Lancaster University
Management School

Index Policies for Promotion Strategies in Reward-based Crowdfunding

Dong Li, Department of Management Science, Lancaster University
Chenguang Wang, Baibing Li, Loughborough Business School, Loughborough University

Workshop on restless bandits, index policies and applications in reinforcement learning
Grenoble, France, Nov 20-21, 2023

Agenda

Lancaster University
Management School

- An introduction to reward-based crowdfunding
- The problem and the model
- Relaxation and decomposition
- Optimal policy for single project problems
- Indexability and index values
- Numerical experiments

An Introduction to Reward-based Crowdfunding

Reward-based crowdfunding is an innovative online financing alternative.

- The fundraisers launch their projects on a crowd-funding platform
- A typical project includes
    1. Information about their products/technology
    2. A pre-specified funding goal
    3. Length of the campaign
    4. A set of reward options that backers can purchase
- The backers choose which project to support, and then decide which reward option to purchase.
- Many platforms follow an **All-or-Nothing** (AoN) scheme.
- The platforms charge a percentage commission of the total funds raised from ***successful*** projects.

£139,918
pledged of £24,238 goal

501
backers

28
days to go

Back this project

Pledge HK$ 852

ABOUT £89

1 x Everything Proof Shirt

→ SUPER EARLY BIRD PRICE
↳ USD $109 / SAVE 36% ($169 Retail)

INCLUDES:
⚡ 1 x Everything Proof Shirt
→ Choose between 3 colors (black-ops, stealth green, and navigator blue).

Pledge HK$ 1,634

ABOUT £169

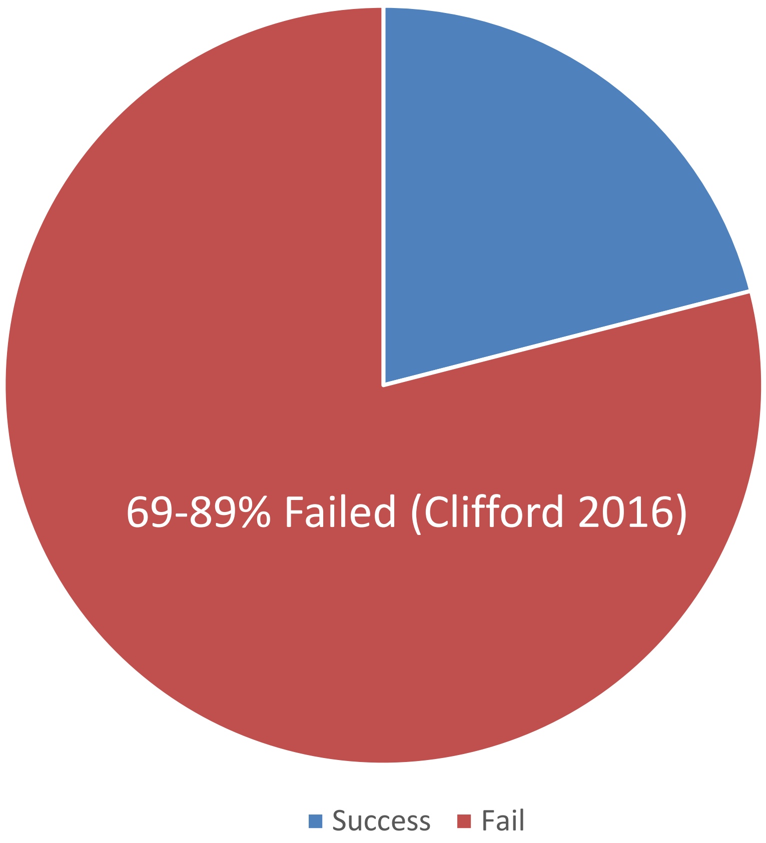
2 x Everything Proof Shirts

→ SUPER EARLY BIRD PRICE
↳ USD $209 / SAVE 38% ($338 Retail)

INCLUDES:
⚡ 2 x Everything Proof Shirts
→ Choose between 3 colors (black-ops, stealth green, and navigator blue).

