What Can be Learned from Behavior? Predictive Ability in Discrete Choice Environments

María José Boccardi

New York University Abu Dhabi
Social Science Division

January 5th 2018

Predictive Approach

- Models are theoretical constructions that simplify a complex reality to aid in the **understanding** and/or **prediction** of economic processes

- Axiomatizations provide **testable implications**
  - How well does the model describe observed behavior?
    - $\Rightarrow$ **Empirical validity**
  - How severe are violations (if any)?
    - $\Rightarrow$ **Extent** of the violations
    - $\Rightarrow$ **Sensitivity** of the test to detect violations $\rightarrow$ degree of identification of the underlying preference functional
  - How useful is the model to predict behavior?
    - **Predictive Ability** depends on both: (i) degree of identification of preferences and (ii) extent of observed deviations
  - What can we say about **competing models**?

- **What is the connection between theory and observable data?**

Why Predictive Ability?

- Empirical validity it is not the only relevant criterion
- For example: Utility Maximization

$A \Rightarrow$	$\{a, b, c\}$	$\{a, c\}$	$\{a, b\}$	$\{b, c\}$
C(A)		a		

- Assuming utility maximization, only implication is $C(\{a, b, c\}) \neq c$
- Obviously, this data is **rational** but **no power**
- Consider now the case where the sample is extended....

Why Predictive Ability?

- Empirical validity it is not the only relevant criterion
- For example: Utility Maximization

$A \Rightarrow$	$\{a, b, c\}$	$\{a, c\}$	$\{a, b\}$	$\{b, c\}$
$\mathbf{C(A)}$	$\neq c$	a		

- Given choices the only empirical implication is $C(\{a, b, c\}) \neq c$
- Only chance to falsify the model is to observe $\mathbf{C(\{a, b, c\})}$
- (Cond.) Probability to detect a violation under uniform behavior is $\frac{1}{3}$
- Assume $C(\{a, b, c\}) = \{a\}$

Why Predictive Ability?

- Empirical validity it is not the only relevant criterion
- For example: Utility Maximization

$A \Rightarrow$	$\{a, b, c\}$	$\{a, c\}$	$\{a, b\}$	$\{b, c\}$
$\mathbf{C(A)}$	a	a		

- Given choices the only empirical implication is $C(\{a, b, c\}) \neq c$
- Only chance to falsify the model is to observe $\mathbf{C(\{a, b, c\})}$
- (Cond.) Probability to detect a violation under uniform behavior is $\frac{1}{3}$
- Assume $C(\{a, b, c\}) = \{a\}$

- Empirical validity it is not the only relevant criterion
- For example: Utility Maximization

$A \Rightarrow$	$\{a,b,c\}$	$\{a,c\}$	$\{a,b\}$	$\{b,c\}$
$\mathbf{C(A)}$	a	a		

- This data is **consistent** with rationality
- Implications of observed data $C(\{a,b\}) = \{a\}$
- (Cond.) Probability to detect a violation on $\{a,b\}$ is $\frac{1}{2}$
- However no power if $\{b,c\}$ is added to the sample

Why Predictive Ability?

- Empirical validity it is not the only relevant criterion
- For example: Utility Maximization vs Limited consideration

$A \Rightarrow$	$\{a, b, c\}$	$\{a, c\}$	$\{a, b\}$	$\{b, c\}$
$\mathbf{C}(\mathbf{A})$	a	a		
$\mathbf{C}'(\mathbf{A})$	a	a		

- **Consistent** BUT also with models of bounded rationality
- If utility maximization, then $C(\{a, b\}) = \{a\}$
- However, $C(\{a, b\}) = \{b\}$ is consistent with limited consideration, where $\Gamma(\{a, b, c\}) = \{b, c\}$ and $\Gamma(\{A\}) = A$ for all $A \neq X$

Why Predictive Ability?

- Empirical validity it is not the only relevant criterion
- For example: Utility Maximization vs Limited consideration

$A \Rightarrow$	$\{a, b, c\}$	$\{a, c\}$	$\{a, b\}$	$\{b, c\}$
$\mathbf{C}(\mathbf{A})$	a	a	a	
$\mathbf{C}'(\mathbf{A})$	a	a	b	

- **Consistent** BUT also with models of bounded rationality
- If utility maximization, then $C(\{a, b\}) = \{a\}$
- However, $C(\{a, b\}) = \{b\}$ is consistent with limited consideration, where $\Gamma(\{a, b, c\}) = \{b, c\}$ and $\Gamma(\{A\}) = A$ for all $A \neq X$

Why Predictive Ability?

- Empirical validity it is not the only relevant criterion

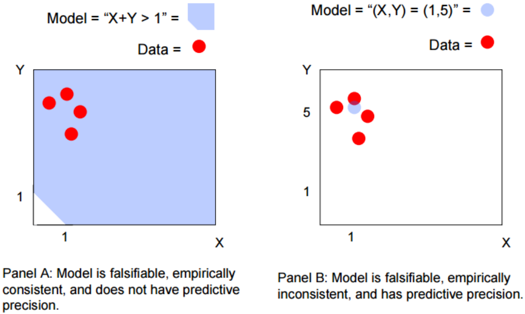

Figure: **Relation between empirical consistency, falsifiability and predictive ability**. Empirical consistency -fit- is important but does not implies usefulness of the model. Even when falsifiable it may not restrict behavior enough to allow for learning about underlying behavior. **Gabaix and Laibson (2008)**

- Lack of power/predictive ability may arise just because of **limited/incomplete data sets**

Intuition:

1. Choices $\Rightarrow$ information about underlying preferences
2. If incomplete data $\Rightarrow$ inferred preferences may be incomplete
3. If transitive closure of revealed preference is incomplete $\Rightarrow$ predictions for unobserved menus may not be unique
4. # Distinct predictions reflects # distinct preference relations that rationalize data $\Rightarrow$ **Model Uncertainty**
5. If RP has cycles $\Rightarrow$ errors to rationalize behavior $\Rightarrow$ Errors should be reflected on predictions $\Rightarrow$ **Error Uncertainty**
6. **Predictive distribution**: Distribution over $\succ$ consistent with RP + "error distribution" $\Rightarrow$ Dispersion of predictive distribution measures **predictive uncertainty**

Predictive Ability: Set-Up

- $\mathbf{X}$ finite set of alternatives, set of menus $2^{\mathbf{X}}_{*} = \{A \in 2^{\mathbf{X}} : |A| \geq 2\}$
- $f(a, A)$: the observed frequency of $a \in A$ chosen from $A \subseteq 2^{\mathbf{X}}_{*}$
- Model $\Rightarrow$ imposes restrictions on $\{f(a, A)\}_{a \in A, A \in 2^{\mathbf{X}}_{*}}$
  - For example: Utility maximization
  - Acyclicality of $R(f)$,

    $R(f) \in \mathbf{X} \times \mathbf{X} : (x, y) \in R(f) \Leftrightarrow f(x, A) > 0$ for $x, y \in A$ and some $A \in \mathcal{A}$

$A \Rightarrow$	$\{a, b, c\}$	$\{a, c\}$	$\{a, b\}$	$\{b, c\}$
C(A)	a	a		

  - $R(f) = \{(a, b); (a, c)\} \in \mathcal{R}\checkmark$, $\mathcal{R}$ LO on X
  - $\mathcal{R}^{ext}(f)$ set of all possible complete, transitive and antisymmetric preference relations, here
    1. $a \succ b \succ c$
    2. $a \succ c \succ b$

Revealed Preference Information

- Then

  1. If $|\mathcal{R}^{ext}(f)| = 1$, preferences are **'just identified'** or **'uniquely identified'**;
  2. If $|\mathcal{R}^{ext}(f)| > 1$, then preferences are **'not uniquely identified'**; and
  3. If $|\mathcal{R}^{ext}(f)| = 0$, observed data is **'not rationalizable'** or **'inconsistent'**.

