Sequential Approval: A Model of “Likes”, Paper Downloads and Other Forms of Click Behaviour.

Paola Manzini* Marco Mariotti† Levent Ülkü‡§
University of Sussex and IZA Queen Mary University of London ITAM‡§

This version: July 2019

Abstract

We consider a model of behaviour that generates a dataset of “approvals” for items in a list. Approval is distinct from choice in that it does not guarantee a final selection (e.g. filling a virtual shopping cart) or may not involve a final selection at all (e.g. online “Likes” or post sharing). We study the criteria that a list designer who wishes to maximise approval should follow to manipulate approver behaviour. Furthermore, we study more standard questions such as the identification, comparative statics and characterisation of the model. The primitives of the model are shown to be substantially pinned down by observed approval behaviour. The key property that drives the model is supermodularity in the item’s quality and position in the list.

J.E.L. codes: D0.

Keywords: Approval behaviour, List design.

*Corresponding author. Department of Economics, University of Sussex Business School, University of Sussex, Falmer, Brighton, BN1 9SL, UK, email:p.manzini@sussex.ac.uk
†School of Economics and Finance, Queen Mary University of London, Mile End Road, London E1 4NS, UK, email: m.mariotti@qmul.ac.uk.
‡CIE and Department of Economics, ITAM, Camino de Sta. Teresa #930, Héroes de Padierna, 10700 CDMX, email: levent.ulku@itam.mx.
§We thank for useful comments and suggestions to Miguel Angel Ballester, Kalyan Chatterjee, Olivier Compte, Mira Frick, Yuhta Ishii, Tomasz Strzalecki, Kemal Yildiz, and seminar audiences at Université d’Aix-Marseille, Bilkent University, Oxford University, University of Helsinki, Université de Montreal, BRIC at Columbia, D-TEA at Paris University, GSE Barcelona Summer School, WWET at Southampton. Levent Ülkü acknowledges financial support from the Asociación Mexicana de Cultura. Manzini and Mariotti are grateful to ITAM for their generous hospitality during several visits.
1 Introduction

Consider an agent who:

- progressively fills her Amazon shopping cart or wish list;
- “likes” or shares posts going down her Facebook or Instagram feed;
- scans headlines from the BBC newsfeed and reads some articles;
- views abstracts and downloads some papers from the NBER new papers email list;
- matches with potential partners on Tinder.

Many online technologies both encourage and make observable a category of behaviour that can be summarised as “approval”, as distinct from “choice”. The main feature shared by the disparate situations in the examples is that the agent is not making a final selection from a set of items. In the online shopping, dating and scientific research examples, the items that are clicked may comprise only a preliminary selection, with the final selection possibly postponed to a subsequent stage. In this case approval can be seen as the first stage of a consideration set model of choice.\(^1\) Approval implies no choice commitment, to the point that the entire set of clicked items may be easily discarded: no-choice in the form of cart-abandonment is typical of online shopping but not of physical shopping.\(^2\) In other cases the issue of a final selection does not even arise: the agent will not award a prize to the best post or article she has read.

We emphasise two additional features of the examples: order and menu-richness. Items are scanned by the approver in a sequential manner - they can be seen as a list. This can be interpreted either in the sense that items come themselves as a list that is followed by the approver (as in Horan [8], Kovach and Ülkü [9], Rubinstein and Salant [13], and Yildiz [15]); or that, while the items are not displayed in list form, the approver scans the items in some order that is observable (as in Caplin et al. [4]). Secondly, the set of approvable items is typically large, sometimes even effectively “open-ended” from the point of view of the individual. There is not a readily identifiable menu that the approver takes in at once; rather, she sequentially examines and possibly approves the available items until


\(^2\)Cart-abandonment rates are thought to be as high as 70%. See e.g. https://baymard.com/lists/cart-abandonment-rate.
the exploration is stopped according to some rule expressing time, cognitive or budget constraints.

These features stand in contrast with the standard demand or choice theory framework, in which observable behaviour is interpreted as either a final selection (in the case of a choice/demand function) or a set of selections (in the case of choice/demand correspondences) that are equally good for the chooser. While at the formal level approval could be captured by a non-standard choice function (or, more precisely, correspondence), the distinction between choice behaviour and approval behavior is conceptual: the factors that affect approval are different from the factors that affect final choice. Approved items can differ significantly in terms of preference ranking. Approval behaviour calls for a tailored model and generates new questions. This paper constitutes a first attempt in this direction.

We will further depart from classical demand/choice theory by viewing behaviour as a function exclusively of the order in which the items are examined (i.e. of the list), taking the menu as fixed. While data in reality will likely offer both menu and list variation, it is the latter that constitutes the novel observable feature of online datasets. We wish to isolate the way in which the order aspect of data, rather than their set aspect, provides information.

Because there is no need to make a final selection, and therefore no need of storing in memory previously approved items, unlike in a search process choice memory constraints do not seem of significant importance in an approval process. With this in mind, we analyse two models of sequential approval in lists subject to a capacity constraint.

We assume that the (unobservable) psychological determinants of approval are:

(a) a preference relation on the fixed set of items;

(b) an approval threshold, above which an item that is encountered along the list is approved;

(c) a capacity constraint, acting either (c.1) on the number of items that can be approved; or (c.2) on the number of items that can be examined.

The approver scans the items and approves all items that pass the threshold, until the capacity constraint is reached. Two submodels are generated, depending on whether case

---

3In some of the examples we have given, for a proper interpretation in terms of list variation, the items should be thought of in terms of characteristics, such as the type of news (sport, politics, local, etc.) or of posts (professional, personal, funny animal videos, etc.).

4Dardanoni et al. [5] argue the related case that identification results in typical choice theoretical frameworks are predicated on an amount of menu variation that is unrealistically rich outside of experimental settings.
c.1 or c.2 above applies. We also assume that while preferences are stable for the purpose of the analysis, threshold and capacity are subject to stochastic variation. This is because, for example, mood may affect the judgment on whether or not an item is worthy of approval; and random factors like the arrival of e-mails or other interruptions and distractions may affect time constraints and the opportunity cost of time. Given the procedure, the joint probability distribution over capacities and thresholds, in conjunction with the preference, gives rise to a stochastic dataset of approvals, which records the probability or frequency with which each item is approved in any observed list.

We show first that the model is useful to address a much discussed issue, namely how online behaviour can be manipulated by interested parties. To this effect, we introduce the notion of list design. Consider an entity who can not only observe but also manipulate lists to pursue an objective. What is the optimal list given the objective? This raises some complex tradeoffs for the designer. For example, to pursue the objective, should positions in the list be used as complements of, or substitutes for, the quality of an item? And if items have different weights in the objective function, are the weights complements or substitutes of their quality? We are able to provide sharp answers to such questions for a broad class of objective functions.

Next we move to more standard questions, such as identification. Assuming that approvals are generated by the model, to what extent can an observer of approvals and lists identify the preferences and the joint probability distribution on approval capacity and thresholds? We show that our models have excellent identification properties: preferences are uniquely identified by behaviour, and the joint probability distribution is uniquely identified either in full generality or under an independence assumption (depending on whether c.1 or c.2 above applies). What is more, the identification can be performed using only a small fraction of the possible lists.

Thirdly, we consider comparative statics: how are the primitives associated with “more approving” behaviour? We link such comparative behaviour to first order stochastic dominance and the mean of a summary statistic of the joint capacity-threshold distribution.

Finally, we address characterisation: which exact constraints on behaviour do the models pose? We show that, beside the “expected” properties (such as monotonicity of the probability of an item being approved with the position in the list), a key property exhibited by an approver in our models is a form of complementarity between the quality and the position of an item. We are able to characterize the model completely under c.1 above, and

---

5Throughout the paper, we will use “quality” to describe the rank of an item in the preference order.
to characterise it partially under c.2. The characterisation under c.2 seems a hard problem which so far remains open.

2 The Models

Let $X$ be a finite set of $n$ items. A list is a strict linear order $\lambda$ on $X$, sometimes denoted $xyz...$ with $x, y, z, ... \in X$, and it describes the order in which the approver scans the items in $X$. For any list $\lambda$ and item $x$, let $\lambda(x)$ be the rank of $x$ in $\lambda$. We denote by $\Lambda$ the collection of all possible lists.

