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

Paola Manzini[1]
University of Sussex and IZA

Marco Mariotti[2]
Queen Mary University of London

Levent Ülkü
ITAM[3][4]

This version: January 2021

[1]Corresponding author. Department of Economics, University of Sussex Business School, University of Sussex, Falmer, Brighton, BN1 9SL, UK, email:p.manzini@sussex.ac.uk

[2]School of Economics and Finance, Queen Mary University of London, Mile End Road, London E1 4NS, UK, email: m.mariotti@qmul.ac.uk

[3]Centro de Investigación Económica, Department of Economics, ITAM, Río Hondo 1, Ciudad de México 01080, Mexico, email: levent.ulku@itam.mx

[4]This version generalises and supersedes the previous version dated July 2019. The models in that version are now removed. We thank for useful comments and suggestions the co-Editor Bart Lipman and three anonymous referees. We also thank Miguel Angel Ballester, Kalyan Chatterjee, Olivier Compte, Mira Frick, Yuhta Ishii, Rubén Martínez, Rich McLean, 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.

Abstract

We consider a model of "approval behaviour" (like Favouriting, Sharing or Wish-listing) for items presented as a list. The approver proceeds to each successive item with a continuation probability that depends on the history of predecessors, approving along the way any item that is considered "acceptable" (e.g., possessing a set of key properties). The procedure is the only one that satisfies Predecessor Monotonicity (the approval probability of an item is non-increasing when the sublist of predecessors lengthens) and Approval Luce Independence (the approval odds for two items when they have the same sublist of predecessors are invariant to the exact identity and listing of the predecessors). The primitives of the model are substantially identified from behaviour. If there are correlations between continuation and acceptability, then Luce Approval Independence is replaced in the characterisation by Monotone Differences (an item's revealed quality and its position in the list are complementary in the production of approval). Finally, we explore the notion of "list design".

J.E.L. codes: D0.

Keywords: Approval behaviour, List design.

1 Introduction

Consider an agent who:

- progressively fills her online shopping cart or wish list;
- "Likes", "Favourites" or shares items in a social network;
- scans headlines from a newsfeed and stores some articles for later reading;
- views abstracts and downloads items from a working paper archive list;
- tags a website page on a social bookmarking site;
- matches with potential partners on a dating site.

Many online technologies both encourage and make observable a category of behaviour that can be summarised as "approval", as distinct from "choice" meant as a final selection. Often the issue of a final selection simply does not arise: the approver needs not award a prize to the best post she has Liked or to the best page she has tagged. In other examples (online shopping, dating and scientific research) the items that are clicked typically comprise only a tentative and preliminary selection, with the final selection possibly postponed to a subsequent stage. In this case approval can be seen as the phase of construction of a "consideration set" during the process of choice.[1] The entire set of clicked items may be easily discarded: wish lists can be ignored and no-choice in the form of cart abandonment is typical of online shopping.[2]

When considering such examples, two issues arise. First, how can we *describe* approval data? Second, how can we *explain (or represent)*[3] approval behaviour? The analogous questions in standard choice theory are answered, respectively, by: "a choice function" and "preference maximisation". The conditions ensuring that a

[1]E.g. Eliaz and Spiegler [9] and Masatlioglu, Nakajima and Ozbay [19] for the deterministic case, and Manzini and Mariotti [18], Brady and Rehbeck [3] and Aguiar [1] for the stochastic case.

[2]Cart-abandonment rates are thought to be as high as 70%. See e.g. https://baymard.com/lists/cart-abandonment-rate.

[3]The issue of whether a formal representation is just a summary of behaviour, or it refers instead to substantive objects with an explanatory value is a large philosophical question that we do not address here, viewing the primitives as an explanation without further discussion. The interested reader is referred to the brilliant treatment by Dietrich and List ([8]).

choice function can be represented as preference maximisation, and the way unobserved preferences can be identified from observed choice, are well-understood. In this paper we aim to carry out a similar general exercise for approval, abstracting from the details of any specific context.

To describe approval data, we introduce the notion of an *approval function*, which maps any list $x_1 x_2 ... x_n$ of items from a fixed menu into a set of arbitrary probabilities p_i, one for each item x_i. The domain of an approval function emphasises the fact that in the examples items typically present themselves to the approver in a sequential manner (or the approver scans the items in some order that is observable, as in Caplin *et al.* [5]).[4] The codomain of an approval function emphasises the fact that items are not "alternatives" to each other. Unlike a stochastic choice function from lists (introduced by Rubinstein and Salant [20]) the p_i are not required to form a probability distribution on the menu. Any intermediate case is possible between adding up to zero (nothing is approved) or to the cardinality of the menu (everything is approved).

To explain approval data, we posit that two broad types of factors motivate a generic approver. First, the drive at any point in the list to continue, depending on what the approver has seen so far. This may reflect time, budget or cognitive constraints, as well as beliefs on what is to come. The second type of factor is the approver's "intrinsic taste" for the objects, namely what she considers as worthy of approval. These two factors are modelled as:

(a) a *continuation function* π, which gives the probability to continue to the next item as a function of the entire sublist of predecessors.

(b) an *acceptability function* σ, which gives the probability of any partition of the items into "worthy" and "not worthy".

This formulation is quite general and capable of encompassing many specific theories of behaviour and objective functions. The mild monotonicity assumption we will make on π (later items are less likely to be reached) is compatible with many plausible psychological hypotheses on how the browsing history affects the desire of an approver to reach further. For example, we admit both "optimism"

[4] In 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.).

(high predecessor quality stimulates further exploration) and "pessimism" (high predecessor quality discourages further exploration). We can handle the typical case of very long lists, effectively infinite from the point of view of the approver, by embedding in π an endogenous or exogenous depth of exploration.

Similarly, the acceptability structure is unrestricted and consistent with a host of different motivations for approving items.[5] What we consider common to all these motivations at an abstract level is that the type of judgement involved in approval is a (binary) *classification* ("shareable", "considerable", "Likeable", and so on) rather than a *ranking*. In other words, we consider the approver as akin to an algorithm that designates emails as "spam" or "not spam", without putting them through a finer quality mesh. A natural model for this is that the approver has in mind a set of properties: possessing these properties qualify an item as acceptable.[6] For example, consider someone who is browsing books to ultimately select one for a gift. With an approver hat on, she decides to put in the online shopping cart a book written by an author she likes, one that has high user ratings, and one that is the best-seller in the genre. Later on, with a chooser hat on, an accurate tradeoff will perhaps have to be made, if she decides to settle on a specific book, but such fine balancing is not necessary during the construction of a consideration set. Or, perhaps, she will discard the entire cart and instead choose a coffee-grinder as the gift. Similarly, a researcher downloads for later study all the new papers in decision theory, and another paper because it is by John Nash. The difference with the gift-giver is that the researcher will not have to choose between papers.

Our model combines the two factors by positing that the approver scans the items along the list, moving from one item to the next with the probability prescribed by the continuation function, and approving all items along the way as prescribed by the acceptability function.

The first issue we tackle is *characterisation*: which exact constraints on behaviour

[5]In the examples, such motivations range widely, from communication (Liking and Favouriting are online "nods" that may signal interest, agreement, and so on) to storage (obviously so for downloading or bookmarking items, but note that Favouriting is a form of bookmarking that facilitates future access to items posted earlier).

[6]At the formal level, this can also be seen as the first step in the construction of a preference or utility, in view of Chipman's [6] representation of utility as a lexicographic sequence of $0-1$ judgements.

does the model pose? The model is characterised by two properties (Theorem 1). The first (Predecessor Monotonicity) simply says that the approval probability of an item does not increase by lengthening the sublist of its predecessors. The second (Approval Luce Independence) is an adaptation of the standard Luce Independence axiom for stochastic choice functions to the approval environment. It says that the approval odds of two items appearing after the same predecessors in two lists do not depend on the identity of the predecessors or on how they are listed. Furthermore, the theorem also shows that an approval function satisfying these two properties is equivalent to one where an item is acceptable whenever it passes a stochastic threshold according to a fixed, deterministic preference order. In this sense the result provides a foundation for *satisficing behaviour* in approval.

Next, we move to *identification*. Assuming that approvals are generated by the model, to what extent can an observer of approvals and lists identify the acceptability and the continuation functions, or the preferences and the thresholds in the satisficing version? We show that our models have excellent identification properties. For the satisficing version of the model, preferences are uniquely identified. The continuation function and the distribution of the satisficing threshold are also uniquely identified under a mild support restriction (Theorem 2). For the general version, under a mild support restriction the continuation function is uniquely identified by behaviour and the acceptability function is identified up to marginalisation, that is, the marginal probability that each item is acceptable is identified (Theorem 3).

Next, we proceed to consider an extension of the model. What if continuation is not independent of what has been approved so far? For example, it might make sense that the agent stops adding to the cart when the cost reaches a certain point. We characterise the model with this type of correlations in its satisficing version. It turns out (Theorem 4) that the only observable difference between the independent and the correlated versions of the model is that the Approval Luce Independence property is replaced by the Monotone Differences property. Roughly, this property can be interpreted as saying that the revealed "quality" of an item and its position in a list are complements in the production of approval probabilities.[7] On the other

[7] Throughout the paper, we will use "quality" to describe the rank of an item in the preference order.

hand, the characterisation of the non-satisficing version (which we show by example not to be equivalent to the satisficing version in this case) is a hard problem which so far remains open.

We show finally 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 some specifications of the models and of the objective functions (Theorem 5 and 6).

Two short discussion sections (in one of which we explore further the relation to the literature and the differences between approval and choice) conclude the paper.

2 The Model

2.1 Notation and definitions

Notation Let X be a finite set of items with cardinality n and let 2^X denote the power set of X including the empty set.[8] A *list* is a linear order[9] λ on X, and a *sublist* is a linear order $\tilde{\lambda}$ on any strict subset of X. Let Λ and $\tilde{\Lambda}$ be the set of all lists and sublists respectively.

Sometimes we view lists and sublists as strings of items, $\lambda = x_1x_2x_3...$, and sometimes we abuse notation and treat lists and sublists as sets. Hence $x \in \tilde{\lambda}$ ($x \notin \tilde{\lambda}$) means that x does (not) appear in sublist $\tilde{\lambda}$, $|\tilde{\lambda}|$ denotes the number of items in $\tilde{\lambda}$, and $\tilde{\lambda} \setminus x$ with $x \in \tilde{\lambda}$ denotes the sublist obtained from $\tilde{\lambda}$ by removing item x and leaving the order on the remaining items intact. For any $A \subseteq X$, $A \cap \tilde{\lambda}$

[8] We use the term "items" rather than "alternatives" to emphasise that, within the capacity constraint, there is no "competition" between the items.

