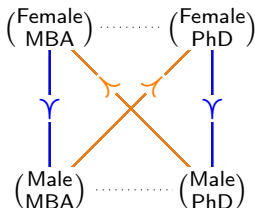


Implicit Preferences Inferred From Choice

Tom Cunningham Jonathan de Quidt

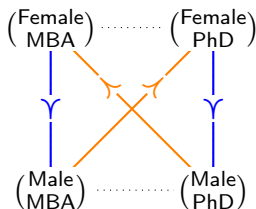
Implicit Preferences

Suppose we observe the following choices:



- ▶ Inconsistent with any U (Gender, Qualification).
- ▶ But intuitively reveals a conflict between two motives
 - ▶ Pro-Female in **vertical** choice sets
 - ▶ Pro-Male in **diagonal** choice sets

Implicit Preferences

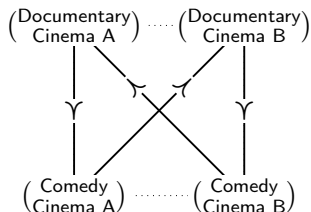


We say: an **explicit** pro-Female and an **implicit** pro-Male preference

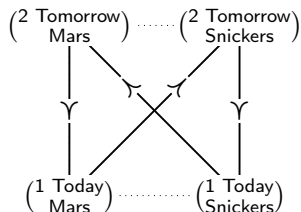
- ▶ **Diagonal** comparisons are **more opaque** than **verticals**
- ▶ **Implicit preference**: stronger in more opaque comparisons

More examples

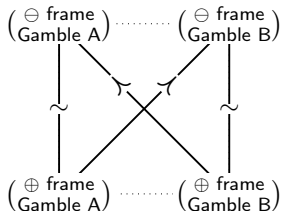
Consumption



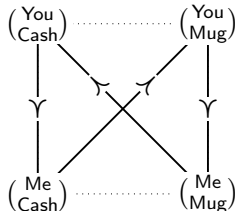
Discounting



Framing



Altruism



Many theories feature two types of motivation

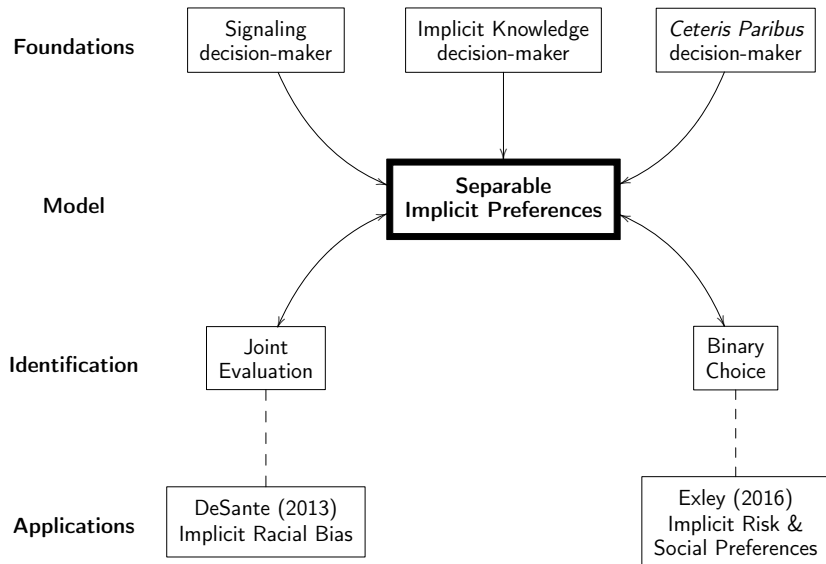
“Signaling”	reputation	vs	intrinsic	Benabou and Tirole (2003, 2006), Norton et al. (2004), Dana et al. (2007), Andreoni & Bernheim (2009), Exley (2016)
	self-image	vs	intrinsic	
	norms	vs	excuses	
“Cognitive”	conscious	vs	unconscious	Greenwald et al. (1998), Greenwald & Krieger (2006), Kahneman (2011), Rand et al. (2012)
	system 2	vs	system 1	
	reflective	vs	impulsive	
“Rule-based”	justifiable	vs	desired	Cherepanov et al. (2013), Ridout (2020)
	restricted	vs	free	
	“Explicit”	vs	“Implicit”	

Our claim: diverse underlying foundations share common behavioral implications, detectable in common ways.

