Cost-Per-Click Pricing for Display Advertising

Sami Najafi-Asadolahi
Leavey School of Business, Santa Clara University, snajafi@scu.edu

Kristin Fridgeirsdottir
Managment Science and Operations, London Business School, kfridgeirsdottir@london.edu

---

Online Supplement

Proofs of Propositions

Proof of Proposition 2

We consider a Markov chain in which the state of the system is defined to be the vector $\mathbf{k} \triangleq (k_1, k_2, ..., k_n) = \sum_{q=1}^{Q} c_q \mathbf{v}_{c_q} \in (\mathbb{N} \cup \{0\})^n$ where $1 \leq Q \leq n$ is the number of slot groups for each of which the remaining number of clicks is the same. For example, if $\mathbf{k} = (1,1,2,2,4)$ then $Q = 3$ because the slots can be considered as three groups that each has the same number of clicks left. In addition, $c_q \geq 0$, $(1 \leq q \leq Q)$ refers to the number of clicks in each group. For example, for $\mathbf{k} = (1,1,2,2,4)$, $c_1 = 1, c_2 = 2$, and $c_3 = 4$. Furthermore, the vector $\mathbf{v}_{c_q} \in (\mathbb{N} \cup \{0\})^n$ is defined as $\mathbf{v}_{c_q} \triangleq \sum_{j \epsilon \mathcal{G}_{c_q}(\mathbf{k})} \mathbf{e}_j^T$ where the set $\mathcal{G}_{c_q}(\mathbf{k})$ is characterized as $\mathcal{G}_{c_q}(\mathbf{k}) \triangleq \{ j \,|\langle \mathbf{k}, \mathbf{e}_j \rangle = c_q \}$ in which $\langle \mathbf{k}, \mathbf{e}_j \rangle$ is the inner product of the two vectors $\mathbf{k}$ and $\mathbf{e}_j$, the jth unit vector. We need to identify all the possible states of the system and obtain the balance equations for every state. We note that each transition equation is a complex multidimensional difference equation for which there is no standard mathematical approach to solve, while in order to find the steady-state probabilities, we need to consider and solve all the transition equations in a single system. Therefore, we use the verification approach, identical to the mathematical induction, to show that the closed-form results hold. The symmetric CPC system has in general 10 distinct transition equations as follows:

i) For $\mathbf{k} = (0, ..., 0) = \mathbf{0}_{n \times 1}$ the flow balance is straightforward to obtain. $\mathbf{k}$ can either go to $\left(\mathbf{k} + \mathbf{x}\mathbf{e}_1^T\right)$ with rate λ or come from the state $\left(\mathbf{k} + \mathbf{e}_1^T\right)$ with rate $\widehat{\mu}$. As a result the flow balance equation becomes: $r\pi_{\mathbf{k}} = \pi_{\mathbf{k}+\mathbf{e}_1^T}$ where $r = \lambda / \widehat{\mu}$.

ii) If $\mathbf{k} = i\mathbf{e}_1^T$ with $1 \le i \le x-1$ then the state $\mathbf{k}$ can go either to sate $\left(\mathbf{k} - \mathbf{e}_1^T\right)$ with rate $\widehat{\mu}$ or to the state $\left(\mathbf{k} + \mathbf{x}\mathbf{e}_2^T\right)$ with rate λ. It also can either come from the state $\left(\mathbf{k} + \mathbf{e}_1^T\right)$ with rate $\widehat{\mu}$ or the state $\left(\mathbf{k} + \mathbf{e}_2^T\right)$ with the rate $\widehat{\mu}/2$. Thus, the balance equation becomes:

$$(1+r)\pi_{\mathbf{k}} = \pi_{\mathbf{k}+\mathbf{e}_1^T} + \frac{1}{2}\pi_{\mathbf{k}+\mathbf{e}_2^T}. \tag{O.1}$$

iii) If $\mathbf{k} = i\mathbf{v}_i$, with $\mathbf{v}_i = \sum_{j=1}^{k} \mathbf{e}_j^T$ where $1 \le k \le n-1$ and $1 \le i \le x-1$ then the state $\mathbf{k}$ makes a transition to state $\left(\mathbf{k} + \mathbf{x}\mathbf{e}_{k+1}^T\right)$ with rate λ. It also makes a transition to the state $\left(\mathbf{k} - \mathbf{e}_1^T\right)$ with rate $\widehat{\mu}$. In addition, the two states $\left(\mathbf{k} + \mathbf{e}_1^T\right)$ and $\left(\mathbf{k} + \mathbf{e}_{k+1}^T\right)$ make transitions to the state $\mathbf{k}$ with rates $\widehat{\mu}/k$ and $\left(\frac{\mathbf{1}_{\{i \neq 1\}} + (k+1)\mathbf{1}_{\{i=1\}}}{k+1}\right)\widehat{\mu}$, respectively. Therefore, the balance equation becomes:

$$(1+r)\pi_{\mathbf{k}} = \frac{1}{k}\pi_{\mathbf{k}+\mathbf{e}_1^T} + \left(\frac{\mathbf{1}_{\{i \neq 1\}} + (k+1)\mathbf{1}_{\{i=1\}}}{k+1}\right)\pi_{\mathbf{k}+\mathbf{e}_{k+1}^T}. \tag{O.2}$$

iv) If $\mathbf{k} = i\mathbf{v}_i$ with $\mathbf{v}_i = \sum_{j=1}^{n} \mathbf{e}_j^T$ and $1 \le i \le x-1$ then $\mathbf{k}$ can either come from the state $\left(\mathbf{k} + \mathbf{e}_1^T\right)$ with the rate $\widehat{\mu}/n$ or go to the state $\left(\mathbf{k} - \mathbf{e}_1^T\right)$ with the rate $\widehat{\mu}$. Therefore, the flow balance equation becomes: $\pi_{\mathbf{k}} = \frac{1}{n}\pi_{\mathbf{k}+\mathbf{e}_1^T}$. (Note that as we do not distinguish between the slots $\mathbf{k} + \mathbf{e}_1^T$,..., $\mathbf{k} + \mathbf{e}_n^T$ all refer to an identical system state, where for notational convenience, here we represent with $\mathbf{k} + \mathbf{e}_1^T$)

v) Define $\vartheta(\mathbf{k}, z) \triangleq \sum_{q=1}^{z} \left|\mathcal{G}_{c_q}(\mathbf{k})\right|$ for any $1 \le z \le Q$ with $\vartheta(\mathbf{k}, 0) = 0$. The the state $\mathbf{k} = \sum_{q=1}^{Q} c_q \mathbf{v}_{c_q}$ with $0 < c_1 < c_2 < ... < c_Q$, $1 \le c_q \le x-1$ where $\mathbf{v}_{c_q} = \sum_{u \epsilon \mathcal{G}_{c_q}(\mathbf{k})} \mathbf{e}_u^T$ and $1 \le \sum_{q=1}^{Q} \left|\mathcal{G}_{c_q}(\mathbf{k})\right| \le n-1$ goes to the state $\mathbf{k} + \mathbf{x}\mathbf{e}_{(\vartheta(\mathbf{k},Q)+1)}^T$ with the transition rate λ, and to one of the sates $\mathbf{k} - \mathbf{e}_{(\vartheta(\mathbf{k},z-1)+1)}^T$, for all $1 \le z \le Q$, with the transition rate $\frac{|\mathcal{G}_{c_z}(\mathbf{k})|}{\vartheta(\mathbf{k},Q)}\widehat{\mu}$. $\mathbf{k}$ also comes from either the state $\mathbf{k} + \mathbf{e}_{(\vartheta(\mathbf{k},Q)+1)}^T$ with rate $\phi_1\widehat{\mu}$ where $\phi_1 = \left(\frac{\mathbf{1}_{\{c_1>1\}} + (\left|\mathcal{G}_{c_1}(\mathbf{k})\right|+1)\mathbf{1}_{\{c_1=1\}}}{\vartheta(\mathbf{k},Q)+1}\right)$ or from one of the states $\mathbf{k} + \mathbf{e}_{(\vartheta(\mathbf{k},z-1)+1)}^T$ with all $1 \le z \le Q$, with rate $\phi_2(z)\widehat{\mu}$ where

$$\phi_2(z) = \left(\frac{\mathbf{1}_{\{c_{z+1}>c_z+1\}} + \mathbf{1}_{\{z=Q\}} + (\left|\mathcal{G}_{c_{z+1}}(\mathbf{k})\right|+1)\mathbf{1}_{\{c_{z+1}=c_z+1\}}}{\vartheta(\mathbf{k},Q)}\right).$$