Model Uncertainty: Predictions for $|\mathcal{R}^{ext}| \geq 1$

- $\{f(a, A)\}_{a \in A, A \in 2^{\mathbf{X}_*}}$ satisfies the utility maximization model, $|\mathcal{R}^{ext}| \geq 1$
- Each $R \in \mathcal{R}^{ext}$ is a complete linear order that rationalizes the data, **predictions**:

$$\hat{f}(a, A) = \hat{f}_{2^{X*}}(A) \times \mathbb{1}[a \; R \; b \quad \forall b \in A] \tag{1}$$

where $\hat{f}_{2^{X*}}(A)$ is the expected frequency of menu A in $\mathcal{O}$. [1]

Definition (Predictions for $2^{\mathbf{X}*}$)

$$\hat{f}^{R(f)}(a, A) = \frac{\gamma_f(a, A)}{r} \times \hat{f}_{2^{\mathbf{X}*}}(A) \quad \forall \;\; a \in A, A \in 2^{\mathbf{X}*} \tag{3}$$

where $\gamma_f(a, A) \equiv \sum_{R_i \in \mathcal{R}^{ext}(f)} \mathbb{1}[a \; R_i \; b \; \forall \; b \in A \setminus \{a\}]$ and $r = |\mathcal{R}^{ext}(f)|$.

-

[1] Unless otherwise stated it is assumed that

$$\hat{f}_{2^{X*}}(A) = \frac{1}{2^{|\mathbf{X}|} - |\mathbf{X}| - 1} = \gamma_{\mathbf{X}} \tag{2}$$

. Alternatively one may assume that $\hat{f}_{2^{X*}}(A) \propto |A|$.

Model Uncertainty: Predictions for $|\mathcal{R}^{ext}| \geq 1$ - Example

- $\{f(a, A)\}_{a \in A, A \in 2^{\mathbf{X}_*}}$ satisfies the utility maximization model, $|\mathcal{R}^{ext}| \geq 1$
- Each $R \in \mathcal{R}^{ext}$ is a complete linear order that rationalizes the data, **predictions**:

$$\hat{f}(a, A) = \hat{f}_{2^{X*}}(A) \times \mathbb{1}[a\ R\ b \ \ \forall b \in A]$$

where $\hat{f}_{2^{X*}}(A)$ is the expected frequency of menu A in $\mathcal{O}$.

Example (Predictions with 'multiplicity' of preferences that rationalize f)

$A \Rightarrow$	$\{a, b, c\}$	$\{a, c\}$	$\{a, b\}$	$\{b, c\}$
C(A)	a $f(a, A) = \frac{1}{2}$	a $f(a, A) = \frac{1}{2}$		
$R_1^{ext}(f) \Rightarrow a \succ b \succ c$	a	a	**a**	**b**
$R_2^{ext}(f) \Rightarrow a \succ c \succ b$	a	a	**a**	**c**

- $R(f) = \{(a, b); (a, c)\}$; $|\mathcal{R}^{ext}(f)| = 2$

- $\{f(a, A)\}_{a \in A, A \in 2^{\mathbf{X}_*}}$ satisfies the utility maximization model, $|\mathcal{R}^{ext}| \geq 1$

- Each $R \in \mathcal{R}^{ext}$ is a complete linear order that rationalizes the data, **predictions**:

$$\hat{f}(a, A) = \hat{f}_{2^{X*}}(A) \times \mathbb{1}\left[a \; R \; b \quad \forall b \in A\right]$$

where $\hat{f}_{2^{X*}}(A)$ is the expected frequency of menu A in $\mathcal{O}$.

Example (Predictions with 'multiplicity' of preferences that rationalize f)

$A \Rightarrow$	$\{a, b, c\}$	$\{a, c\}$	$\{a, b\}$	$\{b, c\}$
C(A)	a $f(a, A) = \frac{1}{2}$	a $f(a, A) = \frac{1}{2}$		
$R_1^{ext}(f) \Rightarrow a \succ b \succ c$	a	a	**a**	**b**
$R_2^{ext}(f) \Rightarrow a \succ c \succ b$	a	a	**a**	**c**

- $\hat{f}^{R_1^{ext}(f)}(b, \{b, c\}) = \hat{f}_{2^{\mathbf{X}_*}}(\{b, c\}) > 0$ while $\hat{f}^{R_2^{ext}(f)}(b, \{b, c\}) = 0$

Model Uncertainty: Predictions for $|\mathcal{R}^{ext}| \geq 1$ - Example

- $\{f(a,A)\}_{a \in A, A \in 2^{\mathbf{X}_*}}$ satisfies the utility maximization model, $|\mathcal{R}^{ext}| \geq 1$

Definition (Predictions for $2^{\mathbf{X}_*}$)

$$\hat{f}^{R(f)}(a,A) = \frac{\gamma_f(a,A)}{r} \times \hat{f}_{2^{\mathbf{X}_*}}(A) \quad \forall \;\; a \in A, A \in 2^{\mathbf{X}_*} \tag{4}$$

where $\gamma_f(a,A) \equiv \sum_{R_i \in \mathcal{R}^{ext}(f)} \mathbb{1}[a \; R_i \; b \; \forall \; b \in A \setminus \{a\}]$ and $r = |\mathcal{R}^{ext}(f)|$.

Example (Predictions Under Ass. 1)

$A \Rightarrow$	$\{a,b,c\}$	$\{a,c\}$	$\{a,b\}$	$\{b,c\}$
C(A)	a $f(a,A) = \frac{1}{2}$	a $f(a,A) = \frac{1}{2}$		
$R_1^{ext}(f) \Rightarrow a \succ b \succ c$	a	a	**a**	**b**
$R_2^{ext}(f) \Rightarrow a \succ c \succ b$	a	a	**a**	**c**
$\hat{f}(a,A)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	-
$\hat{f}(b,A)$	0	-	0	$\frac{1}{8}$
$\hat{f}(c,A)$	0	0	-	$\frac{1}{8}$

Model Uncertainty - Measure

Remark

If $|\mathcal{R}^{ext}(f)| = 1$ then $|\{\hat{f}(a, A) > 0, a \in A\}| = 1$ for all $A \in 2^{\mathbf{X}_*}$. If $|\mathcal{R}^{ext}(f)| > 1$, then $|\{\hat{f}(a, A) > 0, a \in A\}| > 1$ for some A

Definition ('Overall Entropy')

$$H(\hat{f}^{R(f)}) \equiv Entropy(\hat{f}^{R(f)}) = -\sum_{A \in 2^{\mathbf{X}*}} \sum_{a \in A} \hat{f}^{R(f)}(a, A) ln \left(\hat{f}^{R(f)}(a, A)\right)$$

$$\underbrace{=}_{Prop1} E_{\hat{f}_{2^{\mathbf{X}*}}} \left[\underbrace{H(\hat{f}^{R(f)}|A)}_{\text{Menu Entropy}} \right] + H(\hat{f}_{2^{\mathbf{X}*}})$$

Under Ass 1

$$H(\hat{f}^{R(f)}) = \gamma_{\mathbf{X}} \left[\sum_{A \in 2^{\mathbf{X}*}} \sum_{a \in A} \frac{\gamma_f(a, A)}{r} \ln \left(\frac{\gamma_f(a, A)}{r} \right) \right] + \ln \gamma_{\mathbf{X}}$$

Normalized Model Uncertainty

Definition ('Normalized Entropy')

$$NE_{\mathbf{x}}(\hat{f}^{R(f)}) = \frac{H(\hat{f}^{R(f)}) - H(\hat{f}_{2\mathbf{x}*})}{ME_{\mathbf{x}} - H(\hat{f}_{f=\emptyset}(A))} = \frac{E_{\hat{f}_{2\mathbf{x}*}(A)}\left[H(\hat{f}^{R(f)}|A)\right]}{E_{\hat{f}_{f=\emptyset}}(A)\left[H(\hat{f}_{f=\emptyset}|A)\right]}$$

where $ME_{\mathbf{x}} = H(\hat{f}_{f=\emptyset})$ and $\hat{f}_{f=\emptyset}(a, A) = \frac{1}{|A|}\gamma_{\mathbf{x}}$ for all $(a, A) \in \mathcal{O}$.