This paper: A behavioral model of implicit preferences

- ▶ Something we can directly observe in decisions.
- ▶ Analogous to an elasticity, or a complementarity.
- ▶ Can be derived from diverse foundations.

Overview



Foundations (intuition)

Foundations

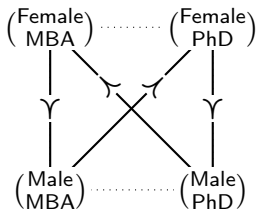
Signaling
decision-maker

Implicit Knowledge
decision-maker

Ceteris Paribus
decision-maker

Signaling decision-maker

- ▶ You care about what you choose **and** what your choice signals.
 - ▶ Observer could be someone else, or yourself (self-signaling).
 - ▶ Formally: preferences over bundles **and** over observer's beliefs.
 - ▶ Observer has independent Gaussian priors over attributes' values.
-



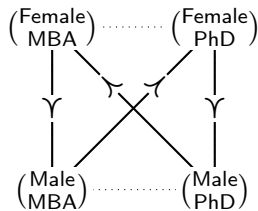
You truly prefer men, but you want the observer to think you prefer women.

Vertical choices transparently reveal your motives.

Diagonal choices are less transparent.

Implicit Knowledge decision-maker

- ▶ Your decisions are influenced by **unconscious associations** that are sometimes useful and sometimes misleading.
- ▶ Formally: 2-systems model of cognition, both stages have private information (Cunningham, 2014)).
- ▶ When considering multiple bundles simultaneously, you learn something about your associations.



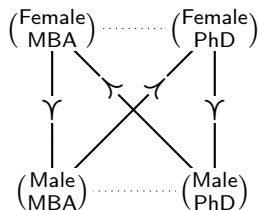
Male candidates give you a “good feeling,” but you don’t know why.

Vertical choices reveal that it’s because of their Gender, and you can override it.

Diagonal choices are less transparent. You partly attribute your good feeling to the Qualification.

Ceteris Paribus decision-maker

- ▶ Decision-maker is constrained by a rule to choose in favor of some attribute when there are no other differences.
- ▶ E.g. “Cannot hire a man over an equally-qualified woman.”

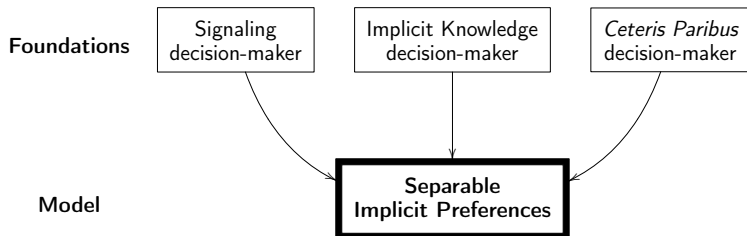


You prefer men.

Vertical choices invoke the rule.

Diagonal choices do not.

Model



- We consider **bundles** of **binary attributes**

$$x \in \mathcal{X} = \{-1, 1\}^n$$

$n = 2$



$n = 3$

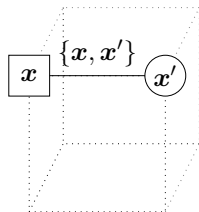


etc

- e.g. {Male,Female}, {Safe,Risky}, {Sooner,Later}, {Green,Blue}

- Utility of bundle x is influenced by a **comparator** x'

$$u(x|\{x, x'\})$$



- Consider two types of decision situation.
 1. **Joint evaluation.** decision-maker reports the utility of x and x' :

$$u(x|\{x, x'\})$$

$$u(x'|\{x', x\})$$

2. **Binary choice.** decision-maker chooses between x and x' :

$$c(\{x, x'\})$$

- Comparison set $\{x, x'\}$ enters preferences through **opacity**

$$\theta_i(\underbrace{|\mathbf{x}' - \mathbf{x}|}_{\substack{\text{Difference} \\ \text{between} \\ \text{bundles}}}) \in [0, 1]$$

Implicit Preference (informal)

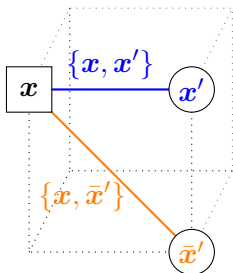
An **implicit preference** for attribute i is expressed **more strongly** when the comparison set is **more opaque** about i