As a result the flow balance equation becomes:

$$(1+r)\pi_{\mathbf{k}} = \phi_1\pi_{\mathbf{k}+\mathbf{e}_{(\vartheta(\mathbf{k},Q)+1)}^T} + \sum_{z=1}^{Q}\phi_2(z)\pi_{\mathbf{k}+\mathbf{e}_{(\vartheta(\mathbf{k},z-1)+1)}^T}. \tag{O.3}$$

vi) If $\mathbf{k} = \sum_{q=1}^{Q} c_q \mathbf{v}_{c_q}$ with $0 < c_1 < c_2 < ... < c_Q$, $1 \le c_q \le x-1$ and $\mathbf{v}_{c_q} = \sum_{u \epsilon \mathcal{G}_{c_q}(\mathbf{k})} \mathbf{e}_u^T$ and $|\mathcal{G}_0(\mathbf{k})| = 0$ the system goes to one the possible states $\mathbf{k} - \mathbf{e}_{(\vartheta(\mathbf{k},z-1)+1)}^T$ with all $1 \le z \le Q$ with the transition rate $\frac{|\mathcal{G}_{c_z}(\mathbf{k})|}{n}\widehat{\mu}$, and comes from one the possible states $\mathbf{k}+\mathbf{e}^T_{(\vartheta(\mathbf{k},z-1)+1)}$ with all $1\leq z\leq Q$ with the transition rate $\phi_3(z)\widehat{\mu}$ where

$$\phi_3(z)=\left(\frac{\mathbf{1}_{\{c_{z+1}>c_z+1\}}+\mathbf{1}_{\{z=Q\}}+(\left|\mathcal{G}_{c_{z+1}}(\mathbf{k})\right|+1)\mathbf{1}_{\{c_{z+1}=c_z+1\}}}{n}\right).$$

As a result the flow balance equation becomes: $\pi_{\mathbf{k}}=\sum_{z=1}^{Q}\phi_3(z)\pi_{\mathbf{k}+\mathbf{e}^T_{(\vartheta(\mathbf{k},z-1)+1)}}$. For example, take $\mathbf{k}=(1,1,2,2,3)$ then $c_1=1$, $c_2=2$, $c_3=3$, $|\mathcal{G}_{c_1}(\mathbf{k})|=2$, $|\mathcal{G}_{c_2}(\mathbf{k})|=2$, $|\mathcal{G}_{c_3}(\mathbf{k})|=1$. Moreover, $c_2=c_1+1$ and $c_3=c_2+1$. In addition, $\mathbf{k}+\mathbf{e}^T_{(\vartheta(\mathbf{k},0)+1)}=(2,1,2,2,3)$, $\mathbf{k}+\mathbf{e}^T_{(\vartheta(\mathbf{k},1)+1)}=(1,1,3,2,3)$, and $\mathbf{k}+\mathbf{e}^T_{(\vartheta(\mathbf{k},2)+1)}=(1,1,2,2,4)$. Thus, we have: $\pi_{\mathbf{k}}=\frac{3}{5}\pi_{\mathbf{k}+\mathbf{e}^T_{(\vartheta(\mathbf{k},0)+1)}}+\frac{2}{5}\pi_{\mathbf{k}+\mathbf{e}^T_{(\vartheta(\mathbf{k},1)+1)}}+\frac{1}{5}\pi_{\mathbf{k}+\mathbf{e}^T_{(\vartheta(\mathbf{k},2)+1)}}$.

vii) For $\mathbf{k}=x\mathbf{v}_x$, where $\mathbf{v}_x=\sum\limits_{u=1}^{|\mathcal{G}_x(\mathbf{k})|}\mathbf{e}^T_u$ and $|\mathcal{G}_x(\mathbf{k})|<n$, the flow balance equation becomes:

$$(1+r)\pi_{\mathbf{k}}=r\pi_{\mathbf{k}-x\mathbf{e}^T_{|\mathcal{G}_x(\mathbf{k})|}}+\left(\frac{\mathbf{1}_{\{x>1\}}+(|\mathcal{G}_x(\mathbf{k})|+1)\,\mathbf{1}_{\{x=1\}}}{|\mathcal{G}_x(\mathbf{k})|+1}\right)\pi_{\mathbf{k}+\mathbf{e}^T_{(|\mathcal{G}_x(\mathbf{k})|+1)}}.$$

viii) For $\mathbf{k}=x\mathbf{v}_x$, with $\mathbf{v}_x=\sum\limits_{u=1}^{n}\mathbf{e}^T_u$ the balance equation can be expressed as: $\pi_{\mathbf{k}}=r\pi_{\mathbf{k}-\mathbf{e}^T_j}$.

ix) If $\mathbf{k}=x\mathbf{v}_x+\sum\limits_{q=1}^{Q}c_q\mathbf{v}_{c_q}$ where $\mathbf{v}_x=\sum\limits_{u=1}^{|\mathcal{G}_x(\mathbf{k})|}\mathbf{e}^T_u$ and $\mathbf{v}_{c_q}=\sum\limits_{u=|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},q-1)}^{|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},q)}\mathbf{e}^T_u$, $|\mathcal{G}_0(\mathbf{k})|>0$ in which $\vartheta(\mathbf{k},0)=0$, and $c_{Q+1}=0$, the system goes to the state $\mathbf{k}+\mathbf{x}\mathbf{e}^T_{(|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},Q)+1)}$ with rate λ or to $\mathbf{k}-\mathbf{e}^T_1$ with rate $\left(\frac{|\mathcal{G}_x(\mathbf{k})|}{|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},Q)}\right)\widehat{\mu}$ or to one of the states $\mathbf{k}-\mathbf{e}^T_{(|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},z-1)+1)}$, $z=1,2,...,Q$ with rate $\left(\frac{\left|\mathcal{G}_{c_{z+1}}(\mathbf{k})\right|}{|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},Q)}\right)\widehat{\mu}$. In addition, the state $\mathbf{k}$ makes a transition either from the state $\mathbf{k}+\mathbf{e}^T_{(|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},Q)+1)}$ with rate $\phi_4\widehat{\mu}$ where $\phi_4=\left(\frac{1+\left|\mathcal{G}_{c_1}(\mathbf{k})\right|\mathbf{1}_{\{c_1=1\}}}{|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},Q)+1}\right)$ or from one of the states $\mathbf{k}+\mathbf{e}^T_{(|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},z-1)+1)}$, $1\leq z\leq Q$ with rate $\phi_5(z)\widehat{\mu}$ where

$$\phi_5(z)=\left(\frac{\mathbf{1}_{\{c_{z+1}>c_z+1\}}+\mathbf{1}_{\{z=Q\}}+(\left|\mathcal{G}_{c_{z+1}}(\mathbf{k})\right|+1)\mathbf{1}_{\{c_{z+1}=c_z+1\}}}{|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},Q)}\right).$$

