Problem Set 2
Cristian Rodriguez

a) Systems that model real world situations that may never have access to a perfect model will benefit from a stochastic policy more than a deterministic one. For example, a game of rock, paper, scissors, where there is no "perfect" decision to make given the current state, and therefore a player using a stochastic policy may make an educated decision, however they may also benefit from getting lucky through a random decision.

b) $$V^{\pi}(s) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t) \mid S_0 = s, a_t = \pi(s_t), s_{t+1} \mid s_t, a_t \sim P\right]$$

* Equation found on CMU 15-780 slides, J. Zico kolter

c) By measuring the results of deterministic policies, one can adapt the algorithm such that it allows for randomness in the action space according to the reward. You can add a random element if you notice that there is no obvious way to predict what your opponent will draw in rock, paper, scissors.

d) Yes