1 Problem description

Consider a duopoly with two stores, A and B, selling the same item. Store A sells the item for a price p while store B sells it for price $p^+ = p + \epsilon$ for some $\epsilon > 0$. Store B also gives a reward R to a customer after making k purchases at B. Our goal is to find the optimal k for store B to maximize its rate of revenue under certain assumptions on customer behavior.

We assume that there is some exogenous probability, λ, during each purchase that forces the customer to go to store B. Let $0 < \beta \leq 1$ denote the discounting factor of future money. (NOTE/QUESTION: in the previous write-up, β was called the expiry factor - is this still the correct interpretaion? If so, it seems like β should be close to zero but in the discounting interpretation β should be close to one) For simplicity, we assume the customer has the following utility as a function of price paid: $u(p) = v > 0$ and $u(p^+) = 0$.

We model the customer's decision problem as a dynamic problem. We index the number of visits the customer makes at store B by i, for $0 \leq i \leq k-1$, and we refer a customer to be in state i after having made i visits to B. At state i, the customer has two possibilities:

1. With probability λ, the customer must visit B, and she is now in state $i+1$.
2. With probability $1-\lambda$, the customer may purchase from A for utility v and remain in state i or purchase from B for no utility but move to state $i+1$.

Let $V(i)$ denote the long term expected reward at state i. Then we may model the decision problem as the following dynamic program.

$$V(i) = \lambda\beta V(i+1) + (1-\lambda)\max\{v + \beta V(i), \beta V(i+1)\} \text{ for } 0 \leq i \leq k-1$$
$$V(k) = R$$

We will show that the decision process exhibits a phase transition; that is prior to some state i_0, the customer will only visit B if she must do so exogenously but after i_0, she always decides to go to B.

Finally, we assume the customer has a look-ahead factor t, which models how many purchases ahead the customer looks ahead when making her current decision. This value will affect the phase transition of the decision process. Consider a distribution T describing the look-ahead factor for consumers. We will focus on threshold distributions; for example, with probability p the look-ahead is t_1 and with probability $1-p$, the look-ahead is t_0.

2 Solving the DP and the phase transition

NEED TO: find conditions under which $V(i)$ is increasing in i and provide proof (these will be different than in the original write-up). The following lemma would be written as a part of this lemma.

Lemma 2.1. *We may write the DP as*

$$V(i) = \max\left\{\frac{\lambda\beta V(i+1) + (1-\lambda)v}{1-(1-\lambda)\beta}, \beta V(i+1)\right\}$$

Proof. We have the following:

$$\begin{aligned} V(i) &= \lambda\beta V(i+1) + (1-\lambda)\max\{v + \beta V(i), \beta V(i+1)\} \\ &= \max\{\lambda\beta V(i+1) + (1-\lambda)(v+\beta V(i)), \beta V(i+1)\} \end{aligned}$$

Assuming $V(i)$ is the left term in the above maximum, we may solve the equation for that term.

$$\begin{aligned} V(i) &= \lambda\beta V(i+1) + (1-\lambda)(v+\beta V(i)) \\ (1-(1-\lambda)\beta)V(i) &= \lambda\beta V(i+1) + (1-\lambda)v \\ V(i) &= \frac{\lambda\beta V(i+1) + (1-\lambda)v}{1-(1-\lambda)\beta} \end{aligned}$$

□

Theorem 2.1. *A phase transition occurs after the consumer makes i_0 visits to firm B, which evaluates to:*

$$\begin{aligned} i_0 &= k - \left\lfloor \log_\beta\left(\frac{v}{R(1-\beta)}\right)\right\rfloor \\ &\equiv k - \Delta \end{aligned}$$

(Note: we may need conditions as we did before - thing to check)

Proof. First we solve for the condition on $V(i+1)$ for us to choose firm B over A willingly.

$$\begin{aligned} \beta V(i+1) &> \frac{\lambda\beta V(i+1) + (1-\lambda)v}{1-(1-\lambda)\beta} \\ \iff \beta V(i+1)\left(1 - \frac{\lambda}{1-(1-\lambda)\beta}\right) &> \left(\frac{1-\lambda}{1-(1-\lambda)\beta}\right)v \\ \iff \beta V(i+1)\left(\frac{1-(1-\lambda)\beta-\lambda}{1-(1-\lambda)\beta}\right) &> \left(\frac{1-\lambda}{1-(1-\lambda)\beta}\right)v \\ \iff \beta V(i+1)\left(\frac{(1-\lambda)(1-\beta)}{1-(1-\lambda)\beta}\right) &> \left(\frac{1-\lambda}{1-(1-\lambda)\beta}\right)v \\ \iff \beta V(i+1) &> \frac{v}{1-\beta} \\ \iff V(i+1) &> \frac{v}{\beta(1-\beta)} \end{aligned}$$

Let i_0 be the minimum state i such that the above holds, so in particular $V(i_0) \leq \frac{v}{\beta(1-\beta)}$ but $V(i_0+1) > \frac{v}{\beta(1-\beta)}$. We know because V is increasing in i (still need to prove), this point is indeed a phase transition: $V(i) > \frac{v}{\beta(1-\beta)}$ for all $i > i_0$, so after this point, the customer always chooses firm B. We may compute $V(i_0)$ easily using this fact.

$$V(i_0) = \beta V(i_0+1) = \cdots = \beta^{k-i_0}V(k) = \beta^{k-i_0}R$$

Thus, we have the following:

$$\begin{aligned}
& \beta^{k-i_0} \leq \frac{v}{R\beta(1-\beta)} < \beta^{k-(i_0+1)} \\
\iff & k - i_0 \geq \log_\beta\left(\frac{v}{R\beta(1-\beta)}\right) > k-(i_0+1) \\
\iff & i_0 \leq k - \log_\beta\left(\frac{v}{R(1-\beta)}\right) + 1 < i_0 + 1 \\
\iff & i_0 = k - \left\lfloor \log_\beta\left(\frac{v}{R(1-\beta)}\right)\right\rfloor \equiv k - \Delta
\end{aligned}$$

□

3 Look-ahead, threshold distribution

Now we assume the look-ahead factor of a customer is drawn from some distribution $t \sim T$. The phase transition of the customer's DP will now depend on t.

$$i_0(t) = \begin{cases} i_0, & \text{if } t \geq \Delta. \\ k - t, & \text{otherwise.} \end{cases}$$

Assuming a customer look-ahead distribution and a fixed reward size, R, we want to choose a k to maximize the revenue per reward cycle. That is we want to maximize the quantity given by revenue of B during the k visits over the total number of purchases (at both A and B) to reach k visits at B. For simplicity we assume $p^+ = 1$, so the revenue of B per reward cycle is $k - R$. The expected total number of purchases per reward cycle is $\frac{i_0(t)}{\lambda} + (k - i_0(t))$, where the first term represents the expected number of visits needed to reach the phase transition with exogonous visits to B and the second term is just the remaining visits to B to receive the reward.

Note that we can think of the length of the reward cycle as a random variable: length of cycle $= \tau + k - i_0(t)$ where τ is a random variable representing the number of visits needed to hit the phase transition $E(\tau) = \frac{i_0(t)}{\lambda}$. Ideally, we like to maximize the following obje