Choice with Endogenous Categorization

Andrew Ellis
LSE

Yusufcan Masatlioglu
Maryland

November 24, 2020

Motivation

> *"All organisms assign objects and events in the environment to separate classes or categories... Any species lacking this ability would quickly become extinct."*
> *–Ashby & Maddox (2005)*

Mental categories used to simplify decision making (Rosch 1978, Ashby and Lee 1993, Loken et al 2008)

Features of categorization:

1. Categories are **context dependent** (Barsalou 1985, Stewart et al 2002, Baird et al 1980)
2. How an object categorized **affects valuation** (Chernev 2011, Mogilner et al 2008, Wanke et al 1998)
3. Categories often take the form of "decision bounds" (Ashby 1992 with many follow ups; Anderson 1991, Love et al 2004)

Motivation

Categories in (Behavioral) Economics:

1. Safe vs. Risky
2. Past vs. Present vs. Future
3. Fair vs. Unfair
4. Equal vs. Unequal
5. Similar to x vs. similar to y vs. ...
6. Gain vs. Loss
7. Unambiguously better than status quo vs. not
8. Attribute 1 important vs. Attribute 2 important vs. ...

Contribution

- Provide & axiomatize model of reference-based categorization
  - Axioms: DM acts "rationally" within a category
- Identify categorization endogenously
- Nests a number of prominent economic models including:
  1. Gain-Loss utility: Tversky and Kahneman (1991)
  2. Status quo effects: Masatlioglu and Ok (2005)
  3. Salience: Bordalo, Gennaioli, and Shleifer (2012,13, etc.; BGS)
  4. Fairness preferences: Fehr and Schmidt (1999)
  5. Distributional preferences: Charness and Rabin (2002)
- Uncover relationships between these and other models
- Apply model to provide a behavioral foundation for Salient Thinking Model (BGS)

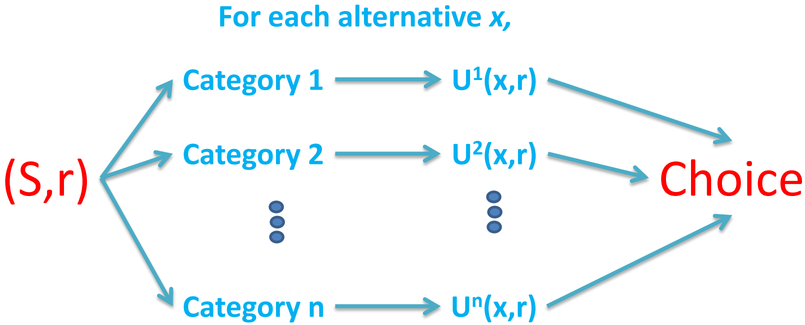
Model

- **Domain:** $X = R^n_{++}$ ($X = R^2_{++}$ for this talk)
  - $x = (x_1, x_2)$
- **Reference point:** $r \in X$
  - does not effect material outcome but may alter choice
  - potentially endogenous, i.e. varies with choice problem S
- **Data:** Either
  1. **complete and transitive** family of relations $\{\succsim_r\}_{r \in X}$, or
  2. **choice correspondence** c on finite subsets S of X

Model

- DM categorizes each alternative according to its characteristics and the reference's
  - Categories "partition" state space
  - Partition described by **category function** that maps r to categories
- Categorization affects preference
  - Preference within category unaffected by reference
  - Preference across categories may be

Categorical Thinking Model



DEFINITION

The family $\{\succsim_r\}_{r \in X}$ conforms to the **Categorical Thinking Model (CTM)** under category function $\mathcal{K} = (K^1, K^2, \ldots, K^m)$ if for each category k there is an additively separable **category utility function** U^k so that when $x \in K^k(r)$ and $y \in K^l(r)$ for some r

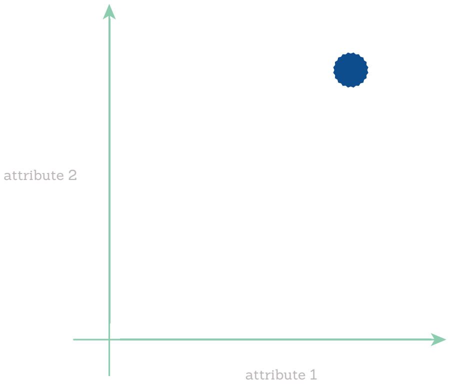
$$x \succsim_r y \iff U^k(x|r) \geq U^l(y|r)$$

and $U^k(\cdot|r)$ is an increasing transformation of $U^k(\cdot)$ for each $r \in X$ and category k.