[9] A linear order is a complete, transitive and antisymmetric binary relation on X.

denotes the subset of A whose members appear in $\tilde{\lambda}$, and $\tilde{\lambda} \cup A$ denotes the subset of X whose members appear in $\tilde{\lambda}$ or in A.

For any $x \in X$ and $\lambda \in \Lambda$, $\lambda(x) \in \{1, ..., n\}$ is the position of x in list λ. Similarly $\tilde{\lambda}(x)$, with $x \in \tilde{\lambda}$, is the position of x in sublist $\tilde{\lambda}$.

Every $x \in X$ and $\lambda \in \Lambda$ induce a sublist $\lambda^x \in \tilde{\Lambda}$ on the strict predecessors of x in λ as follows: $\lambda^x(y) = \lambda(y)$ for every y such that $\lambda(y) < \lambda(x)$. Hence λ^x is the sublist obtained from λ by eliminating x and all its successors. If x is the first item in λ, i.e. if $\lambda(x) = 1$, then λ^x is the empty list, which we denote by $\emptyset$.

We define a partial order $\sqsubseteq$ on $\tilde{\Lambda}$ as follows: $\tilde{\lambda} \sqsubseteq \tilde{\mu}$ if and only if for every $x \in \tilde{\lambda}$, $x \in \tilde{\mu}$ and $\tilde{\mu}(x) = \tilde{\lambda}(x)$. Note that $\sqsubseteq$ is antisymmetric, i.e. if $\tilde{\lambda} \sqsubseteq \tilde{\mu}$ and $\tilde{\mu} \sqsubseteq \tilde{\lambda}$, then $\tilde{\lambda} = \tilde{\mu}$. We denote the strict part of $\sqsubseteq$ by $\sqsubset$, i.e. $\tilde{\lambda} \sqsubset \tilde{\mu}$ if and only if $\tilde{\lambda} \sqsubseteq \tilde{\mu}$ and $\tilde{\lambda} \neq \tilde{\mu}$.

We consider an agent who is presented with several lists on different occasions. The agent scans the list and at each item continues with some probability, deciding whether or not to approve any encountered item. The observed behaviour is recorded by an approval function:

Definition 1. An *approval function* is a map $p : X \times \Lambda \to [0, 1]$.

The quantity $p(x, \lambda)$ is the probability that the agent approves item x in list λ. Note that, unlike for a standard stochastic choice function, we do not impose the adding-up constraint $\sum_{x \in X} p(x, \lambda) = 1$ and we allow for the possibility of approving nothing, i.e. $p(x, \lambda) = 0$ for all $x \in X$. What is more, our domain comprises lists, not menus. The menu X is held fixed in the analysis. The variation comes *only* from lists.[10]

We now describe the (unobserved) primitives we use to explain an approval function.

Definition 2. A *continuation function* is a map $\pi : \tilde{\Lambda} \to [0, 1]$ satisfying

(i) $\pi(\emptyset) = 1$;

(ii) $\pi(\tilde{\lambda}) \geq \pi(\tilde{\mu})$ if $\tilde{\lambda} \sqsubset \tilde{\mu}$.

[10] Yet a different type of dataset would be given by a stochastic *approval correspondence* $P : 2^X \times \Lambda \to [0, 1]$ associating with each list the probability of the possible *approval sets*. Then the adding-up constraint applies over the $P(A, \lambda)$, with $A \in 2^X$. This type of data would incorporate much richer information on the correlations between approvals across lists. Our identification results below would hold a fortiori if an observer had access to this information.

A continuation function tells, for any sublist $\tilde{\lambda}$, the probability that the agent continues after seeing the sublist of predecessors $\tilde{\lambda}$. Restriction (i) says that the agent will examine at least the first item in any list. Restriction (ii) says that the agent follows the list: a later item is never more likely to be reached than an earlier one. Note that the $\pi(\tilde{\lambda})$ are *not* assumed to add up to one.

Definition 3. An *acceptability function* is a probability distribution $\sigma \in \triangle\left(2^X\right)$.

The interpretation is that $\sigma(A)$ is the probability that the agent finds acceptable all the items in $A \subseteq X$ and unacceptable all the others. Thus, σ captures the "tastes" of the agent.

In our model, an approver is a pair (σ, π) of an acceptability and a continuation function.

Definition 4. An approval function p is an *acceptability-continuation (AC) function* if there exists a pair (σ, π) of an acceptability function and a continuation function such that, for all $x \in X$ and $\lambda \in \Lambda$:

$$p(x, \lambda) = \pi(\lambda^x) \sum_{A: x \in A} \sigma(A) \tag{1}$$

In this case we say that p is generated by (σ, π).

2.2 Examples

We provide below three examples of AC functions. In the first two examples the continuation function is obtained endogenously as the solution to an optimisation problem, while in the last one it is a function of exogenous parameters.

Example 1. (Optimal continuation probabilities I: Approving everything that is acceptable) Suppose that each item is independently acceptable with probability $q \in (0, 1)$; and at any point in a list, the agent believes that the next item is acceptable with the same probability q. The aim of the agent is to maximise the probability of approving all items that are acceptable without going any further down the list. More precisely, the agent aims to maximise the probability of stopping exactly at the last acceptable item along any $\lambda \in \Lambda$, approving any acceptable item till then. For example, a reader may wish download all articles that are worthwhile reading, while keeping to a minimum the total number of articles scanned for downloading.

Fix a list $\lambda = x_1 x_2 ... x_n$. If $q > 1/2$, then it is clearly optimal to continue along the list to the end. Suppose that $q \leq \frac{1}{2}$. The solution to this problem is given by an application of the "odds theorem" (Bruss [4]). In our context the theorem prescribes to stop at the first acceptable item x_i, if any, such that the the sum of the odds $\frac{q}{1-q}$ of being acceptable of the remaining items is less than one. Formally, define

$$i^* (q) = \max \left\{ 1, \max \left\{ i : n \geq i \geq 1 \text{ and } \frac{q}{1-q} (n - i + 1) \geq 1 \right\} \right\} \tag{2}$$

where we set $\max \{\emptyset\} = 0$. Then it is optimal to stop at the first acceptable item x_i for which $i \geq i^* (q)$. A corresponding AC function generated by (σ, π) is obtained by setting:

$$\begin{cases} \pi \left(\lambda^{x_i} \right) = 1 & \text{if } i \leq i^* (q) \\ \pi \left(\lambda^{x_i} \right) = 1 - q & \text{if } i > i^* (q) \end{cases} \tag{3}$$

and

$$\sigma (A) = q^{|A|} (1 - q)^{|X \setminus A|} \tag{4}$$

Equations (3) say that the agent reaches for sure the item at the critical position $i^* (q)$, i.e. she continues at any item that belongs to the sublist $\lambda^{x_{i^*(q)}}$ of predecessors of $x_{i^*(q)}$. Then, at each successor item, the agent only continues after any λ^{x_i} if item x_{i+1} (if it exists) is not acceptable, which happens with probability $(1 - q)$. Note that the probability of stopping exactly at item x_i with $i \geq i^* (q)$ is $q (1 - q)^{i - i^*(q)}$: item x_i must be acceptable, which happens with probability q, while all strict predecessors between $i^* (q)$ and i (if any) must be unacceptable, which happens with probability $(1 - q)^{\left|\lambda_{x_i}\right| - \left|\lambda_{x_{i^*(q)}}\right|}$.[11]

Example 2. (Optimal continuation probabilities II: Not missing the best) Sup-

[11] For simplicity we have implicitly assumed that the agent knows the number of items n in solving the problem. Adaptations of the same analysis hold if we let $n = \infty$ (assigning to non-stopping some payoff smaller than that corresponding to the objective), or if we consider a random horizon (see Ferguson [11] and the references therein).

pose that the agent does not want to miss a "key" piece of news, or the best product when constructing a wish list; but is uninterested in finding other acceptable items further down the list. More precisely, the agent has preferences represented by a linear order $\succeq$ on X.[12] He views a list as ordering the items uniformly randomly. Given a list $\lambda = x_1 x_2 ... x_n$, the agent wishes to maximise the probability that the last item x_j she approves is the maximiser of $\succeq$. She finds acceptable any x_i such that $i \leq j$ and $x_i \succeq x^*$ for some fixed $x^* \in X$. The stopping structure is the classical "Secretary problem".[13] It is well known that for any $j = 1, ..., n$ the probability $\Pr\left\{x_j \succ x_i \; \forall i < j\right\}$ that x_j is the best item of all those seen up to that point is $\frac{1}{j}$, and that these events are independent across the x_j. Letting

$$i^* = \max\left\{1, \max\left\{i : n \geq i \geq 1 \text{ and } \sum_{j=i}^{n} \frac{1}{j-1} \geq 1\right\}\right\}$$

it is optimal to stop at the first item x_i (if any) for which $i \geq i^*$ and $x_i \succ x_j$ for all $j < i$.[14] At each x_i for which $i \geq i^*$ he continues with the probability that x_i is not the best item seen so far, which is $1 - \frac{1}{i} = \frac{i-1}{i}$. A corresponding AC function generated by (σ, π) is obtained by setting:

$$\begin{cases} \pi\left(\lambda^{x_i}\right) = 1 & \text{if } i \leq i^* \\ \pi\left(\lambda^{x_i}\right) = \frac{i-2}{i-1} & \text{if } i > i^* \end{cases} \tag{5}$$

and

$$\sigma(A) = 1 \iff A = \{x \in X | x \succeq x^*\}.$$

To calculate $\pi\left(\lambda^{x_i}\right)$ in the second line, recall that the last term of λ^{x_i} is x_{i-1}, with $i > 1$, so that to continue it must be the case that x_{i-1} is not the best item seen so far, which happens with probability $\frac{i-2}{i-1}$.

Example 3. (Independent acceptability and exogenous random depth) Let $\hat{\pi}$: $\{1, 2, ..., n\} \to [0, 1]$ be a probability distribution, where $\hat{\pi}(m)$ is interpreted as the

[12] This is meant in the sense that while the agent does not know the composition of X, he can rank any pair of items presented to him.

[13] See e.g. Ferguson [12] for a comprehensive treatment of this and related models.