Assumption 1: Utility function

$$u(\mathbf{x}|\{\mathbf{x}, \mathbf{x}'\}) = \underbrace{v(\mathbf{x})}_{\text{explicit preferences}} + \overbrace{\sum_{i=1}^n x_i \cdot \underbrace{\kappa_i}_{\substack{\text{implicit} \\ \text{preference} \\ \text{for } i}} \cdot \underbrace{\theta_i(|\mathbf{x}' - \mathbf{x}|)}_{\substack{\text{opacity of} \\ \text{comparison} \\ \text{for } i}}}_{\text{implicit preferences}},$$

Assumption 2: Opacity

$\{x, x'\}$ vs $\{x, \bar{x}'\}$:



1. **Diagonal** more opaque about horizontal attribute

$$\theta_1(|\bar{x}' - x|) \geq \theta_1(|x' - x|)$$

2. **Diagonal** less opaque about vertical attribute

$$\theta_2(|\bar{x}' - x|) \leq \theta_2(|x' - x|)$$

3. **Diagonal** less opaque about depth attribute

$$\theta_3(|\bar{x}' - x|) \leq \theta_3(|x' - x|)$$

- For two comparisons, can rank opacities whenever their differences are ordered $|\{x, x'\}| \geq |\{x, \bar{x}'\}|$ (i.e., we get a semiorder).

Identification of implicit preferences

- ▶ Goal: identify the **sign** of implicit preferences

$$\kappa_i \begin{matrix} \geq \\ \leq \end{matrix} 0$$

- ▶ Two types of data:
 - ▶ Joint evaluation data (“**Scissors**”)
 - ▶ Binary choice data (**Intransitive cycles**)

Identification of implicit preferences

Given a dataset, two possibilities:

1. Falsify the model

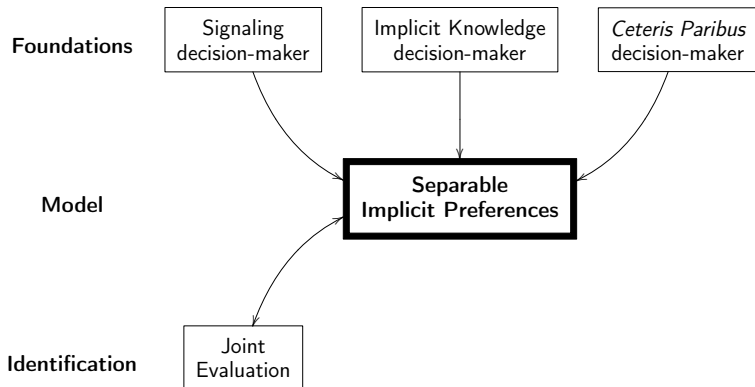
Theorem 1 (evaluation)

Theorem 2 (choice)

2. Identify...

- ▶ **no** implicit preferences
- ▶ **a disjunction**
(e.g. Implicit pro-Male **and/or** pro-MBA)
- ▶ one or more **unique** implicit preferences
(e.g. Implicit pro-Male)

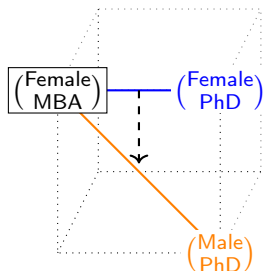
Identification: Evaluation data



Scissor

We call the following pattern a **Scissor**:

$$u \left(\left(\begin{smallmatrix} \text{Female} \\ \text{MBA} \end{smallmatrix} \right) \middle| \left\{ \left(\begin{smallmatrix} \text{Female} \\ \text{MBA} \end{smallmatrix} \right), \left(\begin{smallmatrix} \text{Female} \\ \text{PhD} \end{smallmatrix} \right) \right\} \right) \neq u \left(\left(\begin{smallmatrix} \text{Female} \\ \text{MBA} \end{smallmatrix} \right) \middle| \left\{ \left(\begin{smallmatrix} \text{Female} \\ \text{MBA} \end{smallmatrix} \right), \left(\begin{smallmatrix} \text{Male} \\ \text{PhD} \end{smallmatrix} \right) \right\} \right)$$



- ▶ **Diagonal** is less opaque about Gender
- ▶ **Diagonal** is **more** opaque about Qualification
- ▶ Suppose:

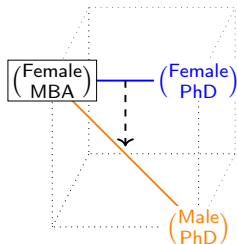
$$u(. | \text{horizontal}) < u(. | \text{diagonal})$$

Conclusion

Implicit pro-Male **and/or** Implicit pro-MBA

Double Scissor

Combine Scissors to obtain unique identification.

