# E APPENDIX: ADDITIONAL RESULTS

**Varying Number of Actions ($|\mathcal{A}|$)**

| | | | |
|---|---|---|---|
| 1. | Synthetic Dataset | Estimating $V(\pi_{\text{bad}})$ using $\mu_{\text{uniform}}$ | Figure 7 |
| 2. | Synthetic Dataset | Estimating $V(\pi_{\text{bad}})$ using $\mu_{\text{good}}$ | Figure 8 |

**Varying Policy-mismatch through $\mu(\beta)$**

| | | | |
|---|---|---|---|
| 1. | Synthetic Dataset | Estimating $V(\pi_{\text{good}})$ | Figure 9 |
| 2. | Synthetic Dataset | Estimating $V(\pi_{\text{bad}})$ | Figure 10 |
| 3. | Movielens Dataset | Estimating $V(\pi_{\text{good}})$ | Figure 11 |
| 4. | Movielens Dataset | Estimating $V(\pi_{\text{bad}})$ | Figure 12 |

**Varying Policy-mismatch through $\pi(\epsilon)$**

| | | | |
|---|---|---|---|
| 1. | Synthetic Dataset | Using $\mu_{\text{uniform}}$ | Figure 13 |
| 2. | Synthetic Dataset | Using $\mu_{\text{good}}$ | Figure 14 |
| 3. | Movielens Dataset | Using $\mu_{\text{uniform}}$ | Figure 15 |
| 4. | Movielens Dataset | Using $\mu_{\text{good}}$ | Figure 16 |

**Varying Bandit Feedback ($|\mathcal{D}|$)**

| | | | |
|---|---|---|---|
| 1. | Synthetic Dataset | Estimating $V(\pi_{\text{good}})$ | Figure 17 |
| 2. | Synthetic Dataset | Estimating $V(\pi_{\text{bad}})$ | Figure 18 |
| 3. | Movielens Dataset | Estimating $V(\pi_{\text{good}})$ | Figure 19 |
| 4. | Movielens Dataset | Estimating $V(\pi_{\text{bad}})$ | Figure 20 |

**Varying Deficient Support**

| | | | |
|---|---|---|---|
| 1. | Synthetic Dataset | Estimating $V(\pi_{\text{good}})$ using $\mu_{\text{uniform}}$ | Figure 21 |
| 2. | Synthetic Dataset | Estimating $V(\pi_{\text{bad}})$ using $\mu_{\text{uniform}}$ | Figure 22 |

**Varying Action Embedding Size**

| | | | |
|---|---|---|---|
| 1. | Synthetic Dataset | Estimating $V(\pi_{\text{good}})$ using $\mu_{\text{uniform}}$ | Figure 23 |
| 2. | Movielens Dataset | Estimating $V(\pi_{\text{good}})$ using $\mu_{\text{uniform}}$ | Figure 24 |

**Data-Driven *vs.* Oracle Action Embeddings**

| | | | |
|---|---|---|---|
| 1. | Synthetic Dataset | Estimating $V(\pi_{\text{good}})$ using $\mu_{\text{uniform}}$ | Figure 25 |
| 2. | Synthetic Dataset | Estimating $V(\pi_{\text{bad}})$ using $\mu_{\text{uniform}}$ | Figure 26 |

**Varying Amount of Pooling**

| | | | |
|---|---|---|---|
| 1. | Synthetic Dataset | Estimating $V(\pi_{\text{good}})$ using $\mu_{\text{good}}$ | Figure 27 |
| 2. | Movielens Dataset | Estimating $V(\pi_{\text{good}})$ using $\mu_{\text{good}}$ | Figure 28 |

**Optimal Amount of Pooling**

| | | | |
|---|---|---|---|
| 1. | Synthetic Dataset | Estimating $V(\pi_{\text{good}})$ using $\mu_{\text{bad}}$ | Figure 29 |
| 2. | Synthetic Dataset | Estimating $V(\pi_{\text{good}})$ using $\mu_{\text{uniform}}$ | Figure 30 |
| 3. | Synthetic Dataset | Estimating $V(\pi_{\text{good}})$ using $\mu_{\text{good}}$ | Figure 31 |

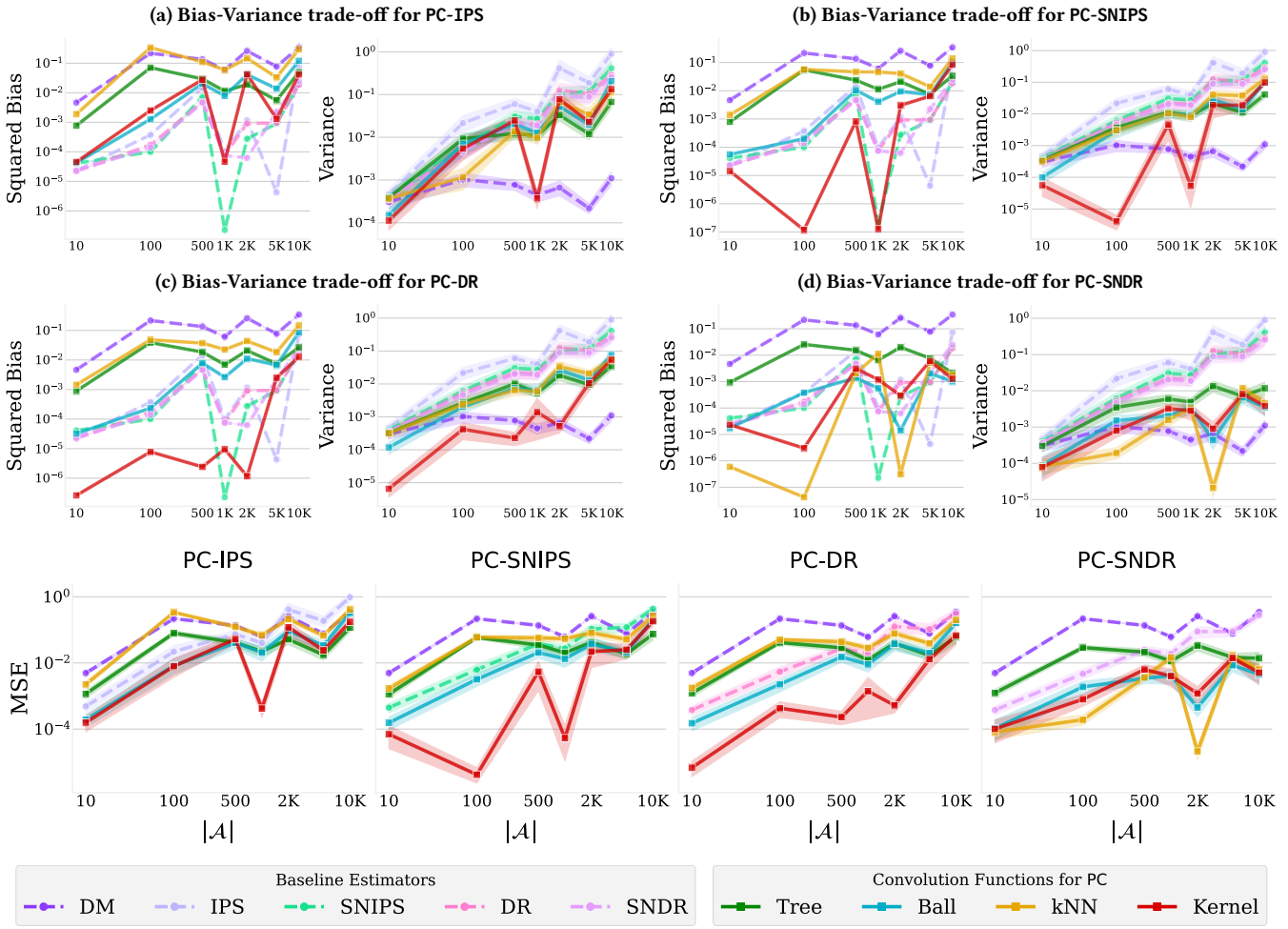


**Figure 7: Change in MSE, Squared Bias, and Variance while estimating $V(\pi_{\text{bad}})$ with varying number of actions (log-log scale) for the synthetic dataset, using data logged by $\mu_{\text{uniform}}$.**

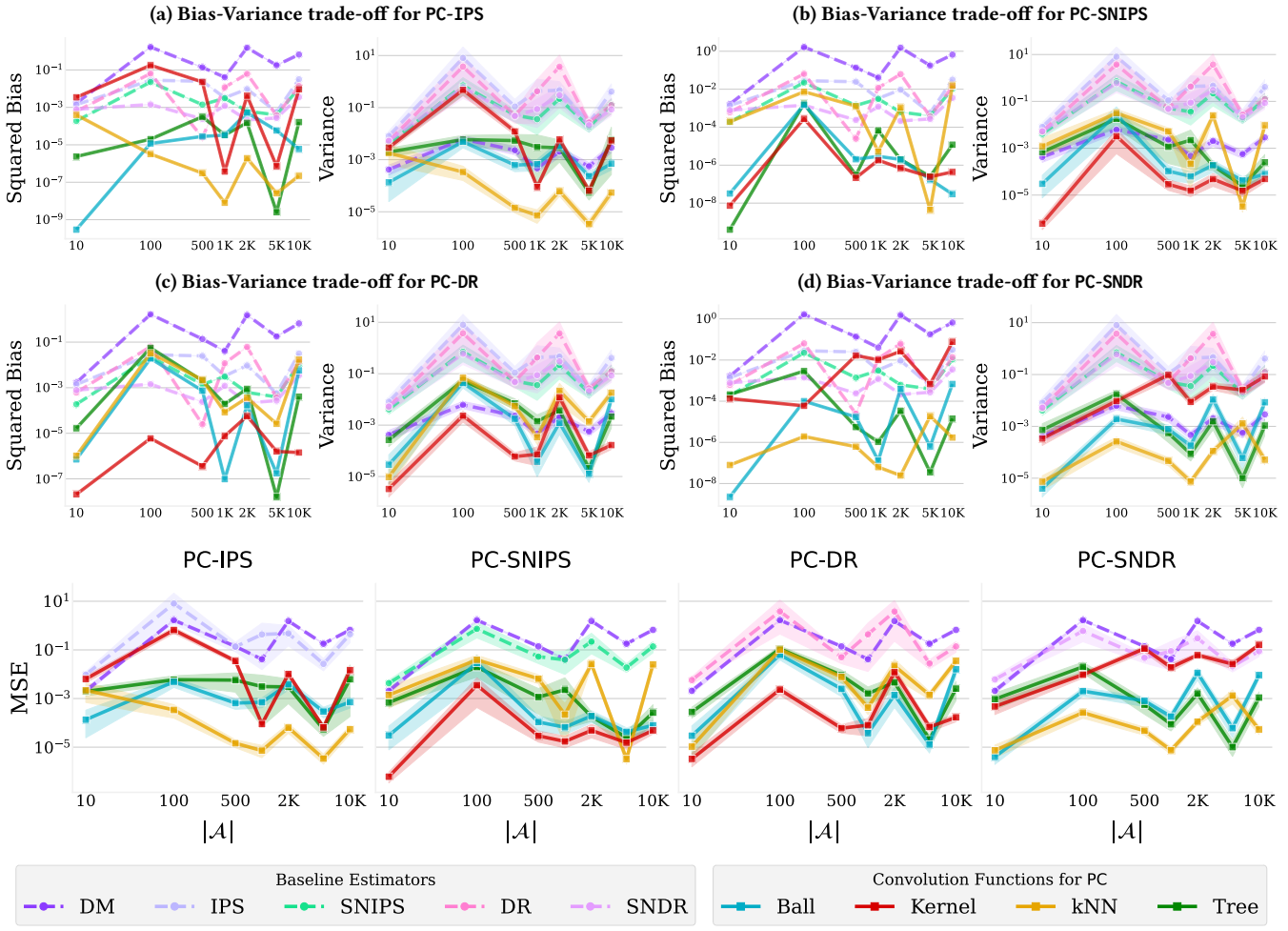


**Figure 8: Change in MSE, Squared Bias, and Variance while estimating $V(\pi_{\text{bad}})$ with varying number of actions (log-log scale) for the synthetic dataset, using data logged by $\mu_{\text{good}}$.**