Hence, the flow balance equation is expressed as:

$$(1+r)\pi_{\mathbf{k}}=r\pi_{\mathbf{k}-\mathbf{e}^T_1}+\phi_4\pi_{\mathbf{k}+\mathbf{e}^T_{(|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},Q)+1)}}+\sum_{z=1}^{Q}\phi_5(z)\widehat{\mu}\pi_{\mathbf{k}+\mathbf{e}^T_{(|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},z-1)+1)}}. \tag{O.4}$$

x) Finally, if $\mathbf{k}=x\mathbf{v}_x+\sum\limits_{q=1}^{Q}c_q\mathbf{v}_{c_q}$, with $\mathbf{v}_x=\sum\limits_{u=1}^{|\mathcal{G}_x(\mathbf{k})|}\mathbf{e}^T_u$, $\mathbf{v}_{c_s}=\sum\limits_{u=|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},q-1)}^{|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},q)}\mathbf{e}^T_u$, $|\mathcal{G}_0(\mathbf{k})|=0$ with an argument similar to those used before, we find the balance equation as:

$$\pi_{\mathbf{k}}=r\pi_{\mathbf{k}-\mathbf{e}^T_1}+\sum_{z=1}^{Q}\left(\frac{\mathbf{1}_{\{c_{z+1}>c_z+1\}}+\mathbf{1}_{\{z=Q\}}+(\left|\mathcal{G}_{c_{z+1}}(\mathbf{k})\right|+1)\mathbf{1}_{\{c_{z+1}=c_z+1\}}}{n}\right)\pi_{\mathbf{k}+\mathbf{e}^T_{(|\mathcal{G}_x(\mathbf{k})|+\vartheta(\mathbf{k},z-1)+1)}}. \tag{O.5}$$

Verifying the closed-form solution with items (i) to (iv), as well as items (vii) and $(viii)$ are immediate. Therefore, our main task is to show that the closed-form solution is true with items (v) to $(vii), (ix)$, and (x).We only verify the solution for item (v) as the rest items are verified similarly. We show item (v) in three stages:

I. First, we consider the case where $c_1 > 1$ and $c_{j+1} > c_j + 1$ for all $1 \leq j \leq Q$. Applying the closed-form solution to the left side of O.3, we establish that $(1+r)\pi_{\mathbf{k}} = \binom{\vartheta(\mathbf{k},Q)}{\left|\mathcal{G}_{c_1}(\mathbf{k})\right| \; \left|\mathcal{G}_{c_2}(\mathbf{k})\right| \ldots \left|\mathcal{G}_{c_Q}(\mathbf{k})\right|}(1+r)r^{\vartheta(\mathbf{k},Q)}$. Next, for convenience, we define $R_1(\mathbf{k}) \triangleq \left(\frac{\mathbf{1}_{[c_1>1]}+(\left|\mathcal{G}_{c_1}(\mathbf{k})\right|+1)\mathbf{1}_{\{c_1=1\}}}{\vartheta(\mathbf{k},Q)+1}\right)\pi_{\mathbf{k}+\mathbf{e}^T_{(\vartheta(\mathbf{k},Q)+1)}}$. Then for the first term in the right side, we get $R_1(\mathbf{k}) = \frac{1}{\vartheta(\mathbf{k},Q)+1}\binom{\vartheta(\mathbf{k},Q)+1}{\left|\mathcal{G}_{c_1}(\mathbf{k})\right| \; \left|\mathcal{G}_{c_2}(\mathbf{k})\right| \ldots \left|\mathcal{G}_{c_Q}(\mathbf{k})\right|}r^{k+1}$. Some simplification yields

$$R_1(\mathbf{k}) = \binom{\vartheta(\mathbf{k},Q)}{\left|\mathcal{G}_{c_1}(\mathbf{k})\right| \; \left|\mathcal{G}_{c_2}(\mathbf{k})\right| \ldots \left|\mathcal{G}_{c_Q}(\mathbf{k})\right|}r^{k+1} = \frac{\vartheta(\mathbf{k},Q)!}{\prod_{q=1}^{Q}\left|\mathcal{G}_{c_q}(\mathbf{k})\right|!}r^{k+1}. \tag{O.6}$$

For the second term in the right side we first define

$$R_2(\mathbf{k},\mathbf{z}) \triangleq \left(\frac{\mathbf{1}_{\{c_{z+1}>c_z+1\}} + \mathbf{1}_{\{z=Q\}} + (\left|\mathcal{G}_{c_{z+1}}(\mathbf{k})\right|+1)\mathbf{1}_{\{c_{z+1}=c_z+1\}}}{\vartheta(\mathbf{k},Q)}\right)\pi_{\mathbf{k}+\mathbf{e}^T_{(\vartheta(\mathbf{k},z-1)+1)}}. \tag{O.7}$$

Next, after applying the assumptions in $R_2(\mathbf{k},\mathbf{z})$ and some algebraic manipulation we find

$$R_2(\mathbf{k},\mathbf{z}) = \left|\mathcal{G}_{c_z}(\mathbf{k})\right| \frac{(\vartheta(\mathbf{k},Q)-1)!}{\prod_{q=1}^{Q}\left|\mathcal{G}_{c_q}(\mathbf{k})\right|!}r^{k}. \tag{O.8}$$

Considering (O.6) and (O.8) , it is easy to show that the whole right hand side of (O.3), i.e., $R.H.S = R_1(\mathbf{k}) + \sum_{z=1}^{Q} R_2(\mathbf{k},\mathbf{z})$, reduces to

$$R.H.S = \frac{\vartheta(\mathbf{k},Q)!}{\prod_{q=1}^{Q}\left|\mathcal{G}_{c_q}(\mathbf{k})\right|!}r^{k+1} + \sum_{z=1}^{Q}\left|\mathcal{G}_{c_z}(\mathbf{k})\right| \frac{(\vartheta(\mathbf{k},Q)-1)!}{\prod_{q=1}^{Q}\left|\mathcal{G}_{c_q}(\mathbf{k})\right|!}r^{k}. \tag{O.9}$$

After some simplifications, knowing that $\sum_{z=1}^{Q}\left|\mathcal{G}_{c_z}(\mathbf{k})\right| = \vartheta(\mathbf{k},Q)$, we find $R.H.S = (1+r)\pi_k$.

II. If $c_1 = 1$ then it is easy to see that $\left(\frac{\mathbf{1}_{\{c_1>1\}}+(\left|\mathcal{G}_{c_1}(\mathbf{k})\right|+1)\mathbf{1}_{\{c_1=1\}}}{\vartheta(\mathbf{k},Q)+1}\right) = \left(\frac{1+\left|\mathcal{G}_{c_1}(\mathbf{k})\right|}{\vartheta(\mathbf{k},Q)+1}\right)$. Hence, after some simplification $R_1(\mathbf{k})$ can be expressed as $R_1(\mathbf{k}) = \vartheta(\mathbf{k},Q)!r^{k+1}/\prod_{q=1}^{Q}\left|\mathcal{G}_{c_q}(\mathbf{k})\right|!$, which is identical to (O.6). Hence, from this point on the rest of the steps are the same.

III. Finally, if for any arbitrary $1 \leq z \leq r-1$, $c_{z+1} = c_z + 1$, then we find

$$R_2(\mathbf{k},\mathbf{z}) = \left(\frac{1+\left|\mathcal{G}_{c_{z+1}}(\mathbf{k})\right|}{\vartheta(\mathbf{k},Q)}\right)\left(\frac{\vartheta(\mathbf{k},Q)}{\left|\mathcal{G}_{c_1}(\mathbf{k})\right|\ \left|\mathcal{G}_{c_2}(\mathbf{k})\right|\dots\left(\left|\mathcal{G}_{c_z}(\mathbf{k})\right|-1\right)\ \left(\left|\mathcal{G}_{c_{z+1}}(\mathbf{k})\right|+1\right)\dots\left|\mathcal{G}_{c_Q}(\mathbf{k})\right|}\right)r^k. \tag{O.10}$$

After some manipulation (O.10) reduces to $R_2(\mathbf{k},\mathbf{z}) = \left|\mathcal{G}_{c_z}(\mathbf{k})\right|(\vartheta(\mathbf{k},Q)-1)!r^k/\prod_{q=1}^{Q}\left|\mathcal{G}_{c_q}(\mathbf{k})\right|!$, which is the same result as in (O.8). As a result, all the steps for verification will be identical afterwards. Hence, the proof becomes complete. ■

Lemma 1 *For any* $n \in \mathbb{N}$*, and* $\rho \in \mathbb{R}^+$, $n-(n+1)\rho+\rho^{n+1} \geq 0$.