Low Success Rate

Lancaster University Management School

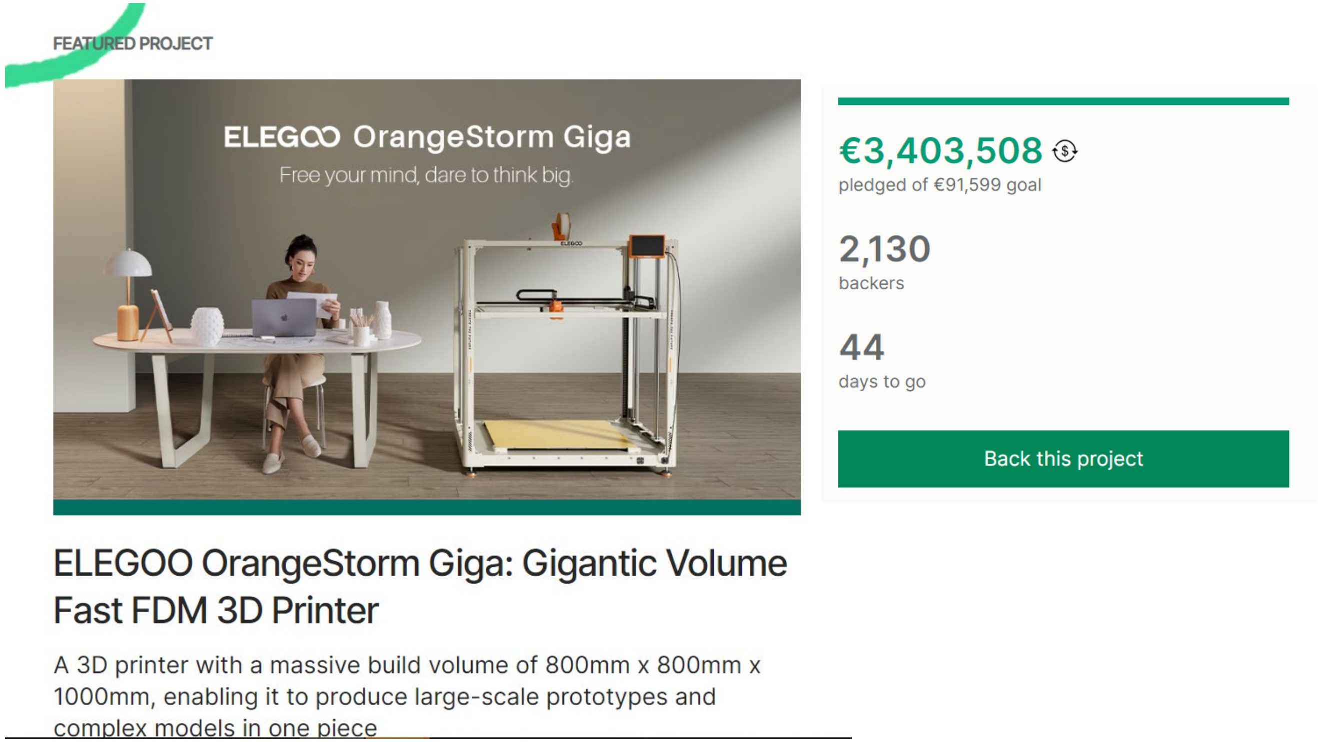


Most researches focus on factors of individual projects

- Fundraisers’ experience and expertise
- Funding goal, duration and reward options
- Information description

Promotion of Projects on the Homepage- Featured Projects

A crowdfunding platform may boost the chance of success of a campaign by highlighting it on the platform's homepage.



Our Research Question

How would platform managers maximise the total revenue by dynamically assigning limited promotion slots to projects?

The Problem Statement

Lancaster University Management School

- J crowdfunding projects seek financial investment from time 0 to T-1. Each project has a funding goal G_j.
- Discretise the time horizon into sufficiently small intervals t, Assume that customers visit the platform according to a Bernoulli process with a probability $\lambda \in (0,1)$ in each time t.
- Upon arrival, each customer either chooses to back one project, say project j, with probability p_j or leave without any purchases.
- Having decided which project to support, the backer chooses one reward option to purchase before leaving.
- At each time period t, the platform chooses one project to promote on its homepage.
- Our aim is to allocate the promotion slot to projects over time to maximise the revenue.