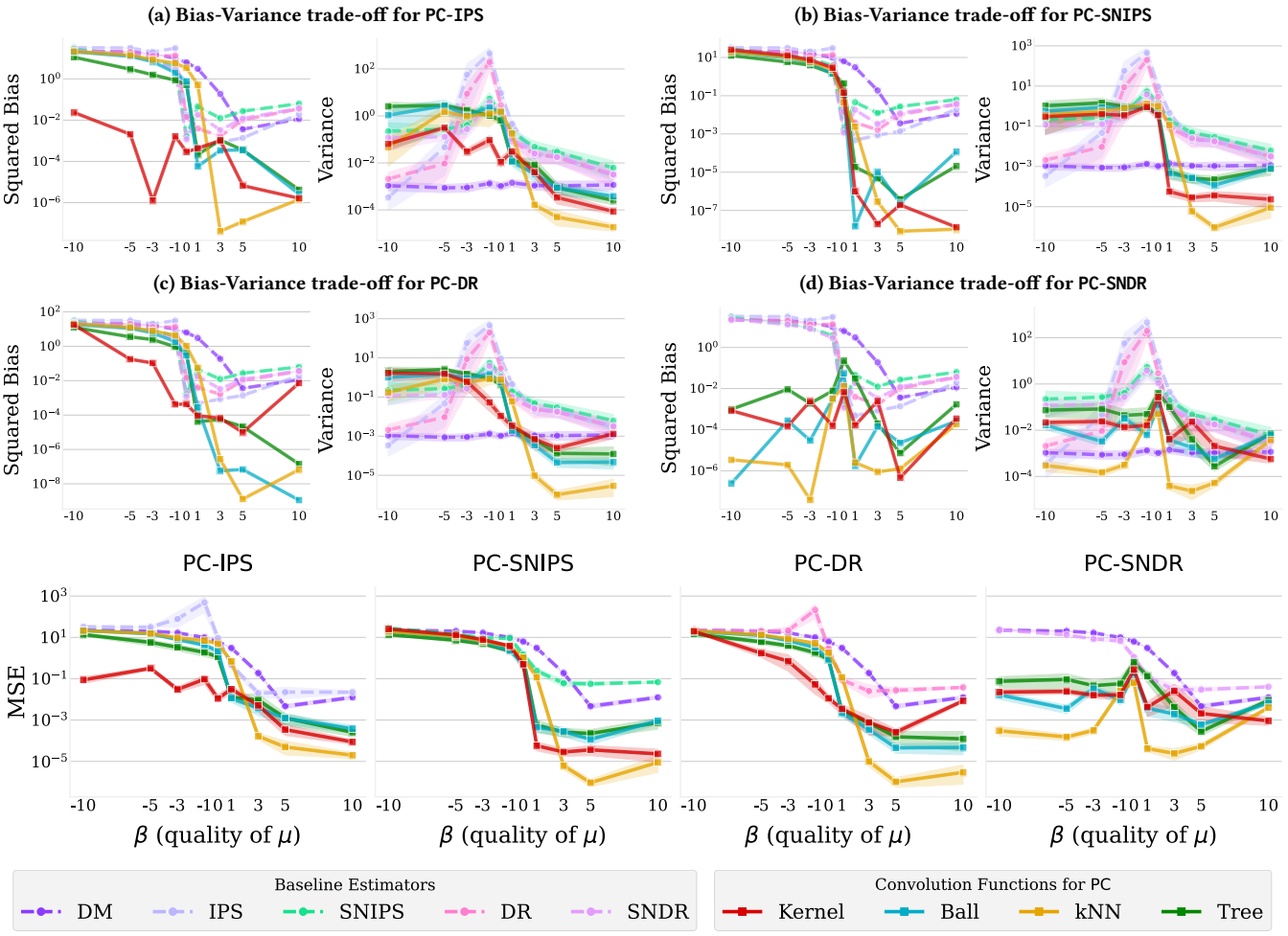


**Figure 9: Change in MSE, Squared Bias, and Variance while estimating $V(\pi_{\text{good}})$ with varying policy-mismatch (log scale) for the synthetic dataset.**

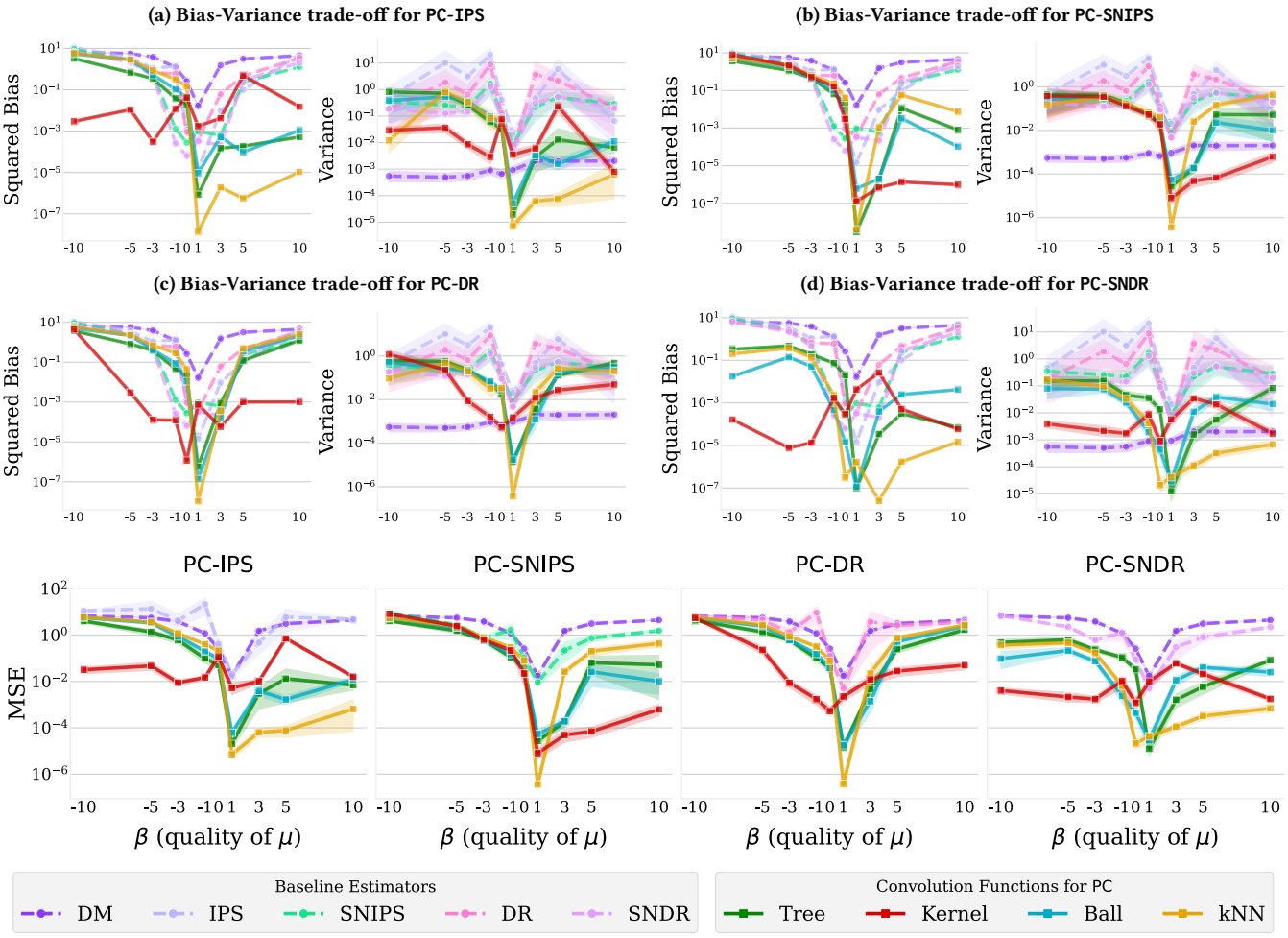


**Figure 10: Change in MSE, Squared Bias, and Variance while estimating $V(\pi_{\text{bad}})$ with varying policy-mismatch (log scale) for the synthetic dataset.**

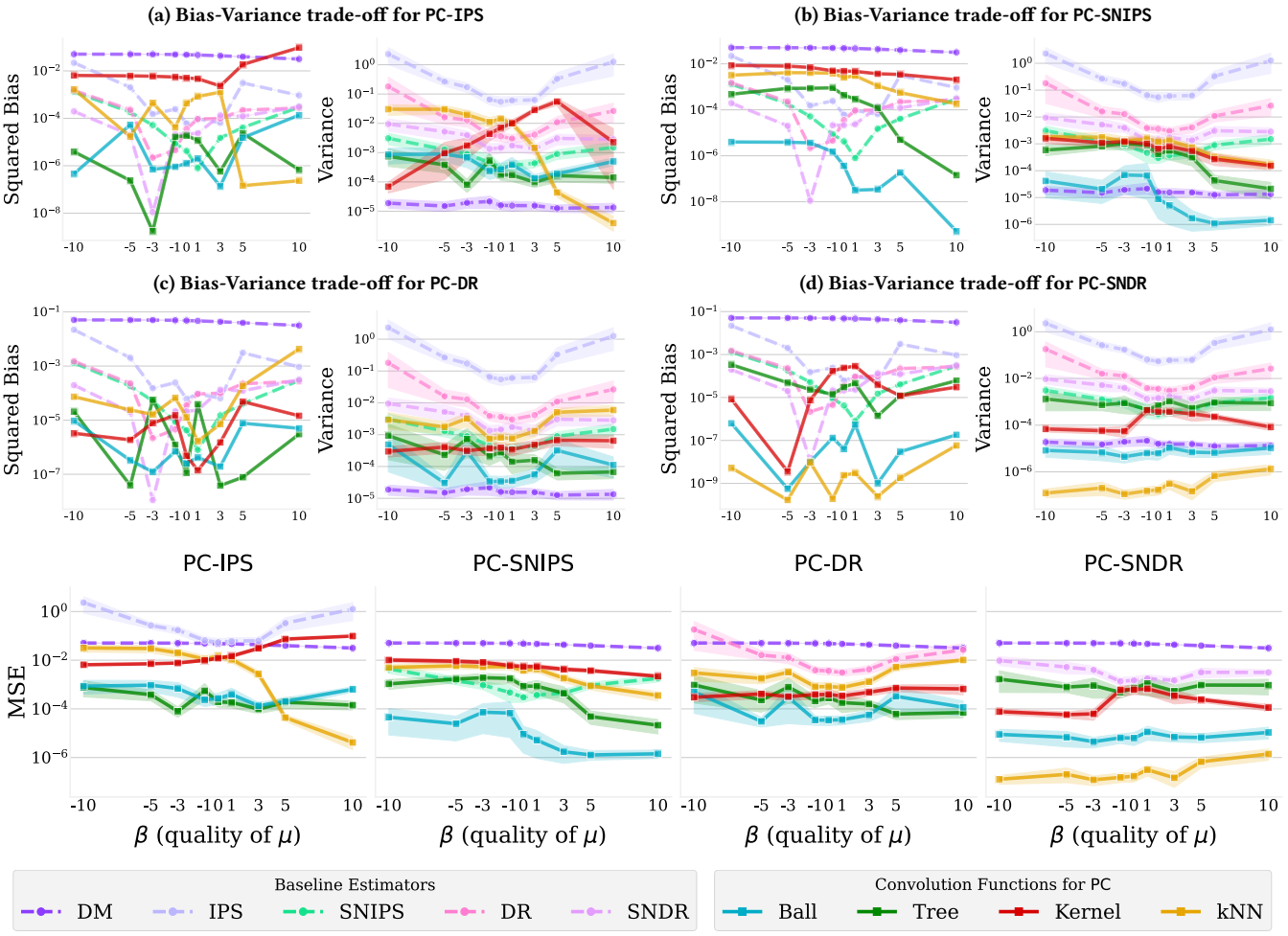


**Figure 11: Change in MSE, Squared Bias, and Variance while estimating $V(\pi_{\text{good}})$ with varying policy-mismatch (log scale) for the movielens dataset.**