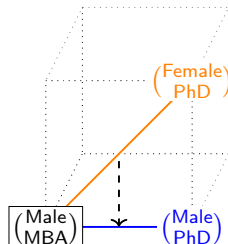


$$u(. | \text{horizontal}) < u(. | \text{diagonal})$$

Implicit pro-Male

and/or

Implicit pro-MBA



$$u(. | \text{horizontal}) > u(. | \text{diagonal})$$

Implicit pro-Male

and/or

Implicit pro-PhD

Conclusion

Gender: Implicit pro-Male.

Qualification: Unknown.

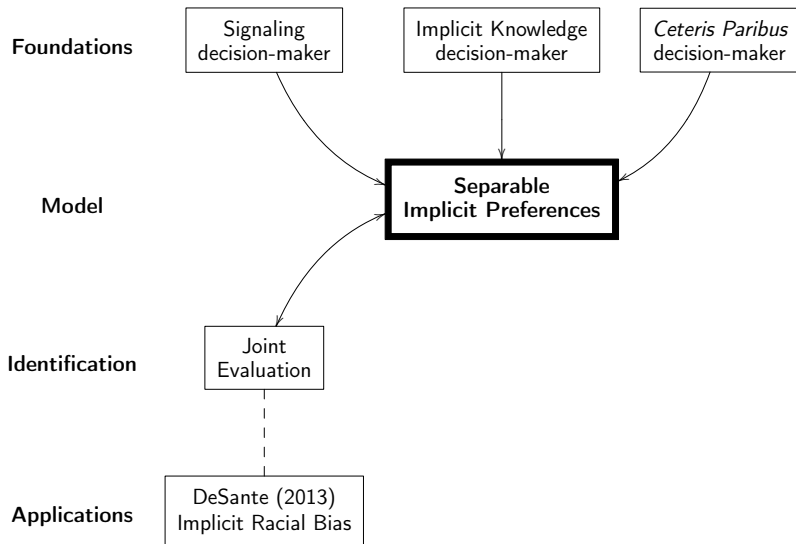
- ▶ A generic Scissor is a pair of evaluations:

$$u(x|\{x, x'\}) \text{ and } u(x|\{x, \bar{x}'\})$$

$$\text{s.t. } |\{x, x'\}| \leq |\{x, \bar{x}'\}|$$

- ▶ **Theorem 1** tells us
 - ▶ When a collection of Scissors falsifies the model
 - ▶ If not, what are the implied implicit preferences.

Application 1



DeSante (AJPS, 2013)

Working Twice as Hard to Get Half as Far: Race, Work Ethic, and America's Deserving Poor

- ▶ A survey experiment on a US representative sample.
- ▶ *Test whether “hard work” is rewarded in a color-blind manner*
- ▶ Participants divide a (hypothetical) budget for state aid between
 1. Applicant 1
 2. Applicant 2
 3. “offset the state’s budget deficit”
- ▶ Applicants differ in:
 - Race $\in \{\text{Black}, \text{White}\}$
 - Work ethic $\in \{\text{Hidden}\}$ or $\{\text{Excellent}, \text{Poor}\}$
 - Other (counterbalanced)
- ▶ We examine how allocations change when comparator’s Race changes

WORK FIRST ASSISTANCE APPLICATION

Applicant Name: **Latoya** [Redacted] Date of Application: [Redacted]
 Address: [Redacted] Telephone: [Redacted]
 County: [Redacted]
 Case No.: [Redacted] District No.: [Redacted]

HOUSEHOLD: List all household members for whom Assistance is being requested:

(Non-applicant household members are not required to provide a social security number, immigrant or citizenship status)

Name	Date of Birth	Sex	Social Security No.	Citizen/ eligible immigrant	Relationship
[Redacted]	08/16/1998	M	[Redacted]	Y	Son
[Redacted]	04/14/2001	F	[Redacted]	Y	Daughter

Does the household include a child who meets the Work First age rule? ☒ Yes ☐ No

Is the child living with an adult who meets the Work First kinship rule? ☒ Yes ☐ No

Has anyone listed on the EA Application ever received EA? ☐ Yes When: [Redacted] ☒ No