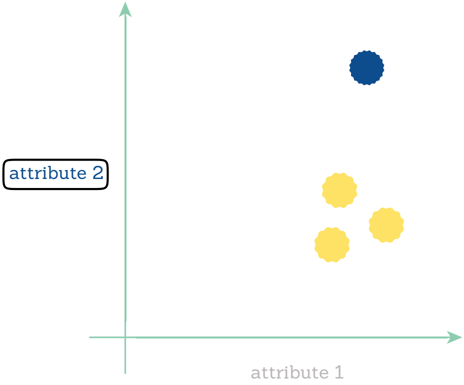
- increasing CTM if U_i^k is increasing in x_i for every category k and dimension i
- Regular CTM if $U^k(\cdot|r)$ is an affine transformation of $U^k(\cdot)$
- Strong CTM if $U^k(\cdot|r) = U^k(\cdot)$

An Illustration

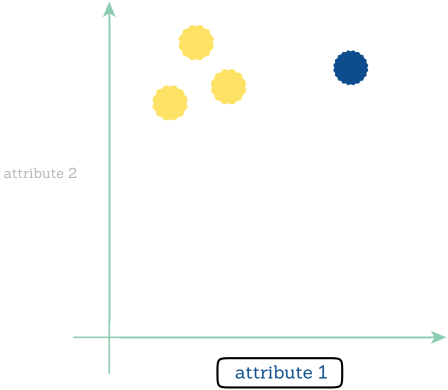
- Bordalo, Gennaioli, and Shleifer [2013]



BGS



BGS

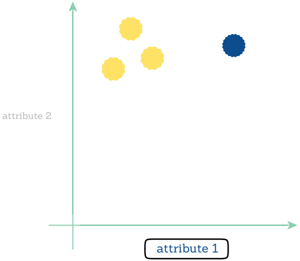


Salience Function

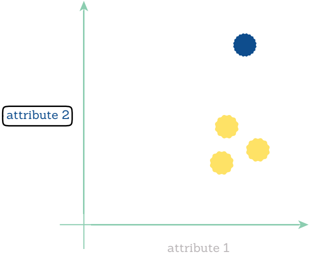
- Salience function $\sigma(x_i, r_i)$
  - measures the salience of x_i compared to r_i
- r_i: the average level

Salience

Attribute 1 is salient for good x if
$\sigma(x_1, \bar{x}_1) > \sigma(x_2, \bar{x}_2)$



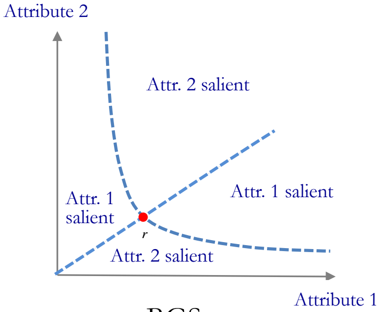
Attribute 2 is salient for good x if
$\sigma(x_1, \bar{x}_1) < \sigma(x_2, \bar{x}_2)$



BGS

A salience function proposed by BGS,

$$\sigma(a,b) = \frac{|a-b|}{a+b}$$



BGS

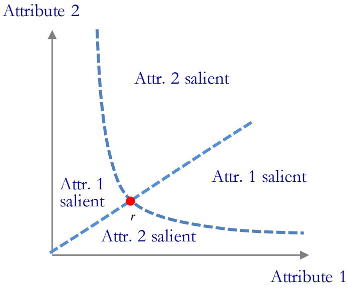
Utility Function

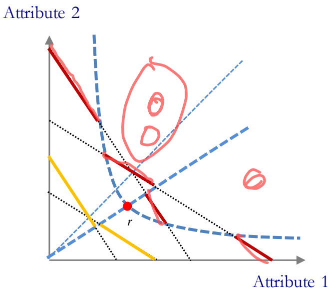
If attribute i "stands out", it receives higher "decision weight"

$$V_{BGS}(x|r) := \begin{cases} wu_1(x_1) + (1-w)u_2(x_2) & \text{if attribute 1 is salient} \\ (1-w)u_1(x_1) + wu(x_2) & \text{if attribute 2 is salient} \end{cases}$$

where $w \in (1/2, 1)$

BGS


Attribute 2
Attr. 2 salient
Attr. 1
salient
Attr. 1 salient
r
Attr. 2 salient
Attribute 1


Attribute 2
r
Attribute 1

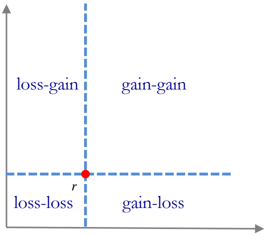
Tversky and Kahneman [1991]

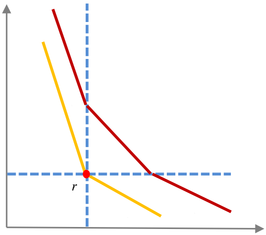
- Extends Prospect Theory to the case of riskless consumption bundles
- The workhorse of modeling behavior in risk-less environment
- Loss Aversion
- 6193 Google citations (as of July 2019)
- Exogenous reference point

Tversky and Kahneman [1991]

$$V_{TK}(x|r) = \begin{cases} (u(x_1) - u(r_1)) + (v(x_2) - v(r_2)) & \text{if gain-gain} \\ \lambda_1(u(x_1) - u(r_1)) + (v(x_2) - v(r_2)) & \text{if loss-gain} \\ (u(x_1) - u(r_1)) + \lambda_2(v(x_2) - v(r_2)) & \text{if gain-loss} \\ \lambda_1(u(x_1) - u(r_1)) + \lambda_2(v(x_2) - v(r_2)) & \text{if loss-loss} \end{cases}$$