Proof The proof is with induction. For $n=1$ the verification is immediate. Next, we assume that for $n=k$, $k-(k+1)\rho+\rho^{k+1} \geq 0$ holds. Then, for $n=k+1$ we have $k+1-(k+2)\rho+\rho^{k+2} = (k-(k+1)\rho+\rho^{k+1})+(\rho-1)(\rho^{k+1}-1)$. Due to the induction assumption, the first term is always positive. For the second term we consider that if $\rho \geq 1$ then $\rho^{k+1} \geq 1$. Hence $(\rho-1)(\rho^{k+1}-1) \geq 0$. Furthermore, if $\rho \leq 1$ then $\rho^{k+1} \leq 1$. Hence, $(\rho-1)(\rho^{k+1}-1) \geq 0$ and the result follows. ■

Proof of Proposition 3

(i) Showing $\partial\mathbb{P}_n/\partial r \geq 0$ is the same as showing $\partial\mathbb{P}_n/\partial\rho \geq 0$ in which $\rho = rx$. From Proposition 2 we have $\mathbb{P}_n = \rho^n(1-\rho)/(1-\rho^{n+1})$. Taking the derivative of $\mathbb{P}_n$ with respect to ρ, we find $\partial\mathbb{P}_n/\partial\rho = \rho^{n-1}(n-(n+1)\rho+\rho^{n+1})/(1-\rho^{n+1})^2$. By Lemma 1 the numerator is positive and the proof follows.

(ii) $\mathbb{P}_n(x+1)-\mathbb{P}_n(x)$ can be simplified as $\mathbb{P}_n(x+1)-\mathbb{P}_n(x) = \sum_{i=0}^{n} r^{n+i}(x^i(x+1)^n - x^n(x+1)^i)/(\sum_{i=0}^{n}(r(x+1))^i\sum_{i=0}^{n}(rx)^i)$. The right hand side is always positive if and only if $x^i(x+1)^n - x^n(x+1)^i \geq 0$. However, this identity always holds since it is simplified to: $x/(x+1) \leq 1$, which always holds. Thus, $P_n(x+1)-P_n(x) \geq 0$.

(iii) Some manipulation yields $\mathbb{P}_{n+1}(x)-\mathbb{P}_n(x) = -(rx)^n(rx-1)^2/(((rx)^{n+1}-1)((rx)^{n+2}-1)) \leq 0$, which is always negative because $((rx)^{n+1}-1)((rx)^{n+2}-1) \geq 0$ and the result follows. ■

Proof of Proposition 4

(i) To show L is increasing in x we note that $L(x+1)-L(x) = \sum_{i=0}^{n} i\left(\mathbb{P}_i(x+1)-\mathbb{P}_i(x)\right)$. After some manipulations it can be shown that $L(x+1)-L(x) = (\sum_{t=0}^{n} G_1(x,t)r^t + \sum_{t=n+1}^{2n} G_2(x,n)r^t)/(\sum_{j=0}^{n} r^j(x+1)^j\sum_{j=0}^{n} r^j x^j)$, where $G_1(x,t) \triangleq \sum_{i=0}^{t}(x+1)^t i((\frac{x}{x+1})^{t-i} - (\frac{x}{x+1})^i)$ and $G_2(x,t) \triangleq \sum_{i=t-n}^{n}(x+1)^t i((\frac{x}{x+1})^{t-i} - (\frac{x}{x+1})^i)$. We separately show that $G_1(x,t)$ and $G_2(x,t)$ are always positive for $t=2k$ and for $t=2k+1$.

(i-1) For $t=2k$, $G_1(x,t)$ reduces to $G_1(x,2k)=(x+1)^{2k}r^{2k}\sum_{i=0}^{k}i(2k-2i)((\frac{x}{x+1})^i-(\frac{x}{x+1})^{2k-i})$, where $(\frac{x}{x+1})^i-(\frac{x}{x+1})^{2k-i}\geq 0$ since $2k-i\geq i$. Hence, the result follows. With a similar argument for $n=2k$, we find $G_2(x,2k)=(x+1)^{2k}r^{2k}\sum_{i=2k-n}^{k}(2k-2i)((\frac{x}{x+1})^i-(\frac{x}{x+1})^n)$, where it is easy to see that $(\frac{x}{x+1})^i-(\frac{x}{x+1})^n\geq 0$. Hence the proof follows.

(i-2) For $t=2k+1$, $G_1(x,t)$ reduces to $G_1(x,2k+1)=(x+1)^{2k+1}r^{2k+1}\sum_{i=0}^{k-1}(2k+1-2i)((\frac{x}{x+1})^i-(\frac{x}{x+1})^{2k+1-i})+k((\frac{x}{x+1})^k-(\frac{x}{x+1})^{k+1})$. It is easy to show that $(\frac{x}{x+1})^i-(\frac{x}{x+1})^{2k+1-i}\geq 0$, since $2k+1-i\geq 0$. In addition, It is clear to see that $(\frac{x}{x+1})^k-(\frac{x}{x+1})^{k+1}\geq 0$ always holds. Hence, $G_1(x,2k+1)\geq 0$. In a similar way, $G_2(x,2k+1)$ can be expressed as $G_2(x,2k+1)=r^{2k+1}(x+1)^{2k+1}\sum_{i=2k-n}^{k-1}i(2k+1-2i)((\frac{x}{x+1})^i-(\frac{x}{x+1})^n)+k((\frac{x}{x+1})^k-(\frac{x}{x+1})^n)$, which is always positive.

(ii) In order to prove L's convexity in x, we note that convexity only holds for $rx>1$. Convexity of L on a real continuum of x implies its convexity over discrete values as well. Thus, assuming x a real variable and after taking twice differentiation of L and making some simplifications, we find that $\partial^2 L/\partial x^2=\sum_{i=0}^{n}\sum_{j=0}^{n}\sum_{k=0}^{n}r^{i+j+k}x^{i+j+k-2}(i^2-ij)\,(i+j-1-2k)\,/(\sum_{j=0}^{n}(rx)^j)^3$, where it is easy to show that for $rx>1$ the numerator is always negative. Hence the proof follows.

(iii) To show L is increasing in r, we note that $L=\sum_{i=0}^{n}i(rx)^i/\sum_{j=0}^{n}(rx)^j$. Taking the derivative of L we obtain $\partial L/\partial r=\left(\sum_{i=1}^{n}\sum_{j=1}^{n}(rx)^{i+j-1}(i^2-ij)\right)/(\sum_{j=0}^{n}(rx)^j)^2$. We note that $\partial L/\partial r\geq 0$ if and only if $\sum_{i=1}^{n}\sum_{j=1}^{n}(rx)^{i+j-1}(i^2-ij)\geq 0$. After applying the new indexing $t=i+j$ and simplifying, we find $\sum_{i=1}^{n}\sum_{j=1}^{n}(rx)^{i+j-1}(i^2-ij)=\sum_{t=2}^{n}\frac{1}{6}t(t+1)(t+2)(rx)^{t-1}+\sum_{t=n+1}^{2n}\frac{1}{6}(2n-t)(2n+1-t)(2n+2-t)(rx)^{t-1}$, which is always positive, and the proof is complete. An argument similar to the one used in part (i) shows L's convexity in r, which is omitted.