A stochastic approval function is a map $p : X \times \Lambda \rightarrow [0, 1]$, where $p(x, \lambda)$ is the probability that $x$ is approved when the decision maker is facing list $\lambda$. Sometimes we will abuse notation and write $p(x, k)$ in place of $p(x, \lambda)$ when $\lambda(x) = k$.

Note that we impose no adding-up-to-one constraint on $\sum_{x \in X} p(x, \lambda)$. Formally, a stochastic approval function could be equivalently defined as a stochastic choice correspondence $P : 2^X \times \Lambda \rightarrow [0, 1]$ associating with each list the probability of the possible approval sets. Then the adding up constraint applies over the $P(A, \lambda)$, with $A \in 2^X$. We prefer to use the formalism above since correspondences in choice theory are traditionally associated with the notion that the elements of the correspondence in a menu are indifferent (or incomparable), which is not the case for approval.

Unlike a standard stochastic choice correspondence, our domain comprises lists, not menus. The menu $X$ is held fixed in the analysis. The variation comes only from lists. What is more, the choice correspondence associated with approval should allow for the possibility of no-choice, i.e. the possibility that $P(\emptyset, \lambda) > 0$ for some $\lambda$.

An approver is a pair $(\succeq, \pi)$ of a preference and a joint capacity-threshold distribution. Unless otherwise specified, a preference is a weak linear order $\succeq$ over $X$. The joint capacity-threshold distribution is a map $\pi : N \times X \rightarrow (0, 1)$, so that $\pi(i, t)$ denotes the (strictly positive) probability that the capacity is $i$ and the (ordinal) approval threshold is $t$. The approver approves all the items along the list that are at least as good as the threshold, subject to the capacity constraint, which can take one of two forms:

1. **Depth constraint**: the constraint acts on the number of items that are examined.

---

6 We use the term “items” rather than “alternatives” to emphasise that, within the capacity constraint, there is no “competition” between the items.

7 A strict linear order is a binary relation $R$ that is transitive and trichotomous, that is such that exactly one of $xRy, yRx, \text{ or } x = y$ is true for all $x, y \in X$.

8 A weak linear order is a complete, transitive and antisymmetric binary relation on $X$. 

2. Approval constraint: the constraint acts on the number of items that are approved.

According to whether (1) or (2) applies we refer to the model as DCM (for Depth Constrained Model) or ACM (for Approval Constrained Model), respectively.

When a preference is given, we will sometimes enumerate the items in descending order of preference, i.e. write \( X \) as \( X = \{x_1, \ldots, x_n\} \) with \( x_i \succ x_j \iff i < j \).

We can now formalise the two models:

**Definition 1.** A stochastic approval function is a DCM, denoted by \( p^D(x, \lambda) \), if there exists a pair \((\succeq, \pi)\) such that

\[
p^D(x, \lambda) = \sum_{x \succeq t} \sum_{i \geq \lambda(x)} \pi(i, t) \tag{1}
\]

**Definition 2.** A stochastic approval function is an ACM, denoted by \( p^A(x, \lambda) \), if there exists a pair \((\succeq, \pi)\) such that

\[
p^A(x, \lambda) = \sum_{x \succeq t} \sum_{i \geq b(\lambda, \lambda(x), t)} \pi(i, t) \tag{2}
\]

where

\[
b(\lambda, j, t) = |\{x \in X : x \succeq t, \lambda(x) \leq j\}|
\]

is the number of items that are in positions earlier than \( j \) in list \( \lambda \) and that pass the threshold \( t \).\(^9\)

When equation (1) (resp., equation (2)) holds we say that \( p^D\) (resp., \( p^A \)) is generated by \((\succeq, \pi)\).

For an approver \((\succeq, \pi)\), \( \pi \) is independent if \( \pi(i, t) = \pi_i \pi_t \) where \( \pi_i \) and \( \pi_t \) denote the marginals of \( \pi \), with \( i = 1, \ldots, n \) and \( t \in X \).

While, as explained, we have in mind applications to a large set of items, for the purpose of understanding the models it is useful to illustrate the structure of the datasets in a simple three-item case, where we assume \( x \succ y \succ z \).

\(^9\)We suppress the dependence of \( b \) on \( \succeq \) for notational simplicity. Also, when no confusion arises we will drop the model-specific superscript and simply write \( p(x, \lambda) \).
DCM

| \( \lambda = xyz \) | \( \sum_{i=1}^{3} \sum_{t \in \{x,y,z\}} \pi (i, t) = 1 \) | \( \sum_{i=2}^{3} \sum_{t \in \{y,z\}} \pi (i, t) \) | \( \pi (3, z) \) |
| \( \lambda = xzy \) | \( \sum_{i=1}^{3} \sum_{t \in \{x,y,z\}} \pi (i, t) = 1 \) | \( \sum_{i \in \{y,z\}} \pi (3, t) \) | \( \sum_{i=2}^{3} \pi (i, z) \) |
| \( \lambda = yxz \) | \( \sum_{i=2}^{3} \pi_i \) | \( \pi_y + \pi_z \) | \( \pi (3, z) \) |
| \( \lambda = yzx \) | \( \sum_{i=3}^{1} \pi_i \) | \( \pi_y + \pi_z \) | \( \sum_{i=2}^{3} \pi (i, z) \) |
| \( \lambda = zxy \) | \( \sum_{i=2}^{3} \pi_i \) | \( \sum_{i \in \{y,z\}} \pi (3, t) \) | \( \pi_z \) |
| \( \lambda = zyx \) | \( \pi_3 \) | \( \sum_{i=2}^{3} \sum_{t \in \{y,z\}} \pi (i, t) \) | \( \pi_z \) |

Table 1: Approval probabilities in the DCM model

ACM

| \( \lambda = xyz \) | \( \sum_{i=1}^{3} \sum_{t \in \{x,y,z\}} \pi (i, t) = 1 \) | \( \sum_{i=2}^{3} \sum_{t \in \{y,z\}} \pi (i, t) \) | \( \pi (3, z) \) |
| \( \lambda = xzy \) | \( \sum_{i=1}^{3} \sum_{t \in \{x,y,z\}} \pi (i, t) = 1 \) | \( \sum_{i=2}^{3} \pi_i + \pi_{3,z} \) | \( \sum_{i=2}^{3} \pi (i, z) \) |
| \( \lambda = yxz \) | \( \pi (1, x) + \sum_{i=2}^{3} \pi_i \) | \( \pi_y + \pi_z \) | \( \pi (3, z) \) |
| \( \lambda = yzx \) | \( \pi (1, x) + \sum_{t \in \{x,y\}} \pi (2, t) + \pi_3 \) | \( \pi_y + \pi_z \) | \( \sum_{i=2}^{3} \pi (i, z) \) |
| \( \lambda = zxy \) | \( \sum_{t \in \{x,y\}} \pi (1, t) + \sum_{i=2}^{3} \pi_i \) | \( \sum_{i=2}^{3} \pi_i + \pi_{3,z} \) | \( \pi_z \) |
| \( \lambda = zyx \) | \( \pi (1, x) + \sum_{t \in \{x,y\}} \pi (2, t) + \pi_3 \) | \( \pi (1, y) + \sum_{i=2}^{3} \sum_{t \in \{y,z\}} \pi (i, t) \) | \( \pi_z \) |

Table 2: Approval probabilities in the ACM model

Comparing the two tables, we observe the key structural difference between the two models. While in the DCM the position of an item in the list is sufficient to determine its approval probability, in the ACM it is the entire set of predecessors that determines approval probabilities. For example, \( p^D (x, zxy) = p^D (x, yxz) \) but \( p^A (x, zxy) > p^A (x, yxz) \).
3 List design

We begin by taking the perspective of an interested “list designer”, who wishes to construct the list so as to maximise some objective function. This is a relevant problem in view of the fact that the issue of “manipulation” of online behaviour has become a significant concern in current public discourse: designing a list in our models to maximise an objective function is an operationalisation of the idea of manipulation. For instance, the list designer may wish to increase the general number of news pieces read or papers downloaded, or just of some particular types of news and papers; or to foster social network involvement through likes and sharing; or to enhance the revenue generated by a shopping cart.\(^\text{10}\) Note that in some, though not all, of these examples, it may be the case that some elements in the list have more value than others from the point of view of the list designer. Feenberg et al. [7] provide specific evidence that, in the case of NBER economics papers email announcement, “even among expert searchers, list-based searches can be manipulated by list placement”.\(^\text{11}\)

To accommodate broadly this type of aims by the designer we consider as the objective function the \textit{weighted sum of the approval probabilities} generated by a list. We concurrently assume that the type \((\succeq, \pi)\) of the approver is known to the designer - the results of the next section will show the kind of experiments a designer can perform to acquire this knowledge.