- Losses hurt: λ_1 and λ_2 are greater than 1
- Constant Loss Aversion



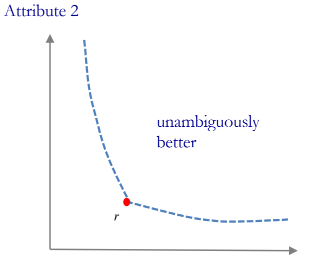


Masatlioglu and Ok [2005,2014]

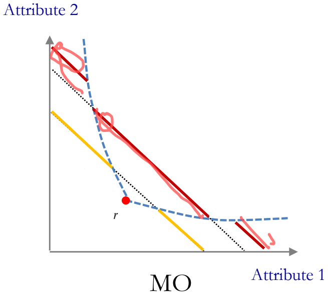
- Status Quo Bias
- Psychologically constrained utility maximization
- The status quo imposes a psychological constrain on decision makers ($\mathcal{Q}(r)$)

Masatlioglu and Ok [2005,2014]

$$V_{MO}(x|r) := \begin{cases} u(x_1) + v(x_2) & \text{if } x \in \mathcal{Q}(r) \\ u(x_1) + v(x_2) - c(r) & \text{otherwise} \end{cases}$$

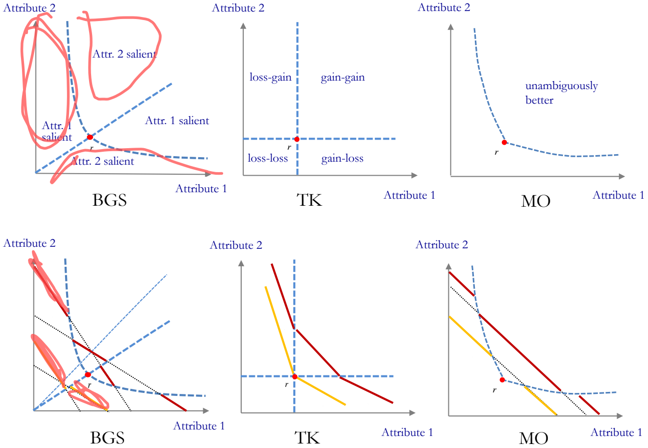


MO



MO

Comparing Models



Categories

- Every category is “well behaved”
- Almost everything is in a category
- Nothing is in two categories
- As the reference point changes, categories change continuously

Social

Categories

Definition

A vector-valued function $\mathcal{K} = (K^1, K^2, \ldots, K^m)$ is a *category function* if each $K^k : X \to 2^X$ satisfies the following properties:

1. $K^k(r)$ is a non-empty, regular open set, and $cl(K^k(r))$ is connected,
2. $\bigcup_{k=1}^m K^k(r)$ is dense,
3. $K^k(r) \bigcap K^l(r) = \emptyset$ for all $k \neq l$, and
4. $K^k(\cdot)$ is continuous.

Categories

- Categories arise from the psychology of the phenomenon to be modeled
- Psychology often makes unambiguous predictions about categorization (Gain-Loss)
- Other times only partial predictions are possible (Salience)
- Solutions:
  - non-choice data
  - **Reveal** categorization from choice

Revealing Categories

- identify categories by looking at how DM makes trade-offs

Definition

$LIS^k(x) = LIS^l(x)$ if there exists a neighborhood O of x so that

$$U^k(y) = U^k(x) \iff U^l(y) = U^l(x) \text{ for all } y \in O;$$

otherwise, $LIS^k(x) \neq LIS^l(x)$.

- $MRS^k(x) = MRS^l(x)$ vs. $MRS^k(x) \neq MRS^l(x)$
- $LIS^k(x) \neq LIS^l(x)$ implies categorization distorts trade-offs

Revealing Categories

Theorem

Let $\{\succsim_r\}_{r\in X}$ be a CTM. For any category k such that $LIS^k(x) \neq LIS^l(x)$ for every $x \in X$ and category $l \neq k$, category k is uniquely identified.

- categories are uniquely identified whenever categorization distorts tradeoffs
- must use discontinuities otherwise (or impossible)

Behavior

Axioms require DM to act "rationally", but only within a category
Within a category, the DM's choices:

1. have no cycles (Reference Irrelevance)
2. respect Monotonicity (Categorical Monotonicity)
3. are "additive" (Categorical Cancellation)
4. are continuous (Categorical Continuity)

Given earlier result, we take categories as given

Categorical Consistency

- For each category i, define $\succsim_i$ so that $x \succsim_i y$ if and only if there exists r such that $x, y \in R_i(r)$ and $x \succsim_r y$.
- $\succsim_i$ captures within category preference

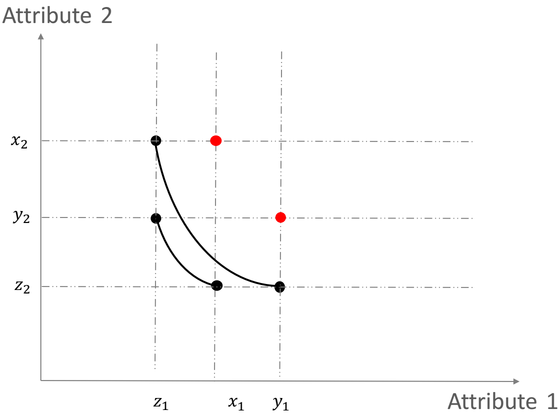
Axiom (Reference Irrelevance)

The relation $\succsim_i$ is acyclic.