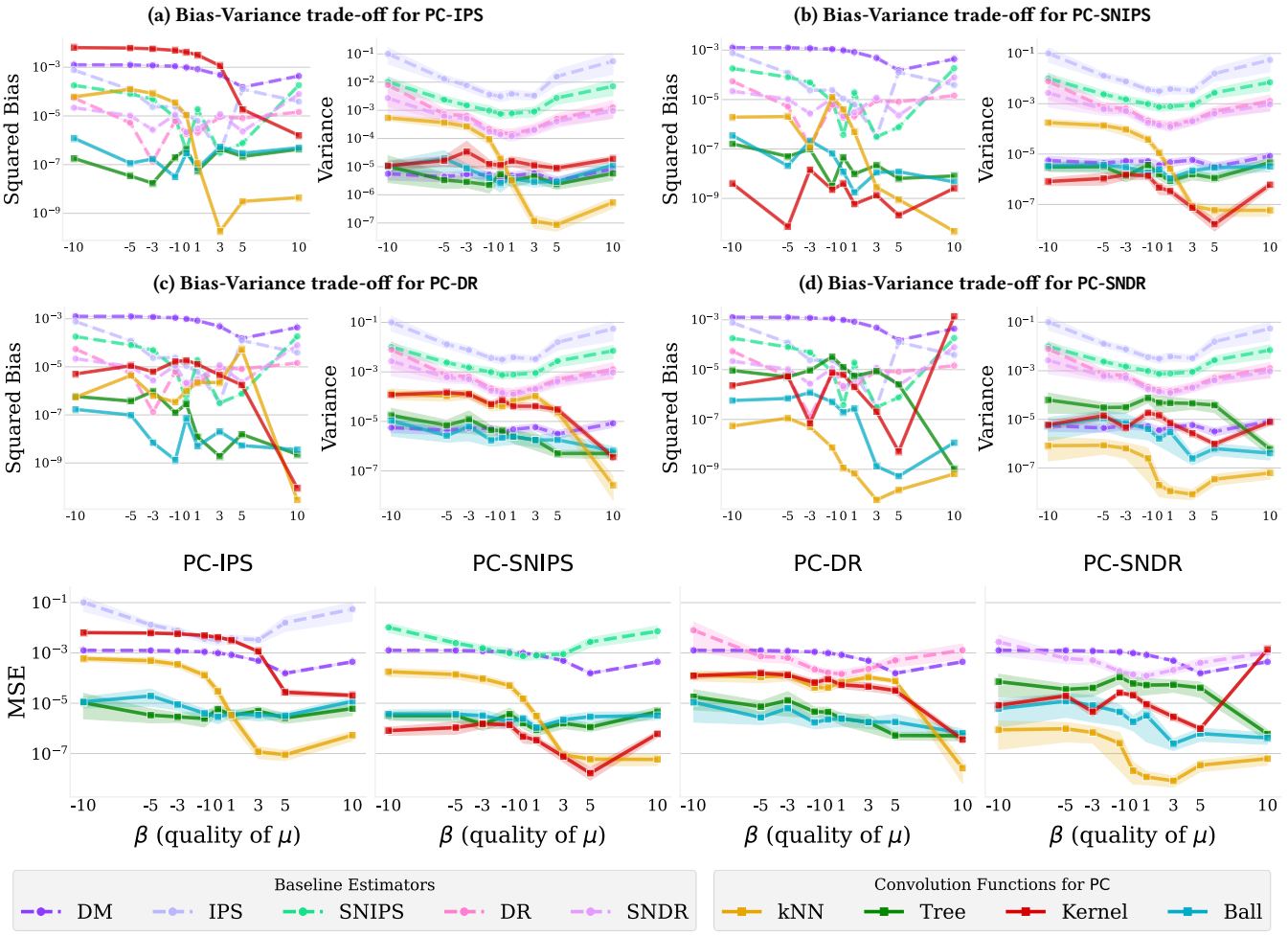


**Figure 12: Change in MSE, Squared Bias, and Variance while estimating $V(\pi_{\text{bad}})$ with varying policy-mismatch (log scale) for the movielens dataset.**

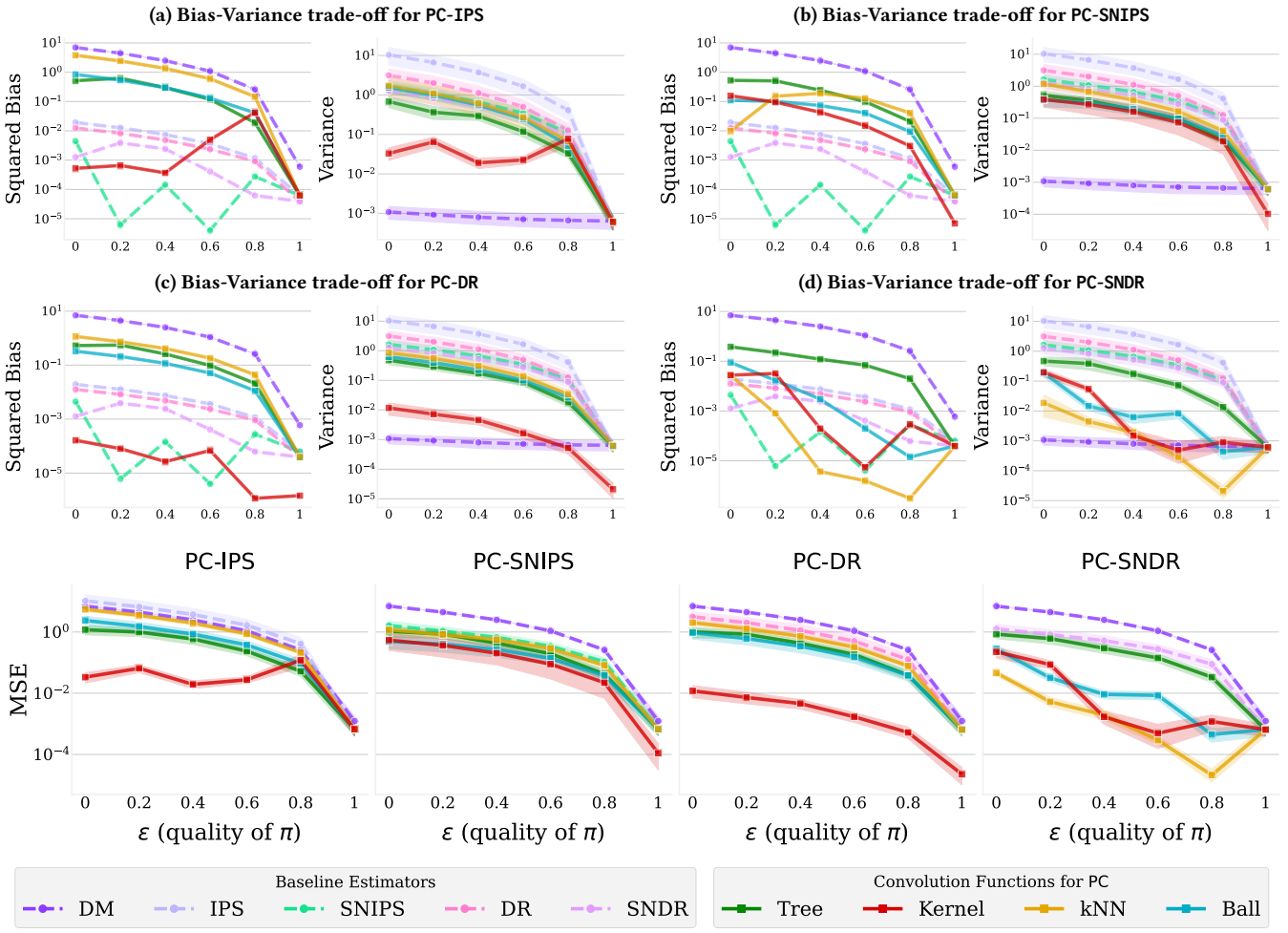


**Figure 13: Change in MSE, Squared Bias, and Variance while estimating various target policies (log scale) for the synthetic dataset, using data logged by $\mu_{\text{uniform}}$.**

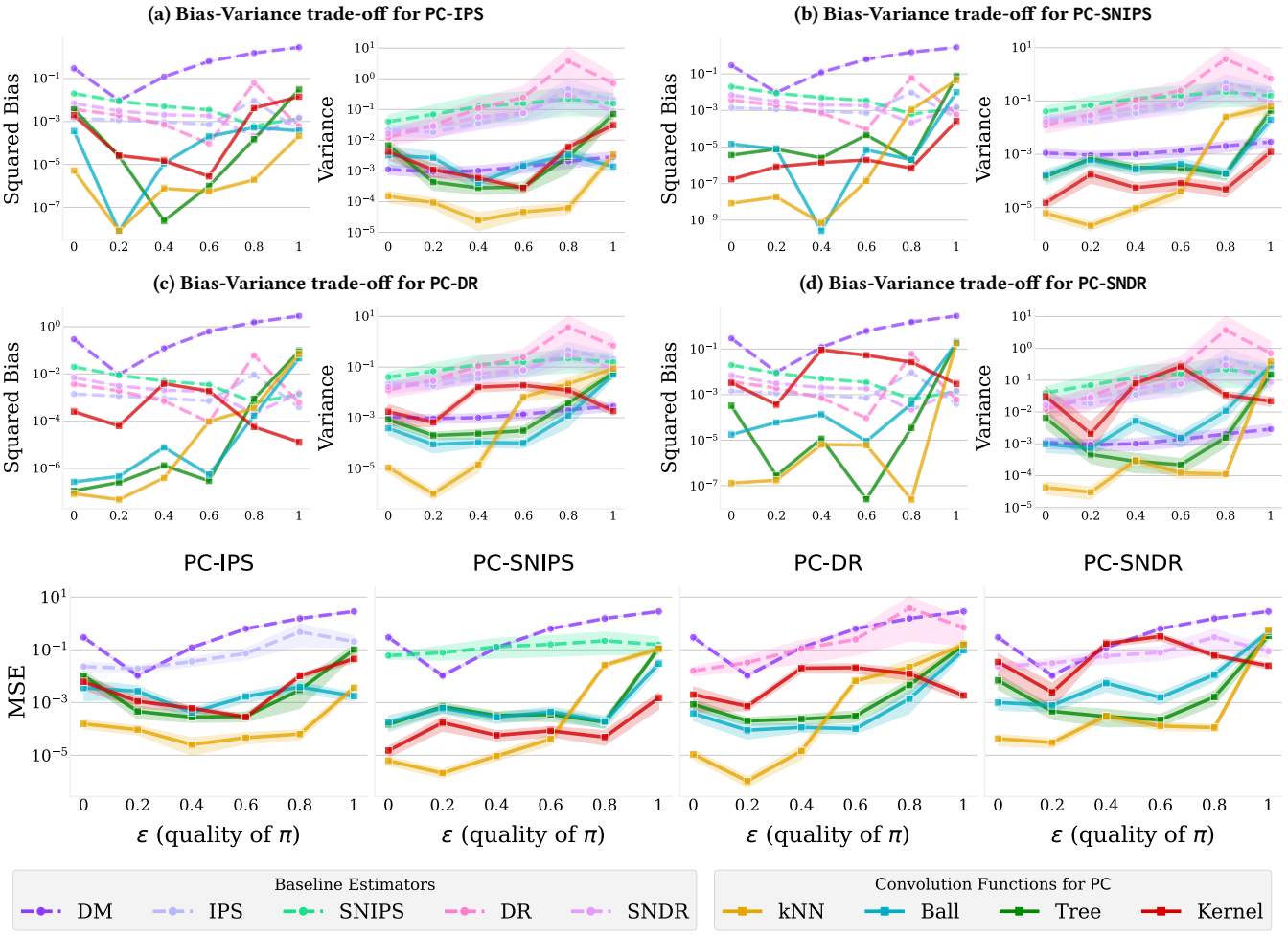


**Figure 14: Change in MSE, Squared Bias, and Variance while estimating various target policies (log scale) for the synthetic dataset, using data logged by $\mu_{\text{good}}$.**

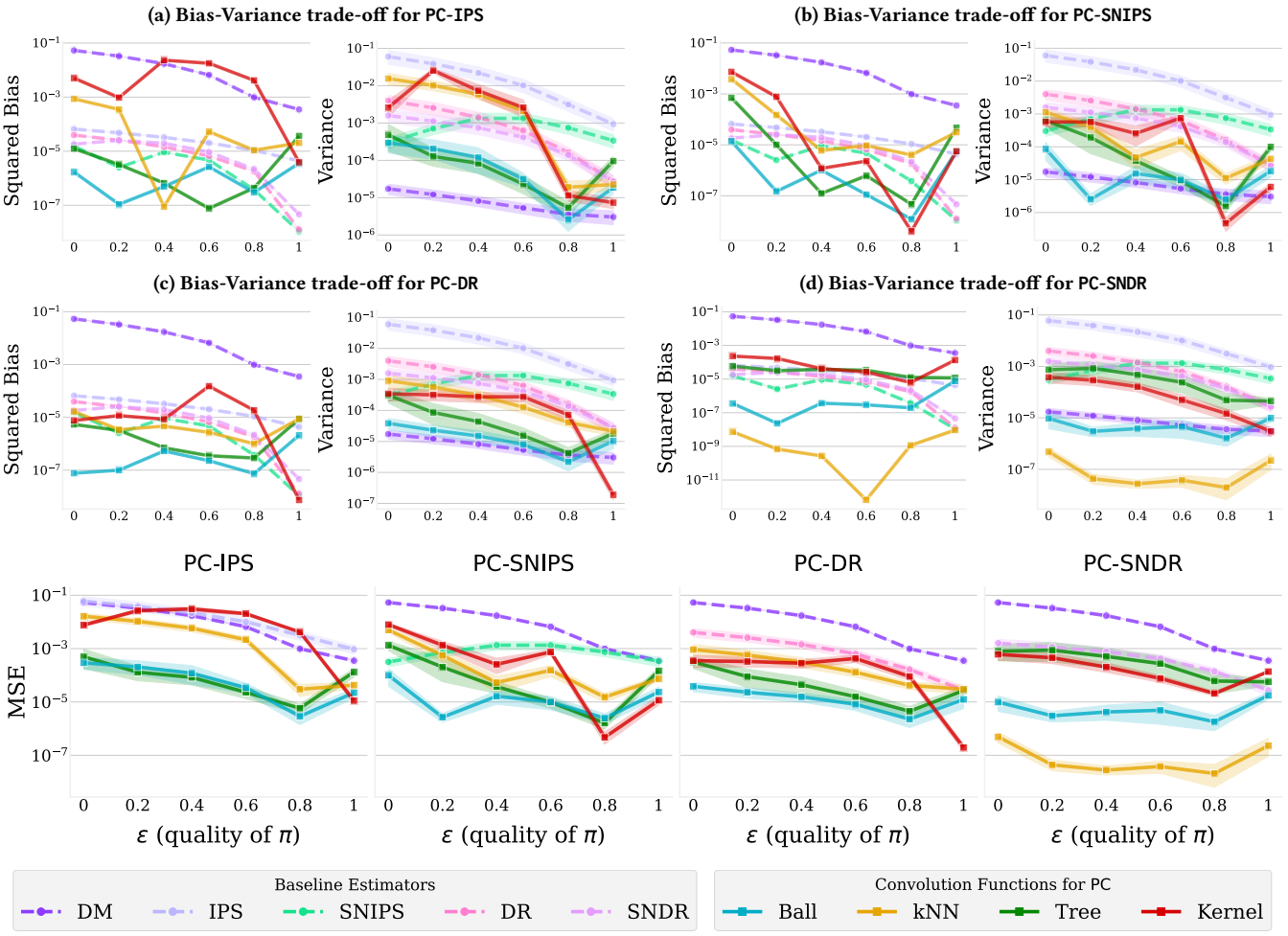


**Figure 15: Change in MSE, Squared Bias, and Variance while estimating various target policies (log scale) for the movielens dataset, using data logged by $\mu_{\text{uniform}}$.**