Letting \(w(x) \in \mathcal{R}_+\) be the weight that the designer associates to item \(x\), the weighted sum of approval probabilities is denoted

\[
W_{\lambda} = \sum_{x \in X} w(x) p(x, \lambda).
\]

\textbf{Definition 3.} A list \(\lambda\) is optimal if \(W_{\lambda} \geq W_{\mu}\) for all \(\mu \in \Lambda\).

Note that we eschew here the case of negative weights. This assumption is in some respects not demanding, as items with negative weights could simply be removed from the list. However it could be potentially limiting in some cases, for example if the designer is compelled by regulation to include loss-making items in the list.

The first dilemma of a designer is caused by the possible discrepancies between the quality and weight rankings: should either criterion be given priority, and if not, how

\(^{10}\)As another example, wish lists, while not implying any promise to buy, are viewed in e-commerce as a way to reduce the probability of cart abandonment by increasing consumer engagement.

\(^{11}\)These authors show that papers listed first in the email announcement for newly listed papers are about 30% more likely to be viewed, downloaded, and subsequently cited.
should the criteria be aggregated? Secondly, given a criterion, it is a priori not clear whether a weaker item according to this criterion should be placed earlier or later in the list. Earlier positions favour approval in both models; should the position be used as a reinforcer of quality or rather as a compensation for the lack of quality in an optimal list? And what about the weight?

Consider for instance the lists \( \lambda = yzx \) and \( \mu = zxy \) in the DCM, and let the preference relation be \( x \succ y \succ z \). Suppose that the weights are the same across all of the items, \( w(a) = \bar{w} \) for all \( a \). The difference in the designer’s objective between the two lists in the DCM is\(^{12}\)

\[
W_\lambda - W_\mu = \bar{w} (-\pi(2, x) + \pi(1, z))
\]  

So the sign of \( W_\lambda - W_\mu \) for two generic lists \( \lambda \) and \( \mu \) is ambiguous and depends on the shape of the distribution. Yet both models help give some stark answers to the design questions. Our first result shows that in the ACM the optimality condition in fact does not depend at all on the joint distribution \( \pi \).

**Theorem 1.** In the ACM, a list \( \lambda \) is optimal if and only if it agrees with the weight ordering, that is if and only if the following condition holds:

\[
w(x) > w(y) \Rightarrow x\lambda y
\]

**Proof.** Note first that an optimal list exists, since there are finitely many lists and thus the map \( \lambda \mapsto \sum_{x \in X} w(x) p(x, \lambda) \) has a maximiser. Consider any \( \lambda, \mu \in \Lambda \) that only differ in the position of two consecutive items. That is, enumerating items by quality, there exist \( x \) and \( y \) such that \( \lambda(y) = \lambda(x) + 1, \mu(x) = \lambda(y), \mu(y) = \lambda(x) \) and \( \lambda(z) = \mu(z) \) for all \( z \in X \setminus \{x, y\} \). It is easy to check from (2) that

\[
p(z, \lambda) = p(z, \mu) \ \forall z \in X \setminus \{x, y\}
\]  

To see the difference in approval probability across the two lists for \( x \) and \( y \), suppose

\[^{12}\text{Using Table 1, } W_\lambda = w(x) \sum_{t \in \{x,y,z\}} \pi(3,t) + w(y) \sum_{i=1}^3 \sum_{t \in \{y,z\}} \pi(i,t) + w(z) \sum_{i=2}^3 \pi(i,z) \text{ and } W_\mu = w(x) \sum_{i=2}^3 \sum_{t \in \{x,y,z\}} \pi(i,t) + w(y) \sum_{t \in \{y,z\}} \pi(3,t) + w(z) \sum_{i=1}^3 \pi(i,z).\]
w.l.o.g. that $x \succ y$. Then the following assertions are easily checked:

1. $b(\mu, \mu(x), x) = b(\lambda, \lambda(x), x)$
2. $b(\mu, \mu(x), y) = b(\lambda, \lambda(x), x) + 1$
3. $b(\mu, \mu(x), z) = b(\lambda, \lambda(x), z) + 1 \forall z \in X : y \succeq z$
4. $b(\mu, \mu(y), z) = b(\lambda, \lambda(y), z) - 1 \forall z \in X : y \succeq z$

As a consequence,

$$p(x, \mu) = p(x, \lambda) - \pi(b(\lambda, \lambda(x), y), y) - \sum_{y \succeq z} \pi(b(\lambda, \lambda(x), z), y)$$  \hspace{1cm} (6)

$$p(y, \mu) = p(y, \lambda) + \pi(b(\lambda, \lambda(x), y), y) + \sum_{y \succeq z} \pi(b(\lambda, \lambda(x), z), y)$$  \hspace{1cm} (7)

Therefore,

$$p(x, \mu) - p(x, \lambda) = p(y, \lambda) - p(y, \mu) < 0$$  \hspace{1cm} (8)

(the inequality holding by the strict positivity of $\pi$), so that:

$$w(x)p(x, \mu) + w(y)p(y, \mu) \geq w(y)p(y, \lambda) + w(x)p(x, \lambda) \iff w(x) \leq w(y)$$  \hspace{1cm} (9)

This means, recalling (5), that if $\lambda$ disagrees with the weight ordering for $x$ and $y$ it is possible to increase $W_\lambda$ to $W_\mu > W_\lambda$ using instead the list $\mu$ that switches two consecutive items $x$ and $y$. Similarly, if $\mu$ disagrees with the weight ordering for $x$ and $y$, it is possible to increase $W_\mu$ to $W_\lambda > W_\mu$. We conclude from this analysis that any list that disagrees with the weight ordering can be improved upon (since any such list must contain at least two consecutive items that disagree with the weight ordering).

Conversely, suppose that a list $\lambda$ agrees with the weight ordering and compare it with an optimal list, say $\mu$. Since $\mu$ is optimal, by the previous argument it must also agree with the weight ordering. So for all $x, y \in X$ for which $w(x) > w(y)$ we have both $x \lambda y$ and $x \mu y$. The lists $\lambda$ and $\mu$ can only disagree on the way they order subsets $B \subset X$ for which $w(x) = w(y)$ for all $x, y \in B$. Let $\lambda_0, \lambda_1, ..., \lambda_K$ be a sequence of lists such that $\lambda_0 = \lambda$, $\lambda_K = \mu$ and any $\lambda_k, \lambda_{k+1}$ differ only in the way they order two consecutive items in $B$. By (9) and (5) $W_{\lambda_k} = W_\lambda$ for all $k = 1, ..., K$, and so $\lambda$ must also be optimal.

One way to read this result is that in the ACM there is no substitutability between the quality and the weight of an item as far as its optimal positioning for list design is concerned. In particular, an optimal list must weakly agree with the weight order. And if two items
tie for weight, they may be listed in either order at the optimum: quality does not function as a tie-breaker either.

An important case is when weights are the same for all items. Then, part (i) of Theorem 2 has the interesting consequence that any list is optimal.

**Corollary 1. (The List Invariance Principle)** Suppose that \( w(x) = w(y) \) for all \( x, y \in X \) and let \( p \) be an ACM. Then \( \sum w(x) p(x, \lambda) = \sum w(y) p(x, \mu) \) for all \( \lambda, \mu \in \Lambda \).

While the result can be obtained as an immediate implication of the statement of Theorem 2, a short direct proof is instructive:

**Proof.**\(^{13}\) Take any capacity-threshold pair \((i, t)\). If \(|\{x : x \succeq t\}| = k \leq i\), then \( k \) items are approved. Otherwise, \( i \) items are approved. Since neither \( k \) nor \( i \) depends on the list, the sum \( \sum_{k=1}^{n} p^L(x_k, \lambda) \) does not depend on the list, and the result follows. \(\square\)

The List Invariance Principle is a key feature of the ACM, and will be used in subsequent results.

