Sequential Approval

Paola Manzini Marco Mariotti Levent Ülkü

Motivation

Consider an individual who:

Scans headlines and reads some articles (or stores away for later reading)

Motivation

Consider an individual who:

Scans headlines and reads some articles (or stores away for later reading)

'Likes' or shares various posts when going down his social media feed

Motivation

Consider an individual who:

Scans headlines and reads some articles (or stores away for later reading)

'Likes' or shares various posts when going down his social media feed

Progressively fills her online shopping cart

Motivation

Consider an individual who:

Scans headlines and reads some articles (or stores away for later reading)

'Likes' or shares various posts when going down his social media feed

Progressively fills her online shopping cart

Progressively 'matches' with potential partners on an online dating site

Motivation - ctnd.

These are diverse examples of behaviour. What do they have in common?

Motivation - ctnd.

These are diverse examples of behaviour. What do they have in common?

i) Objects come *sequentially* to the attention of the agent: they form a *list*

Motivation - ctnd.

These are diverse examples of behaviour. What do they have in common?

i) Objects come *sequentially* to the attention of the agent: they form a *list*

ii) Objects are not quite 'chosen', but merely '*approved*':

- there's not going to be a final choice between articles read or posts Liked or shared.
- an item/partner may or may not be finally selected from the cart/match set
- the whole cart/match set may even be abandoned (they act as 'consideration sets')

Motivation - ctnd.

These are diverse examples of behaviour. What do they have in common?

i) Objects come *sequentially* to the attention of the agent: they form a *list*

ii) Objects are not quite 'chosen', but merely *'approved'*:

- there's not going to be a final choice between articles read or posts Liked or shared.
- an item/partner may or may not be finally selected from the cart/match set
- the whole cart/match set may even be abandoned (they act as 'consideration sets')

iii) The set of potentially approvable objects is very large, even 'endless', so:
- the **order** aspect overrides the **menu** aspect
- some *capacity constraint* is likely to apply (cannot go on forever)

Build a model of sequential approval with features i-iii.:

X is a finite *set of alternatives*, n its cardinality (thought of as being very large) X, and $N = \{1, ..., n\}$.

A *list* is any linear order λ on X, sometimes denoted xyz...

Λ is the set of all lists.

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

The number $p(x, \lambda)$ is the probability that x is approved when the decision maker is facing list λ.

Sequential approval - ctnd.

Have we forgotten the adding up constraint on probabilities?

Sequential approval - ctnd.

Have we forgotten the adding up constraint on probabilities?

No: the sum of the $p(x, \lambda)$ over alternatives is typically not one. This is approval, not choice...

Sequential approval - ctnd.

Have we forgotten the adding up constraint on probabilities?

No: the sum of the $p(x, \lambda)$ over alternatives is typically not one. This is approval, not choice...

Formally, a stochastic approval function could be equivalently defined as a stochastic *correspondence* $C : 2^X \times \Lambda \to [0, 1]$ associating with each list the probability of the possible *approval sets*.

The adding up constraint applies over these objects, *provided* that $C(\emptyset, \lambda)$ is allowed to absorb residual probability.

Sequential approval - ctnd.

Have we forgotten the adding up constraint on probabilities?

No: the sum of the $p(x, \lambda)$ over alternatives is typically not one. This is approval, not choice...

Formally, a stochastic approval function could be equivalently defined as a stochastic *correspondence* $C : 2^X \times \Lambda \to [0, 1]$ associating with each list the probability of the possible *approval sets*.

The adding up constraint applies over these objects, *provided* that $C(\emptyset, \lambda)$ is allowed to absorb residual probability.

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.

Two models of sequential approval - 1 of 4

At the moment of approving the agent is defined by:

1) A *preference* $\succeq$ (a linear order over X).

2) An *approval threshold* (an element of X).

3) A *stopping rule* (a number expressing a capacity constraint - more on this later).

Two models of sequential approval - 2 of 4

We take preference to be the stable element of the agent's psychology.

Approval thresholds and capacity constraints are subject to random shocks.

E.g:

- each morning you may be more or less strict with your FB Likes

- each morning you may have more or less time/patience to go down the list.

Two models of sequential approval - 3 of 4

Let π be a (strictly positive) joint probability distribution over $N \times X$ that describes this randomness.

$\pi(i, t)$ is the joint probability that the approval threshold is t and the capacity constraint is i. An agent is a pair $(\succeq, \pi)$.