[14] While this solution was developed much earlier, it can be seen as an application of the Bruss odds-theorem discussed above.

probability that the agent's depth of list exploration is m: independently of the list faced, all items in positions up to and including i are examined, and no other item is. For example, the agent may stop browsing the list when distracted by emails or other claims to her attention. A corresponding continuation function is defined by setting:

$$\pi\left(\tilde{\lambda}\right) = \sum_{m > |\tilde{\lambda}|} \hat{\pi}\left(m\right)$$

Let $\gamma \in \Delta\left(X\right)$ be a probability distribution on X, and define an acceptability function σ by

$$\sigma\left(A\right) = \prod_{x \in A} \gamma\left(x\right) \prod_{x \notin A} \left(1 - \gamma\left(x\right)\right)$$

With a random depth structure, dependence on the set of examined items boils down to dependence on its *size*. In other words, continuation depends only on the *position* of the item to be reached and not on the the composition of its predecessor set. As for acceptability, each item x has its own probability $\gamma\left(x\right)$ of being acceptable, and the set A is acceptable with the probability the all its members turn out to be acceptable, regarded as independent events.

3 Characterisation and Identification

3.1 Characterisation

In this section we show that, in spite of its generality, the model poses stark restrictions on observable behaviour, and that it always admits a "canonical" preference based interpretation. That is, while there may be various specialisations of the general framework, it is *always* possible to express approval as *preference satisficing* behaviour. Note that such canonical preference emerges from the model, and is not assumed *a priori* in the acceptability function, as we will see below.

We first introduce the two axioms that will be used in the characterisation.

A1. (Predecessor Monotonicity) If $\lambda^x \sqsubseteq \mu^x$, then $p\left(x, \lambda\right) \geq p\left(x, \mu\right)$.

A2. (Approval Luce Independence) For all $x, y \in X$ and all $\lambda, \mu, \alpha, \beta \in \lambda$ for which

$\lambda^x = \alpha^y$, $\mu^x = \beta^y$: if $p(x,\mu) > 0$ and $p(y,\beta) > 0$, then

$$\frac{p(x,\lambda)}{p(x,\mu)} = \frac{p(y,\alpha)}{p(y,\beta)}.$$

A1 says that the probability of approval of an item cannot increase if the set of its predecessors expands, keeping the positions of its original predecessors intact. Note that this is not the same thing as assuming that approval probabilities are monotonic in the *position* of the item in the list. It is important that the later position is obtained by *adding* items to the intact sublist of predecessors. In principle an item can be less likely to be approved when placed in second position after a "discouraging" item than when placed in fourth position after three "encouraging" items.

A2 says (in instances where all approval probabilities are strictly positive) that the approval odds of two items x and y that appear after the same predecessors in two lists λ and α, respectively, do not depend on the identity of these items or on how they are listed. Evidently, this is is a close analog in the list environment of the classical Luce axiom for standard stochastic choice functions. A2 extends this property to some zero-probability cases, by expressing it in the form of the impact that the change of sublists has on different items.

Next, we introduce a seemingly much more restrictive version of our model. For notational convenience, let $x_0 \notin X$ denote a pseudo-item to capture the case when nothing in the list is good enough.

Definition 5. An approval function p is a *satisficing AC (SAC) function* if there exists a triple $(\succsim, \tau, \pi)$, where $\succsim$ is a weak order on X, τ is a probability distribution on $X \cup \{x_0\}$, and π is a continuation function, such that for all $x \in X$ and $\lambda \in \Lambda$,

$$p(x,\lambda) = \pi(\lambda^x) \sum_{t: x \succsim t} \tau(t). \tag{6}$$

In this case we say that p is generated by $(\succsim, \tau, \pi)$.

The weak order $\succsim$ is interpreted as a standard deterministic preference relation over items and t as a random (ordinal) satisfaction threshold, such that a set of items A is acceptable if and only if it consists of all the items that are at least as good as the threshold.

Theorem 1. *The following three statements on an approval function p are equivalent:*

(1) p satisfies A1 and A2.

(2) p is a SAC function.

(3) p is an AC function.

Proof. (1) $\implies$ (2). Suppose that p satisfies A1 and A2. Note that, by A1, $p(x,\lambda) = p(x,\mu)$ if $\lambda^x = \mu^x$ (this must be the case since $\lambda^x \sqsubseteq \mu^x$ implies $p(x,\lambda) \geq p(x,\mu)$, and $\mu^x \sqsubseteq \lambda^x$ implies $p(x,\mu) \geq p(x,\lambda)$). In particular, the approval probability of an $x \in X$ for which $\lambda(x) = 1$ does not depend on how λ lists the remaining items. Thus we can well-define $\rho(x) = p(x,\lambda)$ for any $x \in X$ and $\lambda \in \Lambda$ satisfying $\lambda^x = \emptyset$. We will define the two maps $\tau : X \cup \{x_0\} \to [0,1]$ and $\pi : \tilde{\Lambda} \to [0,1]$ distinguishing two cases. <u>Case 1:</u> $\rho(z) = 0$ for all $z \in X$. By A1, $p(z,\lambda) = 0$ for all $z \in X$ and $\lambda \in \Lambda$. Let $\tau(t) = 0$ for all $t \in X$, let $\pi(\lambda^x) = 1$ for all $x \in X$ and all $\lambda^x \in \tilde{\Lambda}$ and fix an arbitrary weak order $\succsim$ on X. Let $p^{(\tau,\pi,\succsim)}$ be the SAC function generated by these primitives. We have, for all $x \in X$ and $\lambda^x \in \tilde{\Lambda}$,

$$p^{(\tau,\pi,\succsim)}(x,\lambda) = \pi(\lambda^x) \sum_{t:x\succsim t} \tau(t) = 0 = p(x,\lambda),$$

as desired. <u>Case 2:</u> $\rho(z) > 0$ for some $z \in X$. Fix such $z \in X$. For any $\tilde{\lambda} \in \tilde{\Lambda}$, let $\pi(\tilde{\lambda}) = \frac{p(z,\lambda)}{\rho(z)}$ for any $\lambda \in \Lambda$ such that $\lambda^z = \tilde{\lambda}$. It follows that $\pi(\emptyset) = 1$. Furthermore, if $\tilde{\lambda} \sqsubseteq \tilde{\mu}$, then for any $\lambda, \mu \in \Lambda$ such that $\lambda^z = \tilde{\lambda}$ and $\mu^z = \tilde{\mu}$,

$$\pi(\tilde{\lambda}) = \frac{p(z,\lambda)}{\rho(z)} \geq \frac{p(z,\mu)}{\rho(z)} = \pi(\tilde{\mu})$$

where the inequality is due to A1. Next, for any $x, y \in X$, let $x \succsim y$ if and only if $\rho(x) \geq \rho(y)$. Since it has a numerical representation, $\succsim$ is a weak order. Let $\sim_1, ..., \sim_k$ be the indifference classes of $\succsim$ ordered from best to worst, and fix $x_1, ..., x_k$

such that $x_i \in \sim_i$ for every i. Define

$$\tau(t) = \begin{cases} \frac{\rho(x_i)-\rho(x_{i+1})}{|\sim_i|} & \text{for every } t \in \sim_i, i < k, \text{ and} \\ \frac{\rho(x_k)}{|\sim_k|} & \text{for every } t \in \sim_k . \end{cases}$$

Note that $\tau(t) \geq 0$ for every $t \in X$ and $\sum_{t\in X} \tau(t) = \rho(x_1) \leq 1$. Letting $\tau(x_0) = 1 - \rho(x_1)$, we have $\tau \in \triangle(X \cup \{x_0\})$. Note also that if $x \in \sim_i$, then $\sum_{t:x\succsim t} \tau(t) = \rho(x_i)$. Let $p^{(\tau,\pi,\succsim)}$ be the SAC function generated by these primitives. Fix any $x \in X$ and $\lambda \in \Lambda$. Let $x \in \sim_i$ so that $\rho(x) = \rho(x_i)$ and let $\mu \in \Lambda$ satisfy $\mu^z = \lambda^x$. If $\rho(x) > 0$ then

$$\begin{aligned} p^{(\tau,\pi,\succsim)}(x,\lambda) &= \pi(\lambda^x) \sum_{t:x\succsim t} \tau(t) \\ &= \frac{p(z,\mu)}{\rho(z)} \sum_{t:x\succsim t} \tau(t) \\ &= \frac{p(x,\lambda)}{\rho(x)} \rho(x_i) \\ &= p(x,\lambda) \end{aligned}$$

where the third equality is by A2. If $\rho(x) = 0$, on the other hand, then $p(x,\lambda) = 0$ by A1. In this case, $x \in \sim_k$, the lowest indifference class, and furthermore by definition of τ, $\tau(t) = 0$ for every $t \in \sim_k$. It follows that

$$p^{(\tau,\pi,\succsim)}(x,\lambda) = \pi(\lambda^x) \sum_{t:x\succsim t} \tau(t) = \pi(\lambda^x) \sum_{t\in\sim_k} \tau(t) = 0.$$

(2) $\implies$ (3). Let $\sigma(A) = \tau(t)$ if $A = \{x \in X : x \succsim t\}$ for some $t \in X$, $\sigma(\emptyset) = 1 - \sum_{t\in X} \tau(t)$, and $\sigma(A) = 0$ otherwise. It follows that, for all $x \in X$ and $\lambda \in \Lambda$,

$$p(x,\lambda) = \pi(\lambda^x) \sum_{t:x\succsim t} \tau(t) = \pi(\lambda^x) \sum_{A:x\in A} \sigma(A),$$

as desired. (3) $\implies$ (1). For A1, note that if $\lambda^x \sqsubseteq \mu^x$, then $\pi(\lambda^x) \geq \pi(\mu^x)$ and consequently

$$p(x,\lambda) = \pi(\lambda^x) \sum_{A:x\in A} \sigma(A) \geq \pi(\mu^x) \sum_{A:x\in A} \sigma(A) = p(x,\mu).$$

For A2, note that if $\lambda^x = \alpha^y$, $\mu^x = \beta^y$, $p(x,\mu) > 0$ and $p(y,\beta) > 0$, then

$$\frac{p(x,\lambda)}{p(x,\mu)} = \frac{\pi(\lambda^x)}{\pi(\mu^x)} = \frac{\pi(\alpha^y)}{\pi(\beta^y)} = \frac{p(y,\alpha)}{p(y,\beta)}.$$

□

Preference satisficing imposes a strong structure on the acceptability function: it must obey the intervality property that $x \in A, y \succsim x \implies y \in A$ for all $A \subseteq X$ for which $\sigma(A) > 0$. So, for example, with $X = \{x,y,z\}$ if $\sigma(\{x\}) > 0$ then it cannot be the case that $\sigma(\{y\}) > 0$ or that $\sigma(\{z\}) > 0$ or that $\sigma(\{y,z\}) > 0$, since there is no weak order $\succsim$ on X such that the intervality property holds.