In the case of the DCM model, matters are less straightforward. The general optimality condition for a list \( \lambda \) to be optimal can be written as follows:

\[
\sum_{x \in X} \text{sign} (\lambda(x) - \mu(x)) \left[ w(x) \sum_{i = \min\{\lambda(x), \mu(x)\}}^{\max\{\lambda(x), \mu(x)\} - 1} \sum_{t \leq x_i} p(i, t) \right] \leq 0 \forall \mu \in \Lambda. \tag{10}
\]

Condition (10) says that, for any competitor list \( \mu \), the algebraic sum of the losses or gains in approval probabilities arising from moving each item to a different position in \( \mu \) is non-positive. This is a complex condition that in general cannot be expressed in terms of a simple, list-independent metric on items. Nevertheless, we can do so in two important cases (one of which includes the equal-weight case).

**Theorem 2.** Suppose that \( p \) is a DCM generated by \((\succeq, \pi)\).

1) Suppose that weights agree with preferences in the sense that \( x \succ y \Rightarrow w(x) \geq w(y) \). Then the unique optimal list \( \lambda \) agrees with the preferences, i.e. \( x \succ y \iff x \lambda y \).

2) If the probability distributions on thresholds and capacities are independent (with no restriction on the weights), then a list \( \lambda \) is optimal if and only if it satisfies the following condition:

\[
w(x) \sum_{x \geq t} \pi_t > w(y) \sum_{y \geq t} \pi_t \Rightarrow x \lambda y
\]

\(^{13}\)We thank Yuhta Ishii for suggesting this cute proof.
Proof. For both parts, recall that, as noted in the proof of the previous theorem, an optimal list exists. The proof structure here is similar, except that we consider switches between items that are not necessarily consecutive.

1) Take any lists \( \lambda, \mu \) such for some distinct \( x, y \in X, \lambda (x) = \mu (y), \lambda (y) = \mu (x) \) and \( \lambda (z) = \mu (z) \) for all \( z \neq x, y \). From the model it follows immediately that \( p (z, \lambda) = p (z, \mu) \) for all \( z \neq x, y \). Suppose that \( x \succeq y \) and \( y \lambda x \). We can calculate:

\[
\begin{align*}
\omega (x) p (x, \lambda) + \omega (y) p (y, \lambda) &= \omega (x) \sum_{x \leq t} \sum_{i \geq \lambda (x)} \pi (i, t) + \omega (y) \sum_{y \leq t} \sum_{i \geq \lambda (y)} \pi (i, t) \\
&< \omega (x) \sum_{x \leq t} \sum_{i \geq \lambda (x)} \pi (i, t) + \omega (x) \sum_{x \leq t} \sum_{\lambda (x) > i \geq \lambda (y)} \pi (i, t) \\
&+ \omega (y) \sum_{y \leq t} \sum_{i \geq \lambda (y)} \pi (i, t) - \omega (y) \sum_{y \leq t} \sum_{\lambda (x) > i \geq \lambda (y)} \pi (i, t) \\
&= \omega (x) \sum_{x \leq t} \sum_{i \geq \lambda (x)} \pi (i, t) + \omega (y) \sum_{y \leq t} \sum_{i \geq \lambda (x)} \pi (i, t) \\
&= \omega (x) \sum_{x \leq t} \sum_{i \geq \mu (x)} \pi (i, t) + \omega (y) \sum_{y \leq t} \sum_{i \geq \mu (y)} \pi (i, t) \\
&= \omega (x) p (x, \mu) + \omega (y) p (y, \mu).
\end{align*}
\]

where the inequality uses the assumption that \( \omega (x) \geq \omega (y) \) and the strict positivity of \( \pi \). Hence \( \sum_{x \in X} \omega (x) p (x, \mu) > \sum_{x \in X} \omega (y) p (x, \lambda) \). Since the unique list for which switches like the one considered cannot be performed is the one that agrees with \( \succeq \), we have established that this list is the unique maximiser of the objective.

2) Take two lists \( \lambda \) and \( \mu \) that switch some \( x \) and \( y \) as in part (1), with \( y \lambda x \). Then

\[
W_\mu - W_\lambda = \omega (x) \sum_{\lambda (y) \leq i < \lambda (x)} \sum_{x \leq t} \pi_i \pi_t - \omega (y) \sum_{\lambda (y) \leq i < \lambda (x)} \sum_{y \leq t} \pi_i \pi_t \\
= \left( \omega (x) \sum_{x \leq t} \pi_t - \omega (y) \sum_{y \leq t} \pi_t \right) \sum_{\lambda (y) \leq i < \lambda (x)} \pi_i
\]

so that

\[
W_\mu \geq W_\lambda \iff \omega (x) \sum_{x \geq t} \pi_t \geq \omega (y) \sum_{y \geq t} \pi_t.
\] (11)

Hence any list \( \lambda \) for which \( y \lambda x \) but \( \omega (x) \sum_{x \geq t} \pi_t \geq \omega (y) \sum_{y \geq t} \pi_t \) can be improved upon. Conversely, any list \( \lambda \) that satisfies the condition in the statement cannot be improved upon by any other list, by an argument analogous to that in the final part of the proof of Theorem 1.

\[\square\]
The first case of Theorem 2 focuses on the case where any tension between quality and weight has been removed, that is, the weight order does not contradict the preference order. Then the only issue that needs to be resolved is what, if any, degree of substitutability between the position and the quality of items there is for an optimum. Expression (4) of the previous example (which falls under this case) shows that the designer’s ordering of two generic lists may depend on the joint distribution $\pi$. This is because, while each single repositioning of an item either improves or worsens the objective independently of $\pi$, the algebraic sum of these effects does depend on $\pi$. But Theorem 2 shows that in fact the uniquely optimal list must follow the preference order independently of $\pi$. As it turns out, like in the ACM, for achieving an optimum in this case it never pays off to advance in the list a worse item in order to increase the approvals it receives. Unlike in an ACM, however, the List Invariance Principle does not apply: when the weights are the same, quality must be used as a tie-breaker.

The second part of the result tells us that, for general patterns of weights and quality but under independence in the capacity-threshold distribution, there is a specific form of substitutability between the weight and the quality of items for an optimum. This substitutability is expressed by the terms $w(x) \sum_{x \geq t} \pi_t$. In fact, $\sum_{x \geq t} \pi_t$ is a measure of quality - it is larger for items that are higher in the preference ordering.

We now give some intuition for the difference in the optimal design principles between the ACM and the DCM, which might otherwise remain hidden in the formulae. Consider the first part of the statement of Theorem 2. When a better item $x$ is placed in a list $\lambda$ behind a worse item $y$, a switch in the positions of $x$ and $y$ to obtain a new list $\mu$ produces two effects (recall that the approval probabilities of all other items are not affected by the switch).

i) any loss for $y$ is a gain for $x$. If a capacity-threshold realisation $(i, t)$ leads to the approval of $y$ in $\lambda$ but not in $\mu$, then $y$ passes the threshold $t$ and the capacity $i$ is strictly less than $\mu(y)$ but at least as much as $\lambda(i)$. But this means that the realisation $(i, t)$ leads to the approval of $x$ in $\mu$ and not in $\lambda$.

ii) some gain for $x$ is not a loss for $y$. For example, the capacity-threshold realisation $(i, t) = (\lambda(y), x)$ leads to the approval of $x$ in $\mu$ but not in $\lambda$. But it never leads to the approval of $y$.

These effects imply that, if $w(x) \geq w(y)$, a switch that improves the position of the better item always increases the value of the objective function. Therefore, since an optimal list exists, it must be the only list that is not vulnerable to such types of switch, namely
the list that agrees with preferences. Note that, in general, sequences of switches might improve where no single switch improves; hence in the above argument it is crucial that only one candidate list remains.