Two models of sequential approval - 3 of 4

Let π be a (strictly positive) joint probability distribution over $N \times X$ that describes this randomness.

$\pi(i, t)$ is the joint probability that the approval threshold is t and the capacity constraint is i. An agent is a pair $(\succeq, \pi)$.

Consider two possibilities regarding the capacity constraint:

i) *Depth constraint*: the constraint acts on the number of alternatives you *examine*.

ii) *Approval constraint*: the constraint acts on the number of alternatives you *approve*.

Two models of sequential approval - 4 of 4

Let $\lambda(x)$ be the position of x in list λ.

The DCM is represented as

$$p^D(x,\lambda) = \sum_{x \succeq t} \sum_{i \geq \lambda(x)} \pi(i,t)$$

Two models of sequential approval - 4 of 4

Let $\lambda(x)$ be the position of x in list λ.

The DCM is represented as

$$p^D(x,\lambda) = \sum_{x \succeq t} \sum_{i \geq \lambda(x)} \pi(i,t)$$

Let $b(\lambda, j, t)$ be the number of alternatives that are $\succeq t$ and that in list λ are in a position $\leq j$

The ACM is represented as

$$p^A(x,\lambda) = \sum_{x \succeq t} \sum_{i \geq b(\lambda,\lambda(x),t)} \pi(i,t)$$

Related literature

- Rubinstein & Salant TE06 (mostly observable lists, menu variation, choice functions - or correspondences by taking unions of lists)

- Yildiz TE16 (menu variation, rationalisation *by* lists)

- Aguiar, Boccardi & Dean JET16 (menu variation, rationalisation *by* - random - lists)

- Kovach & Ülkü 2017 (menu variation, rationalisation *by* lists, random threshold).

- Caplin, Dean & Martin AER11 (experimental choice process data, infer search order).

Our Questions

1) *Identification*:

2) *List design*:

3) *Characterisation*:

4) *Comparative statics*:

Our Questions

1) *Identification*: Assume that approvals are generated by the model(s). Can an observer of approvals and lists identify the parameters, i.e. preferences $\succeq$ and the joint probabilites $\pi(i, t)$?

2) *List design*:

3) *Characterisation*:

4) *Comparative statics*:

Our Questions

1) *Identification*: Assume that approvals are generated by the model(s). Can an observer of approvals and lists identify the parameters, i.e. preferences $\succeq$ and the joint probabilites $\pi(i, t)$?

2) *List design*: Given an objective (e.g. total number of approvals), which list maximises the objective?

3) *Characterisation*:

4) *Comparative statics*:

Our Questions

1) *Identification*: Assume that approvals are generated by the model(s). Can an observer of approvals and lists identify the parameters, i.e. preferences $\succeq$ and the joint probabilites $\pi(i, t)$?

2) *List design*: Given an objective (e.g. total number of approvals), which list maximises the objective?

3) *Characterisation*: Which exact constraints on observed approval behaviour do the models impose?

4) *Comparative statics*:

Our Questions

1) *Identification*: Assume that approvals are generated by the model(s). Can an observer of approvals and lists identify the parameters, i.e. preferences $\succeq$ and the joint probabilites $\pi(i, t)$?

2) *List design*: Given an objective (e.g. total number of approvals), which list maximises the objective?

3) *Characterisation*: Which exact constraints on observed approval behaviour do the models impose?

4) *Comparative statics*: How are changes in the primitives manifested in behaviour?

3-Alternative example: DCM

Let $x \succ y \succ z$. Denote the marginals of π by π_i and π_t.

DCM	$p(x, \lambda)$	$p(y, \lambda)$	$p(z, \lambda)$

3-Alternative example: DCM

Let $x \succ y \succ z$. Denote the marginals of π by π_i and π_t.

DCM	$p(x, \lambda)$	$p(y, \lambda)$	$p(z, \lambda)$
$\lambda = xyz$			

3-Alternative example: DCM

Let $x \succ y \succ z$. Denote the marginals of π by π_i and π_t.

DCM	$p(x,\lambda)$	$p(y,\lambda)$	$p(z,\lambda)$
$\lambda = xyz$	$\sum_{i=1}^{3} \sum_{t \in \{x,y,z\}} \pi_{i,t} = 1$		

3-Alternative example: DCM

Let $x \succ y \succ z$. Denote the marginals of π by π_i and π_t.