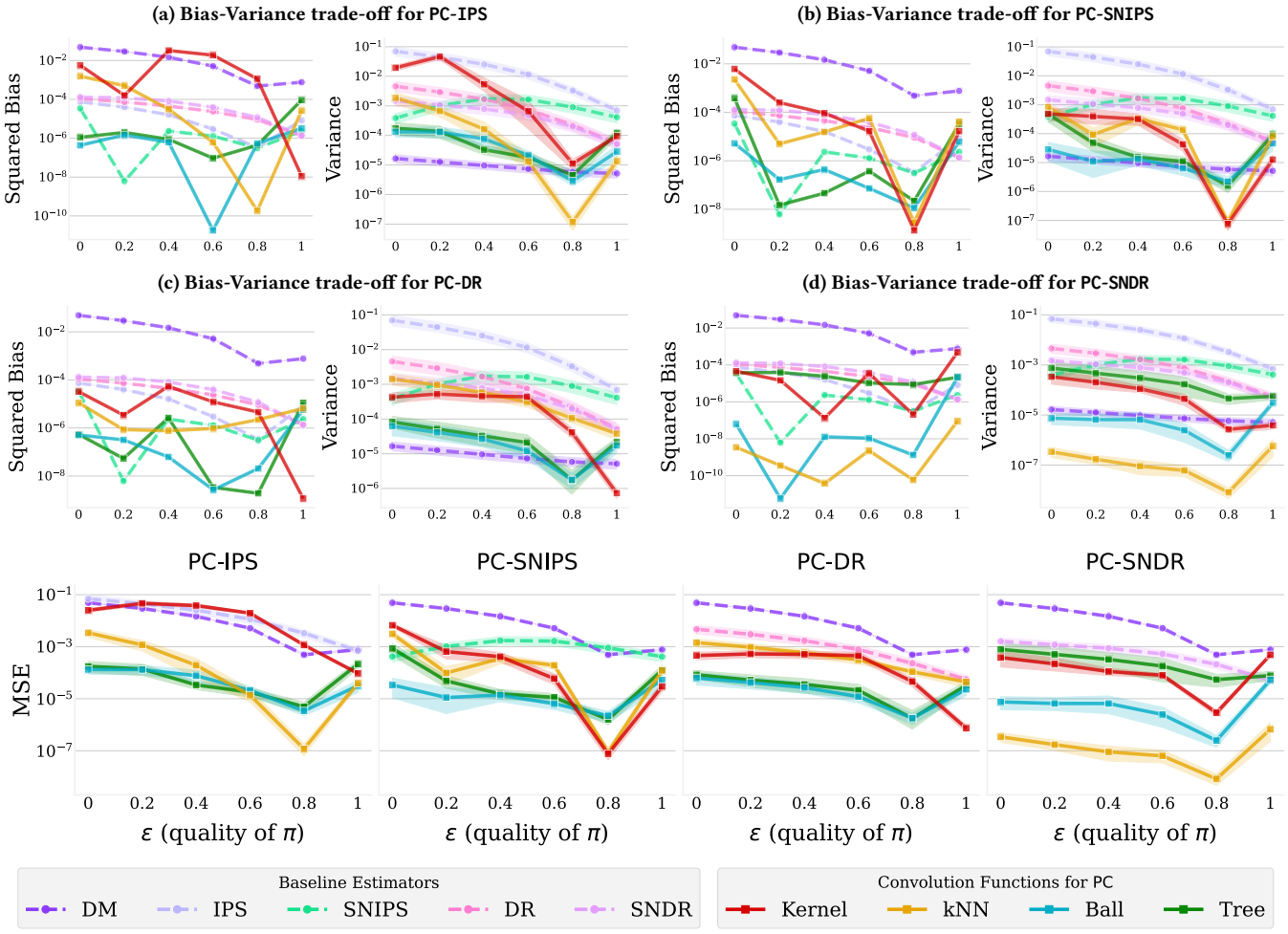


**Figure 16: Change in MSE, Squared Bias, and Variance while estimating various target policies (log scale) for the movielens dataset, using data logged by $\mu_{\text{good}}$.**

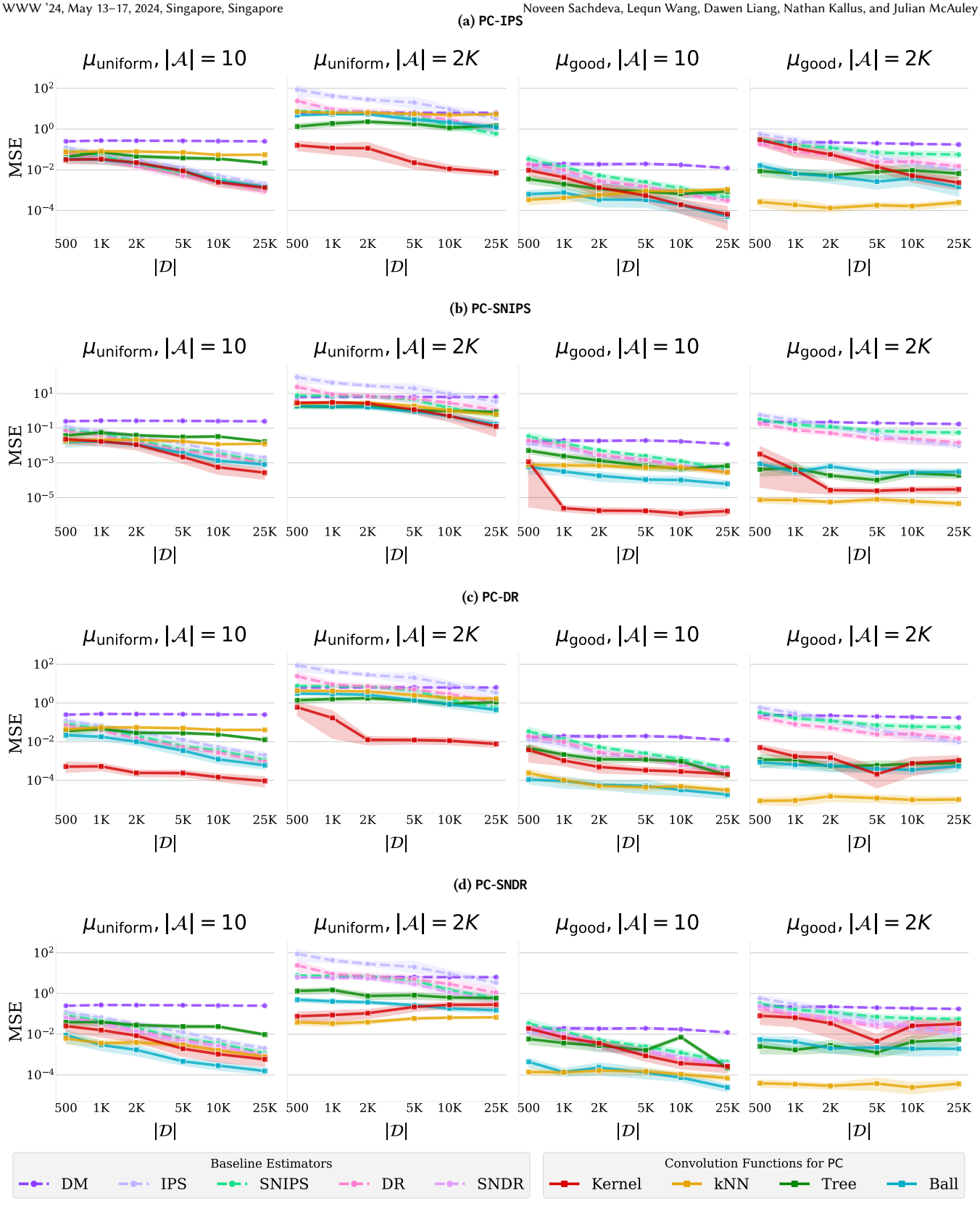


**Figure 17: Change in MSE while estimating $V(\pi_{\text{good}})$ with varying amount of bandit feedback (log-log scale) for the synthetic dataset.**

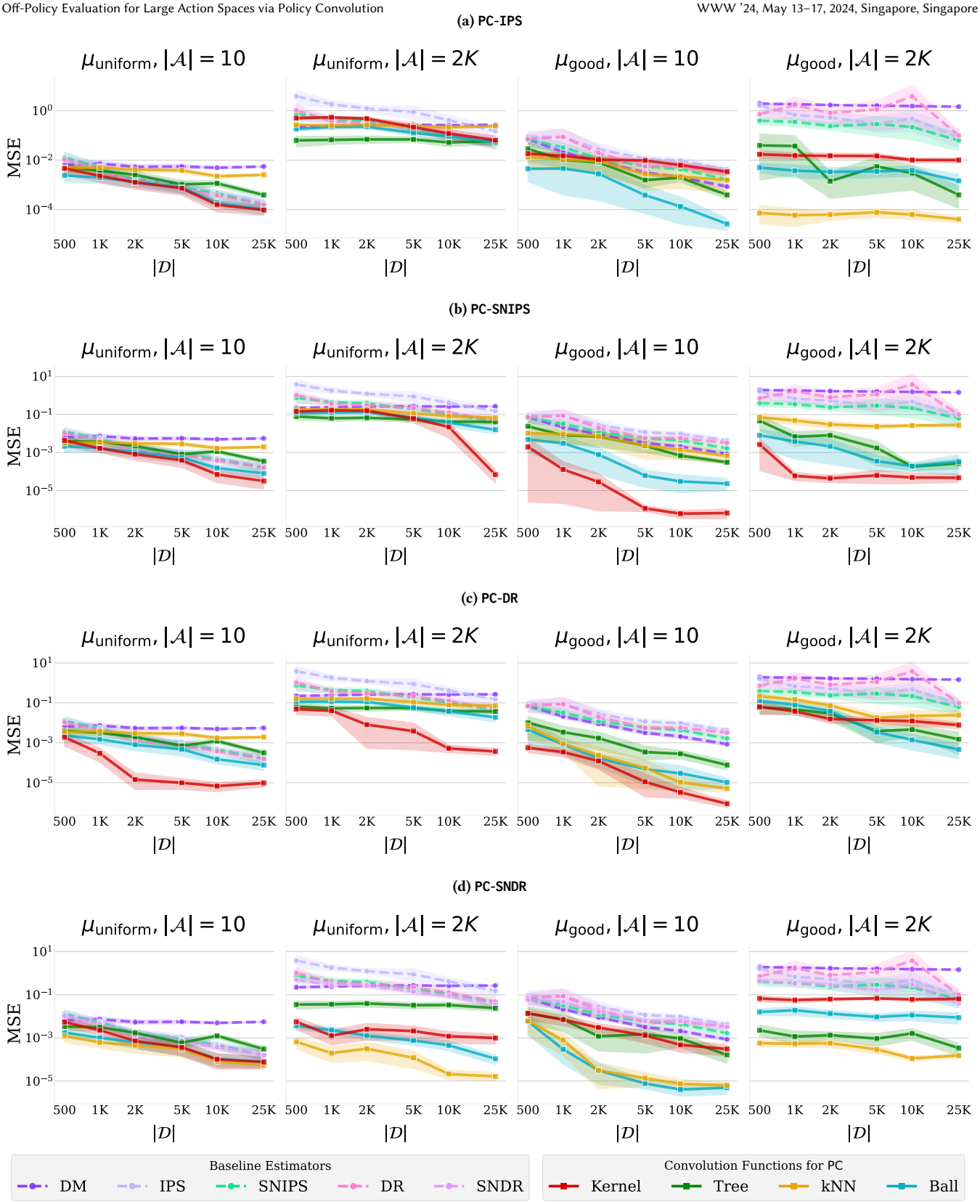


**Figure 18: Change in MSE while estimating $V(\pi_{\text{bad}})$ with varying amount of bandit feedback (log-log scale) for the synthetic dataset.**