This reasoning highlights the fundamental difference between the two models for the purposes of list design. A similar reasoning to the one above could be performed for consecutive items in the ACM, as in the proof of Theorem 1. But, unlike in the DCM, in the ACM y must have absorbed capacity pre-switch for x to gain from the switch. In an ACM there is in fact no special significance for any given position: x gains from advancing one position if, and only if, its predecessor is using the last available unit of capacity. Hence y loses from the switch exactly the approval probability that x gains (this is the content of equation (8). Therefore, since better positions correspond to higher approval probabilities, the objective is increased every time that an item with a larger weight (independently of its quality) is placed before, rather than after, an item with a smaller weight.

4 Identification

We proceed to take the perspective of an observer of a stochastic approval function who would like to retrieve the pair $(\succeq, \pi)$ that generated it, under the assumption that a given model holds. This exercise is methodologically analogous to classical identification exercises in demand and choice theory. The question is evidently of relevance for list designers, online companies and social scientists alike who, for different reasons, are interested in the drivers of approval behaviour.

The following result shows that the primitives of the two models are pinned down by approval probabilities to a surprising extent. The proofs give explicit formulas expressing the psychological primitives in terms of approval probabilities.

**Theorem 3.** Let $p$ be a stochastic approval function.

(i) In a DCM, both preferences and the joint capacity-threshold distribution are identified uniquely. That is, if $p$ is a DCM generated both by $(\succeq, \pi)$ and by $(\succeq', \pi')$, then $(\succeq, \pi) = (\succeq', \pi')$.

(ii) In an ACM, preferences are identified uniquely, and the joint capacity-threshold distribution is identified uniquely if it is independent. That is, if $p$ is an ACM generated both by $(\succeq, \pi)$ and by $(\succeq', \pi')$, then $\succeq = \succeq'$. Moreover, if $\pi$ is independent, it must also be $\pi = \pi'$.

**Proof.** First we prove the following
Lemma 1. (Revelation of preferences) Let \( \lambda \) and \( \lambda' \) be such that \( \lambda (x) = 1 = \lambda' (y) \). Then \( x \succ y \iff p^g (x, \lambda) > p^g (y, \lambda') \), \( g \in \{D, A\} \).

Proof. Observe that in a DCM \( \lambda (x) = 1 \Rightarrow \sum_{i \geq \lambda (x)} \pi (i, t) = \sum_{i = 1}^{n} \pi (i, t) \) for all \( x \in X \). Let \( x \succ y \). Since \( \succ \) is a linear order, \( z \succ x \Rightarrow z \succ y \), so that \( p^D (x, \lambda) = \sum_{x \succ t} \sum_{i \geq 1} \pi (i, t) > \sum_{y \succ t} \sum_{i \geq 1} \pi (i, t) = p^D (y, \lambda') \). For the other direction, let \( p^D (x, \lambda) > p^D (y, \lambda') \). Then there must be some component \( \pi (i, t^*) > 0 \) that appears in the expression (1) for \( p^D (x, \lambda) \) but not in that for \( p^D (y, \lambda') \), so that \( x \succ t^* \) while \( t^* \succ y \), from which \( x \succ y \) follows. The same reasoning applies to the ACM model, noting that \( \sum_{i \geq b (\lambda, \lambda (x), t)} \pi (i, t) = \sum_{i \geq 1} \pi (i, t) \) and \( \sum_{i \geq b (\lambda, \lambda (y), t)} \pi (i, t) = \sum_{i \geq 1} \pi (i, t) \).

Now enumerate the items in descending order of preference (this is well-defined in view of Lemma 1).

i) We will show that, given the preferences, the \( \pi (i, t) \) are uniquely determined by the approval probabilities. The following equalities must hold by the definition of a DCM:
\[
\pi (n, x_n) = p (x_n, n)
\]
\[
\pi (k, x_n) = \sum_{i \geq k} \pi (i, x_n) - \sum_{i \geq k+1} \pi (i, x_n) = p (x_n, k) - p (x_n, k+1) \quad \forall k < n
\]
\[
\pi (n, x_l) = \sum_{j \geq l} \pi (n, x_j) - \sum_{j \geq l+1} \pi (n, x_j) = p (x_l, n) - p (x_{l+1}, n) \quad \forall l < n
\]
\[
\pi (k, x_l) = \sum_{i \geq k} \pi (i, x_l) - \sum_{i \geq k+1} \pi (i, x_l)
\]
\[
= \left( \sum_{j \geq l} \sum_{i \geq k} \pi (i, x_j) - \sum_{j \geq l+1} \sum_{i \geq k} \pi (i, x_j) \right) - \left( \sum_{j \geq l} \sum_{i \geq k+1} \pi (i, x_j) - \sum_{j \geq l+1} \sum_{i \geq k+1} \pi (i, x_j) \right)
\]
\[
= (p (x_l, k) - p (x_{l+1}, k)) - (p (x_l, k+1) - p (x_{l+1}, k+1)) \quad \forall k, l < n.
\]
Together with Lemma 1, these equalities prove (i) in the statement.

ii) Denote the following three types of lists:
\[
\lambda = x_1 x_2 \ldots x_n
\]
\[
\lambda^l = x_l x_1 x_2 \ldots x_{l-1} x_{l+1} \ldots x_n
\]
\[
\lambda_{l,m} = x_1 x_2 \ldots x_{l-1} x_m x_{l+1} \ldots x_{m-1} x_l x_{m+1} \ldots x_n
\]

That is, \( \lambda \) is list that agrees with the preference ordering; \( \lambda^l \) modifies \( \lambda \) by moving \( x_l \) to the first position; and \( \lambda_{l,m} \) modifies \( \lambda \) by switching the positions of \( x_l \) and \( x_m \).
Since \( p(x_l, \lambda^l) = \sum_{j \geq l} \pi_{x_j} \) for all \( l = 1, ..., n \), we have:

\[
\pi_{x_n} = p(x_n, \lambda) \\
\pi_{x_l} = p(x_l, \lambda^l) - \sum_{j \geq l+1} \pi_{x_j} = p(x_l, \lambda^l) - p(x_{l+1}, \lambda^{l+1}) \forall l < n.
\]

Since \( p(x_n, \lambda) = \pi_n \pi_{x_n} \), we have

\[
\pi_n = \frac{p(x_n, \lambda)}{p(x_n, \lambda^n)}.
\]

Since \( p(x_n, \lambda_{n,n-1}) = \pi_{x_n} (\pi_{n-1} + \pi_n) \), we have:

\[
\pi_{n-1} = \frac{p(x_n, \lambda_{n,n-1}) - \pi_{x_n} \pi_n}{\pi_{x_n}} = \frac{p(x_n, \lambda_{n,n-1}) - p(x_n, \lambda)}{p(x_n, \lambda^n)}.
\]

To conclude, since \( p(x_l, \lambda_{l+1,l-1}) = \pi_{l-1} \pi_{x_l} + \sum_{i \geq l} \pi_i \sum_{j \geq l} \pi_{x_j} \), we have

\[
\pi_{l-1} = \frac{p(x_l, \lambda_{l+1,l-1}) - \sum_{i \geq l} \pi_i \sum_{j \geq l} \pi_{x_j}}{\pi_{x_l}} \\
= \frac{p(x_l, \lambda_{l+1,l-1}) - p(x_l, \lambda)}{p(x_l, \lambda^l) - p(x_{l+1}, \lambda^{l+1})} \forall l < n. \quad (12)
\]

Together with Lemma 1, these equalities prove (ii) in the statement.

Theorem (3) establishes that in both the DCM and the ACM preferences can be uniquely retrieved from a dataset of approval probabilities.

For the DCM we also have a unique identification of the joint probability distribution \( \pi \). It is noteworthy that this identification can be implemented using a small number of lists. The proof makes clear that what is needed is a set of lists such that each item in \( X \) appears in each position, which can be obtained with just \( n \) of the \( n! \) lists (choosing any list \( \lambda \), construct the other \( n - 1 \) lists by successively moving the last element to first). Therefore, as the number of items increases, the proportion of lists that need to be observed out of the possible ones falls to zero.

