528±0.46
Near equal	REBIS	0.3	0.3	0.4	0.731±0.02	0.089±0.01	1.805±0.52
Near equal	REBIS	0.3	0.4	0.3	0.731±0.01	0.090±0.01	1.691±0.73
Near equal	REBIS	0.4	0.3	0.3	0.726±0.01	0.096±0.00	1.662±0.32
One lower	REBIS	0.2	0.4	0.4	0.715±0.02	0.085±0.01	1.881±0.57
One lower	REBIS	0.4	0.2	0.4	0.717±0.02	0.092±0.01	1.766±0.52
One lower	REBIS	0.4	0.4	0.2	0.729±0.02	0.097±0.00	1.598±0.21
Progressive	REBIS	0.2	0.3	0.5	**0.738±0.02**	0.076±0.01	2.002±0.49
Progressive	REBIS	0.2	0.5	0.3	0.727±0.02	0.093±0.01	1.872±0.63
Progressive	REBIS	0.3	0.2	0.5	0.731±0.01	0.085±0.01	2.147±0.90
Progressive	REBIS	0.3	0.5	0.2	0.736±0.02	0.097±0.00	1.487±0.20
Progressive	REBIS	0.5	0.2	0.3	0.711±0.02	0.093±0.01	1.637±0.58
Progressive	REBIS	0.5	0.3	0.2	0.705±0.02	0.098±0.00	1.416±0.16
One higher	(Optimal) REBIS	0.2	0.2	0.6	**0.738±0.02**	**0.065±0.01**	**2.232±0.65**
One higher	REBIS	0.2	0.6	0.2	0.730±0.03	0.095±0.01	1.920±1.77
One higher	REBIS	0.6	0.2	0.2	0.698±0.03	0.092±0.01	1.521±0.11

Table 1: The full AUROC, ECE, and adversarial robustness scores for all 19 models examined in this paper. The models using published weights are separated from the REBIS variations. In both sections the highest performing value in each category is highlighted