

Hyperparameter Sweep Results		
Hyperparameter	Sweep Value	Test Cost
Policy Learning Rate	0.0001	17.45 ± 0.22
	0.0003	16.73 ± 0.22
	0.001	16.28 ± 0.23
Value Learning Rate	0.0001	15.99 ± 0.20
	0.0003	16.28 ± 0.23
	0.001	15.70 ± 0.20
Model Scale	10	16.96 ± 0.26
	20	16.28 ± 0.23
	100	15.71 ± 0.20
Gamma	0.95	18.74 ± 0.20
	0.99	16.53 ± 0.20
	0.998	16.28 ± 0.23
GAE Lambda	0.9	15.76 ± 0.21
	0.95	15.21 ± 0.20
	0.99	16.28 ± 0.23
Episode Steps	10000	17.06 ± 0.24
	30000	15.89 ± 0.21
	50000	16.28 ± 0.23

