



B Policy

$$\pi(a | \text{Blue}) \propto \exp \left[\beta \left[\text{Blue} \right] + \log \left[\text{Grey} \right] \right]$$