Modelling of Customers' Choices

Lancaster University
Management School

Random utility function: customers' perceived valuation on project *j*

$$u_j(g_j, a_j) = z_j(g_j, a_j) + \epsilon_j = m_j + \beta_1 a_j + \beta_2 \left(1 - \frac{g_j}{G_j}\right) - \beta_3 + \epsilon_j$$

Multinomial logit model (MNL) – customer's backing probability

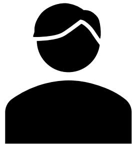
$$p_j(g_j, a_j) = \frac{e^{z_j(g_j, a_j)}}{1 + \sum_{k=1}^{J} e^{z_k(g_k, a_k)}}, 1 \le j \le J$$

Non-purchase probability

$$p_0 = \frac{1}{1 + \sum_{k=1}^{J} e^{z_k(g_k, a_k)}}$$

Customer's pledging

A customer will purchase a reward r_j with a known probability $F_j(r_j)$ where $r_j \in \{1, \dots, R_j\}$



Which project?

Which reward?

G_j : the funding goal of project j

g_j : the shortfall to the funding goal of project j

m_j : the overall attraction of project j

$a_j \in \{0,1\}$: promotion indicator

β_1 : promotion power

β_2 : herding effect

β_3 : side effect, e.g., market saturation

The Model - a Dynamic Program

Lancaster University Management School

- States: $\boldsymbol{g} = (g_1, \dots, g_J)$, a vector of shortfalls for all projects. Denote the state space at time t by Ω_t

- Action: $\boldsymbol{a} = (a_1, \dots, a_J)$, a vector of actions for all projects.
  Action space: $A = \{\boldsymbol{a}\colon\ a_j \in \{0,1\}, \sum_j a_j = 1\}$

- A policy $\pi\colon \Omega_t \to A, \forall 0 \le t \le T-1$, a decision rule to choose the project for promotion after observing the state at each time epoch

- Immediate reward in each time t under policy π:

$$h_t(\boldsymbol{g}, \pi(\boldsymbol{g})) = \lambda \sum_{j=1}^{J} p_j(\boldsymbol{g}, \pi(\boldsymbol{g})) \sum_{r_j=1}^{R_j} r_j F_j(r_j)$$

- Our **objective** is to find a policy that maximises the overall revenue

The Bellman Equation

Lancaster University
Management School

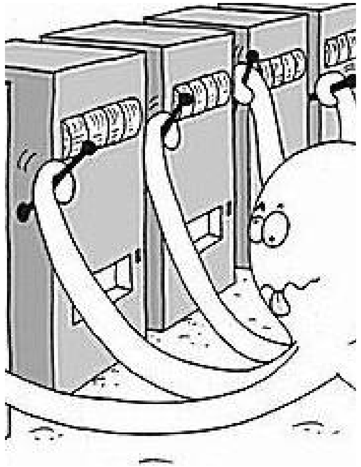
Denote by $V_t(\boldsymbol{g})$ the value function, i.e., the maximal expected fund still obtainable from time t onwards, given the system occupies state $\boldsymbol{g}$ at time t.

$$V_t(\boldsymbol{g}) = \max_{\boldsymbol{a}\in A}\left\{\lambda \sum_{j=1}^{J} p_j(\boldsymbol{g},\boldsymbol{a}) \sum_{r_j=1}^{R_j} F_j(r_j)\big(r_j + V_{t+1}(\tilde{\boldsymbol{g}}\,)\big) + \big(1-\lambda+\lambda p_0(\boldsymbol{g},\boldsymbol{a})\big)V_{t+1}(\boldsymbol{g})\right\}$$

where $\tilde{\boldsymbol{g}} = \boldsymbol{g} - \boldsymbol{r}$ and $\boldsymbol{r}$ is a J-dimensional vector that takes value of r_j on the j-th component and zero elsewhere.

$V_T(\boldsymbol{g}) = \sum_{j=1}^{J} h_T(g_j)$, where $h_T(g_j) = \begin{cases} -(G_j - g_j), & \text{if } g_j > 0 \\ 0, & \text{if } g_j \leq 0 \end{cases}$

Whittle's Restless Bandits Method- In a Nutshell



Restless Bandits

- Each fundraising project is a restless bandit, which always evolves regardless being promoted or not.

Relaxation and Decomposition

- Relaxation 1: allow multiple projects to be promoted simultaneously, but require on average the resource consumed is not more than one.
- Relaxation 2: associate a non-negative Lagrangian multiplier *W* (a fee for promotion) to the constraint and incorporated it into the objective function
- Decomposition: these relaxations allow the problem to be decomposed into a collection of single bandit/project problems.