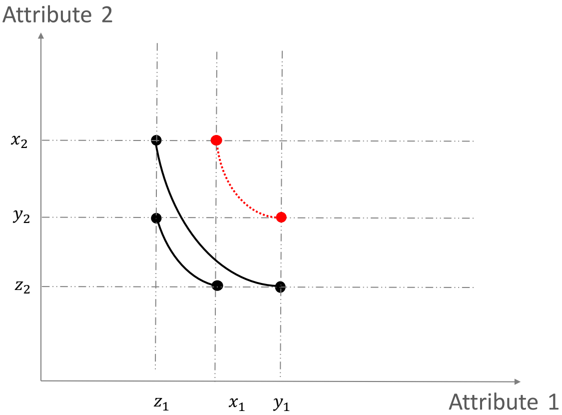
That is, if $x^1 \succsim_{r^1} x^2$ and $x^1, x^2 \in K^i(r^1)$,
$x^2 \succsim_{r^2} x^3$ and $x^2, x^3 \in K^i(r^2)$, ..., and
$x^m \succsim_{r^m} x^{m+1}$ and $x^m, x^{m+1} \in K^i(r^m)$,
then $x^{m+1} \not\succ_{r^{m+1}} x^1$ whenever $x^1, x^{m+1} \in K^i(r^{m+1})$.

- Let $\succsim_i^*$ be the transitive closure if $\succsim_i$

Double Cancellation



Double Cancellation



Categorical Double Cancellation

Axiom

Restricted to $K^k(r)$, $\succsim_r$ satisfies Double Cancellation.

Categorical Monotonicity

Axiom

For any $x, y, r \in X$: if $x \geq y$ and $x \neq y$, then $y \not\succsim_k^* x$ for any category k. In particular, if $x, y \in K^k(r)$, then $x \succ_r y$.

Categorical Continuity

Axiom

For any $r \in X$ and any $x \in \bigcup_i K^i(r)$, the sets

$$UC^j(x) = \{y \in K^j(r) : y \succ_r x\}$$

and

$$LC^j(x) = \{y \in K^j(r) : x \succ_r y\}$$

are open.
Moreover, the set

$$\{x \in \bigcup_i K^i(r) : UC^j(x) \bigcup LC^j(x) = K^j(r)$$
$$\text{and } UC^j(x) \neq K^j(r)$$
$$\text{and } LC^j(x) \neq K^j(r)\}$$

has an empty interior.

Categories

Assumption

For any category i, the following sets are connected:

- $E^i = \bigcup_{r \in X} K^i(r)$,
- $\{x \in E^i : x_j = s\}$ for all dimensions j and scalars s, and
- $\{y \in E^i : x \sim_i^* y\}$ for all $x \in E^i$

The BGS, TK, and MO categories all satisfy the Structure Assumption. Indeed, $E^i = \mathbb{R}^n_{++}$ for every category in these models.

Representation Theorem

Theorem

Under the Structure Assumption, the family $\{\succ_r\}_{r\in X}$ satisfies Reference Irrelevance as well as Categorical Monotonicity, Categorical Double Cancellation, and Categorical Continuity for $\mathcal{K}$ if and only if it conforms to increasing CTM under $\mathcal{K}$.

Representation Theorem

Proof outline:

1. Construct a utility function within each category
   - Each $\succsim_k^*$ is continuous and complete on E^k
   - Categorical cancellation implies $\succsim_k^*$ is "locally" additive
   - Chateauneuf & Wakker (1992) shows each has an additive representation
     - structure assumption allows application
2. Combine category utilities to get an overall representation
   - This modified to get affine or strong CTM

Skip to Summary

List A

- Many models of salience/attention
  - Gabaix and Laibson [2006]
  - Koszegi and Szeidl [2013]
  - Bhatia and Golman [2013]
  - Bordalo, Gennaioli, and Shleifer [2013] (a.k.a. BGS)
  - Cunningham [2013]
  - Gabaix [2014]
  - Schwartzstein [2014]
  - Bushong, Rabin and Schwartzstein et al. [2015]

List B

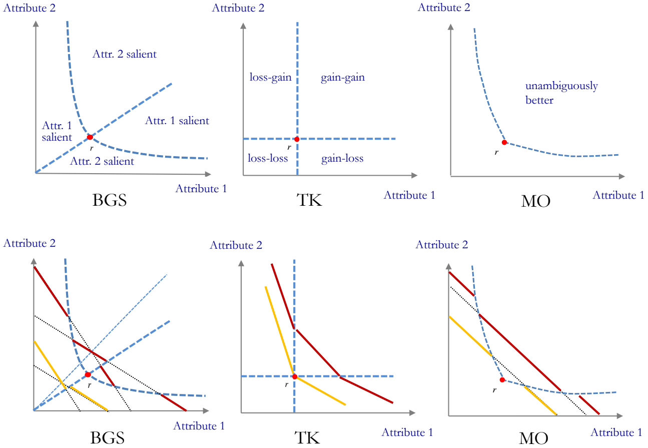
- Classical Theory
- Tversky and Kahneman [1991] (TK)
- Koszegi and Rabin [2006] (KR) (without expectation)
- Masatlioglu and Ok [2005] (MO)

List B: CTM
List A: Not CTM

Comparing Models

- Under additional assumptions (in paper), $V(\cdot|r)$ restricted to $R^k(r)$ is a positive affine transformation of U^k and the U^k's are reweightings of each other
- TK, MO, and BGS are special cases of such a model
- What distinguishes them?