(iv) In order to show L is increasing in n, we need to show that $L_{n+1}(x)-L_n(x)=H(\rho)/G(\rho)$, $\rho=rx$ where $H(\rho)=1+n-2\rho-n\rho+\rho^{n+2}$ and $G(\rho)=1-\rho^{n+1}+\rho^{2n+3}-\rho^{n+2}$. To show the positivity, we show that for all $\rho\geq 0$, $H(\rho)\geq 0$ and $G(\rho)\geq 0$. After some easy differentiations, we find $H'_\rho=-(n+2)+\rho^{n+1}(n+2)$ and $H''_\rho=(n+1)(n+2)\rho^n\geq 0$. From H'_ρ and H''_ρ, it is easy to see that $H(\rho)$ is convex with a global minimum at $\rho=1$. Hence, for all $\rho\geq 0$, $H(\rho)\geq H(1)=0$. Using a similar approach it is easy to verify that $G(\rho)$ has a unique global minimum at $\rho=1$. Hence, $G(\rho)\geq G(1)=0$, and the proof follows. ■

Proof of Proposition 5

(i) The first derivative of $R(\lambda)$ with respect to λ is $\partial R(\lambda)/\partial\lambda=(\partial\Gamma(\lambda)/\partial\lambda)p(\lambda)+(\partial p(\lambda)/\partial\lambda)\Gamma(\lambda)$. Thus, the second derivative is expressed as $\partial^2 R(\lambda)/\partial\lambda^2=(\partial^2\Gamma(\lambda)/\partial\lambda^2)p(\lambda)+$

$2(\partial \Gamma(\lambda)/\partial\lambda)(\partial p(\lambda)/\partial\lambda) + (\partial^2 p(\lambda)/\partial\lambda^2)\Gamma(\lambda)$. To show $\partial^2 R(\lambda)/\partial\lambda^2 \leq 0$, as $\partial p(\lambda)/\partial\lambda \leq 0$ and $\partial^2 p(\lambda)/\partial\lambda^2 \leq 0$, we need only to show $\Gamma(\lambda)$ is increasing concave in λ. That is, $\partial\Gamma(\lambda)/\partial\lambda \geq 0$ and $\partial^2\Gamma(\lambda)/\partial\lambda^2 \leq 0$. By definition, we have $\Gamma = \lambda x(1-\mathbb{P}_n(\lambda)) = \widehat{\mu} r x(1-\mathbb{P}_n(\lambda))$. In order to show the result, since Γ is a function of λ only through r and $\widehat{\mu}$ is constant, we take the derivative of Γ with respect to r. As $\widehat{\mu}$ is constant, for convenience, we scale $\widehat{\mu}$ to $\widehat{\mu}=1$. Then, based on Lemma 3, showing $\partial\Gamma/\partial r \geq 0$ and $\partial^2\Gamma/\partial r^2 \leq 0$ is immediate. Hence, the result follows.

(ii) The proof of part (ii) is immediate from $\partial R(\lambda)/\partial\lambda$ in item (i). ■

Lemma 2 *For any* $n \in \mathbb{N}$, $r \in \mathbb{R}$, $F(x) \triangleq \sum_{i=0}^{n}\sum_{j=0}^{n}(rx)^i(n-i+1)(n+i-2j) \geq 0$.

Proof Some rearrangement gives $F(x) = \sum_{i=0}^{n}(rx)^i[\sum_{j=0}^{n}(n-i+1)(n+i-2j)]$. Next, It is easy to show $\sum_{j=0}^{n}(n-i+1)(n+i-2j) = i(n+1)(n+1-i) \geq 0$. Hence, the proof follows. ■

Lemma 3 *For any* $x \in \mathbb{N}$, $n \in \mathbb{N}$, $\mu, \beta \in \mathbb{R}_+$, $\Gamma = \widehat{\mu} r x(1-\mathbb{P}_n(r))$ *is increasing concave in* r.

Proof Taking the first derivative of $\Gamma = \widehat{\mu}\sum_{i=1}^{n}(rx)^i/\sum_{j=0}^{n}(rx)^j$ and simplifying yields $\partial\Gamma/\partial r = \widehat{\mu}\sum_{i=1}^{n} i(rx)^i/(r(\sum_{j=0}^{n}(rx)^j)^2) \geq 0$. In order to show the convexity, taking the second derivative of Γ with respect to r and simplifying yields $\partial^2\Gamma/\partial r^2 = \widehat{\mu}\sum_{i=1}^{n}\sum_{j=0}^{n}(rx)^{i+j} i(i-1-2j)/(r^2(\sum_{j=0}^{n}(rx)^j)^3)$. After the reindexing $t = i+j$, $\partial^2\Gamma/\partial r^2$ can be expressed as

$$\frac{\partial^2\Gamma(r)}{\partial r^2} = \widehat{\mu}\frac{\sum_{t=1}^{n}(rx)^t\sum_{i=1}^{t} i(3i-2t-1) + \sum_{t=n+1}^{2n}(rx)^t\sum_{i=t-n}^{n} i(3i-2t-1)}{r^2\left(\sum_{j=0}^{n}(rx)^j\right)^3} \leq 0, \tag{O.11}$$

which proves the proposition. To see the reason that $(O.11)$ holds, it is easy to show $\sum_{i=1}^{t} i(3i-2t-1) = 0$, $1 \leq t \leq n$, and $\sum_{i=t-n}^{n} i(3i-2t-1) = (n+1)(2n+1-t)(n-t) \leq 0$; $n+1 \leq t \leq 2n$. Hence the result follows. ■

Proof of Proposition 6

Taking the first derivative using the chain rule, we find $dR(\lambda^*)/dn = \partial R(\lambda^*)/\partial n + \partial R(\lambda^*)/\partial\lambda \times \partial\lambda^*/\partial n$. However, as λ^* satisfies $\partial R(\lambda^*)/\partial\lambda = 0$, we find $dR(\lambda^*)/dn = \partial R(\lambda^*)/\partial n$ (*Envelope Theorem).* Next, we get $\partial R(\lambda^*)/\partial n = 1/\widehat{\mu} \times [\partial\Gamma/\partial n \times p(\lambda^*, n, x) + \Gamma \times \partial p(\lambda^*, n, x)/\partial n]$. We show that $\partial R(\lambda^*)/\partial n = 0$ has a single root. To see the reason, we note that $\partial\Gamma/\partial n > 0$ and $\partial^2\Gamma/\partial n^2 < 0$ (see Lemma 4), $\partial p(\lambda^*, n, x)/\partial n \leq 0$, $\partial^2 p(\lambda^*, n, x)/\partial n \leq 0$. Hence, $\partial^2 R(\lambda^*)/\partial n^2 < 0$, i.e., $\partial R(\lambda^*)/\partial n$ is strictly decreasing in n. Therefore, it crosses the zero line at a single point, namely, n^*. If $n^* \geq 0$ then for all $n \leq n^*$, $R(\lambda^*)$ is increasing in n and for all $n \geq n^*$ it is decreasing in n. If $n^* < 0$ then for all $n \geq 0$, $R(\lambda^*)$ is decreasing in n. ■

Lemma 4 *Given* $x\in\mathbb{N}$, $n\in\mathbb{N}$, $\lambda\in\mathbb{R}_+$, $\Gamma=rx(1-\mathbb{P}_n(r))$ *is increasing concave in* x, *and* n.

Proof In order to show $\Gamma_x\le\Gamma_{x+1}$, it is sufficient to show $\partial\Gamma_x/\partial x\ge 0$, since if Γ_x is increasing in real-valued x, it is clearly increasing in discrete-valued x, which implies the monotonicity result we look for. However, $\partial\Gamma_x/\partial x=\frac{r[1-(rx)^n+n(rx)^n(rx-1)]}{((rx)^{n+1}-1)^2}\ge 0$. Hence, the result follows. Similarly, to show Γ's concavity in x, it is enough to show that Γ_x is concave in real-valued x, i.e., $\partial^2\Gamma_x/\partial x^2\le 0$. However, this result is clear because the second derivative can be expressed as $\partial^2\Gamma_x/\partial x^2=-\frac{(1+n)r(rx)^n\left(-2rx(-1+(rx)^n)+n(-1+rx)\left(1+(rx)^{1+n}\right)\right)}{x\left(-1+(rx)^{1+n}\right)^3}$, which is negative based on Lemma 5.