Indexability and index values

- Prove the ***indexability*** to each project
- Calculate the index values (or fair charges) for each project in each state

Index policies

- Always choose to promote the project with the largest index value.

Relaxations

Lancaster University Management School

Relaxation 1: $\tilde{A} = \{\boldsymbol{a}:\ a_j \in \{0,1\}\}, \tilde{\pi}: \Omega_t \to \tilde{A}, \forall 0 \leq t \leq T-1$. We require

$$E\left[\sum_{t=0}^{T-1}\left(1-\sum_{j=1}^{J}\tilde{\pi}_t\left(\boldsymbol{g}(t)\right)_j\right)\right] \geq 0$$

Relaxation 2: Associate a non-negative Lagrangian multiplier W to the constraint above, and add it to the objective function (1)

$$\hat{V}_0(\boldsymbol{G}) = max_{\tilde{\pi}} E\left[\sum_{t=0}^{T-1} \boxed{h_t\left(\boldsymbol{g}(t), \tilde{\pi}_t\left(\boldsymbol{g}(t)\right)\right)} + \sum_{j=1}^{J} h_T\left(g_j(T)\right) + W\sum_{t=0}^{T-1}\left(1-\sum_{j=1}^{J}\tilde{\pi}_t\left(\boldsymbol{g}(t)\right)_j\right)\right] \quad (2)$$

However, due to the MNL, the problem (2) is not yet decomposable.

$$h_t(\boldsymbol{g}, \pi(\boldsymbol{g})) = \lambda \sum_{j=1}^{J} \boxed{p_j(\boldsymbol{g}, \pi(\boldsymbol{g}))} \sum_{r_j=1}^{R_j} r_j F_j(r_j)$$

Approximation of MNL by BNL

Lancaster University Management School

We further relax the problem by approximating the MNL-based backing probabilities with the following *J* Binomial Logit Models (BNLs), one for each project *j*:

$$p_j^{a_j}(g_j) = \frac{\exp\left(m_j + \beta_1 a_j + \beta_2 \left(1 - g_j/G_j\right)\right)}{1 + \exp\left(m_j + \beta_1 a_j + \beta_2 \left(1 - g_j/G_j\right)\right)}$$

It can be understood that each project faces the entire arrival stream, of which each arriving customer makes a binary choice of either backing this project or not, based on the BNL model above.

Problem (2) can now be decomposed by project.

Single Project Problems

Lancaster University Management School

$$v_0^W(G) = max_{\pi}\left\{\sum_{t=0}^{T-1}\left(\lambda p^{\pi_t(g)}(g)\sum_{r=1}^{R} rF(r) - W\pi_t(g)\right)\right\} \quad (3)$$

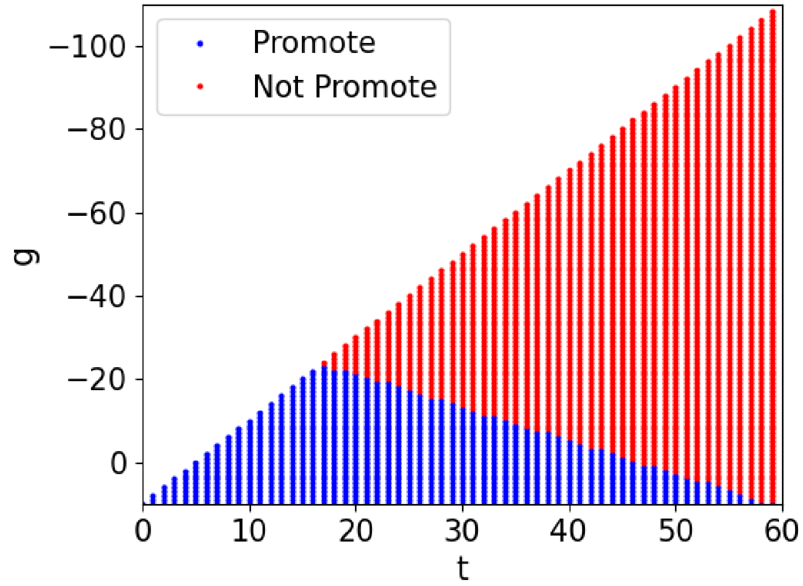
, where we still use π for the single project policy.

In each single-project problem, the project has a dedicated promotion space, and the action is whether or not to use the space for promotion at each decision epoch.

