The Model You Know: Generalizability and Predictive Power of Models of Choice Under Uncertainty

B. Douglas Bernheim Christine Exley Jeffrey Naecker Charles Sprenger

1/24/2019

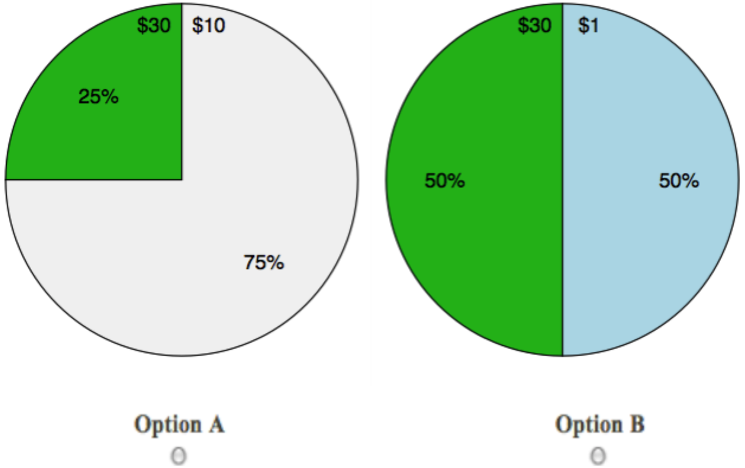
Motivation

- Two important features of models:
  - Interpretability/parsimony
  - Generalizability/predictive power
- Risk preference models
  - Certainly interpretable and parsimonious
  - Known to fit well in sample but may be issues with out-of-sample prediction (eg, Camerer 1992)

Our Contribution

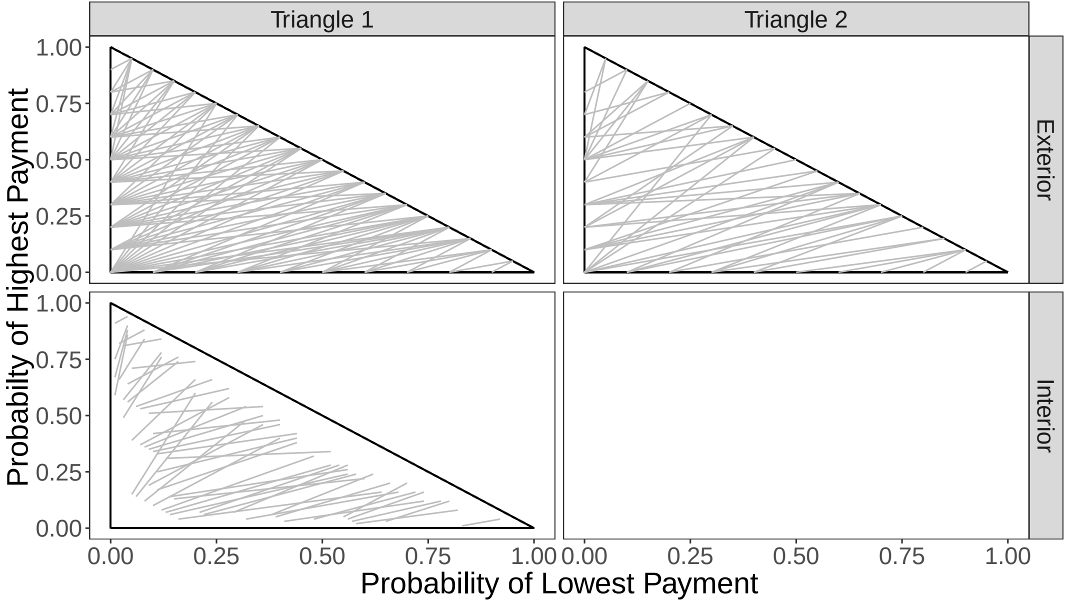
- Test out-of-sample performance of utility models in two settings:
  - Changing stakes
  - Increasing complexity of gambles
- Provide alternative data and methods to
  1. Make more accurate predictions out-of-sample
  2. Get better estimates of treatment effects