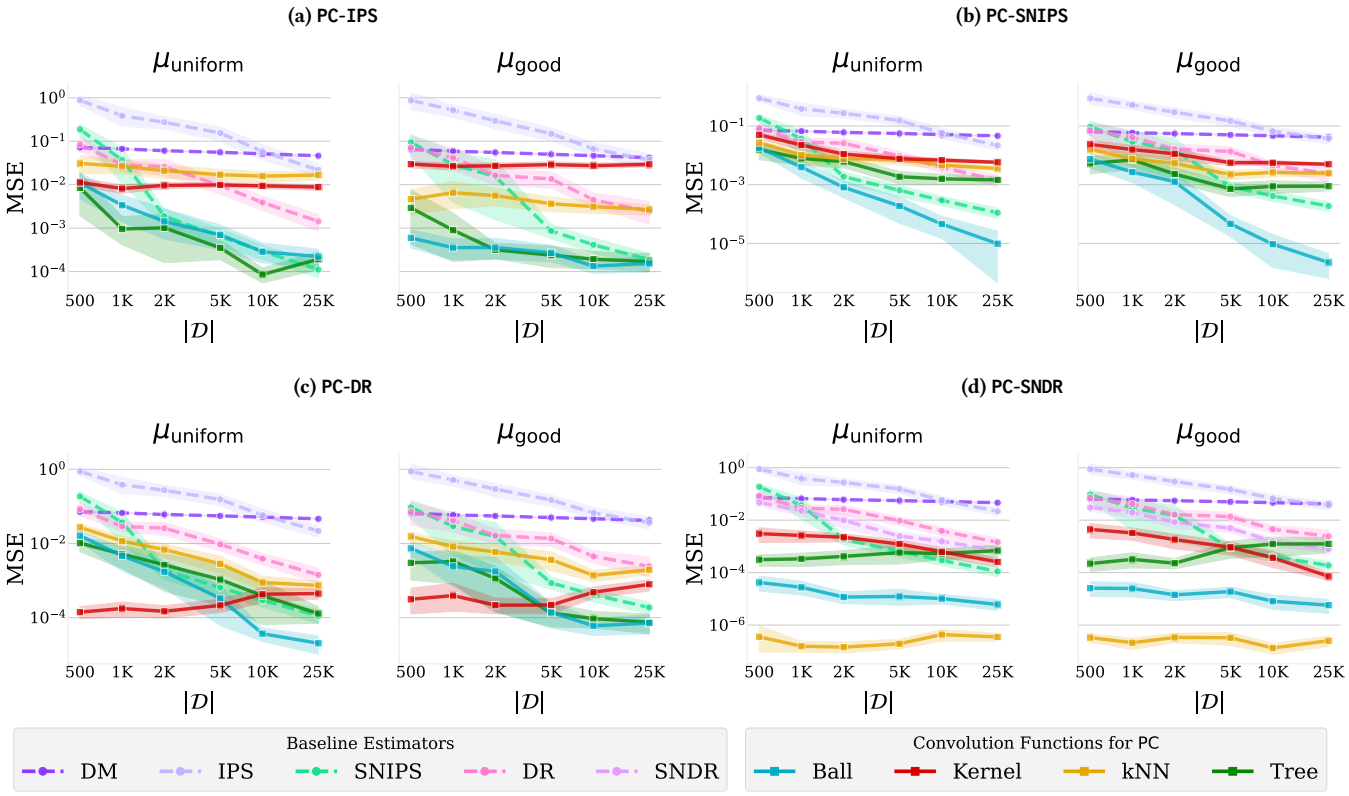


**Figure 19: Change in MSE while estimating $V(\pi_{\text{good}})$ with varying amount of bandit feedback (log-log scale) for the movielens dataset.**

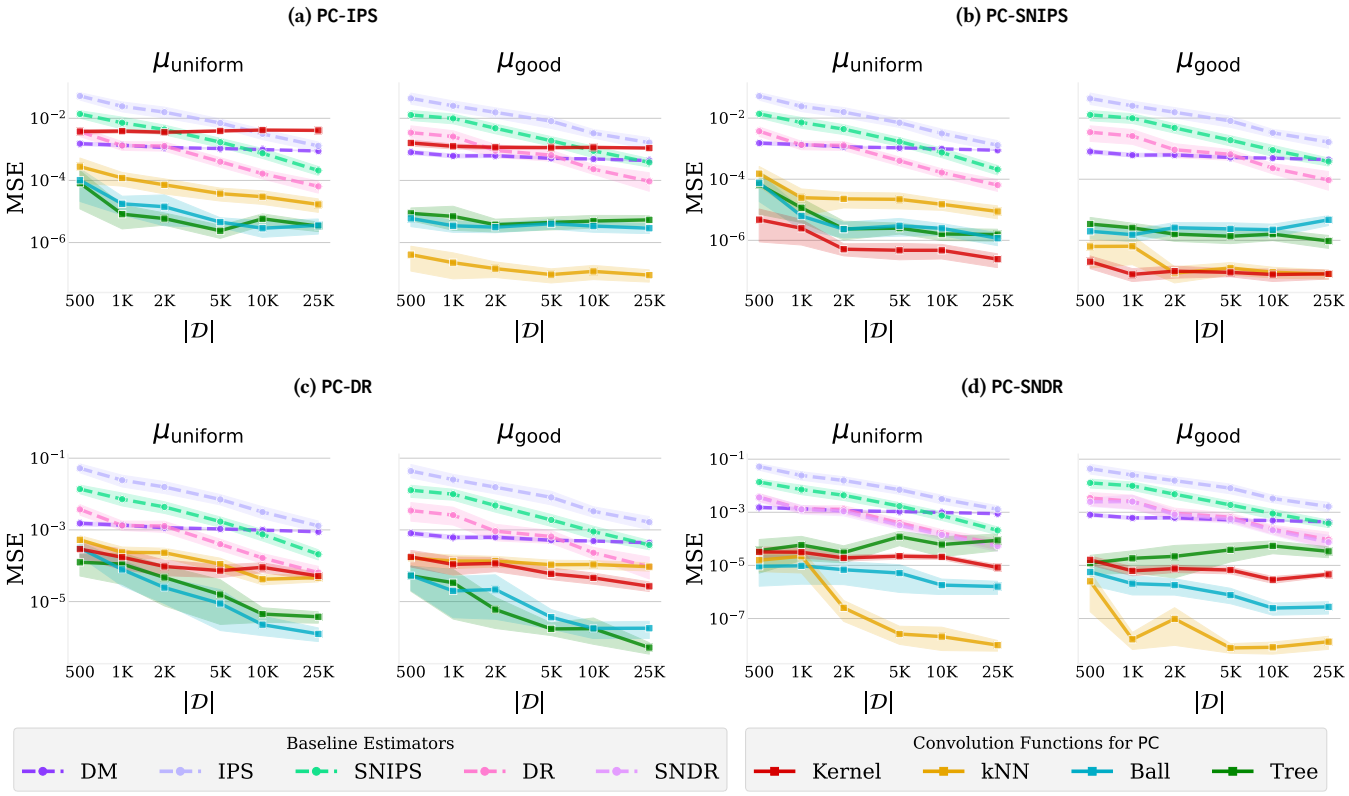


**Figure 20: Change in MSE while estimating $V(\pi_{\text{bad}})$ with varying amount of bandit feedback (log-log scale) for the movielens dataset.**

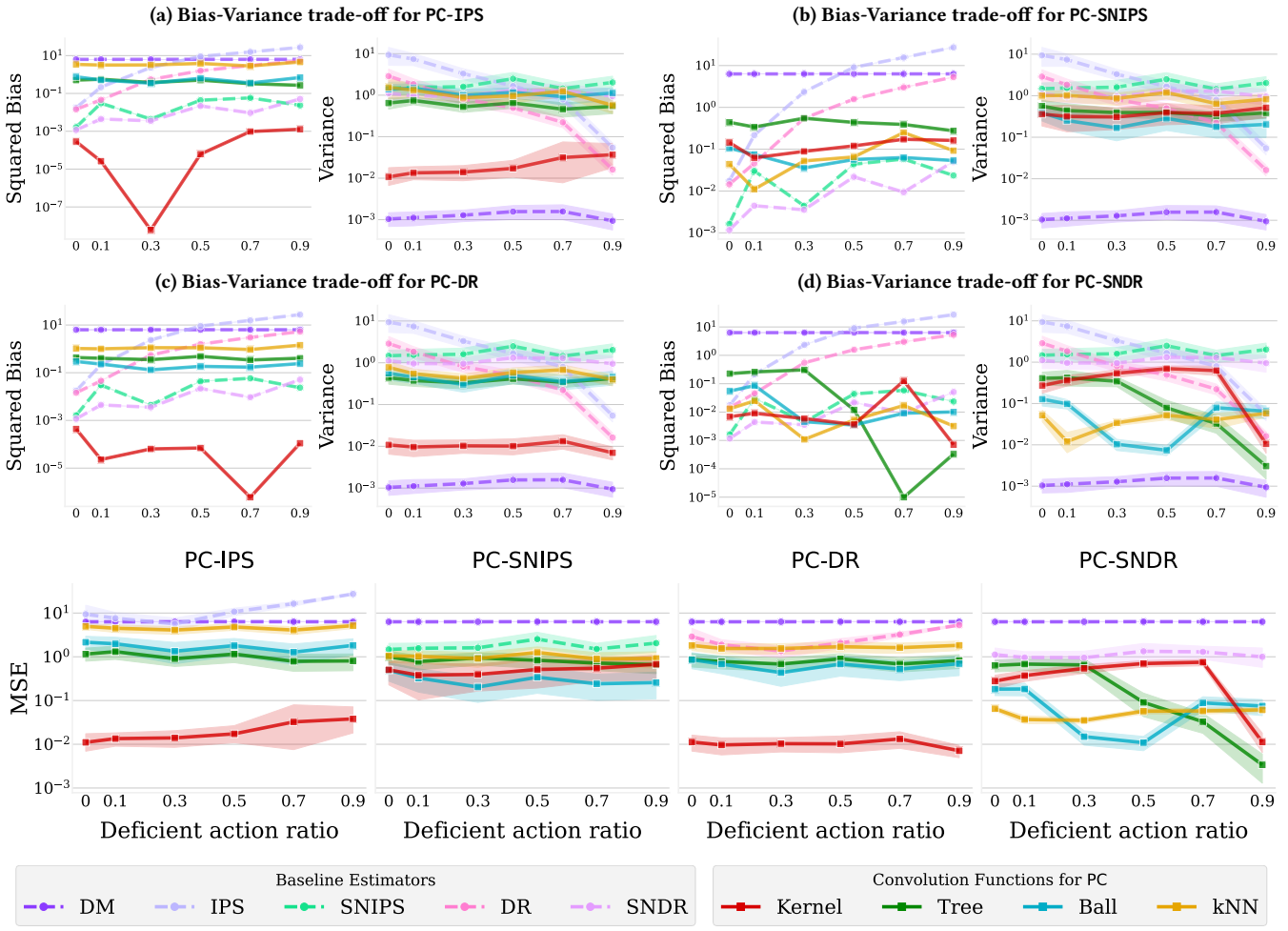


**Figure 21: Change in MSE, Squared Bias, and Variance while estimating $V(\pi_{\text{good}})$ with varying support (log-log scale) for the synthetic dataset (with 2000 actions), using data logged by $\mu_{\text{uniform}}$.**