DCM	$p(x,\lambda)$	$p(y,\lambda)$	$p(z,\lambda)$
$\lambda = xyz$	$\sum_{i=1}^{3} \sum_{t \in \{x,y,z\}} \pi_{i,t} = 1$	$\sum_{i=2}^{3} \sum_{t \in \{y,z\}} \pi_{i,t}$	

3-Alternative example: DCM

Let $x \succ y \succ z$. Denote the marginals of π by π_i and π_t.

DCM	$p(x, \lambda)$	$p(y, \lambda)$	$p(z, \lambda)$
$\lambda = xyz$	$\sum_{i=1}^{3} \sum_{t \in \{x,y,z\}} \pi_{i,t} = 1$	$\sum_{i=2}^{3} \sum_{t \in \{y,z\}} \pi_{i,t}$	$\pi_{3,z}$

3-Alternative example: DCM

Let $x \succ y \succ z$. Denote the marginals of π by π_i and π_t.

DCM	$p(x,\lambda)$	$p(y,\lambda)$	$p(z,\lambda)$
$\lambda = xyz$	$\sum_{i=1}^{3}\sum_{t\in\{x,y,z\}}\pi_{i,t} = 1$	$\sum_{i=2}^{3}\sum_{t\in\{y,z\}}\pi_{i,t}$	$\pi_{3,z}$
$\lambda = xzy$	$\sum_{i=1}^{3}\sum_{t\in\{x,y,z\}}\pi_{i,t} = 1$	$\sum_{t\in\{y,z\}}\pi_{3,t}$	$\sum_{i=2}^{3}\pi_{i,z}$
$\lambda = yxz$	$\sum_{i=2}^{3}\pi_i$	$\pi_y + \pi_z$	$\pi_{3,z}$
$\lambda = yzx$	π_3	$\pi_y + \pi_z$	$\sum_{i=2}^{3}\pi_{i,z}$
$\lambda = zxy$	$\sum_{i=2}^{3}\pi_i$	$\sum_{t\in\{y,z\}}\pi_{3,t}$	π_z
$\lambda = zyx$	π_3	$\sum_{i=2}^{3}\sum_{t\in\{y,z\}}\pi_{i,t}$	π_z

Only the position in the list matters.

ACM	$p(x,\lambda)$	$p(y,\lambda)$	$p(z,\lambda)$
$\lambda = xyz$	$\sum_{i=1}^{3}\sum_{t\in\{x,y,z\}}\pi_{i,t} = 1$	$\sum_{i=2}^{3}\sum_{t\in\{y,z\}}\pi_{i,t}$	$\pi_{3,z}$
$\lambda = xzy$	$\sum_{i=1}^{3}\sum_{t\in\{x,y,z\}}\pi_{i,t} = 1$	$\sum_{i=2}^{3}\pi_{i,y} + \pi_{3,z}$	$\sum_{i=2}^{3}\pi_{i,z}$
$\lambda = yxz$			

3-Alternative example: ACM

ACM	$p(x,\lambda)$	$p(y,\lambda)$	$p(z,\lambda)$
$\lambda = xyz$	$\sum_{i=1}^{3} \sum_{t \in \{x,y,z\}} \pi_{i,t} = 1$	$\sum_{i=2}^{3} \sum_{t \in \{y,z\}} \pi_{i,t}$	$\pi_{3,z}$
$\lambda = xzy$	$\sum_{i=1}^{3} \sum_{t \in \{x,y,z\}} \pi_{i,t} = 1$	$\sum_{i=2}^{3} \pi_{i,y} + \pi_{3,z}$	$\sum_{i=2}^{3} \pi_{i,z}$
$\lambda = yxz$	$\pi_{1,x} + \sum_{i=2}^{3} \pi_{i}$		

3-Alternative example: ACM

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

Here *predecessor set matters* for approval probs, not just position

Identification (DCM)

The identification news regarding DCM is good:

Identification (DCM)

The identification news regarding DCM is good:

Theorem

In the DCM, preferences and joint probabilities $\pi(i, t)$ *are uniquely identified by approval probabilities.*

Identification (DCM)

The identification news regarding DCM is good:

Theorem

In the DCM, preferences and joint probabilities $\pi(i, t)$ *are uniquely identified by approval probabilities.*

The result hinges strongly on the fact that here the position of an alternative in a list determines the approval probability.