Comparing Models



Comparing Models

- Strong Reference Irrelevance
  - Reference point does not affect ranking, only the categorys
  - If $x \in K^i(r) \bigcap K^i(r')$ and $y \in K^j(r) \bigcap K^j(r')$

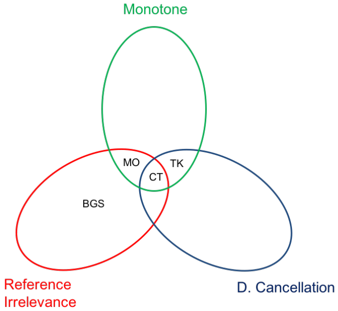
    $$x \succsim_r y \text{ iff } x \succsim_{r'} y$$
- Monotonicity
  - More is better
  - If $y \geq x$, then $y \succsim_r x$, strictly whenever $y \neq x$.
- Double Cancellation (not Categorical)

Comparing Models

	?	?	?	?
Monotonicity	✓	✗	✓	✓
Strong Reference Irrelevance	✓	✓	✗	✓
Double Cancellation	✓	✗	✓	✗

Comparing Models

	Classical	BGS	TK	MO
Monotonicity	✓	✗	✓	✓
Strong Reference Irrelevance	✓	✓	✗	✓
Double Cancellation	✓	✗	✓	✗



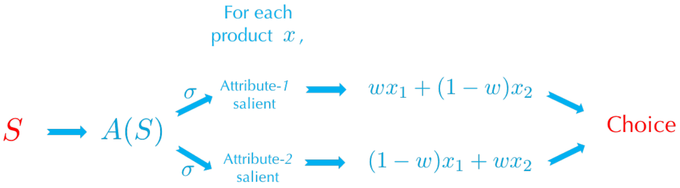
Preference-based CTM: Wrap up

In paper:

1. Full characterization of TK and BGS in this setting
2. CTM, Monotonicity, and SRI
3. New model based on simplification and similarity
4. Discussion of extensions like Fehr & Schidt (1999)

Conclusion

BGS



where $A(S)$ is average levels in S

- Now: only S and $c(S)$ observed

BGS

Roadmap:

1. Find BGS categories for given r
2. Since know mapping from S to $A(S)$, know categories!
3. Use this to provide axiomatization
   - See paper or Appendix
   - Versions of earlier axioms adapted to c

Categories in BGS

- $r = (r_1, r_2)$ the reference good
- $K^k(r)$ be the set of products that are k-salient for the reference point r

- Given a salience function σ, what are the properties of categories, $K^1(r)$ and $K^2(r)$?

Salience Function

- σ is symmetric, continuous, $\sigma(x,x) = \sigma(y,y)$ for all x, y, and increases in contrast: For $\epsilon, \epsilon' \geq 0$ with $\epsilon + \epsilon' > 0$
  If $x > \bar{x}$, then $\sigma(x+\epsilon, \bar{x}-\epsilon') > \sigma(x, \bar{x})$
  If $x < \bar{x}$, then $\sigma(x-\epsilon, \bar{x}+\epsilon') > \sigma(x, \bar{x})$
- Two additional properties
  - Diminishing Sensitivity: Salience decreases as the value of an attribute uniformly increases for all goods: For $\epsilon > 0$,

$$\sigma(x+\epsilon, \bar{x}+\epsilon) \leq \sigma(x, \bar{x})$$

  - Homogeneity: For all $\alpha > 0$,

$$\sigma(\alpha x, \alpha \bar{x}) = \sigma(x, \bar{x})$$

Categories in BGS

- S0 no bundle is both 1-salient and 2-salient, and almost every bundle is either one or the other.
- S1 making a bundle's less salient attribute closer to the reference point does not change the salience of the bundle.
- S6 if every attribute of a good differs from the reference point by the same percentage, then none of the attributes stands out.

Categories in BGS

S0 (Basic) For any $r \in X$: $K^1(r) \bigcap K^2(r) = \emptyset$, $K^1(r) \bigcup K^2(r)$ is dense in X, K^1, K^2 are continuous at r, and $K^1(r), K^2(r)$ are regular open sets.

S1 (Moderation) For any $\lambda \in [0,1]$ and $r \in X$: if $x \in K^k(r)$, $y_k = x_k$, and $y_{-k} = \lambda x_{-k} + (1-\lambda) r_{-k}$, then $y \in K^k(r)$.

S6 (Equal Salience) For any $x, r \in X$: if $\frac{x_1}{r_1} = \frac{x_2}{r_2}$ or $\frac{x_1}{r_1} = \frac{r_2}{x_2}$, then $x \notin K^k(r)$ for $k = 1, 2$.

Categories in BGS

Theorem

*The category function satisfies **S0**, **S1**, and **S6** if and only if it is generated by a homogeneous salience function σ. Any homogeneous salience function generates the same categories.*

- no need for more functional form assumptions
- all homogeneous salience functions lead to the same categories
- our figure independent of the salience function

More general salience functions

What if categories are not homogeneous? Can we identify salience function?