Example (Cont. Example)

$A \Rightarrow$	$\{a, b, c\}$	$\{a, c\}$	$\{a, b\}$	$\{b, c\}$
$\hat{f}(a, A)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	-
$\hat{f}(b, A)$	0	-	0	$\frac{1}{8}$
$\hat{f}(c, A)$	0	0	-	$\frac{1}{8}$

$$NE_{\mathbf{x}}(\hat{f}^{R(f)}) = \frac{3\ln 1 + \ln\frac{1}{2}}{\ln\frac{1}{3} + 3\ln\frac{1}{2}} = \frac{-\ln 2}{-\ln 3 - 3\ln 2} = \frac{\ln 2}{\ln 3 + 3\ln 2} \approx 0.2182$$

Normalized Model Uncertainty

Definition ('Normalized Entropy')

$$NE_{\mathbf{X}}(\hat{f}^{R(f)}) = \frac{H(\hat{f}^{R(f)}) - H(\hat{f}_{2^{\mathbf{X}*}})}{ME_{\mathbf{X}} - H(\hat{f}_{f=\emptyset}(A))} = \frac{E_{\hat{f}_{2^{\mathbf{X}*}}(A)}\left[H(\hat{f}^{R(f)}|A)\right]}{E_{\hat{f}_{f=\emptyset}}(A)\left[H(\hat{f}_{f=\emptyset}|A)\right]}$$

where $ME_{\mathbf{X}} = H(\hat{f}_{f=\emptyset})$ and $\hat{f}_{f=\emptyset}(a, A) = \frac{1}{|A|}\gamma_{\mathbf{X}}$ for all $(a, A) \in \mathcal{O}$.

Remark

Under Assumption 1

$$NE_{\mathbf{X}}(\hat{f}^{R(f)}) = \frac{\sum_{A \in 2^{\mathbf{X}*}} H(\hat{f}^{R(f)}|A)}{\sum_{n=2}^{|\mathbf{X}|} C_n^{|\mathbf{X}|} \ln(n)} \text{ with } C_n^{|\mathbf{X}|} = \frac{|\mathbf{X}|!}{n!\,(|\mathbf{X}| - n)!}$$

Normalized Model Uncertainty

Proposition

- $NE_{\mathbf{X}}(\hat{f}^{R(f)}) \in [0,1]$
- $\frac{\partial NE_{\mathbf{X}}(\hat{f}^{R(f)})}{\partial H(\hat{f}^{R(f)}|A)} = \frac{1}{E_{\hat{f}_{f=\emptyset}}(A)\left[H(\hat{f}_{f=\emptyset}|A)\right]^2} > 0$ *for all* $A \in 2^{\mathbf{X}*}$

Example (Cont. Example- No Information)

$A \Rightarrow$	$\{a,b,c\}$	$\{a,c\}$	$\{a,b\}$	$\{b,c\}$
$\hat{f}(a,A)$	$\frac{1}{12}$	$\frac{1}{8}$	$\frac{1}{8}$	-
$\hat{f}(b,A)$	$\frac{1}{12}$	-	$\frac{1}{8}$	$\frac{1}{8}$
$\hat{f}(c,A)$	$\frac{1}{12}$	$\frac{1}{8}$	-	$\frac{1}{8}$

$$NE_{\mathbf{X}}(\hat{f}^{R(f)}) = \frac{\ln\frac{1}{3} + 3\ln\frac{1}{2}}{\ln\frac{1}{3} + 3\ln\frac{1}{2}} = 1$$

Normalized Model Uncertainty

Proposition

- $NE_{\mathbf{X}}(\hat{f}^{R(f)}) \in [0,1]$
- $\frac{\partial NE_{\mathbf{X}}(\hat{f}^{R(f)})}{\partial H(\hat{f}^{R(f)}|A)} = \frac{1}{E_{\hat{f}_{f=\emptyset}}(A)\left[H(\hat{f}_{f=\emptyset}|A)\right]^2} > 0$ *for all* $A \in 2^{\mathbf{X}*}$

Example (Cont. Example- Some Information)

$A \Rightarrow$	$\{a,b,c\}$	$\{a,c\}$	$\{a,b\}$	$\{b,c\}$
	$f(a,A)=1$			
$\hat{f}(a,A)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	-
$\hat{f}(b,A)$	0	-	0	$\frac{1}{8}$
$\hat{f}(c,A)$	0	0	-	$\frac{1}{8}$

$$NE_{\mathbf{X}}(\hat{f}^{R(f)}) = \frac{3\ln 1 + \ln\frac{1}{2}}{\ln\frac{1}{3} + 3\ln\frac{1}{2}} = \frac{-\ln 2}{-\ln 3 - 3\ln 2} = \frac{\ln 2}{\ln 3 + 3\ln 2} \approx 0.2182$$

Normalized Model Uncertainty

Proposition

- $NE_{\mathbf{X}}(\hat{f}^{R(f)}) \in [0,1]$
- $\frac{\partial NE_{\mathbf{X}}(\hat{f}^{R(f)})}{\partial H(\hat{f}^{R(f)}|A)} = \frac{1}{E_{\hat{f}_{f=\emptyset}}(A)\left[H(\hat{f}_{f=\emptyset}|A)\right]^2} > 0$ *for all* $A \in 2^{\mathbf{X}*}$

Example (Cont. Example- Full Information)

$A \Rightarrow$	$\{a,b,c\}$	$\{a,c\}$	$\{a,b\}$	$\{b,c\}$
	$f(a,A) = \frac{1}{2}$			$f(b,A) = \frac{1}{2}$
$\hat{f}(a,A)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	-
$\hat{f}(b,A)$	0	-	0	$\frac{1}{4}$
$\hat{f}(c,A)$	0	0	-	0

$$NE_{\mathbf{X}}(\hat{f}^{R(f)}) = \frac{4\ln 1}{\ln\frac{1}{3} + 3\ln\frac{1}{2}} = 0$$

What if behavior is not consistent with Utility Maximization?

- Errors are common
  - Lack of attention
  - Misunderstanding of the problem or environment
  - Cognitive limitation
  - Inability to implement the optimal choice

- Identifying Assumption: decreasing probability assign to worse ranked alternatives

Assumption (Error structure)

Let $g^{\mathbf{R}}(\cdot, A) : A \to \{1, \ldots, |A|\}$ *be the (injective) rank function, i.e.* $g^{\mathbf{R}}(a, A) = 1 \Leftrightarrow a = m(\mathbf{R}, A)$, $g^{\mathbf{R}}(a, A) = 2 \Leftrightarrow a = m\left(\mathbf{R}, A \setminus m(\mathbf{R}, A)\right)$, *etc; where* $x = m(R, X) \Leftrightarrow xRa$ *for all* $a \in X \setminus \{x\}$. *Then, the* ***'error process'*** *is s.t.*

$$g^{\mathbf{R}}(a, A) < g^{\mathbf{R}}(b, A) \Rightarrow \eta^{\mathbf{R}}_{(a,A)} \geq \eta^{\mathbf{R}}_{(b,A)}$$

Then,

$$E(f(a, A)^{\mathbf{R}|model} | A) = \eta^{\mathbf{R}}_{(a,A)}$$

with $\sum_{a \in A} \eta^{\mathbf{R}}_{(a,A)} = 1$ *for all* $A \in 2^{\mathbf{X}*}$.

What if behavior is not consistent with Utility Maximization?

- Identifying Assumption: decreasing probability assign to worse ranked alternatives

$$
\begin{aligned}
\varepsilon &= 0 \text{ with probability } \eta^{\mathbf{R}}_{(m(\mathbf{R},A),A)} \\
&\vdots \\
&= g^{\mathbf{R}}(a,A) - 1 \quad \text{with probability } \eta^{\mathbf{R}}_{(a,A)} \\
&\vdots \\
&= |A| - 1 \text{ with probability } \eta^{\mathbf{R}}_{(b,A)} \text{ where } a\ \mathbf{R}\ b \text{ for all } a \in A \setminus \{b\}.
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}_{\hat{f}_{2\mathbf{X}*}}\left[\varepsilon_{(a,A)}\right] &= \sum_{A \in 2^{\mathbf{X}*}} \hat{f}_{2\mathbf{X}*}(A) \sum_{a \in A} \left[g^{\mathbf{R}}(a,A) - 1\right] \eta^{\mathbf{R}}_{(a,A)} \\
&= \sum_{(a,A) \in \mathcal{O}} |\{x \in A : x\mathbf{R}a\}| \left(\hat{f}_{2\mathbf{X}*}(A) \times \eta^{\mathbf{R}}_{(a,A)}\right)
\end{aligned}
$$

What if behavior is not consistent with Utility Maximization?

