$$V^\pi(s) = \sum_{t=1}^{\infty} E\left\{\gamma^{t-1} r_t \mid s_0 = s\right\} \quad (1)$$

$$= \sum_a \pi(s,a)\left[R(s,a) + \gamma \sum_{s'} P(s,s',a) V^\pi(s')\right] \quad (2)$$

$$d^\pi(s') = \lim_{t\to\infty} Pr\left\{s_t = s' \mid s_0, \pi\right\} \qquad \text{(does not depend on } s_0\text{)} \quad (3)$$

$$= \sum_s d^\pi(s) \sum_a \pi(s,a) P(s,s',a) \quad (4)$$

$$\rho^\pi = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} r_t \qquad \text{(does not depend on } s_0\text{)} \quad (5)$$

$$= \sum_s d^\pi(s) \sum_a \pi(s,a) R(s,a) \quad (6)$$

In trying to form an overall discounted performance measure for π, can we use $J(\pi) = \sum_s d^\pi(s) V^\pi(s)$? It turns out we then end up with no effect of the discounting:

$$J(\pi) = \sum_s d^\pi(s) V^\pi(s) \quad (7)$$

$$= \sum_s d^\pi(s) \sum_a \pi(s,a)\left[R(s,a) + \gamma \sum_{s'} P(s,s',a) V^\pi(s')\right] \quad (8)$$

$$= \rho^\pi + \gamma \sum_s d^\pi(s) \sum_a \pi(s,a) \sum_{s'} P(s,s',a) V^\pi(s') \quad (9)$$

$$= \rho^\pi + \gamma \sum_{s'} V^\pi(s') \sum_s d^\pi(s) \sum_a \pi(s,a) P(s,s',a) \quad (10)$$

$$= \rho^\pi + \gamma \sum_{s'} V^\pi(s') d^\pi(s') \quad (11)$$

$$= \rho^\pi + \gamma J(\pi) \quad (12)$$

$$= \rho^\pi + \gamma \rho^\pi + \gamma^2 J(\pi) \quad (13)$$

$$= \rho^\pi + \gamma \rho^\pi + \gamma^2 \rho^\pi + \gamma^3 \rho^\pi + \cdots \quad (14)$$

$$= \frac{1}{1-\gamma} \rho^\pi \quad (15)$$

which is basically a scaled ρ^π, with no effect of discounting.