Proof (sketch)

Preferences between any two alternatives x and y are identified by the ranking of approval probabilities in any lists in which x and y are in the same position (see example).

Proof (sketch)

Preferences between any two alternatives x and y are identified by the ranking of approval probabilities in any lists in which x and y are in the same position (see example).

Relabel alternatives in decreasing order of preference: $x_1, x_2, ..., x_n$.

Proof (sketch)

Preferences between any two alternatives x and y are identified by the ranking of approval probabilities in any lists in which x and y are in the same position (see example).

Relabel alternatives in decreasing order of preference: $x_1, x_2, ..., x_n$.

The approval prob of the worst alternative x_n in a list where it is in last position identifies $\pi(n, x_n)$.

Proof (sketch)

Preferences between any two alternatives x and y are identified by the ranking of approval probabilities in any lists in which x and y are in the same position (see example).

Relabel alternatives in decreasing order of preference: $x_1, x_2, ..., x_n$.

The approval prob of the worst alternative x_n in a list where it is in last position identifies $\pi(n, x_n)$.

If x_n is moved up one position, then:
- x_n still chosen when the threshold is x_n and the capacity is maximal;
- but now also chosen with the same threshold when capacity is only $n-1$.

Hence the difference in approval when x_n is in last position and when it is in position $n-1$ identifies $\pi(n-1, x_n)$:

Proof (sketch)

Preferences between any two alternatives x and y are identified by the ranking of approval probabilities in any lists in which x and y are in the same position (see example).

Relabel alternatives in decreasing order of preference: $x_1, x_2, ..., x_n$.

The approval prob of the worst alternative x_n in a list where it is in last position identifies $\pi(n, x_n)$.

If x_n is moved up one position, then:
- x_n still chosen when the threshold is x_n and the capacity is maximal;
- but now also chosen with the same threshold when capacity is only $n-1$.

Hence the difference in approval when x_n is in last position and when it is in position $n-1$ identifies $\pi(n-1, x_n)$:

Pushing x_n up in the list notch by notch then pins down $\pi(n-2, x_n), ..., \pi(1, x_n)$.

Proof (sketch) - ctnd.

Consider the difference in approval between the k^{th} best alternative and the $(k+1)^{th}$ best alternative when they are last.

The only event in which x_k is approved while x_{k+1} is not is when the threshold is x_k and the capacity is n (if capacity $< n$ or threshold $< x_k$ then neither is approved).

Hence the difference pins down $\pi(n, x_k)$ for all $k < n$.

Proof (sketch) - and finally...

Fix a position $j < n$ and for any $k < n$ compare the approval probs of x_k and x_{k+1} at position j.

The difference in approval probs is the prob of all the events in which the threshold is x_k and the capacity is *at least* j. Namely
$\pi(j, x_k) + ... + \pi(n, x_k)$

Proof (sketch) - and finally...

Fix a position $j < n$ and for any $k < n$ compare the approval probs of x_k and x_{k+1} at position j.

The difference in approval probs is the prob of all the events in which the threshold is x_k and the capacity is *at least* j. Namely $\pi(j, x_k) + ... + \pi(n, x_k)$

Repeat the exercise with position $j + 1$: now the difference is $\pi(j + 1, x_k) + ... + \pi(n, x_k)$

Proof (sketch) - and finally...

Fix a position $j < n$ and for any $k < n$ compare the approval probs of x_k and x_{k+1} at position j.

The difference in approval probs is the prob of all the events in which the threshold is x_k and the capacity is *at least* j. Namely
$\pi(j, x_k) + ... + \pi(n, x_k)$

Repeat the exercise with position $j+1$: now the difference is
$\pi(j+1, x_k) + ... + \pi(n, x_k)$

As good applied economists, now take a diff-in-diff to identify $\pi(j, x_k)$.

Identification (ACM)

In ACM, we need an additional assumption for full identification.

Identification (ACM)

In ACM, we need an additional assumption for full identification.

Theorem

In the ACM, preferences are uniquely identified. Moreover, if the probability distributions on thresholds and capacities are independent, they are uniquely identified by approval probabilities.

Identification (ACM)

In ACM, we need an additional assumption for full identification.

Theorem

In the ACM, preferences are uniquely identified. Moreover, if the probability distributions on thresholds and capacities are independent, they are uniquely identified by approval probabilities.