Does anyone live in the home that is not listed on the EA Application? ☐ Yes ☒ No
 If yes, is the individual(s) a roomer/boarder? ☐ Yes ☐ No

Total assessed monthly need: \$ 900.00

APPLICANT 1 Worker Quality Assessment (circle one):
Poor Average Excellent

Applicant Statement: I understand that it is against the law for me to make false statements and that I am subject to prosecution if I do. I certify that the information I have provided is a true and complete statement of facts according to my best knowledge and belief. I certify, under penalty of perjury, that all persons for whom I am applying are U.S. citizens or qualified immigrants. I declare under penalty of perjury (and being subject to prosecution under 28 U. S. C. § 1746) that the foregoing is true and correct. I give the agency permission to verify any information necessary to determine my eligibility for Emergency Assistance.

Chris D. DeLoe
 Witness's Signature

Applicant's/Representative's Signature

Date

WORK FIRST ASSISTANCE APPLICATION

Applicant Name: **Keisha** [Redacted] Date of Application: [Redacted]
 Address: [Redacted] Telephone: [Redacted]
 County: [Redacted]
 Case No.: [Redacted] District No.: [Redacted]

HOUSEHOLD: List all household members for whom Assistance is being requested:

(Non-applicant household members are not required to provide a social security number, immigrant or citizenship status)

Name	Date of Birth	Sex	Social Security No.	Citizen/ eligible immigrant	Relationship
[Redacted]	05/07/2005	M	[Redacted]	Y	Son
[Redacted]	03/20/2007	F	[Redacted]	Y	Daughter

Does the household include a child who meets the Work First age rule? ☒ Yes ☐ No

Is the child living with an adult who meets the Work First kinship rule? ☒ Yes ☐ No

Has anyone listed on the EA Application ever received EA? ☐ Yes When: [Redacted] ☒ No

Does anyone live in the home that is not listed on the EA Application? ☐ Yes ☒ No
 If yes, is the individual(s) a roomer/boarder? ☐ Yes ☐ No

Total assessed monthly need: \$ 900.00

APPLICANT 2 Worker Quality Assessment (circle one):
 Poor Average **Excellent**

Applicant Statement: I understand that it is against the law for me to make false statements and that I am subject to prosecution if I do. I certify that the information I have provided is a true and complete statement of facts according to my best knowledge and belief. I certify, under penalty of perjury, that all persons for whom I am applying are U.S. citizens or qualified immigrants. I declare under penalty of perjury (and being subject to prosecution under 28 U. S. C. § 1746) that the foregoing is true and correct. I give the agency permission to verify any information necessary to determine my eligibility for Emergency Assistance.

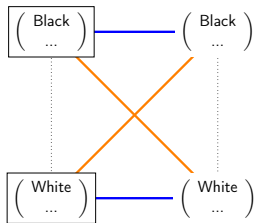
Chris D. DeLoe
 Witness's Signature

Applicant's/Representative's Signature

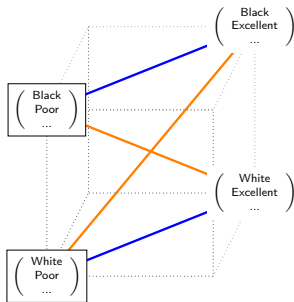
Date

Analysis: $3 \times$ Double Scissors in Race:

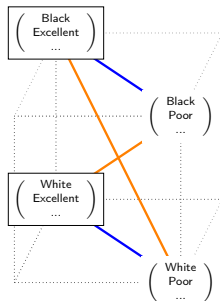
1. *Hidden* work ethic



2. Poor work ethic



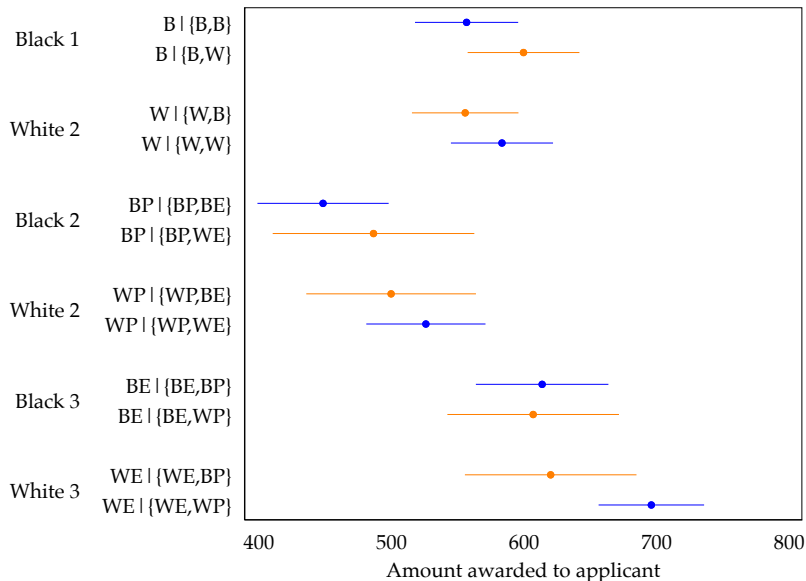
3. Excellent work ethic



Diagonals less opaque about race. If implicit pro-White:

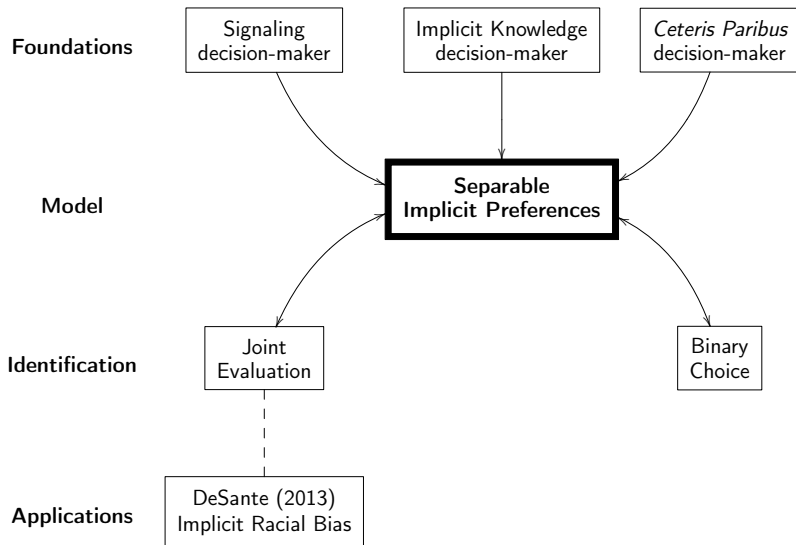
- ▶ $u(\text{Black})$ higher on **diagonals**
- ▶ $u(\text{White})$ lower on **diagonals**

DeSante (2013)



Joint test: $u(x|\{x, x'\}) = u(x|\{x, \bar{x}'\}) \forall x, p = 0.08$.

Identification: Choice data



What choices should we look at?

- ▶ Unit of analysis in choice data is an **intransitive cycle**.
- ▶ Focus on choices where we expect DM is close to indifferent.

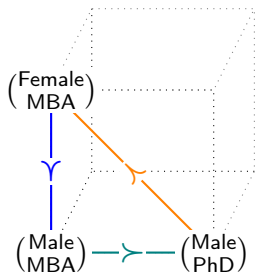
Statistical motivation:

- ▶ Detecting intransitivities is hard
(Tversky, 1969; Regenwetter et al. 2011, Muller-Trede et al. 2015)
- ▶ Strong explicit preferences (v) conceal implicit preferences (κ).

Theoretical motivation:

- ▶ Signaling model applied to choice requires that observer **expects DM to be indifferent** ▶ Why?

Right triangle



- ▶ **Diagonal** more opaque than **vertical** about Gender
- ▶ **Diagonal** more opaque than **horizontal** about Qualification

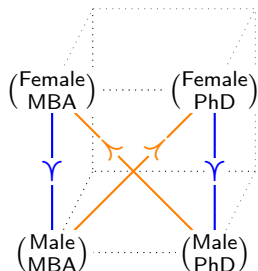
Conclusion

Implicit pro-Male **and/or** Implicit pro-PhD

As with evaluation, observing multiple triangles can reveal a unique implicit preference.

Figure-8

We call the following pattern a **Figure-8**:

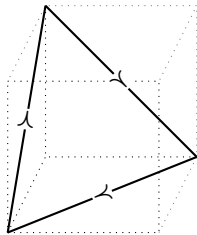


- ▶ **Diagonal** more opaque than **vertical** about Gender
- ▶ Any implicit preferences over Qualification **do not affect choice** on the **verticals**, and **push in opposing directions** on the **diagonals**.