If the action is to promote ($\pi(g) = 1$), the project will be highlighted on the homepage with a cost of W. If the decision is not to promote ($\pi(g) = 0$), the project will not be highlighted and no cost is incurred.

Monotonicity of the Optimal Policy to the Single Project Problem

– under the condition of sufficiently long duration



Optimal policy for project 2 when W = 0.205

Proposition 1 (Monotonicity of the optimal policy) For any $W \geq 0$, the optimal policy π^* satisfies:

- $\pi_t^{*,W}(g) \geq \pi_t^{*,W}(g-r), \forall g \in \Omega_t, 0 \leq t \leq T-1$
- $\pi_t^{*,W}(g) \geq \pi_{t+1}^{*,W}(g), \forall g \in \Omega_t, 0 \leq t \leq T-1$

Table 1: The Setting of an Example

Project	β_1	β_2	β_3	T	λ	G	m	$F(r=1)$	$F(r=2)$
1	1	0.01	1.5	60	0.7	8	0.01	0.73	0.27
2						10	0.1	0.7	0.3

Indexability

Lancaster University Management School

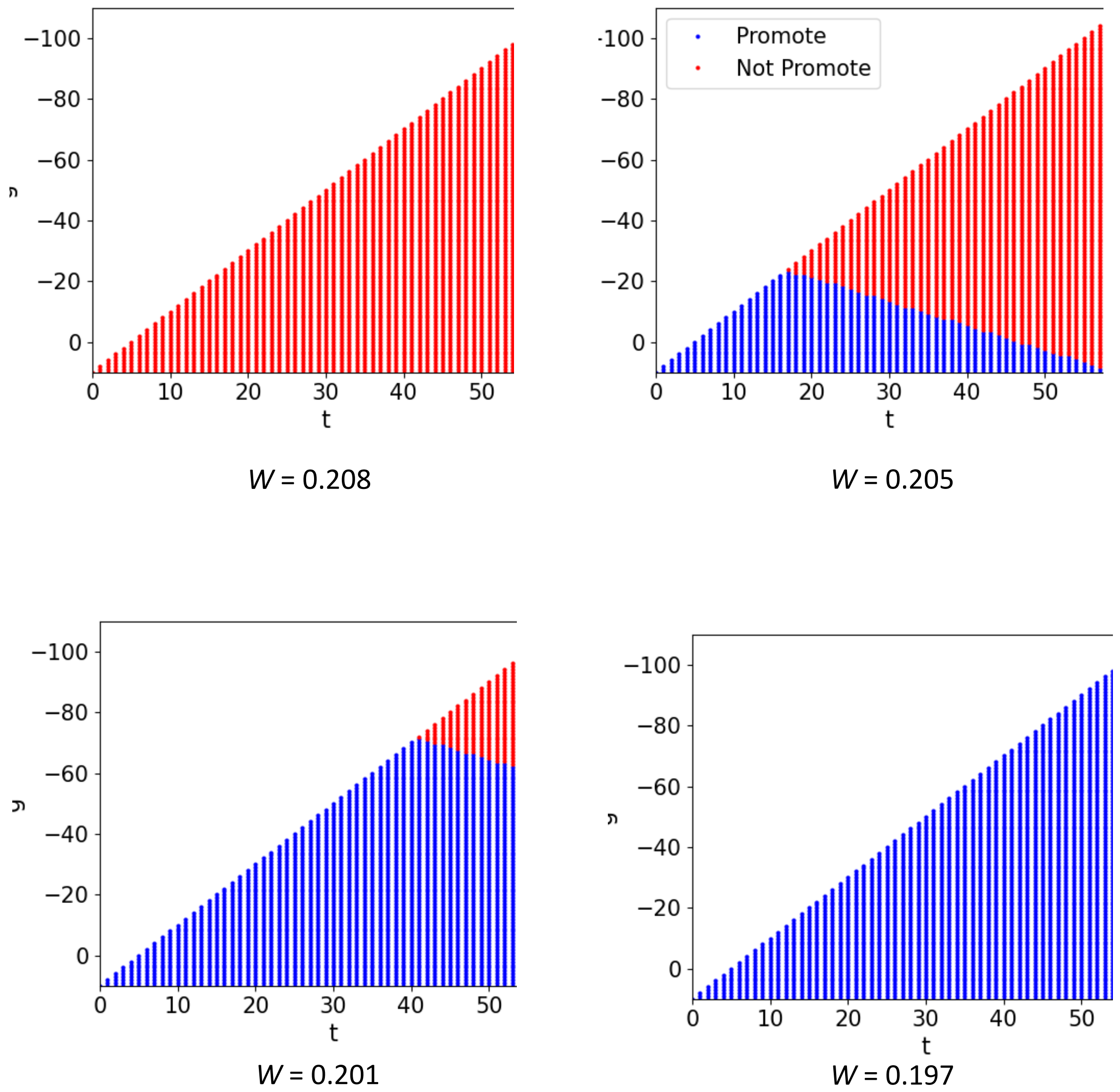
For any W and t, define the optimal promotion set as

