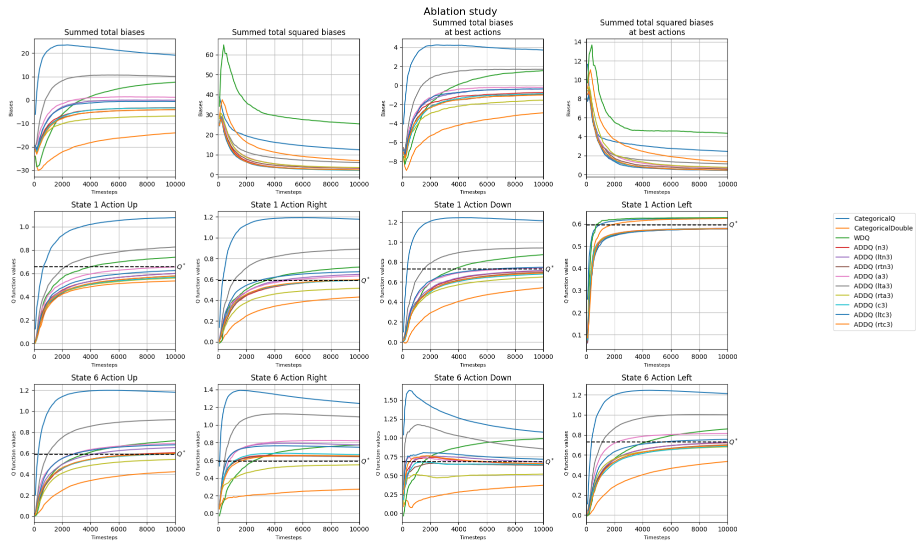
1 Ablation Study



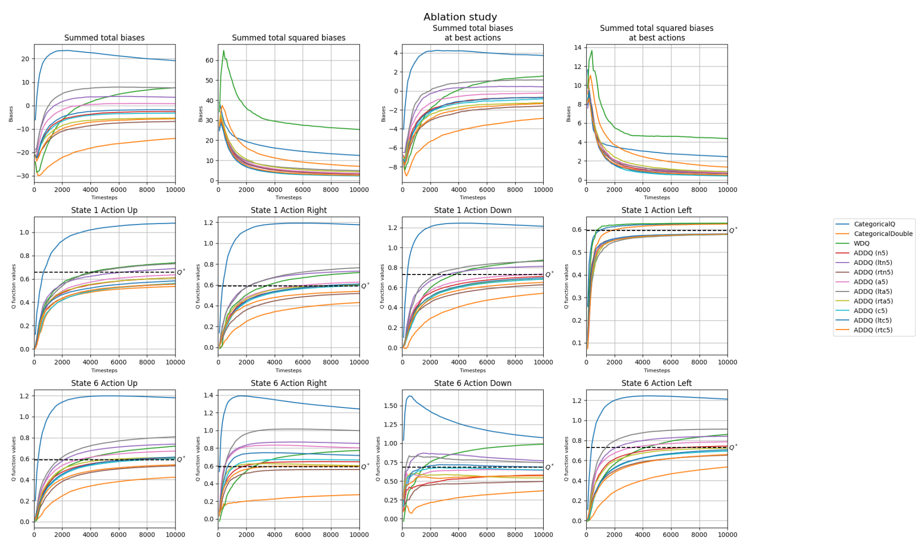


Figure 1: The effect of hyperparameter choice of β is small with respect to the Bias improvement. Compared to Q, DQ, and WDQ the Bias is much lower. Conservative choices seem to work especially well

The choices of beta are named in the following way:

1. (Optional) First two letters: Left-tilted (lt), Right-tilted (rt)
2. First/Third letter: Neutral (n), Aggressive (a), Conservative (c)
3. Final digit: Refers to the number of intervals in the definition of Beta (3 or 5)