3.2 Identification

We proceed to take the perspective of an observer of an approval function who would like to retrieve the primitives that generated it, under the assumption that a given model holds. Both an AC and a SAC function are generated by a large number of parameters. Nevertheless, we will show that both models can be substantially identified.

Theorem 2. *Let p be a SAC function generated both by $(\succsim,\tau,\pi)$ and by $(\succsim',\tau',\pi')$. Suppose that $\tau(t) > 0, \tau'(t) > 0$ for all $t \in X$. Then:*

(i) $\succsim=\succsim'$;

(ii) $\pi = \pi'$;

(iii) For all $t \in X$: $\sum_{z\sim t} \tau(x) = \sum_{z\sim' t} \tau'(z)$.

Proof. (i) As in the proof of Theorem (1), define $\rho(x) = p(x,\lambda)$ for any $x \in X$ and $\lambda \in \Lambda$ satisfying $\lambda^x = \emptyset$. We have $x \succsim y \implies \rho(x) = \sum_{t:x\succsim t} \tau(t) \geq \sum_{t:y\succsim t} \tau(t) = \rho(y)$. Conversely, by the assumption on τ, $\sum_{t:x\succsim t} \tau(t) \geq \sum_{t:y\succsim t} \tau(t) \implies x \succsim y$. Since the same implications hold with $\succsim'$ in place of $\succsim$, $x \succsim y \iff \rho(x) \geq \rho(y) \iff x \succsim' y$ for all $x,y \in X$. (ii) For any $\tilde{\lambda} \in \tilde{\Lambda}$, $x \in X$ and $\lambda \in \Lambda$

such that $\lambda^x = \tilde{\lambda}$, we have $p(x,\lambda) = \pi(\tilde{\lambda}) \sum_{t:x \succsim t} \tau(t) = \pi(\tilde{\lambda})\rho(x)$. Since by the assumption on τ $\rho(x) > 0$, $\pi(\tilde{\lambda})$ is uniquely determined by $\pi(\tilde{\lambda}) = \frac{p(x,\lambda)}{\rho(x)}$.
(iii) Let $\sim_1, ..., \sim_k$ be the indifference classes of $\succsim = \succsim'$ ordered from best to worst, which in view of (i) correspond to the equality classes of ρ. For any $x_k \in \sim_k$ we have $p(x_k, \lambda) = \pi(\lambda^{x_k}) \sum_{t:x_k \succsim t} \tau(t) = \pi(\lambda^{x_k}) \sum_{t \in \sim_k} \tau(t)$. Since $\pi(\lambda^{x_k}) = \frac{p(z,\mu)}{\rho(z)}$ for any $z \in X$ and $\mu \in \Lambda$ for which $\lambda^{x_k} = \mu^z$, $\sum_{t \in \sim_k} \tau(t) = \rho(z) \frac{p(x_k,\lambda)}{p(z,\mu)}$. What is more, for $x_i \in \sim_i$ with $i < k$, $\sum_{t \in \sim_i} \tau(t)$ is determined recursively by $p(x_i, \lambda) = \pi(\lambda^{x_i}) \sum_{t:x_i \succsim t} \tau(t) = \frac{p(z,\mu)}{\rho(z)} \left(\sum_{t \in \sim_i} \tau(t) + \sum_{t:x_i \succ t} \tau(t) \right)$, for any $z \in X$ and $\mu \in \Lambda$ for which $\lambda^{x_i} = \mu^z$, and with $\sum_{t:x_i \succ t} \tau(t)$ taken as known. □

Thus, under a mild restriction, preferences and the continuation function are uniquely identified by the data. The threshold probabilities of the satisficing representation are essentially fully identified: the total probability of the items in any indifference class is pinned down uniquely - only the allocation of the probability mass among indifferent items remains free.

Finally, for all $x \in X$ and given an acceptability function σ, let $\sigma_x = \sum_{x \in A} \sigma(A)$. With similar arguments to those used for Theorem (2) it is easy to show:

Theorem 3. *Let p be an AC function generated both by (σ, π) and by (σ', π'). Suppose that $\sigma_x > 0, \sigma'_x > 0$ for all $x \in X$. Then:*
(i) $\pi = \pi'$
(ii) For all $x \in X$, $\sigma_x = \sigma'_x$.

The assumption on σ in the statement says that every item belongs to at least one acceptable set that occurs with positive probability. Like for the case of the satisficing version, with an AC function the continuation function is fully identified, while all the marginal approval probabilities of each item can be identified.

4 Correlations

So far we have allowed the probability of reaching an item to depend on the entire predecessor sublist, but we have considered this event as independent of what is acceptable. However, some plausible contexts exhibit interdependence between continuation and acceptability, as shown in the following example.

Example 4. (Limited capacity for costly approvals) Consider an online shopper who stops when the cost of her wish list exceeds a certain value, or a reader who stops when the expected reading time of the stored articles exceeds a certain limit. Formally, let $\kappa : X \to \mathbb{R}$ indicate the cost $\kappa(x)$ associated with each approved item x, and given a realised acceptable set A, define the cost of a sublist $\tilde{\lambda}$ as $K\left(\tilde{\lambda}, A\right) = \sum_{x \in \tilde{\lambda} \cap A} \kappa(x)$. Let $\hat{\Pi} : \mathbb{R} \to [0,1]$ be a cumulative distribution function, with $\hat{\Pi}(r)$ indicating the probability that the agent's "approval budget" is at most r. Then define

$$\pi\left(\tilde{\lambda}, A\right) = 1 - \hat{\Pi}\left(K\left(\tilde{\lambda}, A\right)\right).$$

The function $\pi : \tilde{\Lambda} \times 2^X \to [0,1]$ defined above can be seen as an approval-dependent continuation function, which determines an approval function

$$p(x, \lambda) = \sum_{x \in A} \sigma(A)\, \pi(\lambda^x, A).$$

An interesting specialisation of this example, to which we will return later, has a number-of-approvals type of constraint, letting $\kappa(x) = 1$ for all x (so that $K\left(\tilde{\lambda}, A\right) = \left|\tilde{\lambda} \cap A\right|$). For example, the agent may become less motivated to proceed as more acceptable items are encountered, considering it less likely that better items will be found.

4.1 Satisficing Correlations

We extend the satisficing model in order to describe this type of contexts.

Definition 6. An approval function p is a *satisficing correlated AC (SCAC) function* if there exists a triple $\left(\succsim, \tau, \{\pi_t\}_{t \in X}\right)$ where $\succsim$ is a weak order on X, τ is a probability distribution on $X \cup \{x_0\}$, and for every $t \in X$, π_t is a continuation function such that for every $x \in X$ and $\lambda \in \Lambda$,

$$p(x, \lambda) = \sum_{t : x \succsim t} \tau(t)\, \pi_t(\lambda^x).$$

In other words, we are now considering a joint probability distribution over satisfaction thresholds and the continuation event conditional on a given sublist of predecessors. To characterise this model, we introduce a new property:

A3. (Monotone Differences) For all $\lambda, \mu, \alpha, \beta \in \Lambda$ and $x, y \in X$ for which $\lambda^x = \mu^y \sqsubset \alpha^x = \beta^y$: if $p(x,\lambda) \geq p(y,\mu)$, then $p(x,\lambda) - p(y,\mu) \geq p(x,\alpha) - p(y,\beta) \geq 0$.

A3 says that the differences in approval probabilities between two items that follow the same predecessors in two lists is monotonic in the predecessor sublist (and this difference does not change sign). The difference is enhanced by moving the items to earlier positions in the respective lists. In other words, there is *complementarity* between quality (as revealed by differential approval probabilities) and position in the production of approval probabilities. It turns out that the only observable difference between the independent and the correlated versions of the model is that the Approval List Independence property is replaced by the Monotone Difference property.

Theorem 4. *An approval function is a SCAC function if and only if it satisfies A1 and A3.*

Proof. "Only if" Let p be a SCAC function generated by $(\succsim, \sigma, \{\pi_t\}_{t\in X})$. If $\lambda^x \sqsubseteq \mu^x$, then

$$p(x,\lambda) - p(x,\mu) = \sum_{t:x\succsim t} \tau(t)\left(\pi_t(\lambda^x) - \pi_t(\mu^x)\right) \geq 0$$

where the inequality follows from the monotonicity of the continuation functions and the fact that $\tau(t) \geq 0$. This establishes A1. To see that A3 is satisfied, suppose that $\lambda^x = \mu^y \sqsubset \alpha^x = \beta^y$ and that $p(x,\lambda) - p(y,\mu) \geq 0$. There are two cases. If $x \succ y$, then

$$p(x,\lambda) - p(y,\mu) = \sum_{t:x\succsim t\succ y} \tau(t)\,\pi_t(\lambda^x)$$

and

$$p(x,\alpha) - p(y,\beta) = \sum_{t:x\succsim t\succ y} \tau(t)\,\pi_t(\alpha^x).$$

It follows that $p(x,\alpha) - p(y,\beta) \geq 0$. What is more,

$$p(x,\lambda) - p(y,\mu) - (p(x,\alpha) - p(y,\beta)) = \sum_{t:x\succsim t\succ y} \tau(t)\left(\pi_t(\lambda^x) - \pi_t(\alpha^x)\right) \geq 0$$

by the monotonicity of the continuation functions. If $y \succsim x$, on the other hand, then

$$0 \leq \sum_{t:x \succsim t \succ y} \tau(t)\, \pi_t(\lambda^x) = p(y,\mu) - p(x,\lambda) \leq 0$$

yielding $p(x,\lambda) = p(y,\mu)$ and, in particular, $\tau(t)\,\pi_t(\lambda^x) = 0$ for every $t \in X$ such that $x \succsim t \succ y$. Since the continuation functions are monotone, $\tau(t)\,\pi_t(\alpha^x) = 0$ for every $t \in X$ such that $x \succsim t \succ y$ as well. Hence $p(x,\alpha) = p(y,\beta)$ and p satisfies A3. *"If"*. Suppose that p satisfies A1 and A3. Using A1, which implies that $p(x,\lambda) = p(x,\mu)$ if $\lambda^x = \mu^x$, abuse notation and write $p(x,\lambda) = p\left(x,\tilde{\lambda}\right)$ where $\tilde{\lambda} = \lambda^x$. Define a binary relation $\succsim$ on X as follows: $x \succsim y$ iff $p(x,\varnothing) \geq p(y,\varnothing)$. Note that $\succsim$ is a weak order since it has a numerical representation. Enumerate items so that $X = \{x_1, ..., x_n\}$ with $x_1 \succsim ... \succsim x_n$. Define a function $\tau : X \cup \{x_0\} \to [0,1]$ as follows:

$$\tau(t) = \begin{cases} p(x_n,\varnothing) & \text{if } t = x_n \\ p(x_i,\varnothing) - p(x_{i+1},\varnothing) & \text{for every } i = 1,...,n-1 \\ 1 - p(x_1,\varnothing) & \text{if } t = x_0 \end{cases}$$