- Identifying Assumption: decreasing probability assign to worse ranked alternatives

$$\begin{aligned} \varepsilon &= 0 \text{ with probability } \eta^{\mathbf{R}}_{(m(\mathbf{R},A),A)} \\ &\vdots \\ &= g^{\mathbf{R}}(a,A) - 1 \quad \text{with probability } \eta^{\mathbf{R}}_{(a,A)} \\ &\vdots \\ &= |A| - 1 \text{ with probability } \eta^{\mathbf{R}}_{(b,A)} \text{ where } a \;\mathbf{R}\; b \text{ for all } a \in A \setminus \{b\}. \end{aligned}$$

$$\begin{aligned} \mathbb{E}_{\hat{f}_{2\mathbf{X}*}}\left[\varepsilon_{(a,A)}\right] &= \sum_{A \in 2^{\mathbf{X}*}} \hat{f}_{2\mathbf{X}*}(A) \sum_{a \in A} \left[g^{\mathbf{R}}(a,A) - 1\right] \eta^{\mathbf{R}}_{(a,A)} \\ &= \sum_{(a,A) \in \mathcal{O}} \left|\{x \in A : x\mathbf{R}a\}\right| \left(\hat{f}_{2\mathbf{X}*}(A) \times \eta^{\mathbf{R}}_{(a,A)}\right) \end{aligned}$$

- Recover underlying preferences by minimizing swaps index by Apesteguia and Ballester (JPE 2015)

Jump to Error Uncertainty

What if behavior is not consistent with Utility Maximization?

- Identifying Assumption: decreasing probability assign to worse ranked alternatives

Example

$\mathbf{X} = \{a, b, c\}$ and $a\mathbf{R}b\mathbf{R}c$. Assume

$$\eta^{\mathbf{R}}_{(a,\{a,b,c\})} = \frac{5}{8};\ \eta^{\mathbf{R}}_{(b,\{a,b,c\})} = \frac{1}{4};\ \eta^{\mathbf{R}}_{(c,\{a,b,c\})} = \frac{1}{8};$$

$$\eta^{\mathbf{R}}_{(a,\{a,b\})} = \eta^{\mathbf{R}}_{(b,\{b,c\})} = \eta^{\mathbf{R}}_{(a,\{a,c\})} = \frac{5}{7};\ \eta^{\mathbf{R}}_{(b,\{a,b\})} = \eta^{\mathbf{R}}_{(c,\{b,c\})} = \eta^{\mathbf{R}}_{(a,\{a,c\})} = \frac{2}{7}$$

Then

$$\begin{aligned}\mathbb{E}_{\hat{f}_{2\mathbf{X}*}(A)}\left[\varepsilon_{(a,A)}\right] &= \hat{f}_{2\mathbf{X}*}(\{a,b,c\})\left[0 \times \frac{5}{8} + 1 \times \frac{1}{4} + 2 \times \frac{1}{8}\right] \\ &+ \hat{f}_{2\mathbf{X}*}(\{a,b\})\left[0 \times \frac{5}{7} + 1 \times \frac{2}{7}\right] \\ &+ \hat{f}_{2\mathbf{X}*}(\{b,c\})\left[0 \times \frac{5}{7} + 1 \times \frac{2}{7}\right] \\ &+ \hat{f}_{2\mathbf{X}*}(\{a,c\})\left[0 \times \frac{5}{7} + 1 \times \frac{2}{7}\right] = \frac{19}{46}\end{aligned}$$

What if behavior is not consistent with Utility Maximization?

- Identifying Assumption: decreasing probability assign to worse ranked alternatives

Example (Continued)

$$\mathbb{E}_{\hat{f}_{2\mathbf{X}*}}\left[\varepsilon_{(a,A)}\right] = \sum_{(a,A)\in\mathcal{O}} |\{x \in A : x\mathbf{R}a\}| \left(\hat{f}_{2\mathbf{X}*(A)} \times \eta^{\mathbf{R}}_{(a,A)}\right)$$

$$= \frac{1}{4}\left[\eta^{\mathbf{R}}_{(b,\{a,b,c\})} + 2\eta^{\mathbf{R}}_{(c,\{a,b,c\})} + \eta^{\mathbf{R}}_{(b,\{a,b\})} + \eta^{\mathbf{R}}_{(c,\{b,c\})} + \eta^{\mathbf{R}}_{(c,\{a,c\})}\right]$$

Assume instead $a\mathbf{R}'c\mathbf{R}'b$,

$$\mathbb{E}_{\hat{f}_{2\mathbf{X}*}}\left[\varepsilon_{(a,A)'}\right] = \frac{1}{4}\left[2\eta^{\mathbf{R}}_{(b,\{a,b,c\})} + \eta^{\mathbf{R}}_{(c,\{a,b,c\})} + \eta^{\mathbf{R}}_{(b,\{a,b\})} + \eta^{\mathbf{R}}_{(b,\{b,c\})} + \eta^{\mathbf{R}}_{(c,\{a,c\})}\right]$$

$$\mathbb{E}_{\hat{f}_{2\mathbf{X}*}(A)}\left[\varepsilon_{(a,A)'} - \varepsilon_{(a,A)}\right] = \frac{1}{4}\left[\underbrace{\eta^{\mathbf{R}}_{(b,\{a,b,c\})} - \eta^{\mathbf{R}}_{(c,\{a,b,c\})}}_{\geq 0 \text{ Under Ass. } \downarrow\eta} + \underbrace{\eta^{\mathbf{R}}_{(b,\{b,c\})} - \eta^{\mathbf{R}}_{(c,\{b,c\})}}_{\geq 0 \text{ Under Ass. } \downarrow\eta}\right]$$

What if behavior is not consistent with Utility Maximization?

- **Recover preference information** by minimizing the empirical analog of the expectation of errors, i.e.

$$I(f) = \min_{\mathbf{R}} \sum_{(a,A)\in\mathcal{O}} f(a,A)\,|\{x \in A : x\mathbf{R}a\}| \qquad (5)$$

$$\hat{\mathbf{R}}_{\varepsilon}(f) \in arg\min_{\mathbf{R}} \sum_{(a,A)\in\mathcal{O}} f(a,A)\,|\{x \in A : x\mathbf{R}a\}| \qquad (6)$$

Example (Example continued)

Let R_{ijk} be the LO s.t. $iRjRk$ with $i,j,k \in \mathbf{X}$. Then,

$$I(f)|R_{abc} = \frac{19}{46}; \qquad I(f)|R_{acb} = \frac{107}{124} > \frac{19}{46};$$

$$I(f)|R_{bac} = \frac{121}{124} > \frac{19}{46}; \qquad I(f)|R_{bca} = \frac{173}{124} > \frac{19}{46};$$

$$I(f)|R_{cab} = \frac{159}{124} > \frac{19}{46}; \qquad I(f)|R_{cba} = \frac{204}{124} > \frac{19}{46};$$

therefore $\hat{\mathbf{R}}_{\varepsilon}(f) = R_{abc}$.

Error Uncertainty

- **Recover preference information** by minimizing the empirical analog of the expectation of errors, i.e.

$$I(f) = \underbrace{\min_{\mathbf{R}} \sum_{(a,A)\in\mathcal{O}} f(a,A)\,|\{x \in A : x\mathbf{R}a\}|}_{\text{Measure of Inconsistency}} \tag{7}$$

- If $I(f) > 0$ then $|\mathcal{R}(f)| = 0$
- If $\eta_{(m(R,A),A)} \geq \frac{1}{2}$ then it satisfies P-monotonicity and axiomatization from Apesteguia and Ballester (2015) follows.[2]

Remark

$$E(I(f)) = \sum_{n=2}^{|\mathbf{X}|} \frac{\beta_n}{\eta_n^{\mathbf{X}}} \sum_{i=1}^{n} (i-1)\eta^{\mathbf{X}}(i-1) \tag{8}$$

where $\beta_n = \sum_{A\in 2^{\mathbf{X}*}|f(A)>0} \mathbb{1}\,[|A| = n]$ and $\eta_n^{\mathbf{X}} = \sum_{i=1}^{n} \eta^{\mathbf{X}}(i-1)$.
That is, if $\eta^X \to U_X$ then $E(I(f)) \uparrow$.

[2] Satisfies continuity, rationality, concavity, piecewise linearity, ordinal consistency, disjoint composition, and neutrality.