That some kind of restriction is needed to obtain a complete identification in the ACM can be seen from a very simple example. Let \( X = \{x, y\} \) and let \( x \succ y \). The stochastic approval function can be represented as an ACM as follows:
Here, it is impossible to disentangle the sum \( \pi(1, x) + \pi(2, x) \) and this leads to only a partial identification of the parameters. In general, for each \( x_i \), only the partial sums \( \pi(i, x) + \pi(i + 1, x) + \ldots + \pi(n, x) \) can be identified, since any reallocation of probability mass within the sums is behaviourally undetectable. On the other hand, assuming independence the sums can be distentangled. In the example above we have:

\[
\pi_y = p(y, xy), \text{ thus identifying } \pi_y \text{ (and therefore } \pi_x), \text{ and } \\
\pi_y\pi_2 = p(y, xy), \text{ thus identifying } \pi_2 \text{ (and therefore } \pi_1). \\
\]

5 Comparative statics

For a given preference \( \succeq \) consider probability distributions \( \pi_a \) and \( \pi_b \) and their associated approval functions \( p_a \) and \( p_b \), respectively. When can we say that one distribution is more approving than the other? There are at least two plausible criteria:

Definition 4. Say that \( a \) is more approving than \( b \) iff, for any list \( \lambda \), the total number of approvals by \( a \) in \( \lambda \) is greater than that by \( b \), i.e. \( \sum_x p_a(x, \lambda) \geq \sum_x p_b(x, \lambda) \).

Definition 5. Say that \( a \) is strongly more approving than \( b \) iff \( p_a(x, \lambda) \geq p_b(x, \lambda) \) for all items \( x \) and lists \( \lambda \).

That is, while the first definition looks at the total number of approvals in any list, the second one considers dominance not only in each list but also for each item.

We will show that the concepts of strongly more approving in the DCM and of more approving in the ACM can be fully characterised in terms of the primitives. In a DCM the strongly more approving partial order is related to a first order stochastic dominance partial order derived from \( \pi \). In an ACM the more approving order is related to a simple statistics of \( \pi \) that can be easily interpreted.

Enumerating the items from best to worst, for any probability distribution \( \pi \) on \( N \times X \) denote by \( F_\pi \) the cdf, given by \( F_\pi(i, t) = \pi\left( \{(l, x_j) : l \leq i, j \leq t\} \right) \), and define the “dual” cdf \( \hat{F}_\pi \) on \( \{-n, \ldots, -1\} \times \{-n, \ldots, -1\} \) by

\[
\hat{F}_\pi(-i, -t) = \pi\left( \{l : -l \leq -i, -j \leq -t\} \right)
\]
with \( i, t = 1, \ldots, n \).

Any \( \pi \) defines uniquely a (univariate) random variable \( X_{\pi}^{-} \) on \( \{1, \ldots, n\} \) by setting

\[
\Pr (X_{\pi}^{-} = k) = \pi \left( \{ (i, x) : \min (i, x) = k \} \right)
\]

That is, for each capacity-threshold realisation (with the threshold item expressed as its rank), \( X_{\pi}^{-} \) takes on the minimum of the two values. Denote \( E (X_{\pi}^{-}) \) the mean of this random variable.

**Theorem 4.** (i) In the DCM, \( \pi_a \) is strongly more approving than \( \pi_b \) if and only if \( F_{\pi_a} \) first order stochastically dominates \( F_{\pi_b} \). (ii) In the ACM, \( \pi_a \) is more approving than \( \pi_b \) if and only if \( E (X_{\pi_a}^{-}) \geq E (X_{\pi_b}^{-}) \).

**Proof.** (i) We first show that \( F_{\pi_a} \) first order stochastically dominates \( F_{\pi_b} \) if and only if \( \hat{F}_{\pi_a} \) first order stochastically dominates \( \hat{F}_{\pi_b} \). To see this, we recall the standard property that \( F_{\pi_a} \) first order stochastically dominates \( F_{\pi_b} \) if and only if

\[
\sum_{i, t} u (i, t) \pi_a (i, x_t) \geq \sum_{i, t} u (i, t) \pi_b (i, x_t)
\]

for all increasing functions \( u : \{1, \ldots, n\} \times \{1, \ldots, n\} \to \mathbb{R} \). Defining, for any such \( u \),

\[
\hat{u}_i : \{-n, \ldots, -1\} \times \{-n, \ldots, -1\} \to \mathbb{R} \text{ by } \hat{u} (-i, -t) = -u (i, t) ,
\]

the above statement means that, for all \( \hat{u} \) increasing in \(-i\) and \(-t\),

\[
\sum_{i, t} \hat{u} (-i, -t) \pi_a (i, x_t) \leq \sum_{i, t} \hat{u} (-i, -t) \pi_b (i, x_t)
\]

which is equivalent to \( \hat{F}_{\pi_b} \) first order stochastically dominating \( \hat{F}_{\pi_a} \). Suppose now that \( F_{\pi_a} \) first order stochastically dominates \( F_{\pi_b} \). Then for all \( x_t \) and \( \lambda \) we have

\[
F_{\pi_a} (\lambda (x_t), t) \leq F_{\pi_b} (\lambda (x_t), t) \Leftrightarrow \hat{F}_{\pi_b} (-\lambda (x_t), -t) \leq \hat{F}_{\pi_a} (-\lambda (x_t), -t)
\]

\[
\Leftrightarrow \pi_b \left( \{ l, x_j \} : \lambda (x_t) \leq l, t \leq j \} \leq \pi_a \left( \{ l, x_j \} : \lambda (x_t) \leq l, t \leq j \} \right.
\]

\[
\Leftrightarrow \sum_{j \geq 1} \sum_{l \geq \lambda (x_t)} \pi_a (x_j, l) \geq \sum_{j \geq 1} \sum_{l \geq \lambda (x_t)} \pi_b (x_j, l)
\]

\[
\Leftrightarrow p_{D_a}^D (x_t, \lambda) \geq p_{D_b}^D (x_t, \lambda)
\]

\[14\]For the bivariate case, see e.g. Troels and Østerdal [14] and the references therein.
proving one direction. For the other direction, given \(x_t\) and \(i\), choose any list \(\lambda\) such that \(\lambda(x_t) = i\) and run the above chain of implications from the end to show that \(\hat{F}_{\pi_b}(-i, -t) \leq \hat{F}_{\pi_a}(-i, -t)\) for all \(x_t\) and \(i\).

(ii) By the List Invariance Principle (Corollary 1), the total number of approvals in the ACM is the same for any list \(\lambda\), so we can compute it using the base list. Proceeding in this way, we have that for all \(\lambda\):

\[
\sum_i p^A_i(x_t, \lambda) = \sum_{i=1}^n \sum_{j \geq i} \sum_{l \geq b(\lambda; x_t, x_j)} \pi_a(l, x_j)
\]

\[
= \sum_{i=1}^n \sum_{j \geq i} \pi_a(l, x_j)
\]

\[
= \sum_{i=1}^n \pi_a \left( \{(l, j) : \min(l, j) \geq i\} \right)
\]

\[
= \mathbb{E} \left( X_{\pi_a}^- \right)
\]

where the second equality follows from choosing \(\lambda\) as the base list \(x_1 x_2 \ldots x_n\), and the fourth inequality from expressing the mean with the sum of complementary cdf formula for non-negative random variables.\(^{15}\) Deriving \(\sum_i p^A_i(x_t, \lambda)\) analogously, the claim follows.

Statement (i) in the theorem is intuitive, at least in one direction, if one thinks of first order stochastic dominance as shifting probability mass from lower to higher values of the variables: in our context, this means a generalised lowering of thresholds and raising of capacities, which leads to more approvals in each single list-item cell. The less immediate aspect of the result is that this increase in approvals can only occur through a first order stochastic dominance type of shift.

To understand why the variable \(X_{\pi}^-\) in statement (ii) plays the key role, observe that, as we argued in the proof of Corollary 1, this variable expresses the number of items that are approved at any capacity-threshold realisation. In fact, if fewer items than the capacity pass

\(^{15}\)That is,

\[
\sum_{i=1}^n \pi_a \left( \{(l, j) : \min(l, j) \geq i\} \right) = \sum_{i=1}^n \sum_{k=i}^n \pi_a \left( \{(l, j) : \min(l, j) = k\} \right)
\]

\[
= \sum_{k=1}^n \sum_{i=1}^k \pi_a \left( \{(l, j) : \min(l, j) = k\} \right)
\]

\[
= \sum_{k=1}^n k \pi_a \left( \{(l, j) : \min(l, j) = k\} \right).
\]
the threshold, then all the items that pass the threshold are approved. And if more items than the capacity pass the threshold, then the capacity determines the number of items approved. The result of Theorem 4 becomes intuitive when observing that the mean of $X_{\pi}^{-}$, in this interpretation, is simply the aggregate probability of approval $\sum_x p_a (x, \lambda)$.