As in the paper, the intuition for aggressive, conservative, and neutral remains the same (no interpolation, just choosing which Algorithm's update to take vs. more interpolation, with neutral being in between the two choices.
Left- and Right-tilted refers to the shifted intervals for the relative Variance to fall into while choosing the interpolation coefficient. In the paper, only interval choices centered around 1 were considered, Left-tilted favors the Q update, Right-tilted the DQ update.
The choices are:

$$\text{n3:} \qquad \beta := \begin{cases} 0.75 & : S^2_{rel}(s,a) < 0.75 \\ 0.5 & : S^2_{rel}(s,a) \in [0.75, 1.25] \\ 0.25 & : S^2_{rel}(s,a) > 1.25 \end{cases}$$

$$\text{ltn3:} \qquad \beta := \begin{cases} 0.75 & : S^2_{rel}(s,a) < 1.25 \\ 0.5 & : S^2_{rel}(s,a) \in [1.25, 1.75] \\ 0.25 & : S^2_{rel}(s,a) > 1.75 \end{cases}$$

$$\text{rtn3:} \qquad \beta := \begin{cases} 0.75 & : S^2_{rel}(s,a) < 0.25 \\ 0.5 & : S^2_{rel}(s,a) \in [0.25, 0.75] \\ 0.25 & : S^2_{rel}(s,a) > 0.75 \end{cases}$$

$$\text{a3:} \qquad \beta := \begin{cases} 1 & : S^2_{rel}(s,a) < 0.99 \\ 0.5 & : S^2_{rel}(s,a) \in [0.99, 1.01] \\ 0 & : S^2_{rel}(s,a) > 1.01 \end{cases}$$

$$\text{lta3:} \qquad \beta := \begin{cases} 1 & : S^2_{rel}(s,a) < 1.49 \\ 0.5 & : S^2_{rel}(s,a) \in [1.49, 1.51] \\ 0 & : S^2_{rel}(s,a) > 1.51 \end{cases}$$

$$\text{rta3:} \qquad \beta := \begin{cases} 1 & : S^2_{rel}(s,a) < 0.49 \\ 0.5 & : S^2_{rel}(s,a) \in [0.49, 0.51] \\ 0 & : S^2_{rel}(s,a) > 0.51 \end{cases}$$

$$\text{c3:} \qquad \beta := \begin{cases} 0.6 & : S^2_{rel}(s,a) < 0.6 \\ 0.5 & : S^2_{rel}(s,a) \in [0.6, 1.4] \\ 0.4 & : S^2_{rel}(s,a) > 1.4 \end{cases}$$

$$\text{ltc3:} \qquad \beta := \begin{cases} 0.6 & : S^2_{rel}(s,a) < 1.1 \\ 0.5 & : S^2_{rel}(s,a) \in [1.1, 1.9] \\ 0.4 & : S^2_{rel}(s,a) > 1.9 \end{cases}$$

$$\text{rtc3:} \qquad \beta := \begin{cases} 0.6 & : S^2_{rel}(s,a) < 0.1 \\ 0.5 & : S^2_{rel}(s,a) \in [0.1, 0.9] \\ 0.4 & : S^2_{rel}(s,a) > 0.9 \end{cases}$$

$$\text{n5:} \qquad \beta := \begin{cases} 1 & : S^2_{rel}(s,a) \leq 0.25 \\ 0.75 & : S^2_{rel}(s,a) \in (0.25, 0.75) \\ 0.5 & : S^2_{rel}(s,a) \in [0.75, 1.25] \\ 0.25 & : S^2_{rel}(s,a) \in (1.25, 1.75) \\ 0 & : S^2_{rel}(s,a) \geq 1.75 \end{cases}$$

$$\text{ltn5:} \qquad \beta := \begin{cases} 1 & : S^2_{rel}(s,a) \leq 0.75 \\ 0.75 & : S^2_{rel}(s,a) \in (0.75, 1.25) \\ 0.5 & : S^2_{rel}(s,a) \in [1.25, 1.75] \\ 0.25 & : S^2_{rel}(s,a) \in (1.75, 2.25) \\ 0 & : S^2_{rel}(s,a) \geq 2.25 \end{cases}$$

$$\text{rtn5:} \qquad \beta := \begin{cases} 1 & : S^2_{rel}(s,a) \leq -0.25 \\ 0.75 & : S^2_{rel}(s,a) \in (-0.25, 0.25) \\ 0.5 & : S^2_{rel}(s,a) \in [0.25, 0.75] \\ 0.25 & : S^2_{rel}(s,a) \in (0.75, 1.25) \\ 0 & : S^2_{rel}(s,a) \geq 1.25 \end{cases}$$

$$\text{a5:} \qquad \beta := \begin{cases} 1 & : S^2_{rel}(s,a) \leq 0.99 \\ 0.75 & : S^2_{rel}(s,a) \in (0.99, 0.995) \\ 0.5 & : S^2_{rel}(s,a) \in [0.995, 1.005] \\ 0.25 & : S^2_{rel}(s,a) \in (1.005, 1.01) \\ 0 & : S^2_{rel}(s,a) \geq 1.01 \end{cases}$$

$$\text{lta5:} \qquad \beta := \begin{cases} 1 & : S^2_{rel}(s,a) \leq 1.49 \\ 0.75 & : S^2_{rel}(s,a) \in (1.49, 1.495) \\ 0.5 & : S^2_{rel}(s,a) \in [1.495, 1.505] \\ 0.25 & : S^2_{rel}(s,a) \in (1.505, 1.51) \\ 0 & : S^2_{rel}(s,a) \geq 1.51 \end{cases}$$