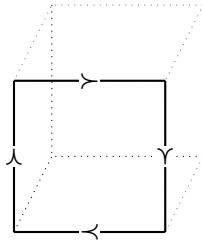
Conclusion

Gender:	Implicit pro-Male.
Qualification:	Unknown.

More cycles

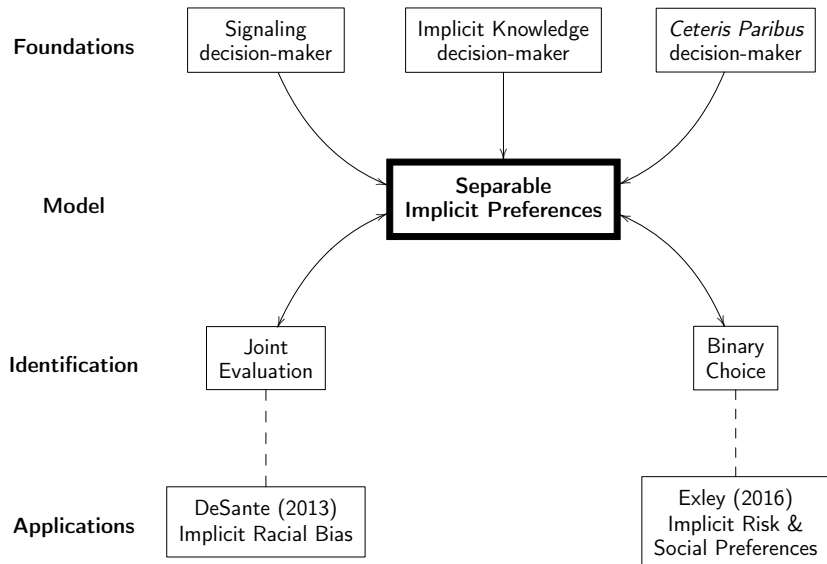


Compatible with any combination of implicit preferences



Cannot be rationalized by any implicit preference

Application 2

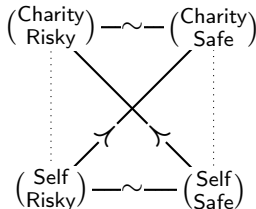


Exley (ReStud, 2016)

Excusing Selfishness in Charitable Giving: The Role of Risk

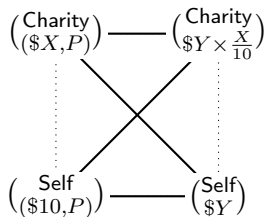
- ▶ Exley studies *the use of risk as an excuse not to give*

- ▶ Basic idea:



Details:

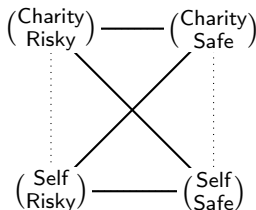
- ▶ First, elicit the $\$X$ for charity just-preferred to $\$10$ for self.
- ▶ Participant then faces choice lists like the following:



- ▶ 7 different lottery pairs, 21 Y increments.

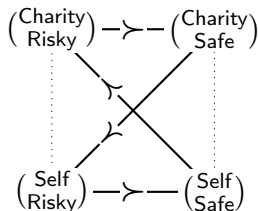
$$\left(\begin{array}{c} \text{Charity} \\ (\$X, P) \end{array} \right), \left(\begin{array}{c} \text{Self} \\ (\$10, P) \end{array} \right) \quad P \in \{.05, .1, .25, .5, .75, .9, .95\}$$

- ▶ We need to transform the data to apply our framework.
- ▶ Exley's design and assumptions naturally imply a binary representation:

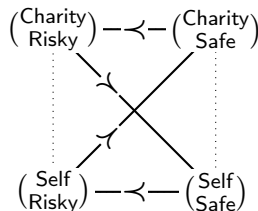


Risk preferences

Without any further assumptions we can identify *Implicit Risk Preferences*



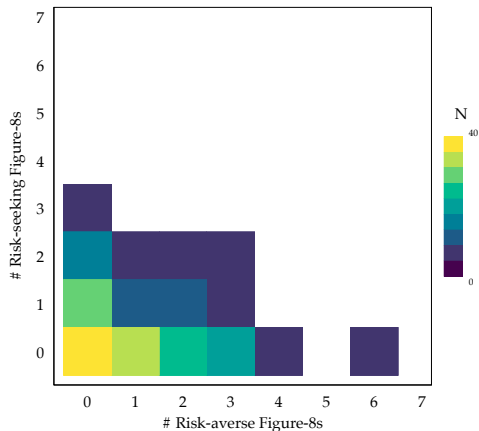
Implicit Risk-averse



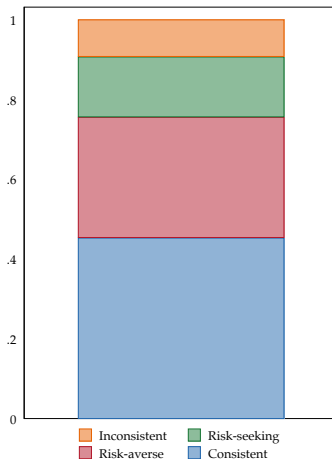
Implicit Risk-seeking

Each participant can exhibit **up to 7** Figure-8s (one per pair of lotteries)

Risk

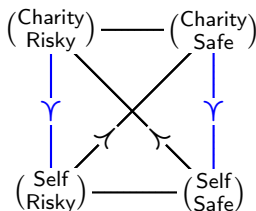


Implicit risk preferences



- ▶ 30% reveal consistent implicit risk-aversion
- ▶ 15% reveal consistent implicit risk-seeking

- Choice data alone do not reveal implicit pro-Self or pro-Charity preferences



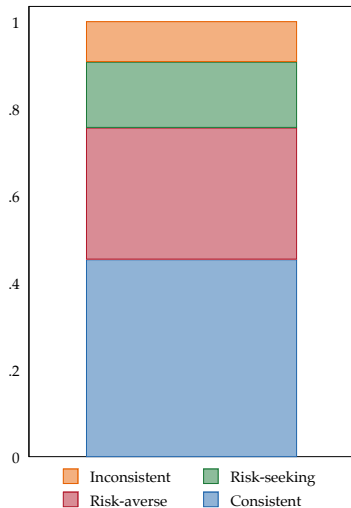
- Exley assumes *linear preferences* over self/charity dollars

$$\begin{pmatrix} Charity \\ \$20 \end{pmatrix} \approx \begin{pmatrix} Self \\ \$10 \end{pmatrix} \Rightarrow \begin{pmatrix} Charity \\ \$10 \end{pmatrix} \approx \begin{pmatrix} Self \\ \$5 \end{pmatrix}$$

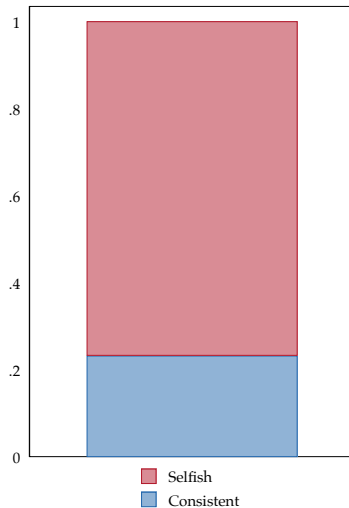
- That allows us to impute the vertical choices and identify a pro-Self Figure-8.

Overall classification

Implicit risk preferences



Implicit selfishness

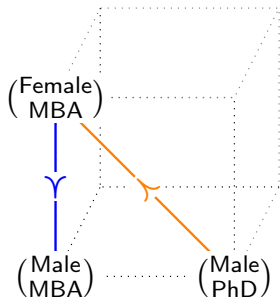


Conclusion

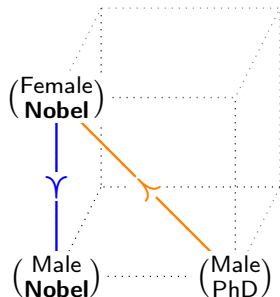
- ▶ Many theories feature a tension between two kinds of motive: explicit/implicit; signaling/intrinsic; conscious/unconscious; constrained/free
- ▶ We formalize a general approach to identifying them in decisions.
- ▶ Broadly applicable:
 - ▶ Implicit discrimination (without an IAT) (e.g. Barron et al., 2020)
 - ▶ Implicit present-bias (e.g. Drucker & Kaufmann, 2020)
 - ▶ Implicit patience (e.g. Cubitt et al., 2018)
 - ▶ Implicit risk and social preferences
 - ▶ Implicit consumption preferences
 - ▶ Framing effects as an implicit preference

Extra material

Signaling without Indifference



Choosing the man is less revealing in **diagonal** than **vertical**



Choosing the man is much **more** revealing in **diagonal** than **vertical**!

► Back