**Remark 1.** The same argument in the proof of (ii) shows that in the DCM, $\pi_b$ is not more approving than $\pi_a$ if $\mathbb{E}(X_{\pi_a}^{-}) > \mathbb{E}(X_{\pi_b}^{-})$. Unlike the case of the ACM, for the DCM the relation might be incomplete (as in principle $\pi_a$ could generate more approvals than $\pi_b$ for one list and fewer with another). But the proof shows that in the DCM if the strict inequality $\mathbb{E}(X_{\pi_a}^{-}) > \mathbb{E}(X_{\pi_b}^{-})$ holds, then $\pi_a$ cannot be dominated for all lists.

### 6 Characterisation

#### 6.1 The Depth Constrained Model

For the DCM, we are able to provide a complete characterisation. Our axiomatisation will highlight that, beside obvious or technical properties, the core of the model is a form of complementarity between the quality and the position of an item. That is, the increase in approvals for an item when it is moved up in a list is larger the higher the quality of the item; or, equivalently, the difference in approvals between a better item and a worse one when they appear in the same position in two lists is larger the better the position.

All the axioms that follow here and in the next subsection are meant for all $x, y \in X$ and $\lambda, \lambda', \lambda'', \lambda''', \mu \in \Lambda$ unless a different quantification is detailed.

**A1. (Only the position matters)** If $\lambda (x) = \mu (x)$, then $p (x, \lambda) = p (x, \mu)$.

**A2. (Monotonicity in positions)** If $\lambda (x) < \mu (x)$, then $p (x, \lambda) > p (x, \mu)$.

**A3. (Supermodularity in quality and position)** If $\lambda (x) = \lambda' (y) = k$, $\lambda'' (x) = \lambda''' (y) = k - 1$ and $p (x, \lambda) > p (y, \lambda')$ then

$$p (x, \lambda'') - p (y, \lambda''') > p (x, \lambda) - p (y, \lambda').$$

Beside these natural properties, we also need to consider some technical properties to take care of the fact that we are working with strict preference orderings and strictly positive depth. We will show below how these technical properties are not really intrinsic to the model, in the sense that they can be completely dropped if these features are eschewed.
A4a. (Dominant item) There exist \( x \in X \) and \( \lambda \in \Lambda \) such that \( p(x, \lambda) = 1 \).

A4b. (Positivity) \( p(x, \lambda) > 0 \).

A4c. (Linearity) If \( p(x, \lambda) = p(y, \mu) \) and \( \lambda(x) = \mu(y) \) then \( x = y \).

**Theorem 5.** A stochastic approval function is a DCM if and only if it satisfies A1-A3 and A4a-A4c.

**Proof.** Necessity. Let \( p \) be a DCM generated by \( (\succeq, \pi) \). Take any \( x \in X \) and \( \lambda \in \Lambda \). A1 follows immediately from the formula for \( p^D(x, \lambda) \). Since \( \pi \) is strictly positive, \( p^D(x, \lambda) \) is strictly decreasing in \( \lambda(x) \) and A2 follows. To see A3, take distinct items \( x, y \) and lists \( \lambda, \lambda', \lambda'', \lambda''' \) such that \( \lambda(x) = \lambda'(y) = k, \lambda''(x) = \lambda'''(y) = k - 1 \) and \( p(x, \lambda) - p(y, \lambda') > 0 \). Then \( x \succ y \). Furthermore

\[
p(x, \lambda) - p(y, \lambda') = \sum_{x \succeq t \succ y} \sum_{i \geq k} \pi(i, t) < \sum_{x \succeq t \succ y} \sum_{i \geq k-1} \pi(i, t) = p(x, \lambda'') - p(y, \lambda''').
\]

A4a follows since if \( x \succeq y \) for all \( y \) and \( \lambda(x) = 1 \), then \( p(x, \lambda) = 1 \). A4b follows from the representation \( p(x, \lambda) = \sum_{x \succeq t} \sum_{i \geq \lambda(x)} \pi(i, t) \) and the assumption that \( \pi \) is strictly positive. A4c follows from the fact that if it was \( x \neq y \) then by the linearity of \( \succeq \) either \( x \succ y \) or \( y \succ x \). Say w.l.o.g. that \( x \succ y \). Then from the representation and \( \pi \) being strictly positive it should be \( p(x, \lambda) > p(y, \mu) \) when \( \lambda(x) = \mu(y) \), since \( y \succeq t \Rightarrow x \succeq t \) for all thresholds \( t \) but there exist \( t \) (e.g. \( t = x \)) for which \( x \succeq t \succ y \).

Sufficiency: Let \( p \) satisfy the axioms. Using A1 abuse notation and write \( p(x, k) = p(x, \lambda) \) for any \( \lambda \) such that \( \lambda(x) = k \). Define \( x \succeq y \) iff \( p(x, n) \geq p(y, n) \) or \( x = y \). By A4c, \( \succeq \) is a linear order.

Enumerate the items in decreasing order of preference (recall we can do this by Lemma 1). Note that the item \( x \) described in A4a is unique, \( \lambda(x) = 1 \) by A2, and \( x = x_1 \). To see this last point, suppose towards a contradiction that \( p(x_i, 1) \geq p(x_1, 1) \) for some \( i > 1 \). Repeated use of A3 yields \( p(x_i, n) \geq p(x_1, n) \), which contradicts the definition of \( \succeq \).

Define \( \pi \) as in the proof of Theorem 3, namely:

\[
\pi(n, x_n) = p(x_n, n)
\]
\[\pi(k, x_n) = p(x_n, k) - p(x_n, k + 1) \forall k < n\]

\[\pi(n, x_l) = p(x_l, n) - p(x_{l+1}, n) \forall l < n\]

\[\pi(k, x_l) = (p(x_l, k) - p(x_{l+1}, k)) - (p(x_l, k + 1) - p(x_{l+1}, k + 1)) \forall k, l < n\]

Observe that:

1. \(\pi(n, x_n) > 0\) by A4b.
2. \(\pi(k, x_n) > 0\) for all \(k < n\) by A2.
3. \(\pi(n, x_l) > 0\) for all \(l < n\) by A4c and the definition of \(\succeq\).
4. \(\pi(k, x_l) > 0\) for all \(k, l < n\) by A4c, the definition of \(\succeq\) and repeated use of A3.

Furthermore, \(\sum_k \pi(k, x_n) = p(x_n, 1)\) and \(\sum_k \pi(k, x_l) = p(x_l, 1) - p(x_{l+1}, 1)\) for all \(l < n\). It follows that \(\sum_l \sum_k \pi(k, x_l) = p(x_1, 1) = 1\). Hence \(\pi\) is a strictly positive joint probability over \(N \times X\) as desired.

To finish, let \(q\) be the DCM generated by \((\succeq, \pi)\). We need to show that \(p = q\). Take any \(x_k\) and \(\lambda\) and suppose that \(\lambda(x_k) = l\). We have

\[q(x_k, \lambda) = \sum_{l' \geq k} \sum_{k' \geq k} \pi(l', x_{k'}) = p(x_k, l) = p(x_k, \lambda)\]

where the second equality follows by construction.

The nature of a DCM is in fact revealed more transparently in a generalisation of the model in which:

1. indifferences are allowed;
2. the domain of \(\pi\) allows for the possibility that the approver examines none of the items;
3. approval probability of an item can therefore be be zero.