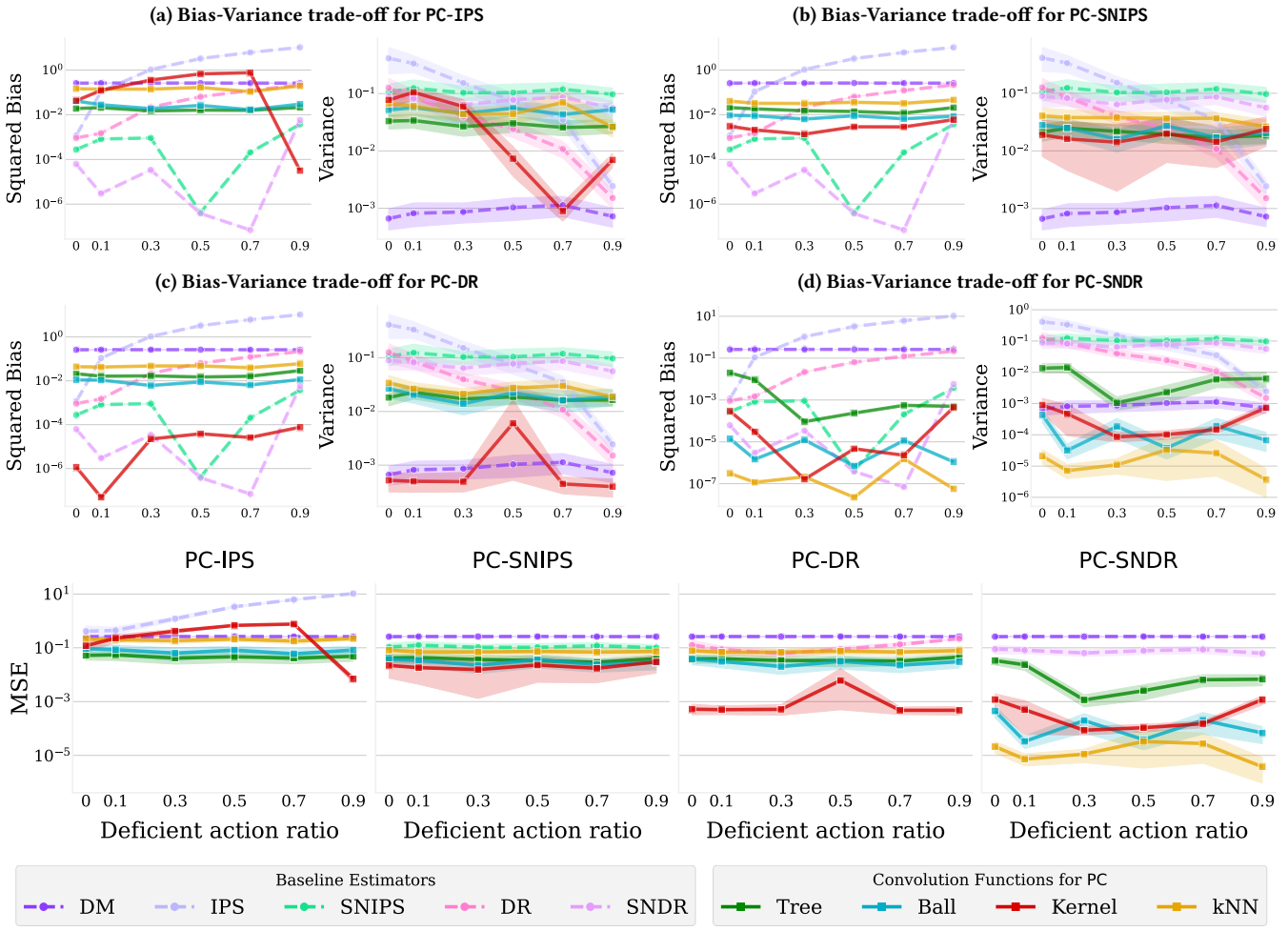


**Figure 22: Change in MSE, Squared Bias, and Variance while estimating $V(\pi_{\text{bad}})$ with varying support (log-log scale) for the synthetic dataset (with 2000 actions), using data logged by $\mu_{\text{uniform}}$.**

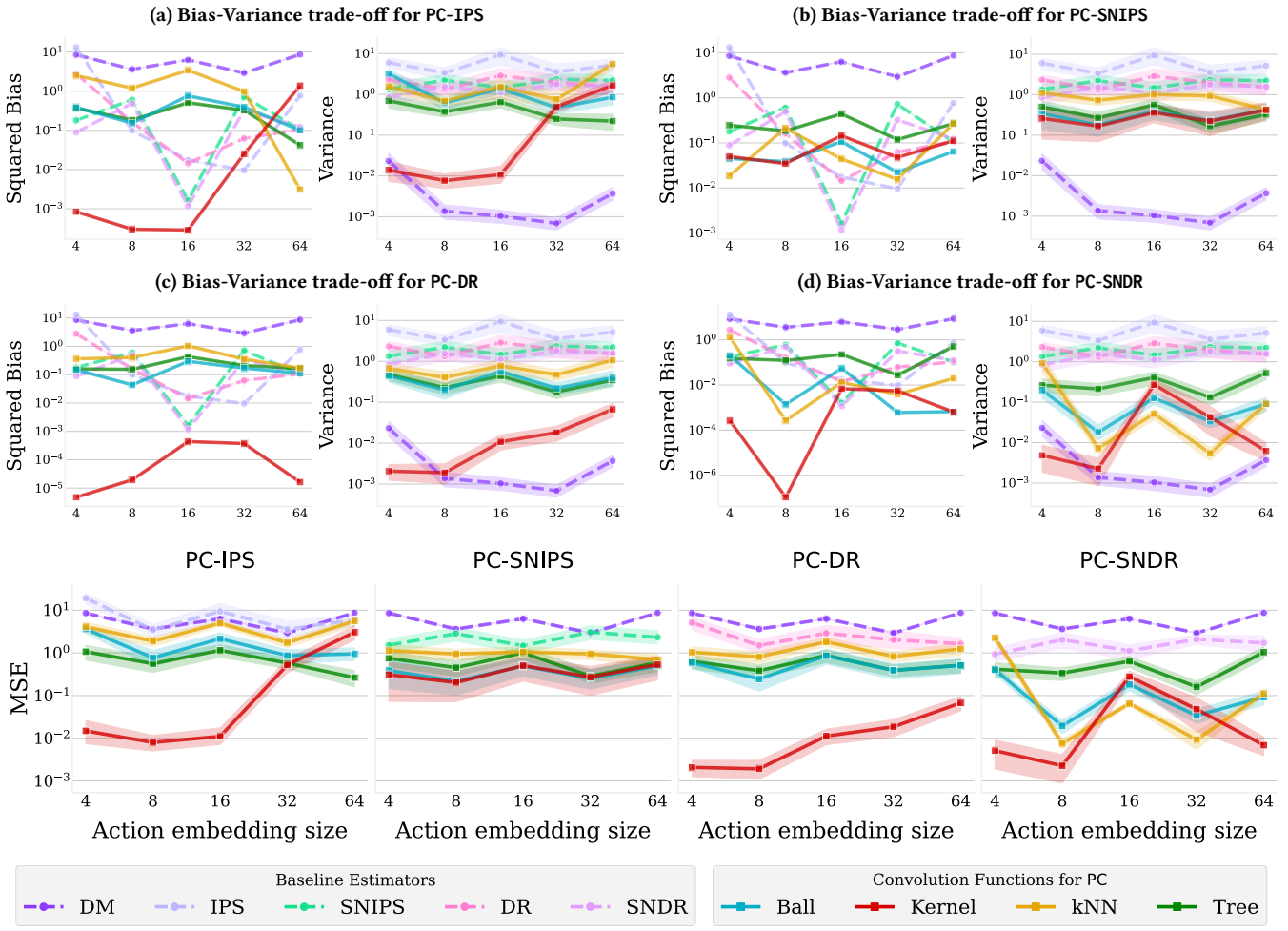


**Figure 23: Change in MSE, Squared Bias & Variance while estimating $V(\pi_{\text{good}})$ with varying action embedding size (log-log scale) for the synthetic dataset (2000 actions), using data logged by $\mu_{\text{uniform}}$.**

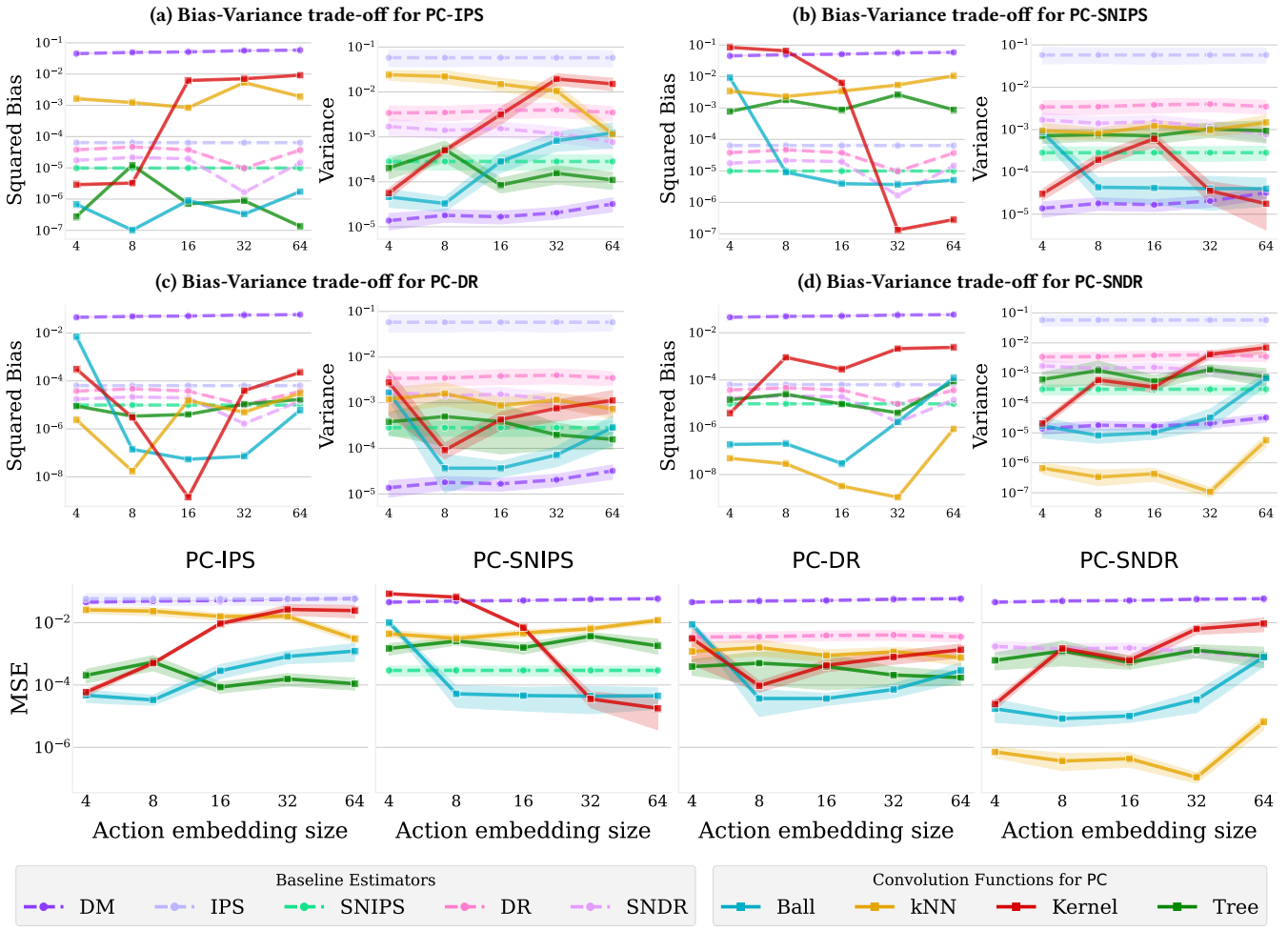


**Figure 24: Change in MSE, Squared Bias & Variance while estimating $V(\pi_{\text{good}})$ with varying action embedding size (log-log scale) for the movielens dataset, using data logged by $\mu_{\text{uniform}}$.**