Extended Predictions to Reflect Error Uncertainty

Definition (Predictive distribution based on 'Stochastic Extension')

Under Let $\widehat{\eta}^{\mathbf{X}}_{R,f}$ be the estimation of the error process. [a]

$$\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta}^{\mathbf{X}}_{R,f}}(a,A) \quad = \quad \hat{f}_{2\mathbf{X}*}(A)\left[\sum_{R\in\mathcal{R}^{ext}_\varepsilon(f)} \frac{1}{|\mathcal{R}^{ext}_\varepsilon(f)|}\frac{\widehat{\eta}^{\mathbf{X}}_{R_i,f}(g^{R_i}(a,A)-1)}{\sum_{j=1}^{|A|}\widehat{\eta}^{\mathbf{X}}_{R_i,f}(j-1)}\right]$$

for all $(a,A)\in\mathcal{O}$ with

$$\mathcal{R}^{ext}_\varepsilon(f) = \{R\in\mathcal{P} : \exists R_\varepsilon(f)\subseteq R \text{ with } R_\varepsilon(f)\in\mathcal{R}_\varepsilon(f)\}$$

where

$$\mathcal{R}_\varepsilon(f) = \{\mathbf{R}\subset\mathbf{X}\times\mathbf{X} : \mathbf{R}\in arg\min_{\mathbf{R}} \sum_{(a,A)\in\mathcal{O}} f(a,A)\,|\{x\in A : x\mathbf{R}a\}|\}$$

[a] Further assumptions may be required to recovered errors with incomplete data, e.g. common relative likelihood of errors (CRLE)

Example

Measuring Predictive Ability with Errors

- Natural extension from measures of model uncertainty,

$$PU_f \equiv H\left(\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta^{\mathbf{X}}_{R,f}}}\right) \underbrace{=}_{\text{Prop.}} E_{\hat{f}_{2\mathbf{X}*}}\left[H\left(\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta^{\mathbf{X}}_{R,f}}} | A\right)\right] + H(\hat{f}_{2\mathbf{X}*})$$

$$NPU_f \equiv \frac{PU_f - H(\hat{f}_{2\mathbf{X}*})}{H(\hat{f}_{f=\emptyset}) - H(\hat{f}_{2\mathbf{X}*})} = \frac{E_{\hat{f}_{2\mathbf{X}*}}\left[H\left(\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta^{\mathbf{X}}_{R,f}}} | A\right)\right]}{\gamma_{\mathbf{X}} \sum_{n=2}^{|\mathbf{X}|} C_n^{|\mathbf{X}|} \ln(n)}$$

Proposition

$$NPU_f \in [0, 1]$$

- Then, predictive ability is the opposite of predictive uncertainty

$$PAM_f = 1 - NPU_f$$

Predictive Ability

- *NPU_f* Reflects **Error Uncertainty:**
  - Reflects the dispersion of the η process

$$\text{'Error Uncertainty Measure'}_A = -\sum_{i=1}^{|A|} \frac{\eta^{\mathbf{X}}(i-1)}{\sum_{j=1}^{|A|} \eta^{\mathbf{X}}(j-1)} \ln\left[\frac{\eta^{\mathbf{X}}(i-1)}{\sum_{j=1}^{|A|} \eta^{\mathbf{X}}(j-1)}\right]$$

$$|\mathcal{R}_{\varepsilon}^{ext}(f)| = 1 \Rightarrow NPU_f = \frac{E_{\hat{f}_2 \mathbf{X}*}\{\text{'Error Uncertainty Measure'}_A\}}{\gamma_{\mathbf{X}} \sum_{n=2}^{|\mathbf{X}|} C_n^{|\mathbf{X}|} \, ln\,(n)}$$

- *NPU_f* Reflects **Model Uncertainty:**
  - Reflects uncertainty wrt underlying preferences

$$\text{'Model Uncertainty Measure'}_A = -\sum_{R_A \in \mathcal{R} \cap A \times A} P(R_A) \ln(R_A)$$

$$\text{with } P(R_A) = \sum_{R \in \mathcal{R}^{ext}(f)} \frac{1}{|\mathcal{R}^{ext}(f)|} \mathbb{1}[R \cap A \times A \equiv R_A]$$

$$\eta^{\mathbf{X}}(0) = 1 \Rightarrow \frac{E_{\hat{f}_2 \mathbf{X}*}\{\text{'Model Uncertainty Measure'}_A\}}{\gamma_{\mathbf{X}} \sum_{n=2}^{|\mathbf{X}|} C_n^{|\mathbf{X}|} \, ln\,(n)}$$

Jump to Application Jump to Conclusions

Simulations

- $\mathbf{X} = \{a, b, c, d, e\}$ such that $a \; R \; b \; R \; c \; R \; d \; R \; e$
- Probability that choices from menu $A \subseteq 2^{\mathbf{X}_*}$, $p \equiv \left\{\frac{1}{10}, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, \frac{9}{10}, 1\right\}$: 6 cases
- Different assumptions on the error process η
  - $\eta_1^X = [.25; .20; .20; .20; .15]$
  - $\eta_2^X = [.35; .25; .20; .15; .05]$
  - $\eta_3^X = [.40; .30; .20; .10; 0]$
  - $\eta_4^X = [.50; .30; .20; 0; 0]$
  - $\eta_5^X = [.70; .20; .075; .025; 0]$
  - $\eta_6^X = [1; 0; 0; 0; 0]$ Rational

Table: Test for differences in means for Predictive Ability Measure with respect to random behavior.

Assumption $\eta^{\mathbf{X}}$	**Probability of observing choices from $A \in 2^{X*}$**					
	0.10	**0.25**	**0.50**	**0.75**	**0.90**	**1.00**
η_1^X [.25;.20;.20;.20;.15]	0.0053 [0.6481]	-0.0109 [0.2866]	-0.0074 [0.2418]	-0.0012 [0.7757]	0.0031 [0.3238]	0.0000 [0.9866]
η_2^X [.35;.25;.20;.15;.05]	0.0086 [0.4585]	0.0064** [0.0136]	0.0091 [0.1692]	0.0165*** [0.0002]	0.0122*** [0.0001]	0.0148*** [0.0000]
η_3^X [.40;.30;.20;.10;0]	0.0078 [0.5043]	0.0201* [0.0545]	0.0165** [0.0123]	0.0280*** [0.0000]	0.0255*** [0.0000]	0.0250*** [0.0000]
η_4^X [.50;.30;.20;0;0]	0.0307*** [0.0089]	0.0435*** [0.0000]	0.0514*** [0.0000]	0.0594*** [0.0000]	0.0671*** [0.0000]	0.0653*** [0.0000]
η_5^X [.70;.20;.075;.025,0]	0.0531*** [0.0000]	0.1616*** [0.0000]	0.2157*** [0.0000]	0.2407*** [0.0000]	0.2534*** [0.0000]	0.2594*** [0.0000]
η_6^X (Rational)	0.1644*** [0.0000]	0.4794*** [0.0000]	0.7338*** [0.0000]	0.8398*** [0.0000]	0.8726*** [0.0000]	0.8901*** [0.0000]

More Tables Skip to Conclusion

Bounded Rationality Models

- Limited Attention vs Utility Maximization $\Rightarrow$ Depends on the extent of limited attention and 'incompleteness' of data Limited Attention

$$f(a, A) > 0 \Rightarrow a = \arg\max_{P} \Gamma(A) \text{ and } \Gamma(A) \subset B \subset A \Rightarrow \Gamma(B) = \Gamma(A)$$

- Categorization and Rationalization vs Utility Maximization $\Rightarrow$ Depends on the extent of limited attention and 'incompleteness' of data Categorization/Rationalization

$$\Psi(A) \equiv \{b \in B : f(b, B) > 0 | A \subset B, B \in \mathcal{A}, b \in A\}$$

- When dealing with limited datasets is important to understand the empirical content of considered models beyond their fit. More complex models of behavior may help understanding deviations but, the additional complexity may induce additional uncertainty when predicting behavior for unobserved economic environments.

Predictive Utility for GEU

- Relates to the predictive approach from Harless and Camerer (Ecta. 1994)
  - Fit and parsimonia, the latter understood as the number of patterns that the theory allows for
- Theories considered:
  - EU
  - Fan-out/Fan-in/Mixed fanning
  - RD concave/convex IC
  - PT
- Predictive ability measures also trade off fit and parsimonia $\sim$ predictive odds proposed by Harless and Camerer
- Main difference: predictive ability measures penalized for allowed patterns even when they are predicted with zero probability.