The need for some restriction is seen from the simplest example: $x \succ y$. Preferences are identified as in DCM. However...

...although we also identify $\pi(2, y) = p(y, xy)$ and $\pi(1, y) = p(y, yx) - p(y, xy)$

it is impossible to break down $\pi(1, x) + \pi(2, x)$ from $\pi(1, x) + \pi(2, x) + \pi(2, y) = p(x, yx)$ and $1 = p(x, xy)$.

ACM	$p(x, \lambda)$	$p(y, \lambda)$
$\lambda = xy$	$\pi_{1,x} + \pi_{2,x} + \pi_{1,y} + \pi_{2,y}$	$\pi_{2,y}$
$\lambda = yx$	$\pi_{1,x} + \pi_{2,x} + \pi_{2,y}$	$\pi_{1,y} + \pi_{2,y}$

Identification (ACM) - cntd.

In general, for any $k < n$, we can only hope to identify $\sum_{l \geq k} \pi(l, x_k)$.

Identification (ACM) - cntd.

In general, for any $k < n$, we can only hope to identify $\sum_{l \geq k} \pi(l, x_k)$.

On the other hand, assume independence, i.e. $\pi(i, t) = \pi_i \pi_t$. Then in the two alternative case we have:

Identification (ACM) - cntd.

In general, for any $k < n$, we can only hope to identify $\sum_{l \geq k} \pi(l, x_k)$.

On the other hand, assume independence, i.e. $\pi(i, t) = \pi_i \pi_t$. Then in the two alternative case we have:

$\pi_y = p(y, yx)$ identifying π_y (and therefore π_x)

Identification (ACM) - cntd.

In general, for any $k < n$, we can only hope to identify $\sum_{l \geq k} \pi(l, x_k)$.

On the other hand, assume independence, i.e. $\pi(i,t) = \pi_i \pi_t$. Then in the two alternative case we have:

$\pi_y = p(y, yx)$ identifying π_y (and therefore π_x)

$\pi_y \pi_2 = p(y, xy)$, identifying π_2 (and therefore π_1)

Identification (ACM) - cntd.

In general, for any $k < n$, we can only hope to identify $\sum_{l \geq k} \pi(l, x_k)$.

On the other hand, assume independence, i.e. $\pi(i, t) = \pi_i \pi_t$. Then in the two alternative case we have:

$\pi_y = p(y, yx)$ identifying π_y (and therefore π_x)

$\pi_y \pi_2 = p(y, xy)$, identifying π_2 (and therefore π_1)

The fact that π_1 is identified by the approval probs of *only* x_2 does generalise: **the approval probs of $x_{k+1}, ..., x_n$ identify π_k.**

This is the key for the recursive identifying algorithm in the proof (spared).

List Design

Primitives can be learned from the observation of approval behaviour across lists.

List Design

Primitives can be learned from the observation of approval behaviour across lists.

Suppose now you:
(1) can control the lists, and
(2) have an objective you want to maximise (list is a choice variable).

List Design

Primitives can be learned from the observation of approval behaviour across lists.

Suppose now you:
(1) can control the lists, and
(2) have an objective you want to maximise (list is a choice variable).

Which list maximises the objective?

List Design

Suppose you want to max the **total number of approvals**
(with the 'large number' assumption that approval probs are identified with the fraction of times, over a large total of times, that an alternative is approved).

List Design

Suppose you want to max the **total number of approvals**
(with the 'large number' assumption that approval probs are identified with the fraction of times, over a large total of times, that an alternative is approved).

This objective makes sense in several instances:
- maximise the number of clicks;
- maximise the number of news pieces read;
- maximise social network involvement through Likes and sharing;
- maximise the size of an online shopping cart, etc.

List Design

Suppose you want to max the **total number of approvals**
(with the 'large number' assumption that approval probs are identified with the fraction of times, over a large total of times, that an alternative is approved).

This objective makes sense in several instances:
- maximise the number of clicks;
- maximise the number of news pieces read;
- maximise social network involvement through Likes and sharing;
- maximise the size of an online shopping cart, etc.

We consider a (much) more general objective: maximize a **weighted sum** of the $p(x, \lambda)$. The weights $w(x)$ allow to include objectives such as revenue per click or favouring some specific alternatives.

Two benchmark results in list design

The two models have very contrasting implications for list design in respect of the stated objective.

Two benchmark results in list design

The two models have very contrasting implications for list design in respect of the stated objective.