Typical Choice Problem



Choice Environment

- Choose between two lotteries, A and B
- Represent in two Machina triangles:
  - Triangle 1: outcomes \$1, \$10, \$30
    - exterior: up to two outcomes possible in any lotter
    - interior: up to three outcomes possible in any lottery
  - Triangle 2: outcomes \$0, \$5, \$20
    - exterior only
- 199 lottery pairs total
- Participants see random set of 80 pairs, shown sequentially
- Lottery A along legs of triangle, while lottery B is along hypotenuse



Triangle 1
Triangle 2
Exterior
Interior
Probabilty of Highest Payment
Probability of Lowest Payment
0.00
0.25
0.50
0.75
1.00

Treatments

Treatment	**Question(s)**
Real	Which option do you prefer? [1 = option A, 0 = option B]
Hypothetical	Hypothetically, which option do you prefer?
Hypothetical likelihood	Hypothetically, how likely would you be to choose Option A over Option B? [1-5]
Vicarious hypothetical	How likely would a typical Stanford undergraduate student be to choose Option A over Option B?
Subjective	Choosing which option would indicate a greater willingness to take risks? Choosing which option would indicate better judgment? Which option is more difficult to evaluate?

Utility Models

1. Expected utility with constant relative risk aversion:

$$U(p, x) = \sum_i p_i x_i^\alpha$$

2. Cumulative prospect theory from Kahneman and Tversky (1992):

$$\begin{aligned} U(p, x; g) = & \left(\pi(p_3, g) - \pi(0, g)\right)x_3^\alpha \\ & + \left(\pi(p_2 + p_3, g) - \pi(p_3, g)\right)x_2^\alpha \\ & + \left(\pi(p_1, g) - \pi(p_2 + p_3, g)\right)x_1^\alpha \end{aligned}$$

where

$$\pi(p, g) = \frac{p^g}{(p^g + (1-p)^g)^{(1/g)}}$$

Errors

Luce decision error formulation:

$$P(\text{choose A}) = \frac{U(A)^{\frac{1}{\mu}}}{U(A)^{\frac{1}{\mu}} + U(B)^{\frac{1}{\mu}}}$$

- $\mu \to 0$: no mistakes (ie all probabilities $= 0$ or 1)
- $\mu \to \infty$: flip a coin (ie all probabilities $= \frac{1}{2}$)

Parameter estimates

Non-Choice Data Methods: Univariate Models

- Regress real choice frequency on hypothetical in triangle 1 exterior at choice problem level:

$$\mathit{real}_{1i} = \alpha + \beta\,\mathit{hyp}_{1i} + \varepsilon$$

- Then use estimated coefficients to predict real in triangle 2 exterior from hypothetical in triangle 2 exterior:

$$\widehat{\mathit{real}_{2i}} = \hat{\alpha} + \hat{\beta}\,\mathit{hyp}_{2i}$$

- Repeat with vicarious hypothetical likelihood mean as predictor
- Same procedure to predict to triangle 1 interior

Non-Choice Data Methods: LASSO

- Large number of predictors:
  - Means for all hypothetical and subjective questions
  - For all Likert-scale questions, fraction of responses $= 1, \leq 2, \leq 3$, etc
- Use regularized regression (LASSO):

$$min_\beta \sum_i (y_i - \beta x_i)^2 + \lambda ||\beta||_2$$

  - Regularization parameter λ set using cross-validation
- Estimation and prediction as with univariate OLS models

Prediction Metrics

- Bias (average prediction error):