Tables

Conclusion

- Meaningful **trade off between fit and power** by measuring the complement of the uncertainty to predict behavior based on the predictive distribution

- **Assuming** that the DM may make **errors**, but still 'most preferred' alternative is chosen with higher probability

  - Number of distinct preference functionals that rationalizes behavior (with errors if necessary) $\Rightarrow$ **Model Uncertainty** – Reflects parsimonia a la Harless and Camerer (1994)

  - Error required to rationalized behavior $\Rightarrow$ **Error Uncertainty**

- Provides an assessment of the performance of a behavioral model even with **incomplete data** $\Rightarrow$ Allowing for **model comparison**

Thank You!

Example ('Model uncertainty')

Assume the conditions of Example with errors and let $\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta}^{\mathbf{X}}_{R,f}}(a, A) = f$ with R such that $aRbRc$ and

$$\widehat{\eta}^{\mathbf{X}}_{R,f}(i) = \begin{cases} \frac{5}{8} & \text{if } i = 0 \\ \frac{1}{4} & \text{if } i = 1 \\ \frac{1}{8} & \text{if } i = 2 \end{cases}$$

and $\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta}^{\mathbf{X}}_{R,f}}(a, A) = f$.

If the observed collection is given by $f'(a, \{a, b\}) = \frac{5}{7}$, $f'(b, \{a, b\}) = \frac{2}{7}$, $f'(a, \{a, c\}) = \frac{5}{7}$ and $f'(c, \{a, c\}) = \frac{2}{7}$; then further assumptions are necessary to estimate $\widehat{\eta}^{\mathbf{X}}_{R,f}$. The preference order recovered is $R_{f'}$ such that $aR_{f'}b$ and $aR_{f'}c$. There are two possible complete and transitive extensions $R^{ext1}_{f'}$ and $R^{ext2}_{f'}$ such that $aR^{ext1}_{f'}bR^{ext1}_{f'}c$ and $aR^{ext2}_{f'}cR^{ext2}_{f'}b$.

Example (Contd - Further structure (I))

Assumption 1 $\widehat{\eta}^{\mathbf{x}}_{R,f}(2) = 0$

$$\widehat{\eta}^{\mathbf{x}}_{R,f}(i) = \begin{cases} \frac{5}{7} & \text{if } i = 0 \\ \frac{2}{7} & \text{if } i = 1 \\ 0 & \text{if } i = 2. \end{cases}$$

Under this assumption, the predictions for the unseen menus are given by

$$\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta}^{\mathbf{x}}_{R,f}}(i, \{a, b, c\}) = \begin{cases} \frac{5}{7} f(\{a, b, c\}) & \text{if } i = a \\ \frac{1}{7} f(\{a, b, c\}) & \text{if } i = b \\ \frac{1}{7} f(\{a, b, c\}) & \text{if } i = c \end{cases}$$

and

$$\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta}^{\mathbf{x}}_{R,f}}(i, \{b, c\}) = \begin{cases} \frac{1}{2} f(\{b, c\}) & \text{if } i = b \\ \frac{1}{2} f(\{b, c\}) & \text{if } i = c \end{cases}$$

Example (Contd - Further structure (II))

Ass 2 $\widehat{\eta}^{\mathbf{X}}_{R,f}(2) = \alpha\widehat{\eta}^{\mathbf{X}}_{R,f}(1)$ with $\alpha \in (0,1)$. In this example, if $\alpha = .5$ we recover the exact distribution when data is complete. Then

$$\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta}^{\mathbf{X}}_{R,f}}(i,\{a,b,c\}) = \begin{cases} \frac{5}{8}f(\{a,b,c\}) & \text{if } i=a \\ \frac{3}{16}f(\{a,b,c\}) & \text{if } i=b \\ \frac{3}{16}f(\{a,b,c\}) & \text{if } i=c \end{cases}$$

and

$$\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta}^{\mathbf{X}}_{R,f}}(i,\{b,c\}) = \begin{cases} \frac{1}{2}f(\{b,c\}) & \text{if } i=b \\ \frac{1}{2}f(\{b,c\}) & \text{if } i=c \end{cases}$$

if $\widehat{\eta}^{\mathbf{X}}_{R,f}(2) = \alpha\widehat{\eta}^{\mathbf{X}}_{R,f}(1)$ and $\alpha = .5$.

Back

Simulations

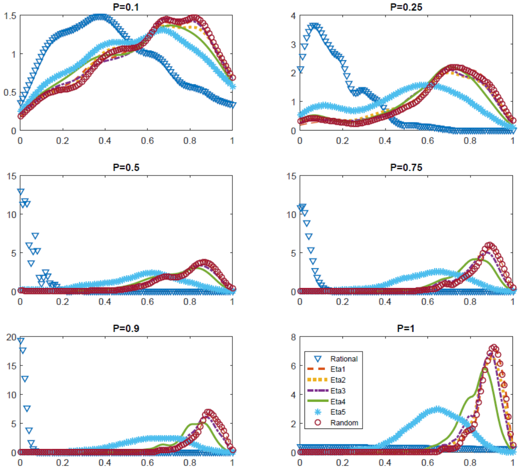

Figure: Distribution of Predictive Uncertainty per level of observability of menus and assumption on the error process.

Simulations

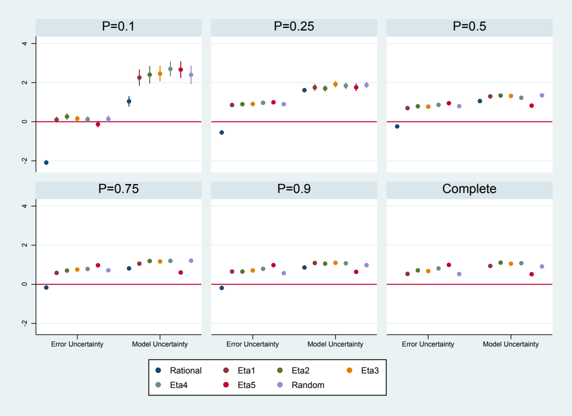

Figure: Correlations Predictive Uncertainty, Model Uncertainty and Error Uncertainty. Confidence intervals at 95% for all treatments.

Back Skip to Conclusion

Example (Example 3 in DR)

A	ae	ef	abd	ade	bde	af	be
$\{x \in A : f(x, A)\}$	e	f	d	a	b	f	b

Limited Attention RP info: (i) aPd, (ii) dPa or bPe and (iii) dPb or aPe. Extends to aPd, bPe and aPe. When predicting choices for $\tilde{A} = \{b, e, f\}$ the model predicts that either b or f can be chosen.

$$PAM_f|A = 1 - \frac{H\left(\hat{f}^{R_\varepsilon(f)}_{\widehat{\mathbf{x}}_{\eta_{R,f}}}|\{b, e, f\}\right)}{H\left(\hat{f}_{f=\emptyset}|\{b, e, f\}\right)} = 1 - \frac{-\left(\frac{1}{2}\ln\frac{1}{2} + \frac{1}{2}\ln\frac{1}{2}\right)}{-\left(\frac{1}{3}\ln\frac{1}{3} + \frac{1}{3}\ln\frac{1}{3} + \frac{1}{3}\ln\frac{1}{3}\right)} = 0.3691$$

Predictive Ability

Boccardi

Example (Example 3 in DR)

Utility Maximization This data cannot be rationalized, for example $f(e, \{a, e\}) > 0$ and $f(a, \{a, d, e\}) > 0$ implies a violation. Since all $R \in \mathcal{R}_\varepsilon(f)$ can be recovered by swapping three times one alternative for a dominated one, pooling observations: $\hat{\eta}_1^{\mathbf{X}} = \frac{4}{7}$ and $\hat{\eta}_2^{\mathbf{X}} = \frac{3}{7}$.

$$\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta^{\mathbf{X}}_{R,f}}}(x, \{b, e, f\}) = \begin{cases} \hat{f}_{2\mathbf{X}*}(\{b, e, f\}) \left[\frac{1}{3}\hat{\eta}_1^{\mathbf{X}} + \frac{2}{3}\hat{\eta}_2^{\mathbf{X}}\right] = \frac{10}{21}\hat{f}_{2\mathbf{X}*}(\{b, e, f\}) & \text{if } x = b \\ \hat{f}_{2\mathbf{X}*}(\{b, e, f\}) \left[\frac{1}{3}\hat{\eta}_2^{\mathbf{X}} + \frac{2}{3}\hat{\eta}_1^{\mathbf{X}}\right] = \frac{11}{21}\hat{f}_{2\mathbf{X}*}(\{b, e, f\}) & \text{if } x = f \\ 0 & \text{if } x = e \end{cases}$$

and the predictive ability measure conditional for this set is given by

$$PAM_f|A = 1 - \frac{H\left(\hat{f}^{\mathcal{R}_\varepsilon(f)}_{\widehat{\eta^{\mathbf{X}}_{\mathcal{R},f}}}|\{b, e, f\}\right)}{H\left(\hat{f}_{f=\emptyset}|\{b, e, f\}\right)} = 1 - \frac{-\left(\frac{10}{21}\ln\frac{10}{21} + \frac{11}{21}\ln\frac{11}{21}\right)}{-\left(\frac{1}{3}\ln\frac{1}{3} + \frac{1}{3}\ln\frac{1}{3} + \frac{1}{3}\ln\frac{1}{3}\right)} = 0.3701$$

Therefore, the predictive ability for the utility maximization model is slightly higher.