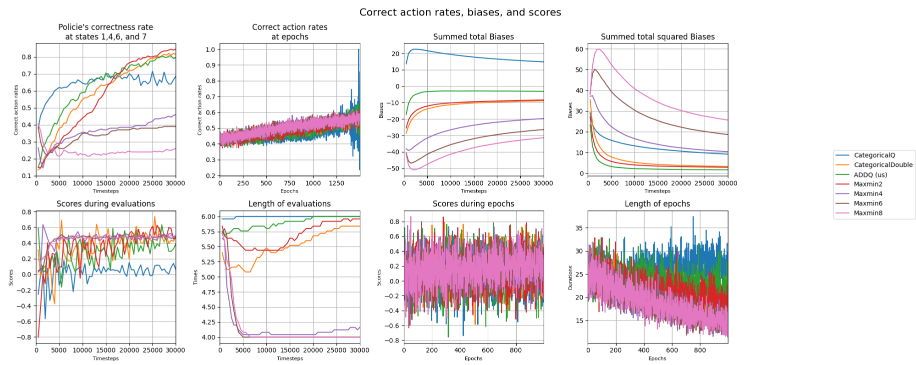
$$\text{rta5:} \qquad \beta := \begin{cases} 1 & : S^2_{rel}(s,a) \leq 0.49 \\ 0.75 & : S^2_{rel}(s,a) \in (0.49, 0.495) \\ 0.5 & : S^2_{rel}(s,a) \in [0.495, 0.505] \\ 0.25 & : S^2_{rel}(s,a) \in (0.505, 0.51) \\ 0 & : S^2_{rel}(s,a) \geq 0.51 \end{cases}$$

$$\text{c5:} \qquad \beta := \begin{cases} 0.7 & : S^2_{rel}(s,a) \leq 0.1 \\ 0.6 & : S^2_{rel}(s,a) \in (0.1, 0.7) \\ 0.5 & : S^2_{rel}(s,a) \in [0.7, 1.3] \\ 0.4 & : S^2_{rel}(s,a) \in (1.3, 1.9) \\ 0.3 & : S^2_{rel}(s,a) \geq 1.9 \end{cases}$$

$$\text{ltc5:} \qquad \beta := \begin{cases} 0.7 & : S^2_{rel}(s,a) \leq 0.6 \\ 0.6 & : S^2_{rel}(s,a) \in (0.6, 1.2) \\ 0.5 & : S^2_{rel}(s,a) \in [1.2, 1.8] \\ 0.4 & : S^2_{rel}(s,a) \in (1.8, 2.4) \\ 0.3 & : S^2_{rel}(s,a) \geq 2.4 \end{cases}$$

$$\text{rtc5:} \qquad \beta := \begin{cases} 0.7 & : S^2_{rel}(s,a) \leq -0.4 \\ 0.6 & : S^2_{rel}(s,a) \in (-0.4, 0.2) \\ 0.5 & : S^2_{rel}(s,a) \in [0.2, 0.8] \\ 0.4 & : S^2_{rel}(s,a) \in (0.8, 1.4) \\ 0.3 & : S^2_{rel}(s,a) \geq 1.4 \end{cases}$$