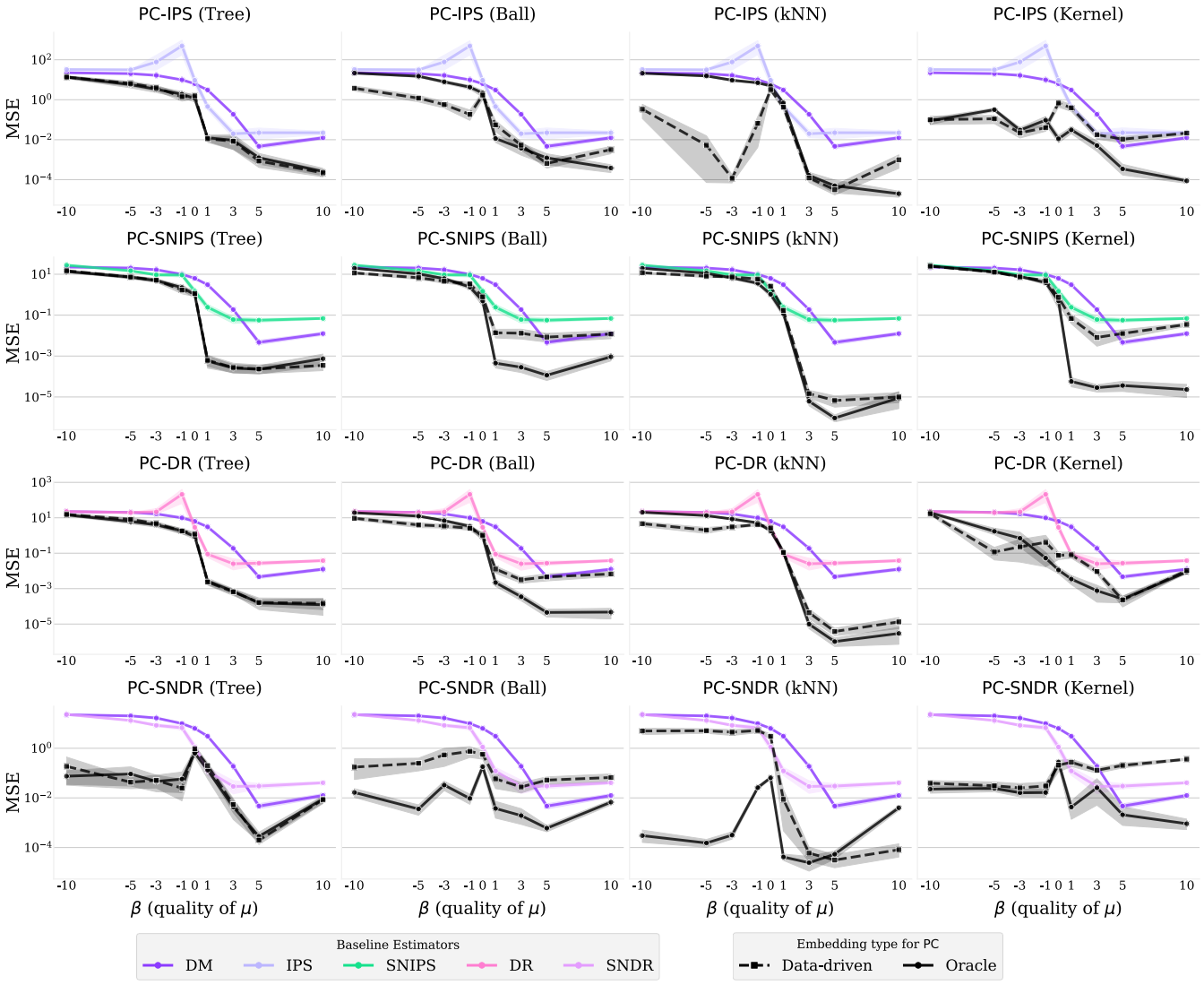


**Figure 25: Change in MSE while estimating $V(\pi_{\text{good}})$ with varying logging policies (log-scale) and using `PC` with `Oracle` *vs.* `Data-driven` action embeddings for the synthetic dataset.**

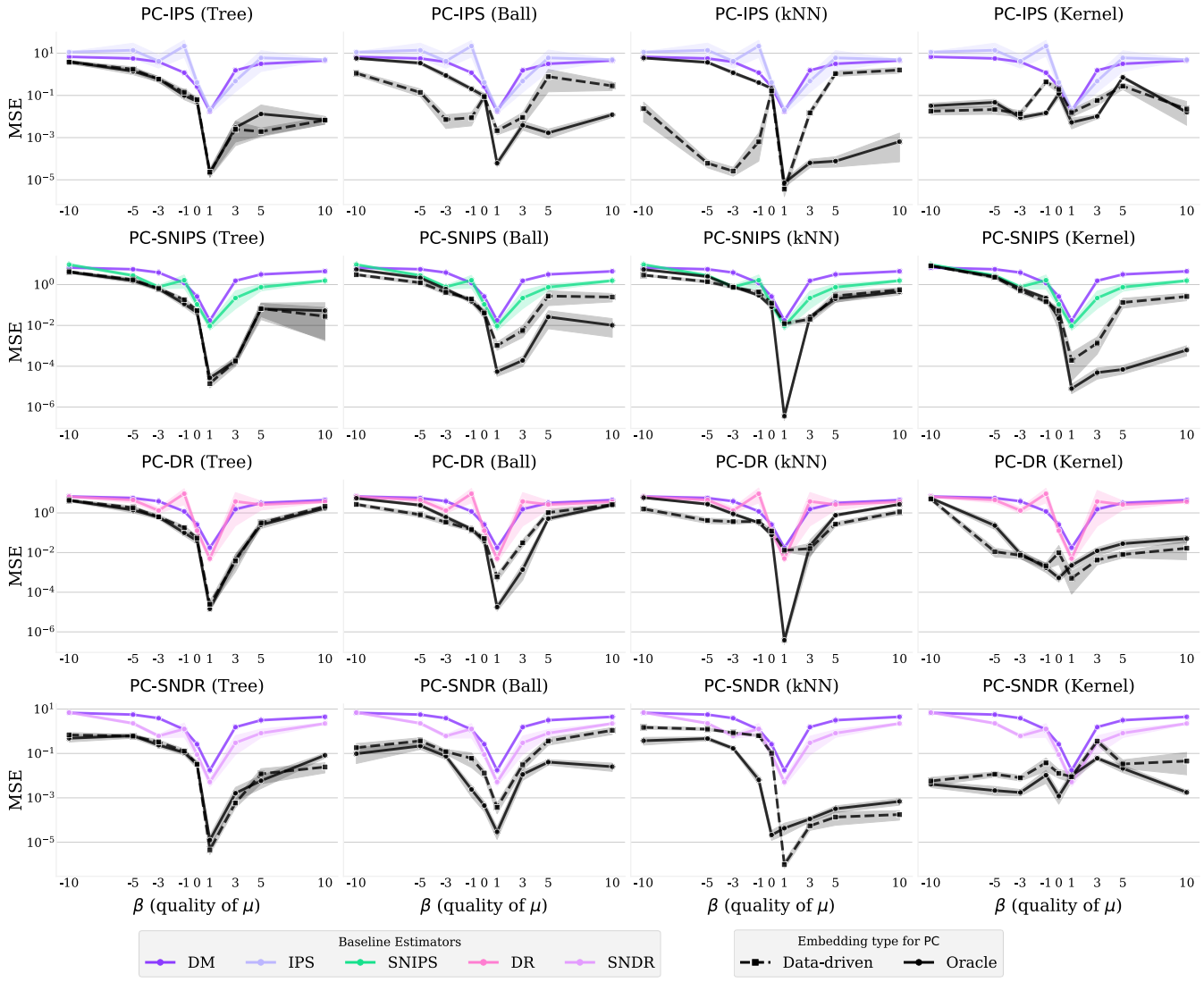


**Figure 26: Change in MSE while estimating $V(\pi_{\text{bad}})$ with varying logging policies (log-scale) and using `PC` with `Oracle` *vs.* `Data-driven` action embeddings for the synthetic dataset.**

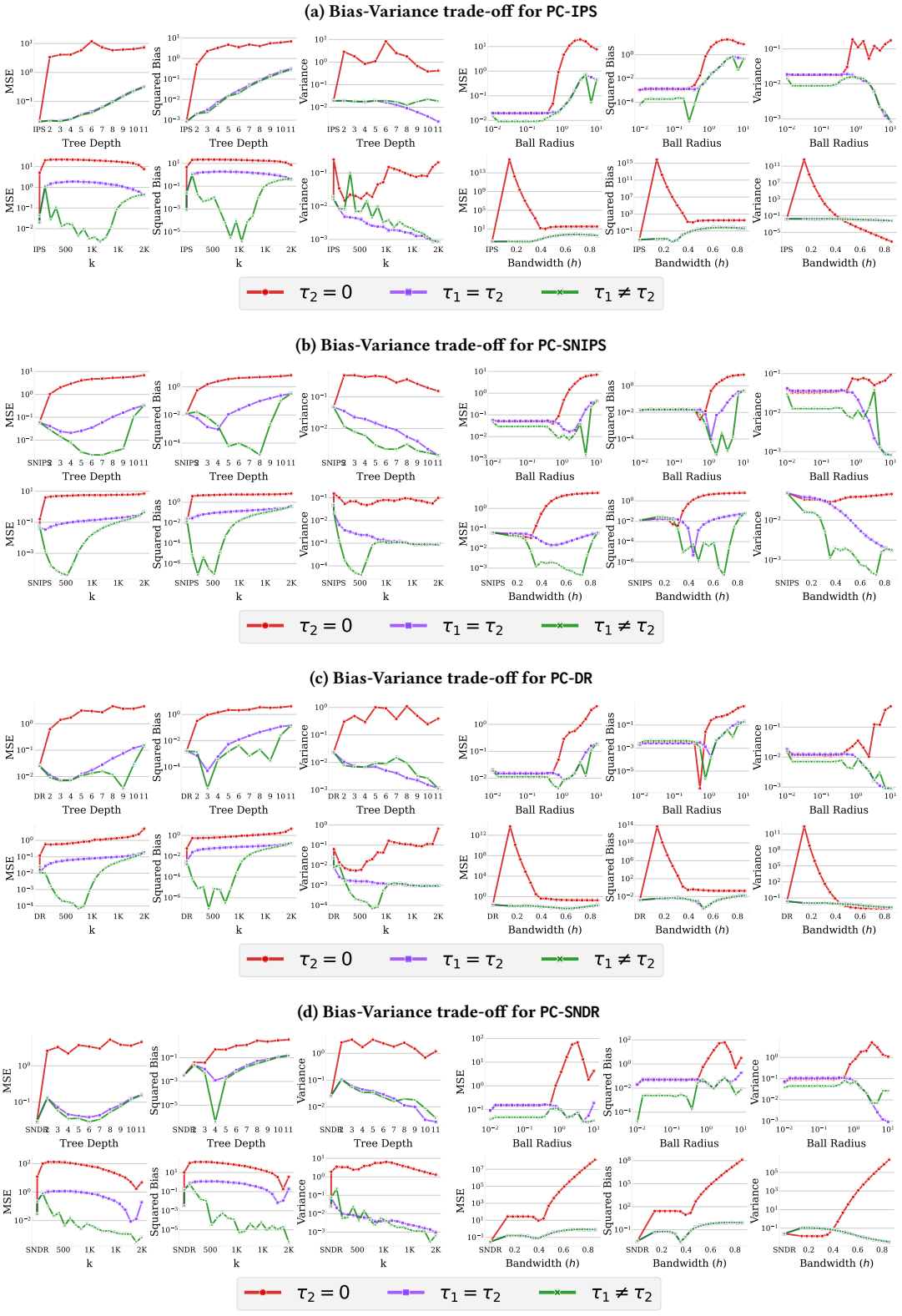


**Figure 27: Visualizing the bias-variance trade-off for the PC estimator for different backbones & pooling strategies, while estimating $V(\pi_{\text{good}})$ with varying amount of pooling (log-log scale) on the synthetic dataset (with 2000 actions), using $\mu_{\text{good}}$ for logging. Note that the respective naïve backbone estimators are the left-most point in each plot, *i.e.*, when there's no pooling.**