$$B_t(W) = \{g: \pi_t^{*,W}(g) = 1, g \in \Omega_t\}$$

From Proposition 1, we have

Indexability: $B_t(W) \subseteq B_t(W')$ for any $W \geq W'$, $\forall 0 \leq t \leq T-1$

Demonstration of Indexability of the Example- project 2



Index Values

Whittle's Index: for an indexable project, the Whittle's index is defined as

$$w(g,t) = argmax_W\{g \in B_t(W)\}$$

Proposition 2: The Whittle index is evaluated as follows:

$$w(g,t) = \lambda\big(p^1(g) - p^0(g)\big)\left(\sum_{r=1}^{R} F(r)r + \Delta v_{t+1}^{\pi^0}(g)\right),$$

where $\Delta v_t^{\pi^0}(g) = \sum_{r=1}^{R} F(r) v_t^{\pi^0}(g-r) - v_t^{\pi^0}(g)$ is the marginal future revenue of an additional purchase under a non-promotion policy π^0.

A closed-form index value approximation: for each state g at time t, we assume that the herding effect from $t+1$ onwards and for all the future states remains the same as it is evaluated at time t for state g, then $v_t^{\pi^0}(g)$ can be approximated by

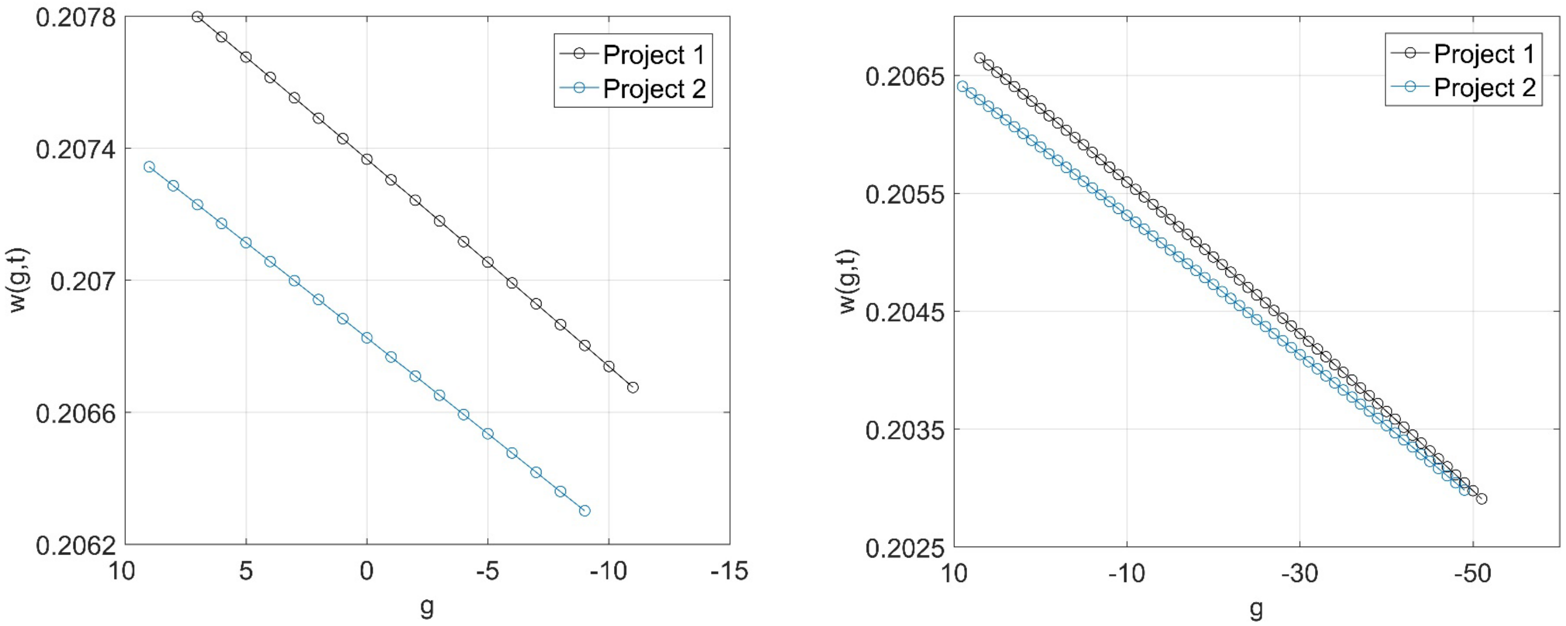
$$\hat{v}_t^{\pi^0}(g) = \lambda p^0(g) \sum_{t}^{T-1} \bar{r} = (T-t)\lambda p^0(g)\bar{r}.$$