S2 Salience depends on the level but not identity of attribute

S3 "More salient than" is transitive

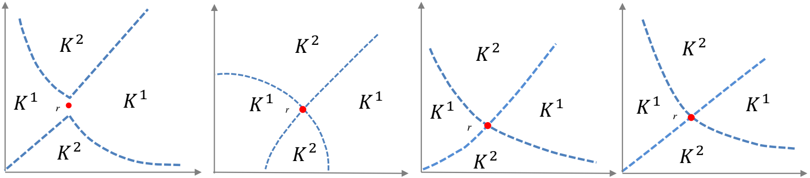
S4 Differences from the reference stand out

S5 An alternative's attribute stands out less when both it and the reference's are increased

More general salience functions

- S2 (Symmetry) If $(a,b) \in K^k(c,d)$, then $(c,d) \in K^k(a,b)$ and $(b,a) \in K^{-k}(d,c)$.
- S3 (Transitivity) If $(a_1,a_2) \notin K^2(r_1,r_2)$ and $(a_2,a_3) \notin K^2(r_2,r_3)$ then $(a_1,a_3) \notin K^2(r_1,r_3)$.
- S4 (Difference) For any x, y, z with $y \neq z$, $(x,y) \in K^2(x,z)$ and $(y,x) \in K^1(z,x)$.
- S5 (Diminishing Sensitivity) For any $x, y, K^1, K^2, \epsilon > 0$, if $(x,y) \notin K^1(r_1,r_2)$, then $(x+\epsilon, y) \notin K^1(r_1+\epsilon, r_2)$

More general salience functions



S0-S3 but not S4-S6 | S0-S4 but not S5-S6 | S0-S5 but not S6 | S0-S6

Figure: Properties **S0-S6** Illustrated

More general salience functions

Theorem

The category function satisfies:

1. ***S0-S4*** *if and only if there exists a salience function* σ *that generates them;*
2. ***S0-S5*** *if and only if the* σ *that generates it has diminishing sensitivity; and*
3. ***S0***, ***S1***, *and* ***S6*** *if and only if it satisfies* ***S0-S6*** *if and only if it is generated by a homogeneous salience function* σ.

More general salience functions

Proposition

Given that $\{\succsim_r\}_{r\in X}$ has a BGS representation, the categories are uniquely identified, namely $K^i(r)$ equals

$$int\left\{x \in X : \exists\epsilon > 0 \ s.t. \ \forall y \in B_\epsilon(x), \ y \sim_r x \iff y \sim_{r_x^i} x\right\}$$

where $r_x^1 = (x_1/2, x_2)$ and $r_x^2 = (x_1, x_2/2)$.

- By S4, $x \in K^1(r_x^1)$
- Compare slope of "IC" through x with r to that with r_x^i

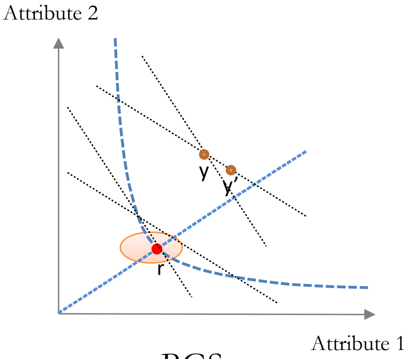
More general salience functions

Proposition

Given that c conforms to BGS with reference equal to average, the categories are uniquely identified.

Conclusion

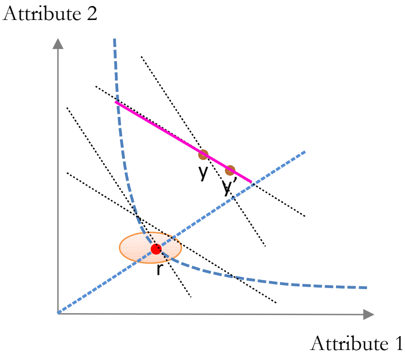
More general salience functions



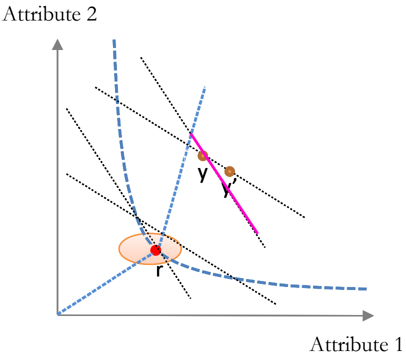
BGS

- Given r and y, pick y' close to y with $U_2(y) = U_2(y') \neq U_1(y')$
- Pick S concentrated around r so $A(\{y, y'\} \bigcup S) \approx A(S) \approx r$