2 Comparison to more Algorithms



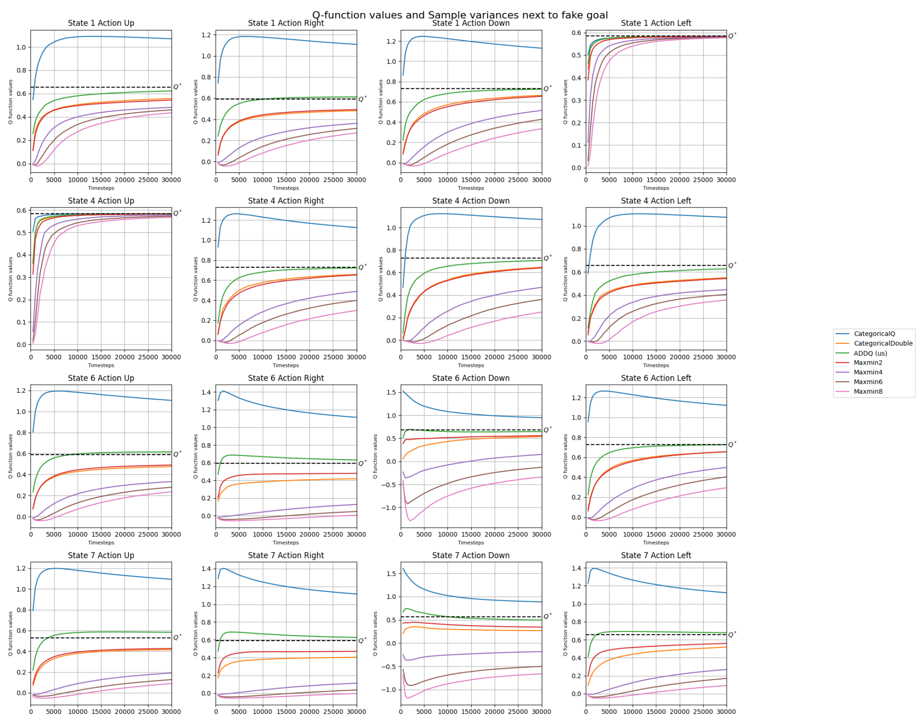


Figure 2: ADDQ compares well to MaxMin Algorithm across different choices of Ensemble sizes. The Bias is significantly lower.

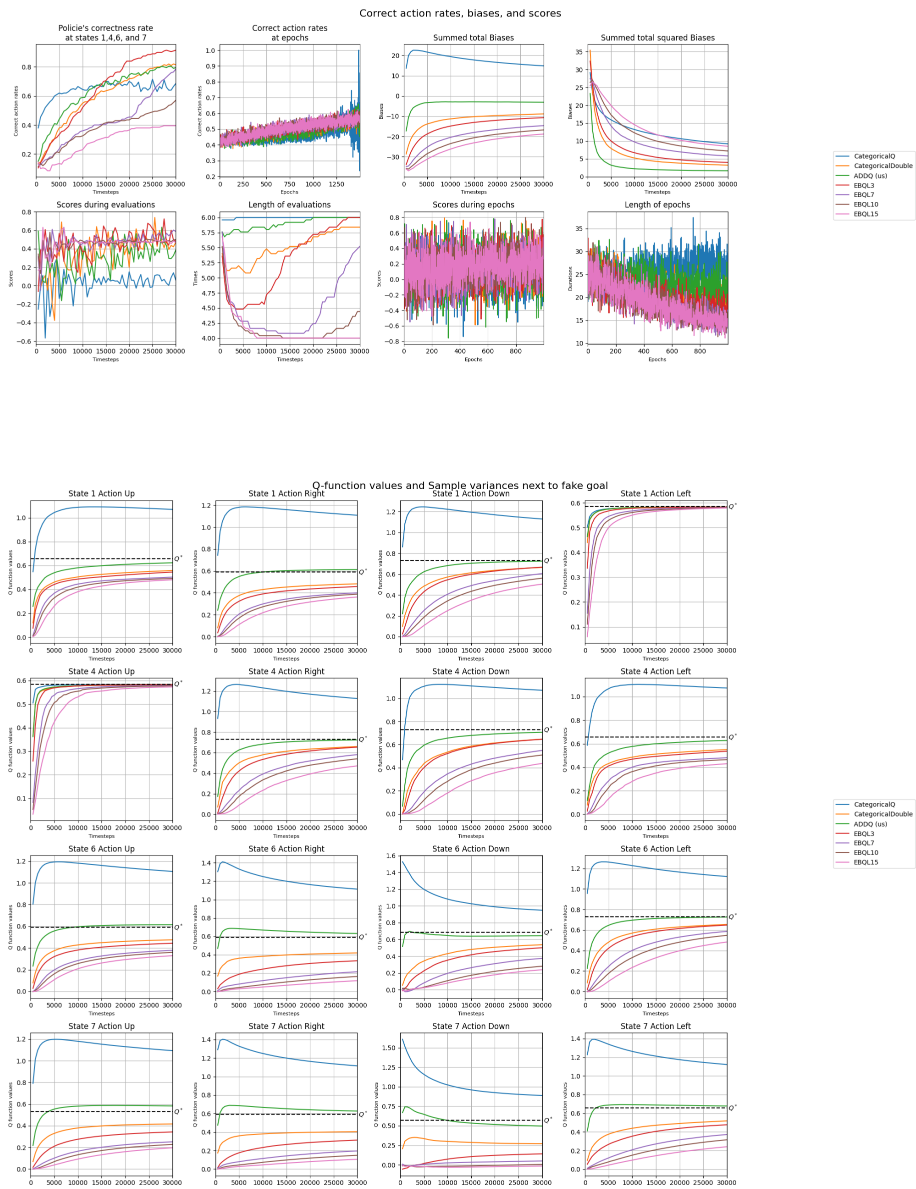
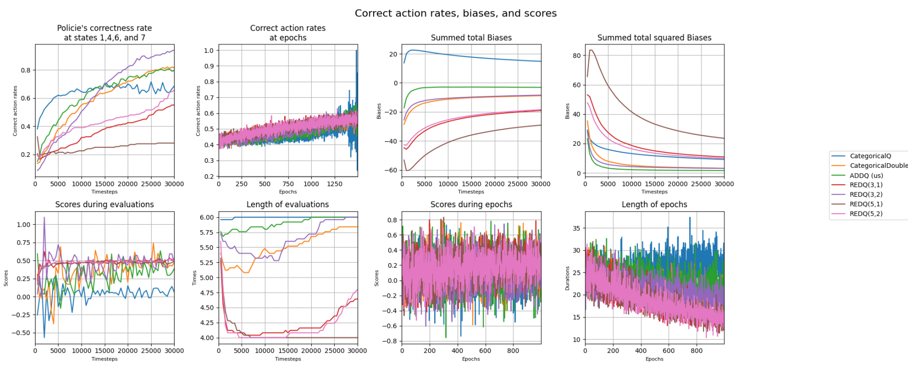