And thus

$$\hat{w}(g,t) = \lambda(p^1(g) - p^0(g))\left[\bar{r} + \lambda\bar{r}(T-t-1)\left(\sum_{r=1}^{R} F(r)(p^0(g-r) - p^0(g))\right)\right].$$

Index Values

Lemma 1: The index value *w(g, t)* (i) increases in state *g*; and (ii) decreases in time *t*.



t=10 *t=30*

Numerical Experiments- The Policies

Smallest/largest shortfall first (**SSF/LSF**):	• always promotes the project with the smallest/largest percentage shortfall
Smallest/largest utility first (**SUF/LUF**):	• always promotes the project with the smallest/largest utility
Greedy policy (**GP**):	• always promotes the unfinished project with the highest funding goal
Conservative policy (**CP**):	• always promotes the unfinished project with the lowest funding goal
Myopic policy (**MP**):	• always promotes the project that leads to the highest immediate reward
Index policy (**IP/IPx**):	• always promotes the project with the largest index value/approximate index value

The Settings

Global parameters

β_1	β_2	β_3	λ
1	0.01	2.5	0.7

Project parameters

		$J = 3$		
Scenario	Project j	1	2	3
Baseline	m	0.04	0.08	0.16
	G	40	80	160
	$F(r=1)$	0.55	0.5	0.48
	$F(r=2)$	0.45	0.5	0.52
	$\bar{r}_b$	1.45	1.5	1.52
Smaller difference btw projects	m_n	0.045	0.08	0.155
	$F_n(r=1)$	0.54	0.5	0.49
	$F_n(r=2)$	0.46	0.5	0.51
	$\bar{r}_n$	1.46	1.5	1.51
Larger difference btw projects	m_e	0.035	0.08	0.165
	$F_e(r=1)$	0.56	0.5	0.47
	$F_e(r=2)$	0.44	0.5	0.53
	$\bar{r}_e$	1.44	1.5	1.53

The Results
– Percentage Revenue Gap between the IP and Other Policies

Scenario	Variation	SSF	LSF	SUF	LUF	GP	CP	MP	IPx
Baseline	-	33.922	140.081	41.428	25.297	25.301	16.358	28.122	-0.061
Global Parameters Sensitivity									
Duration	5% ↑	33.344	70.658	39.744	6.798	4.798	18.731	6.875	0.057
	5% ↓	35.022	219.287	40.707	61.510	61.531	13.582	59.542	0.502
Promotion Power	5% ↑	32.933	110.590	39.939	10.641	10.708	19.162	10.940	-0.097
	5% ↓	35.958	181.861	42.956	51.898	58.308	14.339	47.905	0.342
Herding Effect	5% ↑	34.819	145.683	41.509	24.653	24.747	16.966	25.823	-0.172
	5% ↓	35.174	156.000	40.804	24.246	28.102	16.159	25.216	0.739
Project Parameters Sensitivity									
Attraction	Smaller difference	33.511	158.343	40.395	28.194	27.898	16.746	28.174	0.252
	Larger difference	35.579	150.206	41.155	21.339	23.592	16.640	23.017	-0.580
Pledge	Smaller difference	34.343	149.283	40.302	31.015	30.080	16.642	28.170	0.023
	Larger difference	30.599	140.589	36.521	18.565	19.772	12.545	19.592	0.102

Lancaster University Management School

Thank you for attending, any questions?

dong.li@lancaster.ac.uk