In the capacity-threshold realisation \((0, t)\), no item is examined or approved regardless of the threshold \(t\). Say that \(p\) is a general DCM (g-DCM) if there exists a pair \((\succeq, \pi)\) where \(\succeq\) is a weak order over \(X\), \(\pi\) is a probability over \((N \cup \{0\}) \times X\) and

\[p(x, \lambda) = \sum_{x \geq t} \sum_{i \geq \lambda(x)} \pi(i, t)\]

Consider:

**B1. (Only the position matters, generalised)** If \(\lambda(x) \leq \mu(x)\), then \(p(x, \lambda) \geq p(x, \mu)\).

**B2. (Supermodularity in quality and position, generalised)** If \(\lambda(x) = \lambda'(y) = k\),
\[ \lambda''(x) = \lambda'''(y) = k - 1 \text{ and } p(x, \lambda) \geq p(y, \lambda') \] then

\[ p(x, \lambda'') - p(y, \lambda''') \geq p(x, \lambda) - p(y, \lambda') . \]

**Theorem 6.** A stochastic approval function \( p \) is a \( g \)-DCM if and only if it satisfies B1 and B2:

**Proof:** Since B1 implies A2, we may write \( p(x, k) = p(x, \lambda) \) for any \( \lambda \) such that \( \lambda(x) = k \). Define \( x \succeq y \) iff \( p(x, n) \geq p(y, n) \) and note that \( \succeq \) is a weak order. Enumerate the items in decreasing order of preference as previously. For every \((i, t) \in N \times X\), we define \( \pi(i, t) \) as above. We define \( \pi_0 = 1 - p(x_1, 1) \) and choose \( n \) numbers \( \{\pi(0, x_i)\}_{i=1}^n \) which add up to \( \pi_0 \). Hence \( \pi(i, t) \geq 0 \) for all \((i, t) \in (N \cup \{0\}) \times X \) and \( \sum \pi(i, t) = 1 \) and \( \pi \) is a probability. The rest of the sufficiency argument is exactly the same as the proof of Theorem 5.

To see that the axioms are necessary, note that B1 follows because \( p(x, \lambda) \) depends only on \( \lambda(x) \) and in a (weakly) decreasing way. B2 follows for the same reason that A3 follows in the DCM, with the strict inequality replaced with weak inequality.

Thus, in the general model, even monotonicity in position need not be assumed separately. The model is entirely driven by the fact that only the position of an item matters, and by the complementarity between an item’s position and its quality.

### 6.2 The Approval Constrained Model

Giving a full characterisation of the ACM seems a hard problem and it remains open. We can however provide a partial characterisation, by pointing out that a set of axioms that parallel those used for the characterisation of the DCM is necessary for the ACM. The key is to note that in the ACM, for any alternative \( x \), the set of predecessors of \( x \) determines its approval probability, analogously to the way in which the position of \( x \) determines its approval probability in the DCM. For any \( \lambda \in \Lambda \), let \( P(x, \lambda) \) denote the set of predecessors of \( x \) in list \( \lambda \), that is, \( P(x, \lambda) = \{y \in X : \lambda(y) < \lambda(x)\} \).

**C1. (Only the predecessor set matters)** \( P(x, \lambda) = P(x, \mu) \Rightarrow p(x, \lambda) = p(x, \mu) \).

**C2. (Monotonicity in predecessor set)** \( P(x, \lambda) \subset P(x, \mu) \Rightarrow p(x, \lambda) \geq p(x, \mu) \).

**C3. (Submodularity in quality and predecessor set size)** If \( P(x, \lambda) = P(y, \lambda') = P \supset Q = P(x, \lambda'') = P(y, \lambda'''), \) and \( p(x, \lambda) > p(y, \lambda') \) then

\[ p(x, \lambda'') - p(y, \lambda''') \geq p(x, \lambda) - p(y, \lambda') . \]
It is not difficult to check that C1-C3 are necessary properties for the ACM. It also easy to see that A4a and A4b are satisfied without any change, and that a suitably modified version of A4c must be satisfied. Finally, recall that the List Invariance Principle (Corollary 1) can also be taken as a fundamental necessary property of the ACM.

These conditions are satisfied also by approval procedures that are not an ACM. For example, suppose that each item $x$ is approved with probability $p_x > 0$, depending on $x$ but independently of the list in which $x$ appears; and that $p_x \neq p_y$ for all distinct items $x, y$. This approval procedure satisfies all the properties mentioned above. Yet there can be no ACM that generates these probabilities. In fact, if there was one, the preference relation should agree with the ranking determined by the $p_x$. But then, by adding preferred items to the predecessor set of an $x$, its approval probability should strictly decrease according to the ACM, whereas it remains fixed at $p_x$ in the procedure we have described.

While falling short of a complete characterisation, the properties may nevertheless be useful as they offer simple ways to reject the model on the basis of the data.

7 Concluding remarks

It is a common observation that the nature of data available to social scientists has undergone a profound change in recent years. This is meant both in the sense that behaviours that were not observable before are now observable, and in the sense that people engage in new types of behaviours. While of course many specific studies of “e-activities” exist, our attempt has been to draw together in an abstract model disparate approval activities and datasets that are typical of online life.

While forms of approval are not exclusive to online life, in this paper we have highlighted the online interpretation of the data for three reasons. First, online approval behaviour is much more typical than in physical life, because of the quick and free or near free nature of clicks. Second, approvals as clicks are far more easily observable in a systematic way than than physical approvals. Thirdly, clicks have an economic value to some interested parties beside the final choice; unlike, say, an item placed in a physical cart and then put back on the shelf, or the verbal approval of a friend’s remark (as opposed to a Facebook Like).

A main virtue of our models is that their psychological primitives can be uniquely and non-parametrically identified from observed data, in one case without any restriction.
and on the basis of surprisingly few observations; and in the other case only subject to an independence assumption.

Importantly, this approach has led us to tackle an issue - the “manipulation” of online behaviour by interested parties - for which we mostly lack a formal framework of analysis in spite of it being a key topic in current discourse. In our model this has taken the concrete form of the “list design problem”. We have highlighted several non-obvious tradeoffs and principles of list design. However, the analysis of this issue needs much more investigation and seems a most promising avenue for future research. The objectives of list designers may go beyond the simple weighted sum maximisation we have studied. For example, in a recent prominent case, Facebook was publicly condemned for conducting an experiment in which it manipulated nearly 700,000 users’ news feeds to see whether it would affect their emotions. 16 This type of more sophisticated objectives could be analysed in suitable extensions of our framework. Also, manipulation can leverage on aspects of the environment (such as visual salience) that go beyond lists and their content.

The models themselves call for numerous variations and extensions. Three that seem to us among the most interesting are the following. First, one might consider a version of the setting that brings out explicitly the dynamic content of behaviour. Approval behaviour is captured as a set of pairs \((A_k, \rho_k)\), where \(A_k\) (resp., \(X_k\)) is the set of items approved (resp., encountered) up to stage \(k\) and \(\rho_k : X_k \times A_k \times (X \setminus A_k) \rightarrow [0, 1]\) tells the probability of approving at stage \(k\) any item that has not yet been approved. This framework, which allows for example items to show up repeatedly until they are approved and dynamic thresholds that vary with the items encountered up to any given stage as well as with the stage itself, would substantially enrich the range of behaviours encompassed by the models. Second, with “click-reactions” in mind, approval might be more complex than a binary action: approvals could express a “rating intensity” (e.g. through comments) or have a qualitative dimension (e.g. as in the various types of Facebook or Instagram Likes). A third possible extension could be the introduction of stochastic preferences. It seems sensible, at least as a first step, to treat preferences as being more stable and fundamental than other factors that affect capacity and threshold, such as mood and occasional interruptions. From the perspective of empirical implementation, however, randomness

---

16 Kramer et al. [10] ran an experiment, in agreement with Facebook, by manipulating the News Feed of 689,003 Facebook users to alter their exposure to various emotional expressions (which led to an Editorial Expression of Concern in PNAS). Other manipulations, which Facebook argued it was not explicitly aware of, were operated e.g. by Cambridge Analytica. Exposed by the British newspaper The Guardian, Cambridge Analytica closed down a few months after the scandal broke out (see https://www.theguardian.com/news/series/cambridge-analytica-files).
in preferences would reflect the availability of population rather than individual data and the presence of heterogeneity. It is of note that the List Invariance Principle would continue to hold in this extension of the Approval Constrained Model: given that the sum of the approval probabilities for each preference type is fixed across lists, it must remain fixed when aggregating across preference types.

References