In order to show Γ's monotonicity and concavity in n, as in the previous part, it is easier to show that $\partial\Gamma_n/\partial n\ge 0$, $\partial^2\Gamma_n/\partial n^2\le 0$, since if the two properties hold in the continuos setting, they will hold in the discrete setting as well. It can be seen that $\partial\Gamma/\partial n=\frac{(rx)^{n+1}(rx-1)\ln(rx)}{((rx)^{n+1}-1)^2}>0$ and $\partial^2\Gamma/\partial n^2=-\frac{(rx)^{n+1}(rx-1)(1+(rx)^{n+1})\ln[(rx)]^2}{((rx)^{n+1}-1)^3}<0$, which proves the result. ■

Lemma 5 *Let* $h(\rho)=-2\rho\left(-1+\rho^n\right)+n(-1+\rho)\left(1+\rho^{1+n}\right)$. *Then* $sign(h(\rho))=sign(\rho)$.

Proof It is easy to see that $h^{'}(\rho)=(2+n)\left(1-(1+n)\rho^n+n\rho^{1+n}\right)$ and $h^{''}(\rho)=n(n+1)(n+2)(\rho-1)\rho^{n-1}$. Hence, it can be seen that $h(1)=h^{'}(1)=h^{''}(1)=0$ and $h^{''}(\rho)>0$ for $\rho>1$ and $h^{''}(\rho)<0$ for $\rho<1$ and $h^{'}(\rho)\ge 0$ for all ρ. Hence, $h(\rho)$ is an increasing function with the inflection point at $\rho=1$. Thus, $h(\rho)>0$ if $\rho>1$ and $h(\rho)<0$ if $\rho<1$. ■

Proof of Proposition 7

(i) We need to show that $\frac{\partial\lambda^*}{\partial x}\le 0$. By Implicit Function Theorem, we have

$$\frac{\partial\lambda^*}{\partial x}=-\frac{\frac{\partial^2R(\lambda^*)}{\partial\lambda\partial x}}{\frac{\partial^2R(\lambda^*)}{\partial\lambda^2}}=-\frac{\frac{\partial^2\Gamma(\lambda^*)}{\partial\lambda\partial x}p(\lambda^*,x,n)+\frac{\partial\Gamma(\lambda^*)}{\partial x}\frac{\partial p(\lambda^*,x,n)}{\partial\lambda}+\Gamma(\lambda^*)\frac{\partial^2p(\lambda^*,x,n)}{\partial x\partial\lambda}+\frac{\partial\Gamma(\lambda^*)}{\partial\lambda}\frac{\partial p(\lambda^*,x,n)}{\partial x}}{\frac{\partial^2\Gamma(\lambda^*)}{\partial\lambda^2}p(\lambda^*,x,n)+2\frac{\partial\Gamma(\lambda^*)}{\partial\lambda}\frac{\partial p(\lambda^*,x,n)}{\partial\lambda}+\Gamma(\lambda^*)\frac{\partial^2p(\lambda^*,x,n)}{\partial\lambda^2}},\tag{O.12}$$

in which by First Order Necessary Condition we have $\frac{\partial R(\lambda^*)}{\partial\lambda}=\frac{\partial\Gamma(\lambda^*)}{\partial\lambda}p(\lambda^*,x,n)+\Gamma(\lambda^*)\frac{\partial p(\lambda^*,x,n)}{\partial\lambda}=0$. Note that since x is discrete we are slightly abusing the Implicit Function Theorem. Consider x to be continuous rather than discrete. It is clear that if λ^* is increasing in real-valued x, it is increasing in discrete values of x. Similarly, if $\Gamma(\lambda^*)$ and $\frac{\partial\Gamma(\lambda^*)}{\partial\lambda}$ are increasing (/ decreasing) in any increasing sequence of real values x then the monotonicity holds for any increasing sequence of integer values x. As λ^* is a maximizer of a concave function we have $\frac{\partial^2R(\lambda^*)}{\partial\lambda^2}\le 0$ (Proposition 5), i.e., the denominator is negative. In addition, since $\frac{\partial p(\lambda^*,x,n)}{\partial x}\le 0$ and $\frac{\partial^2p(\lambda^*,x,n)}{\partial x\partial\lambda}\le 0$, we are left with showing $\frac{\partial^2\Gamma(\lambda^*)}{\partial\lambda\partial x}p(\lambda^*,x,n)+\frac{\partial\Gamma(\lambda^*)}{\partial x}\frac{\partial p(\lambda^*,x,n)}{\partial\lambda}\le 0$. Using the *FONC*, $\frac{\partial\Gamma(\lambda^*)}{\partial\lambda}p(\lambda^*,x,n)+$

$\Gamma(\lambda^*)\frac{\partial p(\lambda^*,x,n)}{\partial\lambda}=0$, we need to show $g(\lambda^*)=\frac{\partial\Gamma(\lambda^*)}{\partial x}\frac{\partial\Gamma(\lambda^*)}{\partial\lambda}-\frac{\partial^2\Gamma(\lambda^*)}{\partial\lambda\partial x}\Gamma(\lambda^*)\geq 0$. Without loss of generality, we set $\widehat{\mu}=1$ and thus $\lambda^*=r$. Now, we have $\Gamma(r)=\sum_{i=1}^n(rx)^i/\sum_{j=0}^n(rx)^j$. After some simplification, we find $\partial\Gamma(r)/\partial x=\sum_{i=1}^n i(rx)^i/(x(\sum_{j=0}^n(rx)^j)^2)$. Hence, $(\partial\Gamma(r)/\partial x)(\partial\Gamma(r)/\partial r)=\sum_{i=1}^n\sum_{j=1}^n(rx)^{i+j-1}ij/\left(\sum_{j=0}^n(rx)^j\right)^4$. Moreover, $\frac{\partial^2\Gamma(\lambda^*)}{\partial\lambda\partial x}=\sum_{i=1}^n\sum_{j=0}^n(rx)^{i+j-1}(i^2-2ij)/(\sum_{j=0}^n(rx)^j)^3$. Hence, $\frac{\partial^2\Gamma(\lambda^*)}{\partial\lambda\partial x}\Gamma(\lambda^*)$ can be re-expressed as $\frac{\partial^2\Gamma(\lambda^*)}{\partial\lambda\partial x}\Gamma(\lambda^*)=\sum_{i=1}^n\sum_{j=0}^n\sum_{k=1}^n(rx)^{i+j+k-1}(i^2-2ij)/(\sum_{j=0}^n(rx)^j)^4$. In order to show $g(\lambda^*)\geq 0$ it is enough to show

$$P(x,r,n)=\sum_{i=1}^n\sum_{j=0}^n(rx)^{i+j-1}ij-\sum_{i=1}^n\sum_{j=0}^n\sum_{k=1}^n(rx)^{i+j+k-1}(i^2-2ij)\geq 0.$$

Note that we can represent $P(x,r,n)$ with the following form $P(x,r,n)=\sum_{t=1}^{3n}c_t(rx)^{t-1}$, in which c_t is the appropriate coefficient of the term $(rx)^{t-1}$. We show that $P(x,r,n)\geq 0$ by showing $c_t\geq 0,\ 2\leq t\leq 3n$. We consider the three intervals $2\leq t\leq n$, $n+1\leq t\leq 2n$, and $2n+1\leq t\leq 3n$. We need to show $c_t\geq 0$ holds in each of these intervals separately. Since the procedures of proofs for the three intervals are similar we only show for $2\leq t\leq n$. For every $t\in\{2,...,n\}$ the coefficient of the term $(rx)^{t-1}$ in $\sum_{i=1}^n\sum_{j=0}^n(rx)^{i+j-1}ij$ becomes $\sum_{i=1}^t i(t-i)$, and the coefficient of $(rx)^{t-1}$ in the second sum, $\sum_{i=1}^n\sum_{j=0}^n\sum_{k=1}^n(rx)^{i+j+k-1}(i^2-2ij)$, becomes $\sum_{i=1}^t\sum_{k=1}^{t-i}(i^2-2i(t-i-k))$. Thus, we find $c_t=\sum_{i=1}^t i(t-i)-\sum_{i=1}^t\sum_{k=1}^{t-i}(i^2-2i(t-i-k))=0$. This is true because $\sum_{i=1}^t i(t-i)=\sum_{i=1}^t\sum_{k=1}^{t-i}(i^2-2i(t-i-k))=\frac{1}{6}n(n+1)(n+2)$. Using the same procedure, it can be checked that $c_t\geq 0$ holds for other intervals as well.