Theorem

In the ACM, a list λ *is optimal iff it agrees with order of the weights* $w(x)$*, i.e.* $w(x) > w(y) \Rightarrow x\lambda y$.

Corollary (*List Invariance Principle*): In the ACM, if the weigths are all the same, then any list is optimal.

Two benchmark results in list design

The two models have very contrasting implications for list design in respect of the stated objective.

Theorem

In the ACM, a list λ is optimal iff it agrees with order of the weights $w(x)$, i.e. $w(x) > w(y) \Rightarrow x\lambda y$.

Corollary (*List Invariance Principle*): In the ACM, if the weigths are all the same, then any list is optimal.

Flash proof of Corollary (Credit: Yuhta Ishii): Take any capacity-threshold pair (i, t).
1) If $|\{x : x \succsim t\}| = k \leq i$, then k items are approved.
2) Otherwise, i items are approved.
3) Neither k nor i depends on the list. QED

Theorem

In the DCM:
1) If all weights are the same, then the unique maximiser of the number of approvals is the list that coincides with the preference order.
2) If the weights can differ and the probability distributions on thresholds and capacities are independent, then a list λ is optimal iff

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

3) In general, a list is optimal iff...

Intuition for proof with equal weights (DCM)

Since there are finitely many lists, a maximiser exists.

Suppose $x \succ y$ and look at swaps.

1. Take a list λ in which $\lambda(x) > \lambda(y)$
2. Swap x and y. Let λ' be the same as λ apart from $\lambda'(y) = \lambda(x)$ and $\lambda'(x) = \lambda(y)$

Intuition for proof with equal weights (DCM)

Since there are finitely many lists, a maximiser exists.

Suppose $x \succ y$ and look at swaps.
1. Take a list λ in which $\lambda(x) > \lambda(y)$
2. Swap x and y. Let λ' be the same as λ apart from $\lambda'(y) = \lambda(x)$ and $\lambda'(x) = \lambda(y)$

The approval probs of all z different from x and y are not affected by the swap.

Intuition for proof with equal weights (DCM)

Since there are finitely many lists, a maximiser exists.

Suppose $x \succ y$ and look at swaps.
1. Take a list λ in which $\lambda(x) > \lambda(y)$
2. Swap x and y. Let λ' be the same as λ apart from $\lambda'(y) = \lambda(x)$ and $\lambda'(x) = \lambda(y)$

The approval probs of all z different from x and y are not affected by the swap.

Any loss for y is a gain for x. If (i, t) leads to the approval of y in λ but not in λ', then $y \succeq t$ and $i = \lambda(y)$. Hence (i, t) leads to the approval of x in λ' and not in λ.

Intuition for proof with equal weights (DCM)

Since there are finitely many lists, a maximiser exists.

Suppose $x \succ y$ and look at swaps.
1. Take a list λ in which $\lambda(x) > \lambda(y)$
2. Swap x and y. Let λ' be the same as λ apart from $\lambda'(y) = \lambda(x)$ and $\lambda'(x) = \lambda(y)$

The approval probs of all z different from x and y are not affected by the swap.

Any loss for y is a gain for x. If (i, t) leads to the approval of y in λ but not in λ', then $y \succeq t$ and $i = \lambda(y)$. Hence (i, t) leads to the approval of x in λ' and not in λ.

Some gain for x is not a loss for y. The pair $(i, t) = (\lambda(y), x)$ leads to the approval of x in λ' but not in λ. But it never leads to the approval of y.

Intuition for proof (ACM)

As in DCM, a maximiser exists.

Suppose $x \succ y$.

1. Take a list λ in which $\lambda(x) = \lambda(y) + 1$
2. Swap to λ' - the same as λ apart from $\lambda'(y) = \lambda(x)$ and $\lambda'(x) = \lambda(y)$

Intuition for proof (ACM)

As in DCM, a maximiser exists.

Suppose $x \succ y$.

1. Take a list λ in which $\lambda(x) = \lambda(y) + 1$
2. Swap to λ' - the same as λ apart from $\lambda'(y) = \lambda(x)$ and $\lambda'(x) = \lambda(y)$

Because x and y are adjacent, this can only affect the approval probabilities of x and y.

Intuition for proof (ACM)

As in DCM, a maximiser exists.

Suppose $x \succ y$.