$$\frac{1}{N}\sum_i |\widehat{real}_i - real_i|$$

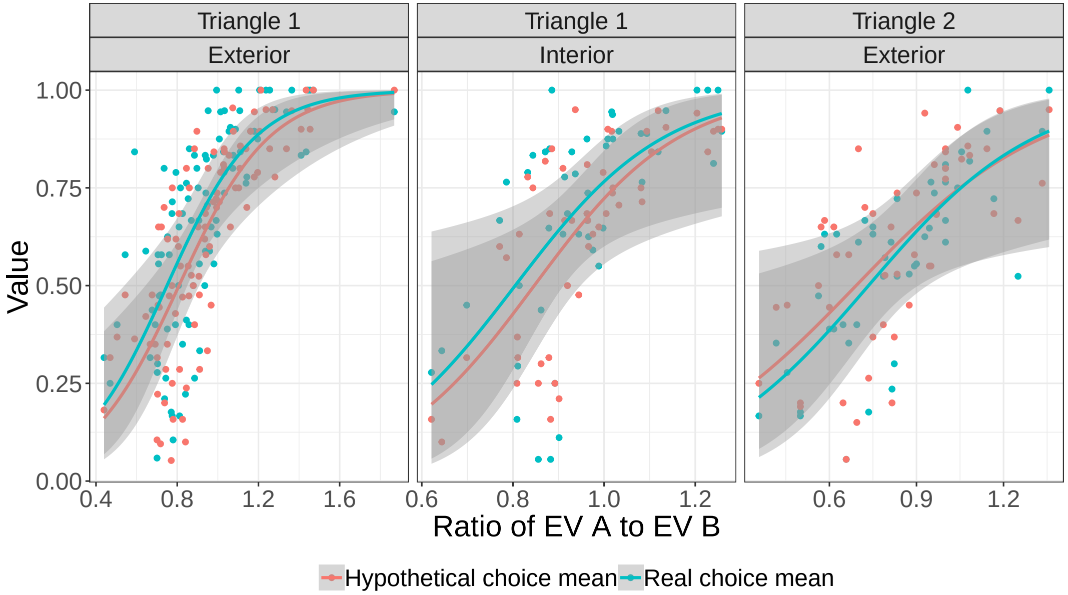
- mean-squared prediction error (MSPE):

$$\frac{1}{N}\sum_i |\widehat{real}_i - real_i|^2$$

- Calibration score is $|\beta - 1|$, with estimated β in the regression equation:

$$real_i = \alpha + \beta\widehat{real}_i + \varepsilon_i$$

Choice Probabilities



Prediction Statistics: Pooled

Label	Bias	Mean Squared Err	Calibration Score
Expected utility: rep agent	0.048	0.035	0.187
Prospect theory: rep agent	0.045	0.033	0.163
Expected utility: hetero agents	-0.024	0.023	0.085
Prospect theory: hetero agents	-0.017	0.024	0.014
Non-choice: all vars	0.012	0.013	0.267
Non-choice: all hyp vars	0.014	0.014	0.319
Non-choice: hyp mean only	0.021	0.016	0.006
Non-choice: vicarious mean only	0.011	0.019	0.016

In-Sample Performance

Label	Bias	Mean Squared Err	Calibration Score
Expected utility: rep agent	0.009	0.014	0.046
Prospect theory: rep agent	0.009	0.013	0.050
Expected utility: hetero agents	-0.054	0.014	0.008
Prospect theory: hetero agents	-0.035	0.016	0.063
Non-choice: all vars	0.000	0.013	0.264
Non-choice: all hyp vars	0.000	0.013	0.336
Non-choice: hyp mean only	0.000	0.015	0.000
Non-choice: vicarious mean only	0.000	0.019	0.000

▸ Visualizations

Out-of-Sample Performance: Interior

Label	Bias	Mean Squared Err	Calibration Score
Expected utility: rep agent	-0.061	0.026	0.366
Prospect theory: rep agent	-0.065	0.027	0.360
Expected utility: hetero agents	-0.103	0.034	0.237
Prospect theory: hetero agents	-0.111	0.041	0.349
Non-choice: all vars	-0.005	0.012	0.305
Non-choice: all hyp vars	-0.007	0.013	0.344
Non-choice: hyp mean only	0.005	0.015	0.060
Non-choice: vicarious mean only	-0.018	0.018	0.079