(ii) The proof is immediate from part (i) and that $\frac{dp(\lambda^*,x,n)}{dx}=\frac{\partial p(\lambda^*,x,n)}{\partial\lambda}\frac{\partial\lambda^*}{\partial x}+\frac{\partial p(\lambda^*,x,n)}{\partial x}$. ■

Lemma 6 *Given* $x\in\mathbb{N}$, $r\in\mathbb{R}_+$, $\Gamma=rx(1-\mathbb{P}_n(r))$, *if* $rx\geq\exp(1)$ *then* $\frac{\partial^2\Gamma(\lambda^*)}{\partial\lambda\partial n}\leq 0$.

Proof Using the formulation of Γ and some simplification gives:

$$\frac{\partial^2\Gamma(\lambda^*,x,n)}{\partial\lambda\partial n}=\frac{(rx)^n(B+A\ln(rx))}{((rx)^{n+1}-1)^3},$$

where $A=(1-rx)(1+\sum_{i=1}^n[(rx)^{n+1}-(rx)^i]+n)$ and $B=(1-rx)(1-(rx)^{n+1})>0$. Since $rx>1$, it is clear to see that $A<0$. Thus, $B+\ln(rx)A<B+A$. On the other hand, it can be shown that $B+A=(1-rx)(1+(rx)^2\sum_{i=1}^{n-2}[(rx)^{n+1}-(rx)^i]+(rx)((rx)-1)^2+n+1)<0$. Hence, $\frac{\partial^2\Gamma(\lambda^*,x,n)}{\partial\lambda\partial n}\leq\frac{(rx)^n(B+A\ln(rx))}{((rx)^{n+1}-1)^3}\leq\frac{(rx)^n(B+A)}{((rx)^{n+1}-1)^3}<0$. Hence, the result follows. ■

Proofs of Proposition 8

(i) We need to show that $\frac{\partial\lambda^*}{\partial n}\leq 0$. By Implicit Function Theorem, we have

$$\frac{\partial\lambda^*}{\partial n} = -\frac{\frac{\partial^2 R(\lambda^*)}{\partial\lambda\partial n}}{\frac{\partial^2 R(\lambda^*)}{\partial\lambda^2}} = -\frac{\frac{\partial^2 \Gamma(\lambda^*)}{\partial\lambda\partial n}p(\lambda^*,x,n) + \frac{\partial\Gamma(\lambda^*)}{\partial n}\frac{\partial p(\lambda^*,x,n)}{\partial\lambda} + \Gamma(\lambda^*)\frac{\partial^2 p(\lambda^*,x,n)}{\partial n\partial\lambda} + \frac{\partial\Gamma(\lambda^*)}{\partial\lambda}\frac{\partial p(\lambda^*,x,n)}{\partial n}}{\frac{\partial^2\Gamma(\lambda^*)}{\partial\lambda^2}p(\lambda^*,x,n) + 2\frac{\partial\Gamma(\lambda^*)}{\partial\lambda}\frac{\partial p(\lambda^*,x,n)}{\partial\lambda} + \Gamma(\lambda^*)\frac{\partial^2 p(\lambda^*,x,n)}{\partial\lambda^2}},$$

in which by First Order Necessary Condition we have $\frac{\partial R(\lambda^*)}{\partial\lambda} = \frac{\partial\Gamma(\lambda^*)}{\partial\lambda}p(\lambda^*,x,n) + \Gamma(\lambda^*)\frac{\partial p(\lambda^*,x,n)}{\partial\lambda} = 0$. Based on Proposition 5, it is clear that $\frac{\partial^2 R(\lambda^*)}{\partial\lambda^2} \le 0$. Based on the proposition assumptions and Lemma 4, it is sufficient to show that $\frac{\partial^2\Gamma(\lambda^*)}{\partial\lambda\partial n} \le 0$. However, for $rx = (\lambda^*/\widehat{\mu})x \ge \exp(1)$, this holds based on Lemma 6. Thus, for $rx = (\lambda^*/\widehat{\mu})x \ge \exp(1)$, $\frac{\partial\lambda^*}{\partial n} < 0$. In addition, it can be shown that when $rx = 1$, i.e., $\lambda^* \to \widehat{\mu}/x$, then $\frac{\partial^2\Gamma(\lambda^*)}{\partial\lambda\partial n} \to \frac{1}{2(n+1)^2} > 0$, $\frac{\partial\Gamma(\lambda^*)}{\partial n} \to \frac{1}{(n+1)^2} > 0$ and $\frac{\partial\Gamma(\lambda^*)}{\partial\lambda} \to \frac{n}{2(n+1)} > 0$. Hence, $\frac{\partial\lambda^*}{\partial n} > 0$ in the neighborhood of $\lambda^* = \widehat{\mu}/x$. Therefore, by Rolle's theorem there exists a point $w \in [\widehat{\mu}/x, \exp(1)\widehat{\mu}/x]$ such that for $\lambda^* = w$, $\frac{\partial\lambda^*}{\partial n} = 0$; for $\lambda^* < w$, $\frac{\partial\lambda^*}{\partial n} > 0$; and for $\lambda^* > w$, $\frac{\partial\lambda^*}{\partial n} < 0$.

(ii) By Implicit Function Theorem:

$$\frac{dp(\lambda^*,x,n)}{dn} = \underbrace{\frac{\partial p(\lambda^*,x,n)}{\partial n}}_{<0} + \underbrace{\frac{\partial p(\lambda^*,x,n)}{\partial\lambda}}_{<0}\frac{\partial\lambda^*}{\partial n}.$$

It is clear that when $\frac{\partial\lambda^*}{\partial n} > 0$ then $\frac{dp(\lambda^*,x,n)}{dn} < 0$. That is, if the publisher decides to attract more advertisers it optimally lowers its price. But, if $\frac{\partial\lambda^*}{\partial n} < 0$ then $\frac{dp(\lambda^*,x,n)}{dn}$ may become either positive or negative. That is, if the publisher sees that it needs fewer advertisers it may decide to increase or decrease the price. The increase or decrease of the price depends on whether the impact of the reduction of λ on the price increase is greater or the impact of the price penalty for adding an extra slot. ■

Proof of Proposition 9

The proof is immediate by the definition of the CTR and the explanation in the text. ■

Proof of Proposition 10

(i) The optimal revenue using the CPM pricing scheme is $R_{cpm}(\lambda^*_{cpm}) = \lambda^*_{cpm}(1 - \mathbb{P}^{cpm}_n(\lambda^*_{cpm};\mu,N,n))Np_{cpm}(\lambda^*_{cpm})$. The optimal revenue of using of an alternative equivalent CPC pricing scheme is $R_{cpc}(\lambda^*_{cpc}) = \lambda^*_{cpc}(1 - \mathbb{P}^{cpc}_n(\lambda^*_{cpc};\mu,N,n))xp_{cpc}(\lambda^*_{cpc})$. As we assume that the pricing schemes are equivalent (generate the same optimal revenues), we have $R_{cpm}(\lambda^*_{cpm}) = R_{cpc}(\lambda^*_{cpc})$. Hence, by dividing the two revenue functions, we find that $\frac{\lambda^*_{cpm}}{\lambda^*_{cpc}} = \frac{p_{cpc}(\lambda^*_{cpc})x(1-\mathbb{P}^{cpc}_n(\lambda^*_{cpc}))}{p_{cpm}(\lambda^*_{cpm})N(1-\mathbb{P}^{cpm}_n(\lambda^*_{cpc}))}$. Thus, $\lambda^*_{cpm}/\lambda^*_{cpc} > 1$ if the right side is greater than one.