Back

Example (Example 4 in DR)

Consider the following observed menus and choices

A	ab	ad	ae	bd	be	de	abd	abe	bde
$\{x \in A : f(x, A)\}$	b	d	a	d	e	e	b	a	d

Rationalization/Categorization Observed choices yield dP^*b, bP^*a, aP^*e, eP^*d, eP^*b and dP^*a which is a complete order. And imposes (i) dPa or dPb from $R = \{a, b, d\}$ and dP^*b, (ii) bPa or bPe, from $R = \{a, b, e\}$ and bP^*a; and (iii) ePb or ePd from $R = \{b, d, e\}$ and eP^*d. The out of sample prediction for $\{a, d, e\}$ is that either a or d can be picked, but e would not be consistent. Predictions for $\{a, d, e\}$

$$PAM_f|A = 1 - \frac{H\left(\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta^{\mathbf{x}}_{R,f}}} | \{b, e, f\}\right)}{H\left(\hat{f}_{f=\emptyset} | \{b, e, f\}\right)} = 1 - \frac{-\left(\frac{1}{2}\ln\frac{1}{2} + \frac{1}{2}\ln\frac{1}{2}\right)}{-\left(\frac{1}{3}\ln\frac{1}{3} + \frac{1}{3}\ln\frac{1}{3} + \frac{1}{3}\ln\frac{1}{3}\right)} = 0.3691$$

Example (Example 4 in DR)

Utility Maximization Cannot be rationalized, since, for example $f(d, \{b, d\}) > 0$ and $f(b, \{a, b, d\}) > 0$. However, the only well-behaved preference order that minimizes the swap index, and it is given by $R_\varepsilon(f)$ such that $dR_\varepsilon(f)bR_\varepsilon(f)aR_\varepsilon(f)e$; and the estimation of the error process delivers $\hat{\eta}_1^{\mathbf{X}} = \frac{5}{9}$ and $\hat{\eta}_2^{\mathbf{X}} = \frac{4}{9}$. Then, the predictive distribution for $\{a, d, e\}$ is given by

$$\hat{f}^{R_\varepsilon(f)}_{\widehat{\eta^{\mathbf{X}}_{R,f}}}(x, \{a, d, e\}) = \begin{cases} \hat{f}_{2\mathbf{X}*}(\{a, d, e\}) \left[\eta_2^{\mathbf{X}}\right] = \frac{4}{9}\hat{f}_{2\mathbf{X}*}(\{a, d, e\}) & \text{if } x = a \\ \hat{f}_{2\mathbf{X}*}(\{a, d, e\}) \left[\hat{\eta}_1^{\mathbf{X}}\right] = \frac{5}{9}\hat{f}_{2\mathbf{X}*}(\{a, d, e\}) & \text{if } x = d \\ 0 & \text{if } x = e \end{cases}$$

and the predictive ability measure conditional for this set is given by

$$PAM_f|A = 1 - \frac{H\left(\hat{f}^{\mathcal{R}_\varepsilon(f)}_{\widehat{\eta^{\mathbf{X}}_{\mathcal{R},f}}}|\{a, d, e\}\right)}{H\left(\hat{f}_{f=\emptyset}|\{a, d, e\}\right)} = 1 - \frac{-\left(\frac{4}{9}\ln\frac{4}{9} + \frac{5}{9}\ln\frac{5}{9}\right)}{-\left(\frac{1}{3}\ln\frac{1}{3} + \frac{1}{3}\ln\frac{1}{3} + \frac{1}{3}\ln\frac{1}{3}\right)} = 0.3747$$

Back

Tables GEU

Table: Table V.- Harless real gains from unit triangle interior. Outcomes: \$0,\$3,\$6. Probabilities: $S_1(.84, .14, .02)$; $R_1(.89, .01, .10)$; $S_3(.04, .94, .02)$; $R_3(.09, .81, .10)$; $S_5(.44, .14, .42)$; $R_5(.49, .01, .50)$; $S_7(.04, .14, .82)$; $R_7(.09, .01, .90)$

Criteria	EU	Fan-out	Fan-in	MF	RDcave	RDvex	PT
				Results			
% Explained	.310	.429	.631	.583	.905	.571	.857
% Allowed	.063	.125	.375	.375	.813	.563	.563
# patterns	2	6	6	13	9	9	5
% Max	.310	.429	.774	.774	.964	.893	.893
Z-stat	5.245	6.051	5.376	4.702	2.922	.870	5.924
Selten	.247	.304	.256	.208	.092	.009	.295
Error	.366	.219	.166	.216	.092	.216	.108
χ^2-stat	59.92	29.22	15.37	29.18	6.88	29.18	8.55
Avg. LOO	.656	1.095	1.411	1.392	2.056	1.567	1.845
NEU	.886	.855	.885	.837	.885	.839	.885

Tables GEU

Table: Table V.- Harless real gains from unit triangle interior. Outcomes: \$0,\$3,\$6. Probabilities: $S_1(.84, .14, .02)$; $R_1(.89, .01, .10)$; $S_3(.04, .94, .02)$; $R_3(.09, .81, .10)$; $S_5(.44, .14, .42)$; $R_5(.49, .01, .50)$; $S_7(.04, .14, .82)$; $R_7(.09, .01, .90)$

Criteria	EU	Fan-out	Fan-in	MF	RDcave	RDvex	PT
				Ranking			
% Explained	7	3	4	1	5	2	6
Expl/Allowed	1	2	4	6	7	5	3
Expl/Max	1	4	5	3	7	2	6
Z-stat	1	3	4	6	7	2	5
Selten	1	3	4	6	7	2	5
Error	7	3	4	1	4	2	4
PO vs EU - HC	1	2	5	6	7	3	4
Avg. LOO	1	2	4	3	7	5	6
Dist. true	4	3	4	1	4	2	4
Dist. rnd	7	3	4	1	4	2	4

Table: Table VI Harless real losses from unit triangle interior. Outcomes: -$4,-$2,$0. Probabilities: $S_1(.8,.18,.02)$; $R_1(.88,.02,.10)$; $S_3(.02,.96,.02)$; $R_3(.1,.8,.1)$; $S_5(.41,.18,.41)$; $R_5(.49,.02,.49)$; $S_7(.02,.18,.80)$; $R_7(.10,.02,.88)$

Criteria	EU	Fan-out	Fan-in	MF	RDcave	RDvex	PT
				Results			
% Explained	.279	.380	.595	.886	.595	.684	.317
% Allowed	.125	.375	.375	.813	.563	.563	.188
# patterns	2	6	6	13	9	9	3
% Max	.279	.646	.646	.949	.823	.823	.392
Z-stat	3.785	.614	4.720	2.505	1.627	3.298	3.420
Selten	.154	.005	.220	.074	.032	.121	.129
Error	.281	.281	.222	.141	.248	.234	.362
χ^2-stat	18.72	18.72	7.46	3.08	11.31	10.19	22.79
Avg. LOO	1.266	1.456	1.866	2.369	1.772	1.962	1.256
NEU	.965	.965	.942	.930	.950	.948	.969