Visualizations

Out-of-Sample Performance: Triangle 2

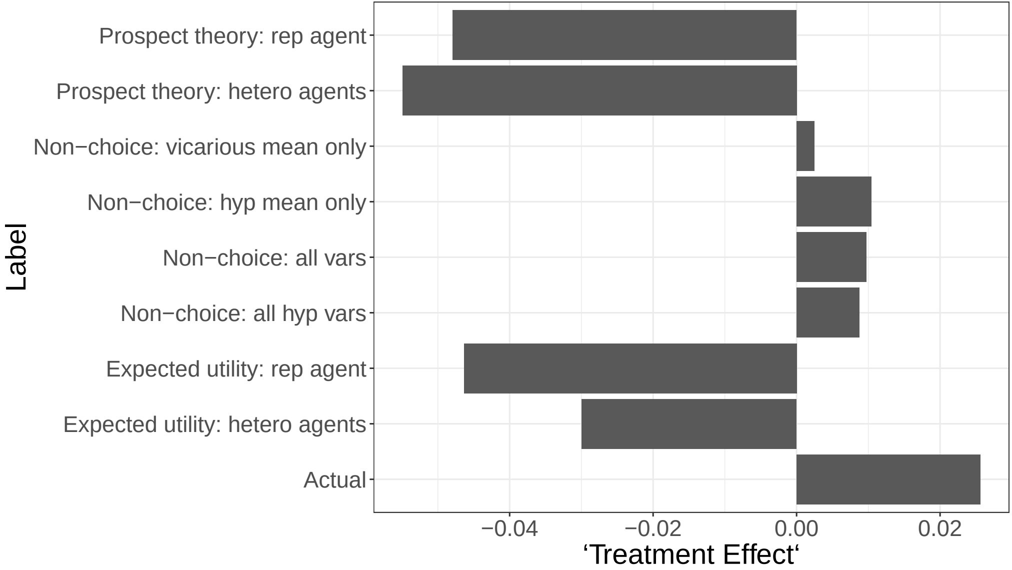
Label	Bias	Mean Squared Err	Calibration Score
Expected utility: rep agent	0.234	0.088	0.342
Prospect theory: rep agent	0.226	0.079	0.291
Expected utility: hetero agents	0.114	0.030	0.182
Prospect theory: hetero agents	0.110	0.024	0.079
Non-choice: all vars	0.050	0.014	0.184
Non-choice: all hyp vars	0.062	0.017	0.208
Non-choice: hyp mean only	0.077	0.020	0.063
Non-choice: vicarious mean only	0.063	0.019	0.050

Visualizations

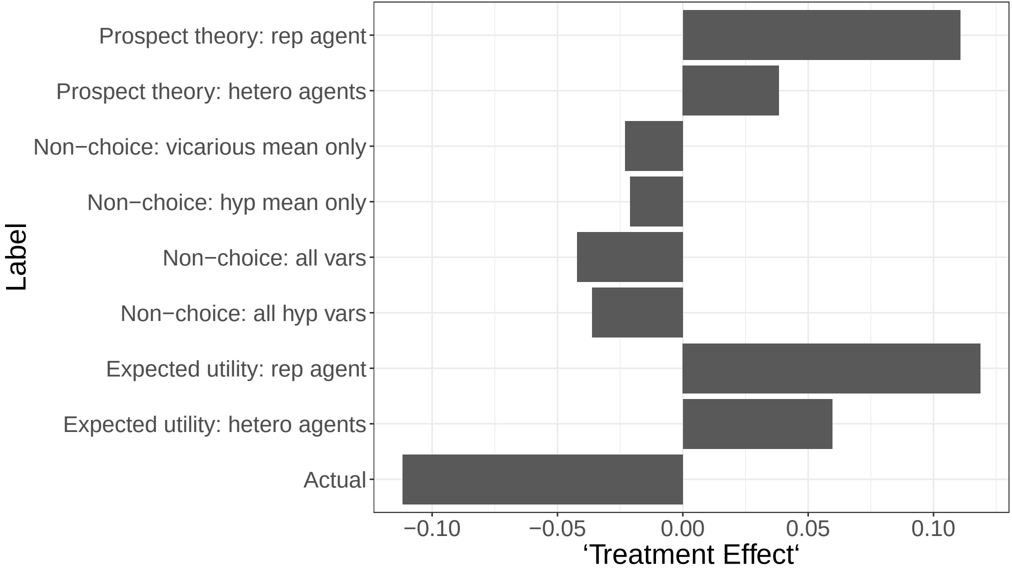
So What?