(ii) It is easy to show that $CTR(\lambda)$ is decreasing in λ, i.e., $\partial CTR/\partial\lambda < 0$. Thus, with $\lambda^*_{cpm} < \lambda^*_{cpc}$, we have $CTR(\lambda^*_{cpm}) < CTR(\lambda^*_{cpc})$. ■

Proof of Proposition 11

In order to derive $\mu^{j,m}$, we note that there are two streams of viewers that consider subsystem (j,m). The first stream consists of viewers who consider subsystem (j,m) as their first choice referred to as $W^1_{j,m} := \mu\varpi_{j,m}$ where μ is the advertisers' arrival rate at the website. From those viewers, $S^1_{j,m} := W^1_{j,m}(1-\mathbb{P}_0^{j,m})$ can see real ads while the rest $B^1_{j,m} := W^1_{j,m}\mathbb{P}_0^{j,m}$ only see filler ads on display in subsystem (j,m). Therefore, $W^{1,(g,h)}_{(j,m)} := \alpha^{g,h}_{j,m}B^1_{j,m}$ of them consider subsystem (g,h) while the rest $B^1_{j,m} - W^{1,(g,h)}_{(j,m)}$ immediately leave the system. From those viewers who consider the subsystem (g,h), $B^{1,(g,h)}_{(j,m)} = W^{1,(g,h)}_{(j,m)}\mathbb{P}_0^{g,h}$ face only filler ads. Therefore, $W^2_{j,m} = \alpha^{j,m}_{g,h}B^{1,(g,h)}_{(j,m)}$ of them consider subsystem (j,m) once again, while the rest $B^{1,(g,h)}_{(j,m)} - W^2_{j,m}$ leave the the publisher's website. In short, $W^2_{j,m}$ is the fraction of the $W^1_{j,m}$ viewers who consider ads in subsystem (j,m) as their first choice, but after experiencing a complete loop would consider subsystem (j,m) again[1]. Note that theoretically this loop of events can be repeated infinitely. However, in practice $\alpha^{g,h}_{j,m}$ or $\alpha^{j,m}_{g,h}$ can be near zero meaning that viewers may leave the website quickly just after once or twice visiting of subsystems with filler ads. Given this, in loop κ, we find

$$S^\kappa_{j,m} = W^\kappa_{j,m}(1-\mathbb{P}_0^{j,m}), \qquad \text{(O.13)}$$
$$W^\kappa_{j,m} = \mu\varpi_{j,m}\left(\alpha^{g,h}_{j,m}\alpha^{j,m}_{g,h}\mathbb{P}_0^{j,m}\mathbb{P}_0^{g,h}\right)^{\kappa-1},\ \kappa = 1,2,...$$

where $S^\kappa_{j,m}$ is the fraction of the $\mu\varpi_{j,m}$ viewers who initially consider subsystem (j,m) in the first loop and after a few unsuccessful attempts eventually visit and consider real ads posted in subsystem (j,m) in loop κ. Hence, the overall number of viewers in the first stream (only in interaction with subsystem (g,h)) is expressed as

$$\begin{aligned} S^{(g,h)}_{(j,m)} &= \sum_{\kappa=1}^{\infty} S^\kappa_{j,m} = \mu\varpi_{j,m}(1-\mathbb{P}_0^{j,m})\sum_{\kappa=1}^{\infty}\left(\alpha^{g,h}_{j,m}\alpha^{j,m}_{g,h}\mathbb{P}_0^{j,m}\mathbb{P}_0^{g,h}\right)^{\kappa-1} \\ &= \frac{\mu\varpi_{j,m}(1-\mathbb{P}_0^{j,m})}{1-\alpha^{g,h}_{j,m}\alpha^{j,m}_{g,h}\mathbb{P}_0^{j,m}\mathbb{P}_0^{g,h}}, \qquad 1\le g\le J,\ 1\le h\le M^g,\ (g,h)\ne(j,m) \end{aligned} \qquad \text{(O.14)}$$

The second stream of viewers includes those who initially consider subsystem (g,h) but finally visit subsystem (j,m). Based on a similar argument it can be shown that

$$S^{\kappa,(j,m)}_{(g,h)} = W^{\kappa,(j,m)}_{(g,h)}(1-\mathbb{P}_0^{j,m}), \qquad \text{(O.15)}$$
$$W^{\kappa,(j,m)}_{(g,h)} = \mu\varpi_{g,h}\mathbb{P}_0^{g,h}\alpha^{j,m}_{g,h}\left(\alpha^{g,h}_{j,m}\alpha^{j,m}_{g,h}\mathbb{P}_0^{j,m}\mathbb{P}_0^{g,h}\right)^{\kappa-1},\ \kappa = 1,2,...$$

[1] Since the subsystems can belong to different pages, the publisher might post new ads in a subsystem before a viewer re-checks it. In addition, some ads may leave the subsystems and give their place to other ads or filler ads.

where $S_{(g,h)}^{\kappa,(j,m)}$ is the fraction of the $\mu\varpi_{g,h}$ viewers who had first selected to consider ads in subsystem (g,h), but shifted to consider ads in subsystem (j,m) instead, in loop κ. Thus, the total number of viewers in the second stream is

$$
\begin{aligned}
S_{(g,h)}^{(j,m)} &= \sum_{\kappa=1}^{\infty} S_{(g,h)}^{\kappa,(j,m)} = \mu\varpi_{g,h}\mathbb{P}_0^{g,h}\alpha_{g,h}^{j,m}(1-\mathbb{P}_0^{j,m})\sum_{\kappa=1}^{\infty}\left(\alpha_{j,m}^{g,h}\alpha_{g,h}^{j,m}\mathbb{P}_0^{j,m}\mathbb{P}_0^{g,h}\right)^{\kappa-1} \\
&= \frac{\mu\varpi_{g,h}\mathbb{P}_0^{g,h}\alpha_{g,h}^{j,m}(1-\mathbb{P}_0^{j,m})}{1-\alpha_{j,m}^{g,h}\alpha_{g,h}^{j,m}\mathbb{P}_0^{j,m}\mathbb{P}_0^{g,h}}, \qquad 1\le g\le J,\ 1\le h\le M^g,\ (g,h)\neq(j,m) \qquad \text{(O.16)}
\end{aligned}
$$

Therefore, the overall number of viewers who successfully visit subsystem (j,m), considering the interactions with all other subsystems, would be

$$
\overline{\mu}^{j,m} = (1-\mathbb{P}_0^{j,m})\mu\frac{\varpi_{j,m}+\sum_{\substack{g=1\\(g,h)\neq(j,m)}}^{J}\sum_{h=1}^{M^g}\varpi_{g,h}\mathbb{P}_0^{g,h}\alpha_{g,h}^{j,m}}{1-\sum_{\substack{g=1\\(g,h)\neq(j,m)}}^{J}\sum_{h=1}^{M^g}\alpha_{j,m}^{g,h}\alpha_{g,h}^{j,m}\mathbb{P}_0^{j,m}\mathbb{P}_0^{g,h}}, \qquad 1\le j\le J,\ 1\le m\le M^j,
$$

while the overall number of viewers who visit subsystem (j,m) (including both successful and unsuccessful attempts) is

$$
\mu^{j,m} = \mu\frac{\varpi_{j,m}+\sum_{\substack{g=1\\(g,h)\neq(j,m)}}^{J}\sum_{h=1}^{M^g}\varpi_{g,h}\mathbb{P}_0^{g,h}\alpha_{g,h}^{j,m}}{1-\sum_{\substack{g=1\\(g,h)\neq(j,m)}}^{J}\sum_{h=1}^{M^g}\alpha_{j,m}^{g,h}\alpha_{g,h}^{j,m}\mathbb{P}_0^{j,m}\mathbb{P}_0^{g,h}}, \qquad 1\le j\le J,\ 1\le m\le M^j.
$$

Hence, the proof is complete. ■