Note that τ is a probability on $X \cup \{x_0\}$. The next step is the identification of $\{\pi_{x_i}\}_{i=1,...,n}$. We introduce another pseudo-item $x_{n+1} \notin X \cup \{x_0\}$, and extend $\succsim$ to $X \cup \{x_{n+1}\}$ so that $x_i \succ x_{n+1}$ for every $i = 1,...,n$. We posit that x_{n+1} does not belong to any list or any sublist. Now extend the domain of p by writing $p\left(x_{n+1},\tilde{\lambda}\right) = 0$ for every $\tilde{\lambda}$. Take any nonempty sublist $\tilde{\lambda}$ and $x_i \in X$. Define

$$i^*\left(\tilde{\lambda}\right) = \max\left\{j \leq i : x_j \notin \tilde{\lambda}\right\}, \text{ and}$$

$$i_*\left(\tilde{\lambda}\right) = \min\left\{j > i : x_j \notin \tilde{\lambda}\right\}.$$

That is, $i^*\left(\tilde{\lambda}\right)$ indexes the last item not in the sublist $\tilde{\lambda}$ (i.e. successors or pseudo-items) and in the same class as, or immediately better than, x_i. Similarly, $i_*\left(\tilde{\lambda}\right)$ indexes the first item not in the sublist $\tilde{\lambda}$ (i.e. successors or pseudo-items) in the immediately worse indifference class than x_i. Note that $0 \leq i^*\left(\tilde{\lambda}\right) \leq i < i_*\left(\tilde{\lambda}\right) \leq n+1$. We distinguish two cases. Case 1: $i^*\left(\tilde{\lambda}\right) = 0$. Note that in this case it must be $x_i \succ x_j$ for all $x_j \notin \tilde{\lambda}$, and $x_i \in \tilde{\lambda}$. Let $\pi_{x_i}\left(\tilde{\lambda}\right) = 0$. Case 2: $i^*\left(\tilde{\lambda}\right) > 0$. Let $\tilde{\lambda}_{-1}$ denote the sublist obtained from $\tilde{\lambda}$ by removing its last item while keeping the positions of the remaining items intact (that is, letting $x \in X$ satisfy $\tilde{\lambda}(x) = \left|\tilde{\lambda}\right|$, $\tilde{\lambda}_{-1} = \tilde{\lambda} \setminus x$). Define recursively

$$\pi_{x_i}\left(\tilde{\lambda}\right) = \begin{cases} 1 & \text{if } \tilde{\lambda} = \emptyset \\ \pi_{x_i}\left(\tilde{\lambda}_{-1}\right) \frac{p\left(x_{i^*\left(\tilde{\lambda}\right)}, \tilde{\lambda}\right) - p\left(x_{i_*\left(\tilde{\lambda}\right)}, \tilde{\lambda}\right)}{p\left(x_{i^*\left(\tilde{\lambda}\right)}, \tilde{\lambda}_{-1}\right) - p\left(x_{i_*\left(\tilde{\lambda}\right)}, \tilde{\lambda}_{-1}\right)} & \text{otherwise} \end{cases}$$

with the convention that $\frac{0}{0} = 0$. We now argue that π_{x_i} is a continuation function. We first establish that $\pi_{x_i}\left(\tilde{\lambda}\right) \in [0,1]$. Clearly, we need only show this when $i^*\left(\tilde{\lambda}\right) > 0$ and $\tilde{\lambda} \neq \emptyset$. Since $i^*\left(\tilde{\lambda}\right) \leq i < i_*\left(\tilde{\lambda}\right)$, it follows that $p\left(x_{i^*\left(\tilde{\lambda}\right)}, \emptyset\right) \geq p\left(x_{i_*\left(\tilde{\lambda}\right)}, \emptyset\right)$. Then, by A3, $p\left(x_{i^*\left(\tilde{\lambda}\right)}, \tilde{\lambda}_{-1}\right) - p\left(x_{i_*\left(\tilde{\lambda}\right)}, \tilde{\lambda}_{-1}\right) \geq 0$. Using A3 again,

$$\frac{p\left(x_{i^*\left(\tilde{\lambda}\right)}, \tilde{\lambda}\right) - p\left(x_{i_*\left(\tilde{\lambda}\right)}, \tilde{\lambda}\right)}{p\left(x_{i^*\left(\tilde{\lambda}\right)}, \tilde{\lambda}_{-1}\right) - p\left(x_{i_*\left(\tilde{\lambda}\right)}, \tilde{\lambda}_{-1}\right)} \in [0,1].$$

This implies that $\pi_{x_i}\left(\tilde{\lambda}\right) \leq \pi_{x_i}\left(\tilde{\lambda}_{-1}\right)$. Noting that $i^*(\tilde{\mu}) > 0$ for every $\tilde{\mu} \sqsubset \tilde{\lambda}$ and substituting sequentially for the continuation functions for the smaller sublists, we conclude that $\pi_{x_i}\left(\tilde{\lambda}\right)$ is obtained by multiplying $\pi_{x_i}(\emptyset)$ by a sequence of numbers, each of which is between 0 and 1. Hence $\pi_{x_i}\left(\tilde{\lambda}\right) \in [0,1]$, as desired. To establish monotonicity, we need to show that $\pi_{x_i}\left(\tilde{\lambda}\right) \leq \pi_{x_i}\left(\tilde{\lambda}_{-1}\right)$ for every i and $\tilde{\lambda} \in \tilde{\Lambda} \setminus \emptyset$. We have already shown this when $i^*\left(\tilde{\lambda}\right) > 0$. If $i^*\left(\tilde{\lambda}\right) = 0$, this follows since $\pi_{x_i}\left(\tilde{\lambda}\right) = 0$ and $\pi_{x_i}\left(\tilde{\lambda}_{-1}\right) \in [0,1]$. We next establish the following fact.

Claim. For all $\tilde{\lambda} \setminus \emptyset$ and all $x_i \in X \setminus \tilde{\lambda}$,

$$\sum_{j=i}^{i_*\left(\tilde{\lambda}\right)-1} \tau\left(x_j\right) \pi_{x_j}\left(\tilde{\lambda}\right) = p\left(x_i, \tilde{\lambda}\right) - p\left(x_{i_*\left(\tilde{\lambda}\right)}, \tilde{\lambda}\right).$$

Proof of the Claim: We will use induction on the size of $\tilde{\lambda}$. To begin, suppose that $\tilde{\lambda} = x_l$. Note that $l \neq i$. There are two cases.

Case 1: If $l \neq i+1$, then $i_*\left(x_l\right) = i+1$ and

$$\begin{aligned}\sum_{j=i}^{i_*(x_l)-1} \tau\left(x_j\right) \pi_{x_j}\left(x_l\right) &= \tau\left(x_i\right) \pi_{x_i}\left(x_l\right) \\ &= \tau\left(x_i\right) \frac{p\left(x_i, x_l\right)-p\left(x_{i+1}, x_l\right)}{p\left(x_i, \emptyset\right)-p\left(x_{i+1}, \emptyset\right)} \\ &= p\left(x_i, x_l\right)-p\left(x_{i+1}, x_l\right).\end{aligned}$$

Note that if $p\left(x_i, \emptyset\right)-p\left(x_{i+1}, \emptyset\right)>0$, then the last equality follows from the definition of $\tau\left(x_i\right)$. If $p\left(x_i, \emptyset\right)-p\left(x_{i+1}, \emptyset\right)=\tau\left(x_i\right)=0$, on the other hand, $p\left(x_i, x_l\right)-p\left(x_{i+1}, x_l\right)=0$ by A3. Hence $\tau\left(x_i\right) \frac{p\left(x_i, x_l\right)-p\left(x_{i+1}, x_l\right)}{p\left(x_i, \emptyset\right)-p\left(x_{i+1}, \emptyset\right)}=0=p\left(x_i, x_l\right)-p\left(x_{i+1}, x_l\right)$, as desired.

Case 2: If $l=i+1$, then $i_*\left(x_l\right)=(i+1)_*\left(x_l\right)=i+2$ and

$$\begin{aligned}\sum_{j=i}^{i_*(x_l)-1} \tau\left(x_j\right) \pi_{x_j}\left(x_l\right) &= \tau\left(x_i\right) \pi_{x_i}\left(x_l\right)+\tau\left(x_{i+1}\right) \pi_{x_{i+1}}\left(x_l\right) \\ &= \left(\tau\left(x_i\right)+\tau\left(x_{i+1}\right)\right) \frac{p\left(x_i, x_l\right)-p\left(x_{i+2}, x_l\right)}{p\left(x_i, \emptyset\right)-p\left(x_{i+2}, \emptyset\right)} \\ &= p\left(x_i, x_l\right)-p\left(x_{i+2}, x_l\right).\end{aligned}$$