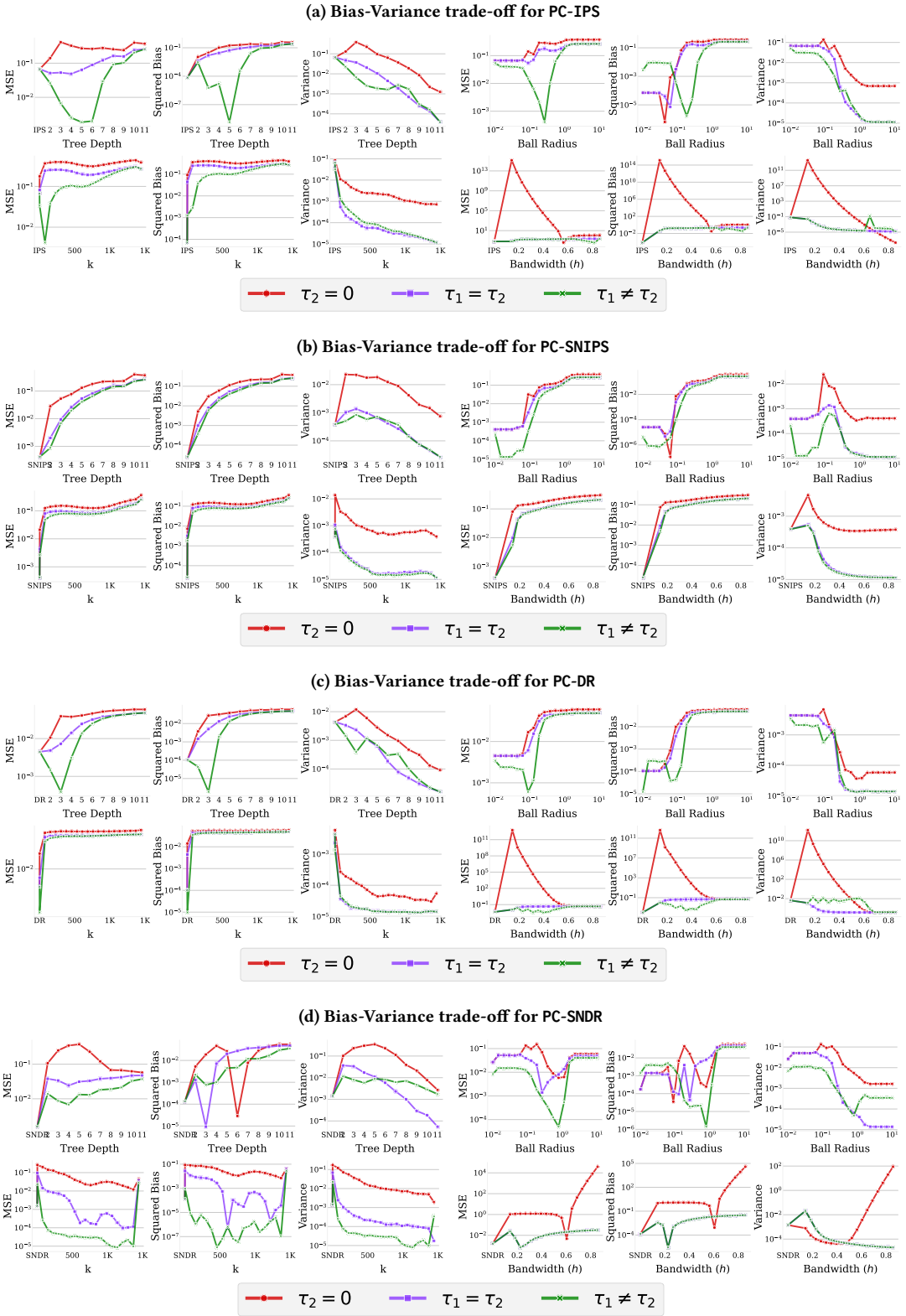


**Figure 28: Visualizing the bias-variance trade-off for the PC estimator for different backbones & pooling strategies, while estimating $V(\pi_{\text{good}})$ with varying amount of pooling (log-log scale) on the movielens dataset, using $\mu_{\text{good}}$ for logging. Note that the respective naïve backbone estimators are the left-most point in each plot, *i.e.*, when there's no pooling.**

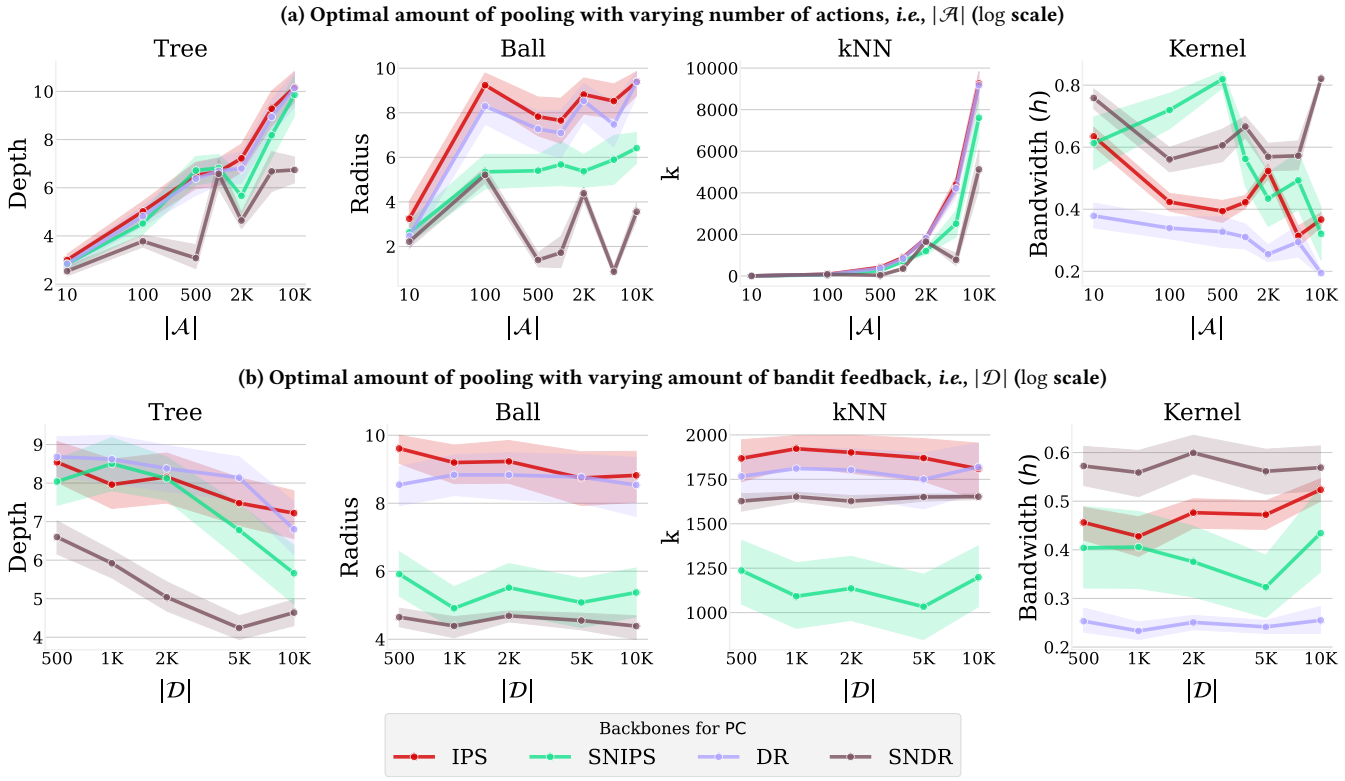


**Figure 29: Change in the optimal amount pooling for PC while estimating $V(\pi_{\text{good}})$ for the synthetic dataset, using data logged by $\mu_{\text{bad}}$.**

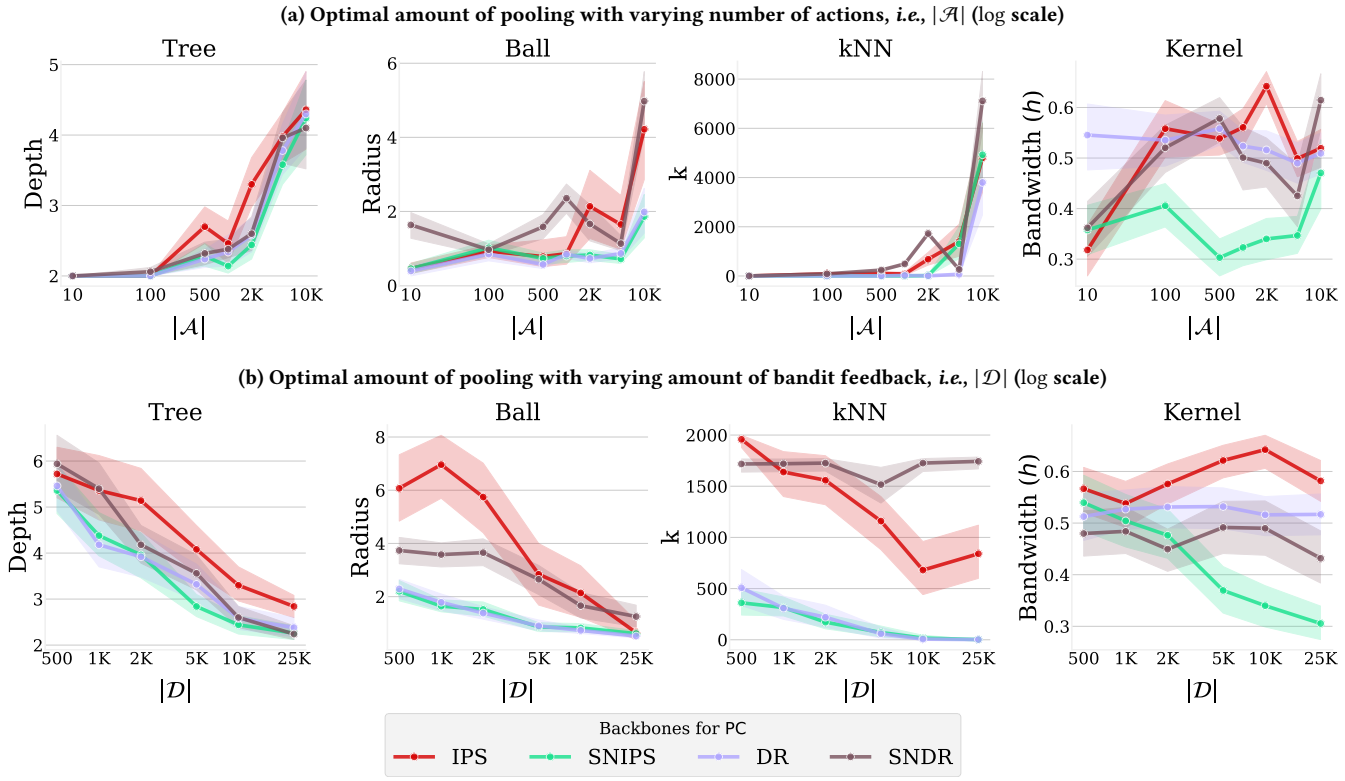


**Figure 30: Change in the optimal amount pooling for PC while estimating $V(\pi_{\text{good}})$ for the synthetic dataset, using data logged by $\mu_{\text{uniform}}$.**

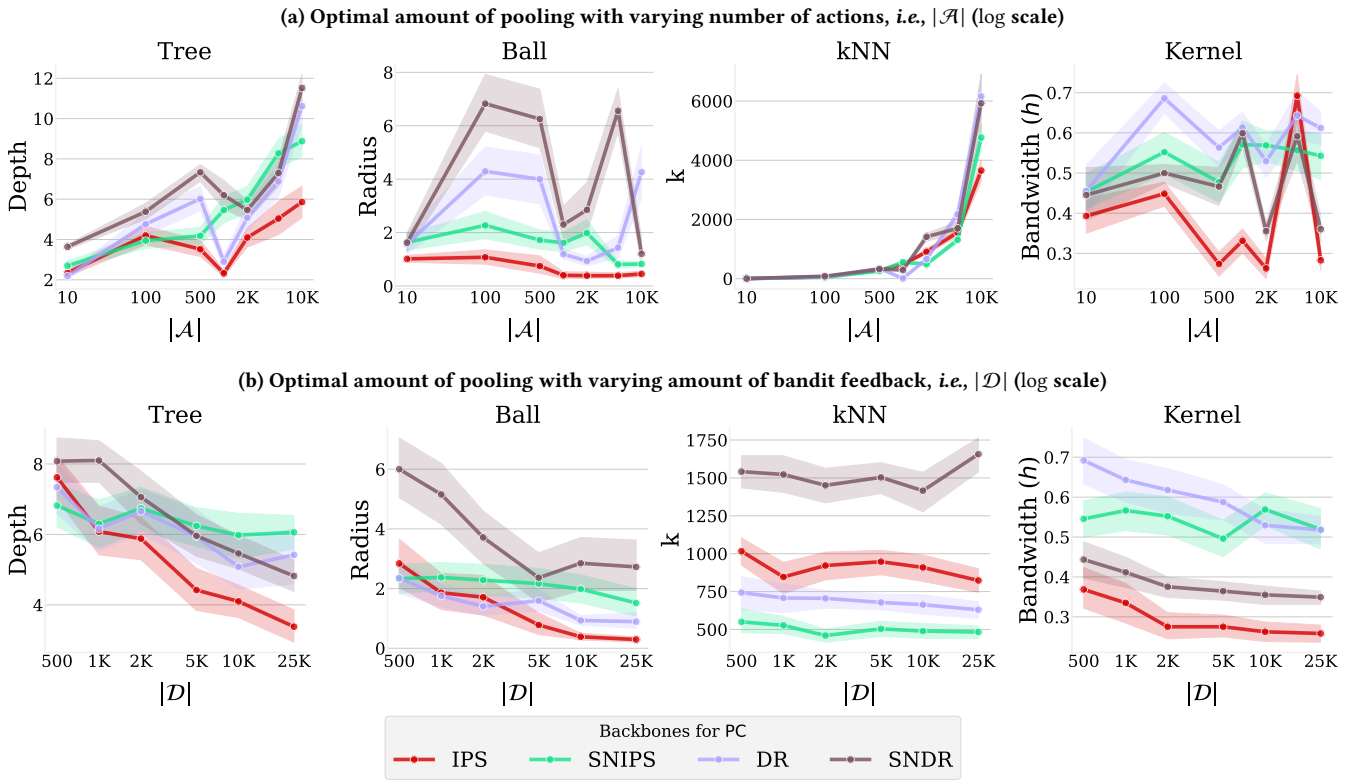


**Figure 31: Change in the optimal amount pooling for PC while estimating $V(\pi_{\text{good}})$ for the synthetic dataset, using data logged by $\mu_{\text{good}}$.**