More general salience functions



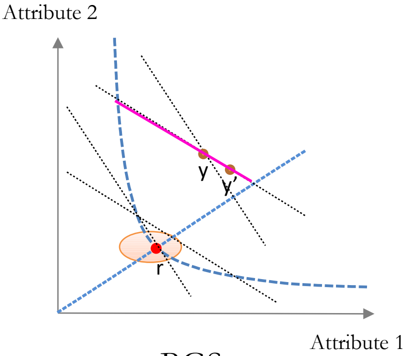
BGS



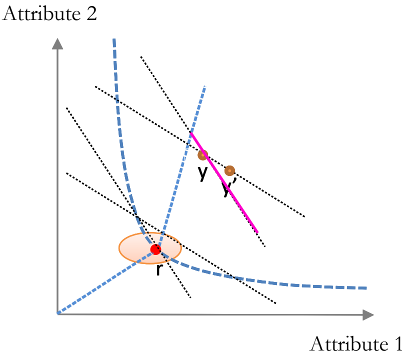
BGS

- Either y and y' chosen from $\{y, y'\} \bigcup S$ so $y \in R_2(r)$
- Or not, in which case $y \in R_1(r)$

More general salience functions



BGS



BGS

- If y and y' chosen, then $\sigma(y_2, r_2) > \sigma(y_1, r_1)$
- Otherwise, $\sigma(y_2, r_2) \leq \sigma(y_1, r_1)$
- Use symmetry to fill in σ

Characterization of BGS

What if reference point differs from the average?

Definition

Function A is a **generalized average** if for any $S = \{x^1, \ldots, x^m\} \in \mathcal{X}$:
(i) the function $x \mapsto A([S \setminus \{x_1\}] \bigcup \{x\})$ is continuous at x_1,
(ii) for any $\epsilon > 0$ and any finite $S' \in \bigcup_i R_i(A(S))$, there exists $S' \in \mathcal{X}$ so that $S^* \supset S \bigcup S'$, $d\left(A\left(S^*\right) - A(S)\right) < \epsilon$, and for any $x' \in S^* \setminus S'$, $\min_{x \in S} d(x', x) < \epsilon^2$, and
(iii) $A(S) \in co(S) \setminus ext(S)$ for all S.

Same result

Literature Review

Reference point

Reference –dependent choice

	Exogenous	Endogenous	
		Semi	Full
$\{U_r\}$	Tversky and Kahneman, 1991 Munro and Sugden, 2003 Sugden, 2003 Sagi, 2006 Salant and Rubinstein, 2008 …	Bodner and Prelec, 1994 Kivetz et al, 2004 Orhun, 2009 Bordalo et al, 2013 Tserenjigmid, 2015 Ellis and Masatlioglu, 2018	Koszegi and Rabin, 2006 Freeman, 2017 Kıbrıs, Masatlioglu, Suleymanov, 2018 Lim, Xi Zhi (RC), 2019
$(U, Q(r))$	Masatlioglu and Ok, 2005, 2014 Apesteguia and Ballester, 2009 Ortoleva, 2010, Masatlioglu and Nakajima, 2013 Dean, Kıbrıs, Masatlioglu, 2017 …		Ok et al, 2015 Kıbrıs, Masatlioglu, Suleymanov, 2018

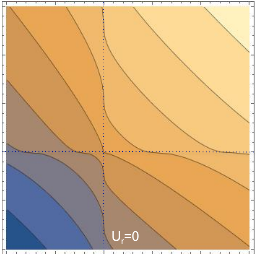
Wrapping up

- Categorization important ingredient in many behavioral models
- Our framework and model help understand
  - differences between modeling approaches to the same phenomenon
  - similarities between models of distinct phenomena
  - testable implications of important models
- categories can be derived endogenously from choice behavior

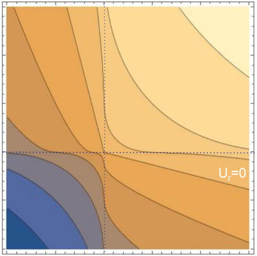
THANK YOU

"In those remote pages it has been written that the animals can be divided into (a) those that belong to the Emperor, (b) embalmed ones, (c) those that are trained, (d) suckling pigs, (e) mermaids, (f) fabulous ones, (g) stray dogs, (h) those that are included in the present classification, (i) those that tremble as if they are mad, (j) innumerable ones, (k) those drawn with a very fine camelhair brush, (l) others, (m) those that have just broken a flower vase, (n) those that look like flies from a long way off."
–Borges (1966), allegedly quoting the Celestial Emporium of Benevolent Knowledge

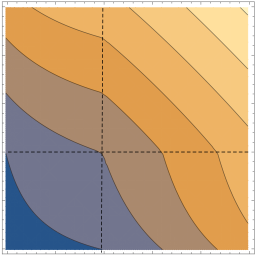
Continuous Salience



KR

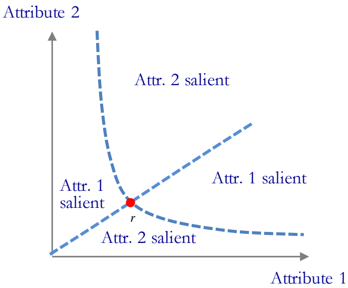


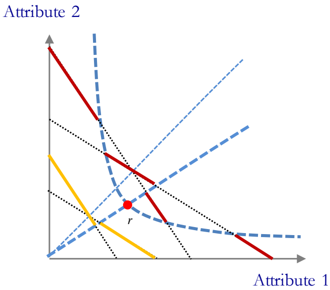
TK



Continuous BGS

BGS and Monotonicity





- What causes BGS to violate monotonicity?
- Is there a model "close" to BGS that satisfies it?