The last equality follows from an argument analogous to that used for Case 1. To wit, if $p\left(x_i, \emptyset\right)-p\left(x_{i+2}, \emptyset\right)>0$, then we note that $\tau\left(x_i\right)+\tau\left(x_{i+1}\right)=p\left(x_i, \emptyset\right)-p\left(x_{i+2}, \emptyset\right)$ and cancel terms. If $p\left(x_i, \emptyset\right)-p\left(x_{i+2}, \emptyset\right)=0$, then $p\left(x_i, \emptyset\right)-p\left(x_{i+2}, \emptyset\right)$ by A3, from which the last equality follows. Now suppose that the statement holds for all $\tilde{\lambda} \in \tilde{\Lambda}$ for which $1 \leq \left|\tilde{\lambda}\right| \leq m$ and take $\tilde{\lambda} \in \tilde{\Lambda}$ such that $\left|\tilde{\lambda}\right|=m+1$ and $x_i \notin \tilde{\lambda}$. Note that for all $j \in \left\{i, \ldots, i_*\left(\tilde{\lambda}\right)-1\right\}$, $j^*\left(\tilde{\lambda}\right)=i$ and $j_*\left(\tilde{\lambda}\right)=i_*\left(\tilde{\lambda}\right)$. Consequently, for all such j,

$$\pi_{x_j}\left(\tilde{\lambda}\right)=\frac{p\left(x_i, \tilde{\lambda}\right)-p\left(x_{i_*\left(\tilde{\lambda}\right)}, \tilde{\lambda}\right)}{p\left(x_i, \tilde{\lambda}_{-1}\right)-p\left(x_{i_*\left(\tilde{\lambda}\right)}, \tilde{\lambda}_{-1}\right)}$$

and therefore

$$\sum_{j=i}^{i_*(\tilde{\lambda})-1} \tau(x_j)\pi_{x_j}\left(\tilde{\lambda}\right) = \frac{p\left(x_i,\tilde{\lambda}\right) - p\left(x_{i_*(\tilde{\lambda})},\tilde{\lambda}\right)}{p\left(x_i,\tilde{\lambda}_{-1}\right) - p\left(x_{i_*(\tilde{\lambda})},\tilde{\lambda}_{-1}\right)} \sum_{j=i}^{i_*(\tilde{\lambda})-1} \tau\left(x_j\right)\pi_{x_j}\left(\tilde{\lambda}_{-1}\right).$$

We will show that

$$\sum_{j=i}^{i_*(\tilde{\lambda})-1} \tau\left(x_j\right)\pi_{x_j}\left(\tilde{\lambda}_{-1}\right) = p\left(x_i,\tilde{\lambda}_{-1}\right) - p\left(x_{i_*(\tilde{\lambda})},\tilde{\lambda}_{-1}\right)$$

from which the claim follows. Let $\tilde{\lambda}\left(x_l\right) = m+1$. That is, x_l is the item removed from $\tilde{\lambda}$ when generating $\tilde{\lambda}_{-1}$. There are two cases. Recalling that x_i is not listed in $\tilde{\lambda}$, suppose first that $l \notin \left\{i+1,...,i_*\left(\tilde{\lambda}\right)-1\right\}$. In this case, $i_*\left(\tilde{\lambda}_{-1}\right) = i_*\left(\tilde{\lambda}\right)$. Consequently, since $\tilde{\lambda}_{-1}$ lists $m-1$ items,

$$\sum_{j=i}^{i_*(\tilde{\lambda})-1} \tau\left(x_j\right)\pi_{x_j}\left(\tilde{\lambda}_{-1}\right) = \sum_{j=i}^{i_*(\tilde{\lambda}_{-1})-1} \tau\left(x_j\right)\pi_{x_j}\left(\tilde{\lambda}_{-1}\right) = p\left(x_i,\tilde{\lambda}_{-1}\right) - p\left(x_{i_*(\tilde{\lambda})},\tilde{\lambda}_{-1}\right)$$

as desired. If, on the other hand, $l \in \left\{i+1,...,i_*\left(\tilde{\lambda}\right)-1\right\}$, then $i_*\left(\tilde{\lambda}_{-1}\right) = l$ and $l_*\left(\tilde{\lambda}_{-1}\right) = i_*\left(\tilde{\lambda}\right)$. Hence we have

$$\begin{aligned}
\sum_{j=i}^{i_*(\tilde{\lambda})-1} \tau\left(x_j\right)\pi_{x_j}\left(\tilde{\lambda}_{-1}\right) &= \sum_{j=i}^{l-1} \tau\left(x_j\right)\pi_{x_j}\left(\tilde{\lambda}_{-1}\right) + \sum_{j=l}^{i_*(\tilde{\lambda})-1} \tau\left(x_j\right)\pi_{x_j}\left(\tilde{\lambda}_{-1}\right) \\
&= \sum_{j=i}^{i_*(\tilde{\lambda}_{-1})-1} \tau\left(x_j\right)\pi_{x_j}\left(\tilde{\lambda}_{-1}\right) + \sum_{j=l}^{l_*(\tilde{\lambda}_{-1})-1} \tau\left(x_j\right)\pi_{x_j}\left(\tilde{\lambda}_{-1}\right) \\
&= p\left(x_i,\tilde{\lambda}_{-1}\right) - p\left(x_{i_*(\tilde{\lambda}_{-1})},\tilde{\lambda}_{-1}\right) + p\left(x_l,\tilde{\lambda}_{-1}\right) - p\left(x_{l_*(\tilde{\lambda}_{-1})},\tilde{\lambda}_{-1}\right) \\
&= p\left(x_i,\tilde{\lambda}_{-1}\right) - p\left(x_{i_*(\tilde{\lambda})},\tilde{\lambda}_{-1}\right)
\end{aligned}$$

where the penultimate equality is by the induction hypothesis. This finishes the proof of the Claim. □

Going back to the proof of the theorem, we finally need to show that $p\left(x_i,\tilde{\lambda}\right) = \sum_{j=i}^{n} \tau\left(x_j\right)\pi_{x_j}\left(\tilde{\lambda}\right)$ for all $x_i \in X$ and all $\tilde{\lambda} \in \Lambda$ such that $x_i \notin \tilde{\lambda}$. Take any such x_i

and $\tilde{\lambda}$. If $\tilde{\lambda} = \emptyset$, then the equality follows from the definition of τ and the fact that $\pi_{x_j}(\emptyset) = 1$ for all $j \in \{i, ..., n\}$. If $\tilde{\lambda} \neq \emptyset$, let $\{x_i, ..., x_{n+1}\} \setminus \tilde{\lambda} = \{x_{k^1}, ..., x_{k^{r+1}}\}$ where $i = k^1 < ... < k^{r+1} = n + 1$. Note that $k_*^s(\tilde{\lambda}) = k^{s+1}$ for all $s = 1, ..., r$. Then

$$\begin{aligned}
\sum_{j=i}^{n} \tau(x_j)\,\pi_{x_j}(\tilde{\lambda}) &= \sum_{s=1}^{r} \sum_{j=k^s}^{k^{s+1}-1} \tau(x_j)\,\pi_{x_j}(\tilde{\lambda}) \\
&= \sum_{s=1}^{r} \sum_{j=k^s}^{k_*^s(\tilde{\lambda})-1} \tau(x_j)\,\pi_{x_j}(\tilde{\lambda}) \\
&= \sum_{s=1}^{r} p(x_{k^s}, \tilde{\lambda}) - p\left(x_{k_*^s(\tilde{\lambda})}, \tilde{\lambda}\right) \\
&= \sum_{s=1}^{r} p(x_{k^s}, \tilde{\lambda}) - p(x_{k^{s+1}}, \tilde{\lambda}) \\
&= p(x_i, \tilde{\lambda}) - p(x_{n+1}, \tilde{\lambda}) \\
&= p(x_i, \tilde{\lambda}).
\end{aligned}$$

where the third equality follows from the Claim. The proof is complete. □

The interdependence between acceptability and continuation introduced with a SCAC function is arbitrary (i.e. the correlation can be positive or negative). Yet, the above characterisation shows that its main observable implication is always a *positive* association, or complementarity, between position and "quality" (the standing of an item in the preference order) when generating approval probabilities.

4.2 Unsatisficing correlations

Interestingly, the equivalence between the satisficing and the general version of the model established in section 3 breaks down once correlations are introduced.

Definition 7. An approval function p is a *correlated AC (CAC) function* if there exists a pair $(\sigma, \{\pi_A\}_{A \subseteq X})$ where σ is an acceptability function, and for every $A \subseteq X$, π_A is a continuation function such that for every $x \in X$ and $\lambda \in \Lambda$,

$$p(x, \lambda) = \sum_{x \in A} \sigma(A)\,\pi_A(\lambda^x)$$

We show that a CAC function may fail the Monotone Differences property A3.

Example 5. Let $X = \{x, y, w, z\}$. Let p be a CAC function generated by $(\sigma, (\sigma, \{\pi_A\}_{A \subseteq X}))$ such that 1. $\sigma(\{x\}) = 0.9$ and $\sigma(\{y\}) = 0.1$. 2. $\pi_{\{x\}}(z) = 0.2$, $\pi_{\{x\}}(zw) = 0.19$, $\pi_{\{y\}}(z) = 0.8$, $\pi_{\{y\}}(zw) = 0.1$. Then:

$$\begin{aligned} p(x, zxyw) - p(y, zyxw) &= \sigma(\{x\}) \pi_{\{x\}}(z) - \sigma(\{y\}) \pi_{\{y\}}(z) \\ &= 0.9 \times 0.2 - 0.1 \times 0.8 \\ &= 0.1 > 0 \end{aligned}$$

and

$$\begin{aligned} p(x, zwxy) - p(y, zwyx) &= \sigma(\{x\}) \pi_{\{x\}}(zw) - \sigma(\{y\}) \pi_{\{y\}}(zw) \\ &= 0.9 \times 0.19 - 0.1 \times 0.1 \\ &= .161 \\ &> p(x, zxyw) - p(y, zyxw) \end{aligned}$$

violating A3.

Therefore, in view of Theorem (4), we have:

Remark 1. There exist CAC functions that are not SCAC functions.

It may be tempting to think that a CAC function can be characterised simply by dropping A3 in the SCAC characterisation Theorem (4), that is, in terms of Predecessor Monotonicity alone (which is clearly necessary). But this is false. To see this, consider $X = \{x, y, z\}$. Writing $p(x, \lambda) = p(x, \lambda^x)$, let $p(x, \emptyset) = p(y, \emptyset) = 1$, $p(z, \emptyset) = 0$, $p(x, z) = 0.8$ and $p(y, z) = 0.7$ and choose the remaining approval probabilities to satisfy Predecessor Monotonicity. If p is a CAC function generated by $(\sigma, \{\pi_A\}_{A \subseteq X})$, then $\sigma(\{x, y\}) = 1$ necessarily. But the approval probabilities following the sublist z then imply the contradiction $0.8 = 1 \times \pi_{\{x,y\}}(z) = 0.7$.

The characterisation of a CAC function remains a non-trivial open problem.

