

$$\sigma_0(c^t, d^t) \sum_{c'=1}^{|C|} \sum_{v'_0=1}^2 R_0(v'_0) Pr(v'_0|c', c^t, a_0^t, d^t) Pr(c'|c^t, d^t) \quad (1)$$

$$Pr(d^t|c^t, a^t) := \frac{\exp(\beta Q_{DM}(c^t, a^t, d^t))}{\sum_{d=1}^{|D|} \exp(\beta Q_{DM}(c^t, a^t, d))} \quad (2)$$

$$Q_{DM}(c^t, a^t, d) := Q_{ext}(c^t, d) + \sum_{\theta=0}^{|A|} \mu_\theta \sigma_\theta(c_t, d) \sum_{c'=1}^{|C|} \sum_{v'_\theta=1}^2 R_\theta(v'_\theta) Pr(v'_\theta|c', c^t, a_\theta^t, d) Pr(c'|c^t, d) \quad (3)$$

$$\mathbb{R}_{mean}(c^t) = \frac{\sum_{d=d_1, d_2} \left[\left(\sum_{c'=1}^{|C|} f(c^t, d, c') \right) \left(\sum_{c'=1}^{|C|} R_{ext}(c'|c^t, d) Pr(c'|c^t, d) \right) \right]}{\sum_{d=d_1, d_2} \left(\sum_{c'=1}^{|C|} f(c^t, d, c') \right)} \quad (4)$$