CTM and Monotonicity

- A version of CTM that
  - satisfies SRI,
  - satisfies Monotonicity
  - permits "salience" to affect preference
- Salience reweighs utilities through categories
- No category uniformly better or worse
- An CTM has **salience utilities** if different slopes in different categorys

CTM and Monotonicity

Proposition

Suppose there exists some x such that $(x, x) \in X$.
If $\{\succsim_r\}_{r \in X}$ is a CTM under $\mathcal{R}$ with at least two regions, has salience utilities and satisfies RI, then $\succsim_r$ ***violates Monotonicity*** *for some r.*

- No way to specify categories and weights that prevents violations of Monotonicity

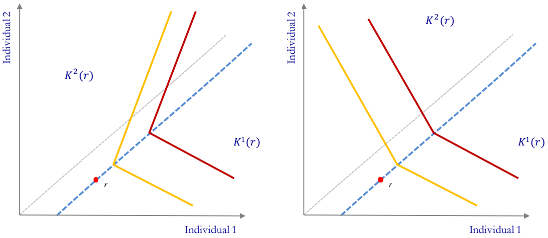


Figure: Left: Fehr Schmidt (1999) and Right: Charness and Rabin (2002)

Inequality Aversion

Categories $\mathcal{K}^{RIA} = (K^E, K^G)$ where

$$K^G(r) = \{x \in X : x_1 - r_1 > x_2 - r_2\},$$

$$K^E(r) = \{x \in X : x_1 - r_1 < x_2 - r_2\},$$

and

$$V_{RIA}(x|r) = \begin{cases} x_1 - \alpha[(x_1 - r_1) - (x_2 - r_2)] & \text{if } x \in K^E(r) \\ x_1 - \beta[(x_2 - r_2) - (x_1 - r_1)] & \text{if } x \in K^G(r) \end{cases}$$

Distributional preferences

$\mathcal{K}^{CR} = (K^1, K^2)$ where

$$K^j(r) = \left\{ x \in X : j = \arg\min_i (x_i - r_i) \right\}$$

and

$$V_{CR}(x|r) = \begin{cases} (1-\lambda)(x_1 - r_1) + \lambda[\delta(x_1 - r_1) + (1-\delta)\sum_k (x_k - r_k)] \\ (1-\lambda)(x_1 - r_1) + \lambda[\delta(x_2 - r_2) + (1-\delta)\sum_k (x_k - r_k)] \end{cases}$$

Back

Axioms for BGS

- First axiom is a version of WARP
- Consider two budget sets S^1 and S^2
- $x^1 \in c(S^1)$ and $x^2 \in S^1$
- $x^2 \in c(S^2)$ and $x^1 \in S^2$
- Then WARP implies $x^1 \in c(S^2)$

Axioms for BGS

- Consider two budget sets S^1 and S^2
- $x^1 \in c(S^1)$ and $x^2 \in S^1$
- $x^2 \in c(S^2)$ and $x^1 \in S^2$
- The salience of products does not change when the menu changes from S^1 to S^2
  - E.g., x^1 is 1-salient in both sets and x^2 is 2-salient in both sets
- Then $x^1 \in c(S^2)$.

Axioms for BGS

- Consider two budget sets S^1 and S^2
- $x^1 \in c(S^1)$ and $x^2 \in S^1$
- $x^2 \in c(S^2)$ and $x^1 \in S^2$
- The salience of products does not change when the menu changes from S^1 to S^2
  - $x^i \in R_k(A(S^1)) \cap R_k(A(S^2))$ for some k
- Then $x^1 \in c(S^2)$.

Axioms for BGS

Axiom (Salience-SARP)

For any finite sequences of pairs $(x^i, S^i)_{i=1}^n$ *such that for every* $i = 1, \ldots, n-1$,
$x^i \in c(S^i)$, $x^{i+1} \in S^i$, *and* $x^{i+1} \in R_k(A(S^i)) \cap R_k(A(S^{i+1}))$ *for some* $k \in \{1, 2\}$*:*
if $x^n \in c(S^n)$, $x^1 \in S^n$, *and* $x^1 \in R_k(A(S^1)) \bigcap R_k(A(S^n))$ *for some* k, *then* $x^1 \in c(S^n)$.

Axioms for BGS

- Simply restrict usual axioms to within categories
  - Categorical Monotonicity
  - Categorical Continuity
  - Categorical Linearity

Axioms for BGS

- The indifference curves in category 1 should be steeper than in category 2

Axiom (Salient Dimension Overvalued (SDO))

For $x, y \in S \bigcap S'$ with $x_k > y_k$ and $y_{-k} > x_{-k}$, if $x, y \in R_k(A(S))$, $x, y \in R_{-k}(A(S'))$, and $y \in c(S)$, then $x \notin c(S')$.

Axioms for BGS

▷ both salience and preference treat attributes symmetrically, permuting the attributes of all objects in the same way does not change rankings.

Axiom (Reflection)

For any $S \in \mathcal{X}$, if $(a, b) \in c(S)$ and T is the reflection of S, then $(b, a) \in c(T)$.

Characterization of BGS

Theorem

A choice correspondence $c(\cdot)$ satisfies Axioms 1-6 if and only if it has a salient thinking representation.

In paper: generalization to additive but not linear utility function