5 Application: Examples of List Design

5.1 The List Design Problem

As an application of our theory of approval, we now study some examples of "list design". The issue of manipulation of online behaviour has become a significant concern in current public discourse. We operationalise the idea of manipulation by assuming that a designer constructs the list so as to maximise some objective function. For instance, the list designer may wish to increase the general number of news pieces read or of papers downloaded, or just of some particular types of news and papers; or to foster social network involvement through likes, favouriting and sharing; or to enhance the revenue generated by a wish list or a shopping cart. 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.* [10] 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".[15]

To accommodate broadly this type of aims by the designer we consider as the objective function the *weighted sum of the approval probabilities* generated by a list. We concurrently assume that the primitives that generated the data are known to the designer (see section 3.2).

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

$$W_\lambda = \sum_{x \in X} w(x)\, p(x, \lambda). \tag{7}$$

Definition 8. A list λ 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

[15]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.

the list.

We will consider two specialisations of our satisficing model, one with correlations and the other without. We assume throughout the section the following strictness conditions: (1) for all $t \in X$, $\tau(t) > 0$, and (2) $\tilde{\lambda} \sqsubset \tilde{\mu} \implies \pi_t(\tilde{\lambda}) > \pi_t(\tilde{\lambda})$.

Model A: p is a SAC function generated by $(\succsim, \tau, \pi)$ such that for all $\tilde{\lambda}, \tilde{\mu} \in \tilde{\Lambda}$:

$$\left|\tilde{\lambda}\right| = |\tilde{\mu}| \implies \pi\left(\tilde{\lambda}\right) = \pi\left(\tilde{\mu}\right). \tag{8}$$

Model B: p is a SCAC function generated by $(\succsim, \tau, \{\pi_t\}_{t \in X})$ such that for all $\tilde{\lambda}, \tilde{\mu} \in \tilde{\Lambda}$ and all $t \in X$:

$$\left|\tilde{\lambda} \cap \{z : z \succeq t\}\right| = |\tilde{\mu} \cap \{z : z \succeq t\}| \implies \pi_t\left(\tilde{\lambda}\right) = \pi_t\left(\tilde{\mu}\right). \tag{9}$$

In other words, in both models the predecessor sublist affects continuation only through the cardinality of a relevant object. In Model A the relevant object is the set of items that have been *examined* so far (so what matters is only how deep the agent is in the list). In Model B the relevant object is the set of items that have been *approved* so far.

While restrictive, these specialisations still allow a rich range of attitudes on the part of the approver. For example, in Model B, having encountered a large number of acceptable items may generate both optimism and pessimism about the quality of the remaining items, as well as a non-monotonic attitude. Similarly, Model A is consistent with any pattern for the rate of change in continuation probabilities.

The first dilemma for a designer is caused by the possible discrepancies between the listing of the items according to the quality or the weight rankings: should either of the conflicting listing criteria be given priority? If not, how should the criteria be aggregated?

Secondly, given a criterion, it is not clear a priori whether a weaker item according to this criterion should be placed earlier or later in the list. Earlier positions favour approval in our 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 Model A, 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 with Model A is

$$\begin{aligned}
&W_\lambda - W_\mu = \\
&= \bar{w}\left(\pi(\emptyset)\sum_{t\in\{y,z\}}\tau(t) + \pi(y)\tau(z) + \pi(yz)\sum_{t\in\{x,y,z\}}\tau(t)\right) \\
&- \bar{w}\left(\pi(\emptyset)\tau(z) + \pi(z)\sum_{t\in\{x,y,z\}}\tau(t) + \pi(zx)\sum_{t\in\{y,z\}}\tau(t)\right) \\
&= \bar{w}\left[\tau(y)(1-\pi(y)) - \tau(x)(\pi(y) - \pi(yz))\right]
\end{aligned}$$

where the last line is obtained using $\pi(\emptyset) = 1$ and conditions (8). Therefore the sign of $W_\lambda - W_\mu$ for two generic lists λ and μ is ambiguous, depending on the shape of the continuation function and the threshold probabilities.

We can clarify the design questions substantially. Our first result shows that in Model B, unlike the example above, the optimality condition in fact *does not depend at all on the primitives of the model*.

Theorem 5. *In Model B a list λ 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, 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 (9) that

$$p(z,\lambda) = p(z,\mu) \;\; \forall z \in X \setminus \{x,y\} \tag{10}$$

We assume w.l.o.g. that $x \succ y$. The differences in approval probability across the two lists for x and y are:

$$p(x,\lambda) - p(x,\mu) = \sum_{t:x \succsim t} \tau(t)\, \pi_t(\lambda^x) - \sum_{t:x \succsim t} \tau(t)\, \pi_t(\mu^x) \tag{11}$$

$$p(y,\mu) - p(y,\lambda) = \sum_{t:y \succsim t} \tau(t)\, \pi_t(\mu^y) - \sum_{t:y \succsim t} \tau(t)\, \pi_t(\lambda^y)\,. \tag{12}$$

Note that, since $\lambda^x = \mu^y$, for all $t \in X$:

$$\pi_t(\lambda^x) = \pi_t(\mu^y) \tag{13}$$

What is more, in view of (9):

$$t \precsim y \prec x \implies \pi_t(\lambda^y) = \pi_t(\mu^x) \tag{14}$$

$$y \prec t \precsim x \implies \pi_t(\lambda^x) = \pi_t(\mu^x) \tag{15}$$

We can now calculate:

$$\begin{aligned} p(x,\lambda) - p(x,\mu) - (p(y,\mu) - p(y,\lambda)) &= \\ \sum_{t:x \succsim t \succ y} \tau(t)\, \pi_t(\lambda^x) - \sum_{t:x \succeq t} \tau(t)\, \pi_t(\mu^x) + \sum_{t:y \succsim t} \tau(t)\, \pi_t(\lambda^y) &= \\ \sum_{t:x \succeq t \succ y} \tau(t)\, \pi_t(\lambda^x) - \sum_{t:x \succsim t} \tau(t)\, \pi_t(\mu^x) + \sum_{t:y \succsim t} \tau(t)\, \pi_t(\mu^x) &= \\ \sum_{t:x \succsim t \succ y} \tau(t)\, \pi_t(\lambda^x) - \sum_{t:x \succsim t \succ y} \tau(t)\, \pi_t(\mu^x) &= 0. \end{aligned}$$

where the first equality is by (11), (12) and (13), the second equality by (14), and the final equality by (15). In sum, we have:

$$p(x,\lambda) - p(x,\mu) = p(y,\mu) - p(y,\lambda) > 0 \tag{16}$$

(the inequality holding by the the strictness assumption on π and τ), so that:

$$w(x)\, p(x,\mu) + w(y)\, p(y,\mu) \geq w(y)\, p(y,\lambda) + w(x)\, p(x,\lambda) \Leftrightarrow w(x) \leq w(y) \tag{17}$$

Take any $\lambda \in \Lambda$ for which $w(x) > w(y)$ and $y \lambda x$ for some consecutive $x, y \in X$.

Then by (10) and (17) it is possible to increase W_λ by switching the items x and y. This shows that $w(x) > w(y) \Rightarrow x\lambda y$ with x and y consecutive is necessary for optimality. This necessary condition can be extended to non-consecutive items x and y by iteratively applying it to the connecting consecutive pairs. That is, suppose $y\lambda x$ and in particular $z_1 = y\lambda z_2 \lambda ... \lambda z_k = z$ with any z_i and z_{i+1} consecutive in λ. The condition for consecutive pairs implies $z_i \lambda z_{i+1} \implies w(z_i) \geq w(z_{i+1})$ for all $i = 1, ..., k-1$. Therefore $y\lambda x \implies w(y) \geq w(x)$ for all $x, y \in X$, as desired.

Conversely, suppose that a list λ agrees with the weight ordering and compare it with an optimal list, say μ. Since μ 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 λ and μ can only disagree on the way they order pairs of items x and y for which $w(x) = w(y)$. Clearly, switching these items cannot affect the value of a list, since such switches cannot change the number of approved predecessors of any item. Therefore, $W_\mu = W_\lambda$ and λ is also optimal. $\square$

One way to read this result is that for Model B *there is no substitutability between the quality and the weight* of an item as far is 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.*

A leading special case is when weights are the same for all items. Then, as an immediate implication of the statement of Theorem 6:

Corollary 1. (**The List Invariance Principle**) *In Model B, suppose that $w(x) = w(y)$ for all $x, y \in X$. Then all lists are optimal: $\sum_{x\in X} w(x)\, p(x, \lambda) = \sum_{x\in X} w(x)\, p(x, \mu)$* for all $\lambda, \mu \in \Lambda$.

Thus, the designer in this case has no power at all to affect the value of a list. To understand this rather surprising result, consider any realisation of the threshold t and of the continuation events (yes or no) associated with every cardinality. Let $|\{z : z \succeq t\}| = k$, and let i be the smallest cardinality for which the approver does not continue. If $k \leq i$, then k items are approved. Otherwise, i items are approved. Since neither k nor i depends on the list, and the argument holds for an arbitrary realisation, the number of approvals does not depend on the list.[16]

[16]We thank Yuhta Ishii for suggesting this argument.

For Model A, matters are very different:

Theorem 6. *In Model A a list λ is optimal if and only if it satisfies the following condition:*

$$w(x)\sum_{t:x\succsim t}\tau(t) > w(y)\sum_{t:y\succsim t}\tau(t) \Rightarrow x\lambda y$$

Proof. 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. Take $\lambda, \mu \in \Lambda$ such that $\lambda(x) = \mu(y)$, $\lambda(y) = \mu(x)$ for some distinct $x, y \in X$, and $\lambda(z) = \mu(z)$ for all $z \neq x, y$.