Figure 3: ADDQ compares well to Ensemble Bootstrapped QL Algorithm across different choices of Ensemble sizes. The Bias is significantly lower.



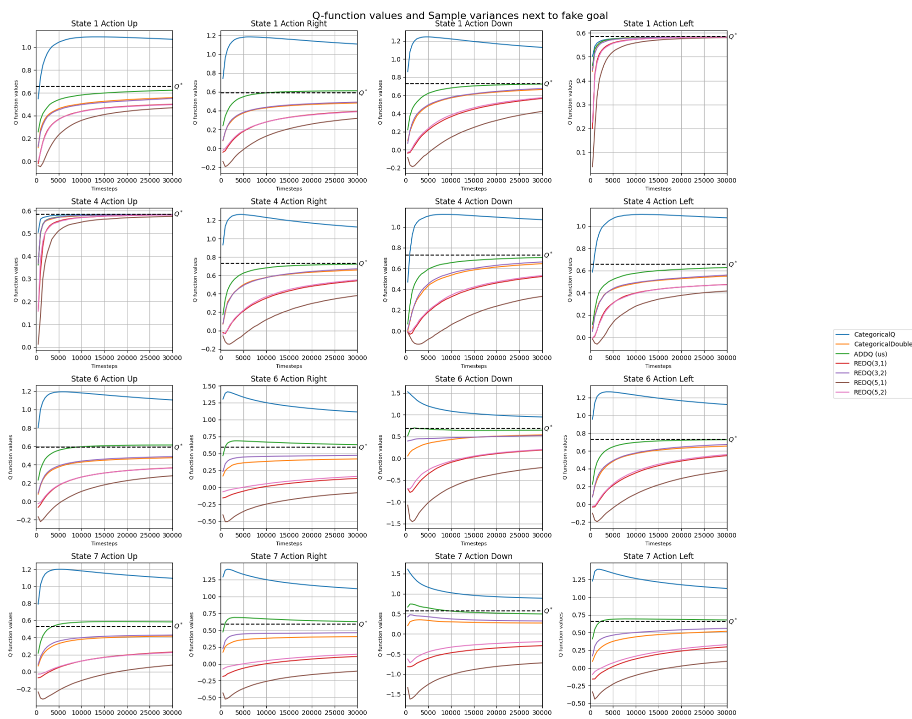


Figure 4: ADDQ compares well to Randomized Ensemble DQL Algorithm across different choices of Ensemble sizes and sizes of the subset to be updated. The Bias is significantly lower.