Tables GEU

Table: Table VI Harless real losses from unit triangle interior. Outcomes: -$4,-$2,$0. Probabilities: $S_1(.8, .18, .02)$; $R_1(.88, .02, .10)$; $S_3(.02, .96, .02)$; $R_3(.1, .8, .1)$; $S_5(.41, .18, .41)$; $R_5(.49, .02, .49)$; $S_7(.02, .18, .80)$; $R_7(.10, .02, .88)$

Criteria	EU	Fan-out	Fan-in	MF	RDcave	RDvex	PT
				Ranking			
% Explained	7	5	3	1	3	2	5
Expl/Allowed	1	7	3	5	6	4	2
Expl/Max	1	7	3	2	6	4	5
Z-stat	2	7	1	5	6	4	3
Selten	2	7	1	5	6	4	3
Error	5	5	2	1	4	3	7
PO vs EU - HC	1	4	2	7	6	5	3
Avg. LOO	2	3	5	7	4	6	1
Dist. true	6	6	2	1	4	3	4
Dist. rnd	6	6	2	1	4	3	5

Tables GEU

Table: Table VII Chew and Waller: \$0,-\$40,\$100. Probabilities: $S_o(0,1,0)$; $R_o(.5,0,.5)$; $S_i(0,1,0)$; $R_i(.05,.9,.05)$; $S_l(.9,.1,0)$; $R_l(.95,0,.05)$; $S_h(0,.1,.9)$; $R_h(.05,0,.95)$

Criteria	EU	Fan-out	Fan-in	MF	RDcave	RDvex	PT
				Results			
% Explained	.212	.697	.323	.879	.374	.838	.798
% Allowed	.125	.375	.375	.813	.563	.563	.625
# patterns	2	6	6	13	9	9	10
% Max	.374	.758	.758	.990	.899	.899	.929
Z-stat	3.07	6.86	-2.33	2.65	-4.80	6.14	4.55
Selten	.087	.322	-.052	.066	-.189	.276	.173
Error	.339	.178	.339	.121	.458	.143	.158
χ^2-stat	81.99	15.16	89.89	10.35	90.19	11.15	14.17
Avg. LOO	1.204	1.824	1.376	2.143	1.763	2.056	1.959
NEU	.985	.866	.985	.858	1.000	.859	.863

Tables GEU

Table: Table VII Chew and Waller: \$0,-\$40,\$100. Probabilities: $S_o(0, 1, 0)$; $R_o(.5, 0, .5)$; $S_i(0, 1, 0)$; $R_i(.05, .9, .05)$; $S_l(.9, .1, 0)$; $R_l(.95, 0, .05)$; $S_h(0, .1, .9)$; $R_h(.05, 0, .95)$

Criteria	EU	Fan-out	Fan-in	MF	RDcave	RDvex	PT
				Ranking			
% Explained	7	4	6	1	5	2	3
Expl/Allowed	2	1	6	5	7	3	4
Expl/Max	5	2	6	3	7	1	4
Z-stat	4	1	6	5	7	2	3
Selten	4	1	6	5	7	2	3
Error	5	4	5	1	7	2	3
PO vs EU - HC	5	1	6	4	7	2	3
Avg. LOO	1	4	2	7	3	6	5
Dist. true	5	4	7	1	6	2	3
Dist. rnd	5	3	6	2	7	1	4

Table: Table VIII - Sopher and Gigliotti - Common consequence Hypothetical large gains on unit triangle boundary: \$0,\$1M,\$5M. Probabilities: $S_1(0, 1, 0)$; $R_1(.01, .89, .1)$; $S_2(.89, .11, 0)$; $R_2(.9, 0, .1)$; $S_3(0, .11, .89)$; $R_3(.01, 0, .99)$; $S_4(.79, .11, .1)$; $R_4(.8, 0, .2)$; $S_5(.01, .89, .1)$; $R_5(.02, .78, .2)$

Criteria	EU	Fan-out	Fan-in	MF	RDcave	RDvex	PT
				Results			
% Explained	.220	.538	.258	.936	.667	.376	.796
% Allowed	.063	.250	.250	.656	.406	.406	.563
# patterns	2	8	8	21	13	13	18
% Max	.333	.747	.747	.968	.887	.887	.952
Z-stat	5.906	8.666	1.230	8.489	8.093	-2.395	7.521
Selten	.158	.288	.008	.279	.260	-.030	.233
Error	.299	.198	.299	.061	.147	.280	.126
χ^2-stat	189.90	97.40	189.90	15.36	53.91	186.11	48.93
Avg. LOO	.949	1.593	1.074	2.535	1.951	1.364	2.174
NEU	.921	.865	.921	.788	.826	.916	.820

Tables GEU

Table: Table VIII - Sopher and Gigliotti - Common consequence Hypothetical large gains on unit triangle boundary: \$0,\$1M,\$5M. Probabilities: $S_1(0, 1, 0)$; $R_1(.01, .89, .1)$; $S_2(.89, .11, 0)$; $R_2(.9, 0, .1)$; $S_3(0, .11, .89)$; $R_3(.01, 0, .99)$; $S_4(.79, .11, .1)$; $R_4(.8, 0, .2)$; $S_5(.01, .89, .1)$; $R_5(.02, .78, .2)$

Criteria	EU	Fan-out	Fan-in	MF	RDcave	RDvex	PT
				Ranking			
% Explained	7	4	6	1	3	5	2
Expl/Allowed	1	2	6	4	3	7	5
Expl/Max	5	4	7	1	3	6	2
Z-stat	5	1	6	2	3	7	4
Selten	5	1	6	2	3	7	4
Error	7	5	7	2	4	6	3
PO vs EU - HC	5	3	6	2	1	7	4
Avg. LOO	1	4	2	7	5	3	6
Dist. true	6	4	6	1	3	5	2
Dist. rnd	6	4	6	1	3	5	2

Table: Table IX. Sopher and Gigliotti - Common consequence Hypothetical large gains on unit triangle interior: \$0,\$1M,\$5M. Probabilities: $S_1(.01, .98, .01)$; $R_1(.02, .87, .11)$; $S_2(.80, .19, .01)$; $R_2(.81, .08, .11)$; $S_3(.01, .19, .80)$; $R_3(.02, .08, .90)$; $S_4(.70, .19, .11)$; $R_4(.71, .08, .21)$; $S_5(.02, .87, .11)$; $R_5(.03, .76, .21)$

Criteria	EU	Fan-out	Fan-in	MF	RDcave	RDvex	PT
				Results			
% Explained	.424	.484	.538	.853	.734	.630	.669
% Allowed	.063	.250	.250	.656	.406	.406	.219
# patterns	2	8	8	21	13	13	7
% Max	.446	.745	.745	.973	.870	.870	.712
Z-stat	10.205	7.491	8.628	6.719	9.465	7.292	12.263
Selten	.361	.234	.288	.197	.327	.224	.450
Error	.186	.184	.150	.104	.111	.173	.119
χ^2-stat	102.60	102.34	81.46	43.79	49.53	95.89	52.56
Avg. LOO	1.041	1.159	1.460	2.129	1.787	1.351	1.683
NEU	.818	.818	.801	.769	.766	.810	.771

Tables GEU

Table: Table IX. Sopher and Gigliotti - Common consequence Hypothetical large gains on unit triangle interior: \$0,\$1M,\$5M. Probabilities: $S_1(.01, .98, .01)$; $R_1(.02, .87, .11)$; $S_2(.80, .19, .01)$; $R_2(.81, .08, .11)$; $S_3(.01, .19, .80)$; $R_3(.02, .08, .90)$; $S_4(.70, .19, .11)$; $R_4(.71, .08, .21)$; $S_5(.02, .87, .11)$; $R_5(.03, .76, .21)$

Criteria	EU	Fan-out	Fan-in	MF	RDcave	RDvex	PT
	Ranking						
% Explained	7	6	5	1	2	4	3
Expl/Allowed	1	4	3	7	5	6	2
Expl/Max	1	7	6	3	4	5	2
Z-stat	2	5	4	7	3	6	1
Selten	2	5	4	7	3	6	1
Error	7	6	4	1	2	5	3
PO vs EU - HC	2	5	4	6	3	7	1
Avg. LOO	1	2	4	7	6	3	5
Dist. true	7	6	4	1	2	5	3
Dist. rnd	6	7	4	2	1	5	3

Back