From (8) it follows immediately that $p(z,\lambda) = p(z,\mu)$ for all $z \neq x, y$. Also, in view of (8), $\pi(\lambda^x) = \pi(\mu^y)$ and $\pi(\lambda^y) = \pi(\mu^x)$. Suppose that $w(x)\sum_{t:x\succsim t}\tau(t) > w(y)\sum_{t:y\succsim t}\tau(t)$ and $y\lambda x$. Then

$$\begin{aligned}
&W_\mu - W_\lambda = \\
&= w(x)\left(\sum_{t:x\succsim t}\tau(t)\,\pi(\lambda^x) - \sum_{t:x\succsim t}\tau(t)\,\pi(\mu^x)\right) - w(y)\left(\sum_{t:y\succsim t}\tau(t)\,\pi(\mu^y) - \sum_{t:y\succsim t}\tau(t)\,\pi(\lambda^y)\right) \\
&= w(x)\left(\sum_{t:x\succsim t}\tau(t)\,\pi(\lambda^x) - \sum_{t:x\succsim t}\tau(t)\,\pi(\mu^x)\right) - w(y)\left(\sum_{t:y\succsim t}\tau(t)\,\pi(\lambda^x) - \sum_{t:y\succsim t}\tau(t)\,\pi(\mu^x)\right) \\
&= w(x)\sum_{t:x\succsim t}\tau(t)\left(\pi(\lambda^x) - \pi(\mu^x)\right) - w(y)\sum_{t:y\succsim t}\tau(t)\left(\pi(\lambda^x) - \pi(\mu^x)\right) \\
&> 0
\end{aligned}$$

where the inequality follows from the strictness assumption on π and τ. Hence the condition on the statement is necessary for a list to be optimal. Conversely, any list λ 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 5. □

This result has the immediate consequence:

Corollary 2. *In Model A, suppose that weights agree with preferences in the sense that $x \succ y \Rightarrow w(x) \geq w(y)$. Then the unique optimal list λ agrees with the preferences, i.e. $x \succ y \Leftrightarrow x\lambda y$.*

Theorem (6) tells us that 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 $w(x)\sum_{t:x\succsim t}\tau(t)$. In fact, $\sum_{t:x\succsim t}\tau(t)$ is a measure of quality: it is larger for items that are higher in the preference ordering. When the two components of the measure agree, there is no need for trade-offs.

Why do Model A and Model B behave so differently? Consider Model A. When a better item x is placed in a list λ behind a worse item y, a switch in the positions of x and y to obtain a new list μ 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. Suppose that a threshold-continuation realisation leads to the approval of y in λ but not in μ. This means that y passes the threshold, and that the smallest cardinality i for which the approver does not continue is strictly less than $\mu(y)$ but at least as much as $\lambda(x)$. But this means that this particular realisation of the random variables leads to the approval of x in μ and not in λ.

ii) some gain for x is not a loss for y. For example, a threshold-continuation realisation such that $t = x \succ y$ and the position $\mu(x)$ is reached while the position $\lambda(x)$ is not reached leads to the approval of x in μ but not in λ. 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. When there is disagreement between $\succsim$ and w a similar argument applies.

A similar reasoning to the one above could be performed for *consecutive* items in Model B, as in the proof of Theorem 5. But, unlike in Model A, in Model B there is 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 approval. Hence *y loses from the switch exactly the approval probability that x gains* (this is the content of equation (16). 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.

The main message from this study of cardinality-based models is the following. If it is the number of approvals that affects the desire to explore the list, then optimal lists must use as information the designer's preferences, ignoring the approver's preferences. On the other hand, if the desire to explore the list is determined by the position in the list, both parties' preferences must be considered.

6 Comments on the related literature

The closest papers to the present one are those that constitute the small literature on choice from lists, or choice from menus where a list (a search order) appears as a primitive of the model. It is useful to make some further comments on the conceptual difference between the act of approval and the act of choice. "Choosing" means selecting one out of a set of feasible alternatives. This is clear in the concept of a choice function. A choice *correspondence* admits multiplicity either in the sense that it describes all that is "choos*able*" (because they are indifferent to each other, as in the standard utility maximisation model, or in the model of choice from list by Horan [13]); or "chos*en* in different circumstances" (because the primitives change across circumstance as e.g. in Salant and Rubinstein [21], where frame variation generates a variety of choices).[17] An approval function, by contrast, describes sets of items that are actually approv*ed* on a *single* circumstance (or realisation of a random variable). Even when many objects are chosen at once (e.g., a set of applicants in the job market) the objects of choice are *sets*, and only one set can be chosen at once. On the contrary, several sets can be approved at once. In general, once an object of choice has been fixed (a portfolio, a consumption bundle, a set of candidates), only one can be meaningfully chosen, while many can be approved.

A stochastic choice function describes the possibility of choosing several alternatives, each one for a different realisation of an unobserved variable that appears random to the researcher. For example, as Rubinstein and Salant [20] put it:

> An intuitive reason for randomness in choice from sets is that although

[17] In their generalisation of Rubinstein and Salant's [20] deterministic model of choice from lists, Koshevoy and Savaglio [14] combine these two aspects, by defining a choice correspondence as the union of choice correspondences from all possible lists of *sets*.

> the decision maker deterministically chooses from lists, there is an underlying random process that transforms sets into lists.

A search order is also the key random variable in Aguiar, Boccardi and Dean (ABD) [2], whereas it is a set of binary comparisons in Yildiz [22] and a threshold in Kovach and Ülkü [15], where *one* alternative is chosen on each realisation.[18] With an approval function, as noted, each realisation of the random variables generates *multiple* approved items. To clarify this point further, it is instructive to quote an example of the stochastic choice procedure in ABD, who describe the decision to buy [choose] a book at an airport before a flight: given a realised search order (list)

> They examine the available books one by one, looking for one which satisfies their requirements...If they find such a book, they immediately go to the checkout and buy it.

In contrast, an approver would *not* immediately go to the checkout. He would continue browsing and approve everything he finds acceptable.

ABD continue:

> If they search the entire selection and don't find a book which matches these criteria then they go back and choose the best of the books that they did see.

In contrast, if an approver has not approved anything by the end of the list (if any), he would just leave with an empty approved set.

For concreteness, recall Example 2, where we considered an agent who maximises the probability of stopping at the best item while approving all items along the way that are above a given quality threshold. This distinguishes clearly the approver both from a standard satisficer, who would select the first above-threshold item; and from a preference maximiser, who would select only the best items among those seen.

[18] Or multiple alternative are choos*able* if preferences are represented by a weak rather than a strict order.

7 Concluding remarks

Approval activities are evidently of considerable economic and social interest, because of the value of the information about users contained in approvals[19], the facilitation of the sale process[20], the generation of new modes of advertising, and the building of online communities. Facebook is worth hundreds of billions of dollars, it bought Instagram for a billion and even a relatively "small" company such as Pinterest is worth dozens of billions and has more than trebled in value in 2020. Parallel to the explosion of companies that focus on social media and user-generated content, the nature of available data has undergone a profound change in recent years.[21] 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 certainly 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. Third, 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 a nod to a friend's remark as opposed to Liking and favouriting.

A main appealing feature of our model is that it can be sharply characterised and its primitives can be uniquely and non-parametrically identified from observed data. Furthermore, our 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

[19]Such information can be both valuable for third parties to which it can be sold, or allow to the company to provide a larger scale and more efficient service (as in the case of dating companies).

[20]See e.g. Lee, Lee, Oh [17] on the role of Facebook Likes for sales growth.

[21]See e.g. https://www.statista.com/markets/424/topic/540/social-media-user-generated-content/

highlighted several non-obvious tradeoffs and principles of list design in some circumscribed contexts. 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.[22] 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.

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 wished to isolate the way in which the order aspect of data, rather than their set aspect, provides information.[23] Future work might be done to incorporate both types of variation.

References

[1] Aguiar, Victor H. (2017) Random categorization and bounded rationality", *Economics Letters* (159): 46-52.

[2] Aguiar, Victor H., María José Boccardi, and Mark Dean (2016), "Satisficing and stochastic choice", *Journal of Economic Theory*, 166: 445-482.

[3] Brady, Richard L., and John Rehbeck (2016). Menu-dependent stochastic feasibility. *Econometrica* 84:1203–1223.

[22]Kramer *et al.* [16] 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).

[23]Dardanoni *et al.* [7] 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.

[4] Bruss, F. Thomas (2000) "Sum the Odds to One and Stop", *The Annals of Probability* 28(3): 1384–1391.

[5] Caplin, Andrew, Mark Dean, and Daniel Martin (2011) "Search and satisficing", *American Economic Review* 101(7): 2899-2922.

[6] J. Chipman (1960) "The foundations of utility", *Econometrica* 28:193–224.

[7] Dardanoni, Valentino, Paola Manzini, Marco Mariotti and Christopher J. Tyson (2020) "Inferring Cognitive Heterogeneity from Aggregate Choices", *Econometrica* 88(3): 1269-1296.

[8] Dietrich, Franz, and Christian List (2016) "Mentalism versus behaviourism in economics: a philosophy-of-science perspective", *Economics and Philosophy* 32(2): 249-281.

[9] Eliaz, Kfir, and Ran Spiegler (2011) "Consideration Sets and Competitive Marketing", *The Review of Economic Studies* 78(1): 235-262

[10] Feenberg, Daniel R., Ina Ganguli, Patrick Gaule and Jonathan Gruber (2017) "It's Good to Be First: Order Bias in Reading and Citing NBER Working Papers", *Review of Economics and Statistics* 99 (1): 32-39.

[11] Ferguson, Thomas S. (2016) "The Sum-the-Odds Theorem with Application to a Stopping Game of Sakaguchi", *Mathematica Applicanda* 44(1): 45-61.

[12] Ferguson, Thomas S. (2020) Optimal Stopping and Applications, https://www.math.ucla.edu/˜tom/Stopping/Contents.html.

[13] Horan, Sean (2011) "Sequential Search and Choice from Lists", working paper, https://sites.google.com/site/smhoran/research.

[14] Koshevoy, Gleb, and Ernesto Savaglio (2017) "Enveloped choice functions and path-independent rationality", working paper, https://econpapers.repec.org/paper/usiwpaper/765.htm.

[15] Kovach, Matthew, and Levent Ülkü, (2020), "Satisficing with a Variable Threshold, *Journal of Mathematical Economics*, 87: 67-76.

[16] Kramer, Adam D.I., Jamie E. Guillory, Jeffrey T. Hancock (2014) "Emotional contagion through social networks", *Proceedings of the National Academy of Sciences* 111 (24): 8788-8790.

[17] Lee, Kyunghee, Byungtae Lee and Wonseok Oh (2015) "Thumbs Up, Sales Up? The Contingent Effect of Facebook Likes on Sales Performance in Social Commerce", *Journal of Management Information Systems* 32(4):109-143.

[18] Manzini, Paola and Marco Mariotti (2014) "Stochastic choice and consideration sets", *Econometrica* 82:1153–1176.

[19] Masatlioglu, Yusufcan, Daisuke Nakajima and Erkut Y. Ozbay (2012) "Revealed Attention", *American Economic Review* 102(5): 2183-2205.

[20] Rubinstein, Ariel, and Yuval Salant (2006), "A model of choice from lists", *Theoretical Economics* 1(1): 3-17.

[21] Salant, Yuval, and Ariel Rubinstein (2008), "(A, f): Choice with Frames", *The Review of Economic Studies* 75(4): 1287–1296.

[22] Yildiz, Kemal (2016), "List-rationalizable choice", *Theoretical Economics*, 11(2), 587-599.