1. Take a list λ in which $\lambda(x) = \lambda(y) + 1$
2. Swap to λ' - the same as λ apart from $\lambda'(y) = \lambda(x)$ and $\lambda'(x) = \lambda(y)$

Because x and y are adjacent, this can only affect the approval probabilities of x and y.

Any loss for y is a gain for x. Suppose (i, t) leads to the approval of y in λ but not in λ'. Then (i, t) cannot lead to the approval of x in λ (capacity is exhausted) but has to lead to the approval of x in λ'.

Intuition for proof (ACM) - ctnd.

Is any gain for x a loss for y?

Intuition for proof (ACM) - ctnd.

Is any gain for x a loss for y?

Yes - *this is the crucial difference* from DCM.

Intuition for proof (ACM) - ctnd.

Is any gain for x a loss for y?

Yes - *this is the crucial difference* from DCM.

If (i, t) leads to the approval of x in λ' but not in λ, then y must have exhausted capacity in λ, meaning $y \succeq t$. Furthermore capacity cannot have been exhausted in λ when the agent reaches y. Hence (i, t) must lead to the approval of y in λ. But since x will consume the last bit of capacity, (i, t) cannot lead to the approval of y in λ'.

Intuition for proof (ACM) - ctnd.

Is any gain for x a loss for y?

Yes - *this is the crucial difference* from DCM.

If (i, t) leads to the approval of x in λ' but not in λ, then y must have exhausted capacity in λ, meaning $y \succeq t$. Furthermore capacity cannot have been exhausted in λ when the agent reaches y. Hence (i, t) must lead to the approval of y in λ. But since x will consume the last bit of capacity, (i, t) cannot lead to the approval of y in λ'.

(In DCM x gains from the swap in events that do not benefit y before the swap. In ACM this cannot happen: y must have absorbed capacity pre-swap for x to gain from the swap. Hence y loses what x gains.)

Comparatives

For a given preference $\succeq$ consider probability distributions π_a and π_b and their associated approval functions p_a and p_b, respectively.

Say that a is ***strongly more approving*** than b iff $p_a(x, \lambda) \geq p_b(x, \lambda)$ for all alternatives x and lists λ.

Comparatives

For a given preference $\succeq$ consider probability distributions π_a and π_b and their associated approval functions p_a and p_b, respectively.

Say that a is ***strongly more approving*** than b iff $p_a(x, \lambda) \geq p_b(x, \lambda)$ for all alternatives x and lists λ.

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

For a given preference $\succeq$ consider probability distributions π_a and π_b and their associated approval functions p_a and p_b, respectively.

Say that a is ***strongly more approving*** than b iff $p_a(x,\lambda) \geq p_b(x,\lambda)$ for all alternatives x and lists λ.

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

Numbering the alternatives from best to worst, any π defines uniquely a (univariate) numerical random variable X_π^- on $\{1,...,n\}$ that gives the *minimum* of any capacity-threshold pair (m,i), i.e.

$$\Pr\left(X_\pi^- = i\right) = \pi\left(\{(m,x_j) : \min(m,j) = i\}\right)$$

Comparatives

Given a π_a , let F_a denote the cdf.

Theorem. In the DCM, a is strongly more approving than b if and only if F_a first order stochastically dominates F_b

Given a π_a , let F_a denote the cdf.

Theorem. In the DCM, a is strongly more approving than b if and only if F_a first order stochastically dominates F_b

Theorem. In the ACM, a is weakly more approving than b if and only if $\mathbb{E}\left(X_{\pi_a}^-\right) \geq \mathbb{E}\left(X_{\pi_b}^-\right)$.

Characterisation

How do we know whether the agent is representable through DCM or ACM?

Characterisation

How do we know whether the agent is representable through DCM or ACM?

A1. If $\lambda(x) = \lambda'(x)$, then $p(x, \lambda) = p(x, \lambda')$ (only the position matters)

Characterisation

How do we know whether the agent is representable through DCM or ACM?

A1. If $\lambda(x) = \lambda'(x)$, then $p(x, \lambda) = p(x, \lambda')$ (only the position matters)

A2. If $\lambda(x) < \lambda'(x)$, then $p(x, \lambda) > p(x, \lambda')$. (higher positions are better)

Characterisation

How do we know whether the agent is representable through DCM or ACM?

A1. If $\lambda(x) = \lambda'(x)$, then $p(x, \lambda) = p(x, \lambda')$ (only the position matters)

A2. If $\lambda(x) < \lambda'(x)$, then $p(x, \lambda) > p(x, \lambda')$. (higher positions are better)

A3. If $\lambda(x) = \lambda'(y) = k$, $\mu(x) = \mu'(y) = k - 1$ and $p(x, \lambda) > p(y, \lambda')$ then

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

(supermodularity in quality and position)

Characterisation

How do we know whether the agent is representable through DCM or ACM?

A1. If $\lambda(x) = \lambda'(x)$, then $p(x, \lambda) = p(x, \lambda')$ (only the position matters)

A2. If $\lambda(x) < \lambda'(x)$, then $p(x, \lambda) > p(x, \lambda')$. (higher positions are better)

A3. If $\lambda(x) = \lambda'(y) = k$, $\mu(x) = \mu'(y) = k - 1$ and $p(x, \lambda) > p(y, \lambda')$ then

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

(supermodularity in quality and position)

A4a. There exists x such that if $\lambda(x) = 1$, then $p(x, \lambda) = 1$. (dominant alternative)
A4b. For all x and λ, $p(x, \lambda) > 0$. (positivity)
A4c. If $p(x, \lambda) = p(y, \lambda')$ and $\lambda(x) = \lambda'(y)$ then $x = y$.
(linearity)

Characterisation (DCM)

Theorem. A stochastic approval function is a DCM if and only if it satisfies A1-A4.

Characterisation (DCM)

Theorem. A stochastic approval function is a DCM if and only if it satisfies A1-A4.

Various extensions are possible with minor variations of the axioms:

- Preferences can be weak orders
- Depth can be zero
- Probabilities can be zero

Characterisation (DCM) - cntd.

B1. If $\lambda(x) \leq \lambda'(x)$, then $p(x, \lambda) \geq p(x, \lambda')$. (higher positions are weakly better)

B2. If $\lambda(x) = \lambda'(y) = k$, $\mu(x) = \mu'(y) = k - 1$ and $p(x, \lambda) \geq p(y, \lambda')$ then

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

(the advantage of better alternatives is weakly enhanced in higher positions)

Characterisation (DCM) - cntd.

B1. If $\lambda(x) \leq \lambda'(x)$, then $p(x, \lambda) \geq p(x, \lambda')$. (higher positions are weakly better)

B2. If $\lambda(x) = \lambda'(y) = k$, $\mu(x) = \mu'(y) = k - 1$ and $p(x, \lambda) \geq p(y, \lambda')$ then

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

(the advantage of better alternatives is weakly enhanced in higher positions)

Theorem

A stochastic approval function is a generalised DCM if and only if it satisfies B1 and B2.

Characterisation (ACM)

Partly but not entirely parallel to that of DCM.

A1' Only predecessor set matters

A2' Smaller predecessor sets are better

A3' Supermodularity in quality and smallness of predecessor set

A4' Positivity, linearity, dominant alternative

A5' If consecutive x and y are switched, y gains exactly what y loses.

Characterisation (ACM)

Wishful Theorem. A stochastic approval function is an ACM only if it satisfies A1-A5', and perhaps also if.

The big problem for a *neat* characterisation is that only the *number* of better alternatives counts in the predecessor set, whereas also the *identity* of worse alternatives counts.

Concluding remarks

A model of approval, not choice.

We have studied situations in which both the menu and the selections are typically 'large' (either 'pre-choice' or 'non-choice').

Concluding remarks

A model of approval, not choice.

We have studied situations in which both the menu and the selections are typically 'large' (either 'pre-choice' or 'non-choice').

Approval is intrinsically related to satisficing behaviour. We have provided two models that seems plausible, but others are possible.

Concluding remarks

A model of approval, not choice.

We have studied situations in which both the menu and the selections are typically 'large' (either 'pre-choice' or 'non-choice').

Approval is intrinsically related to satisficing behaviour. We have provided two models that seems plausible, but others are possible.

In particular, natural to look at non-stationary thresholds.

Concluding remarks

A model of approval, not choice.

We have studied situations in which both the menu and the selections are typically 'large' (either 'pre-choice' or 'non-choice').

Approval is intrinsically related to satisficing behaviour. We have provided two models that seems plausible, but others are possible.

In particular, natural to look at non-stationary thresholds.

List design seems to offer ample scope for further relevant research.

THANK YOU!