- What can we do with predictions?
- One answer: estimate treatment effects without observing treatment
- Two treatments:
    1. Increase complexity
    2. Decrease stakes

Exterior to Interior (Increase Complexity)



Triangle 1 to Triangle 2 (Decrease Stakes)

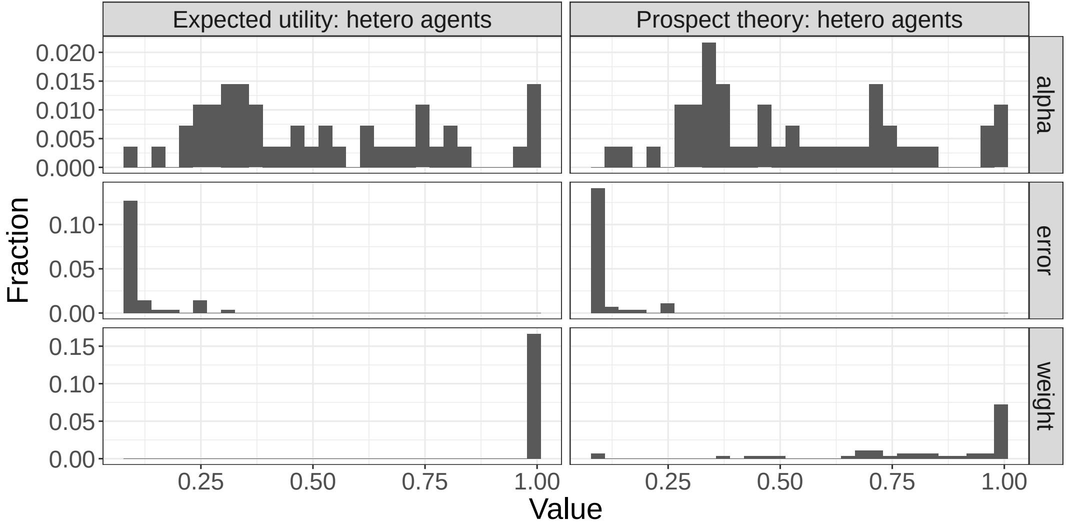


Conclusion

- Utility models may not be best option for predicting treatment effects
- Next step: Adding additional benchmark using methods from Naecker and Peysakhovich (2017)
  - Can suggest improvements to utility models

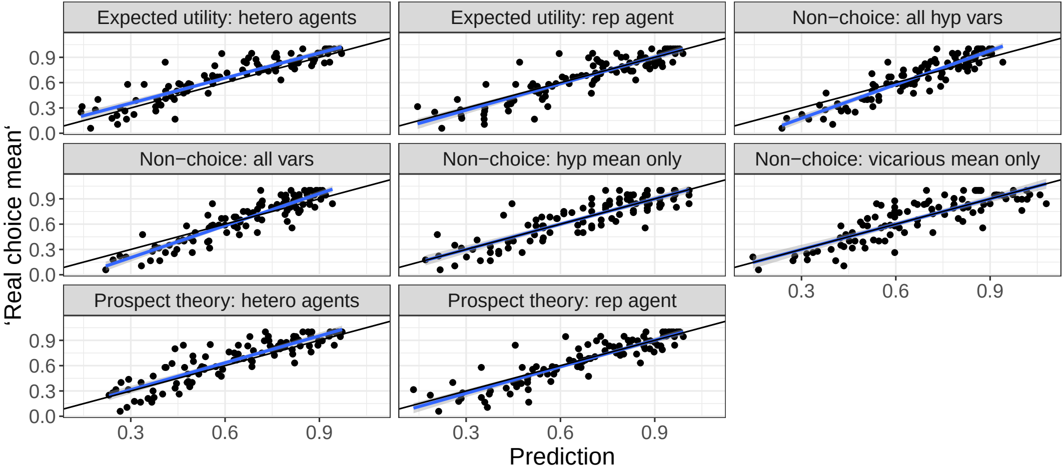
Appendix

Utility Parameter Estimates



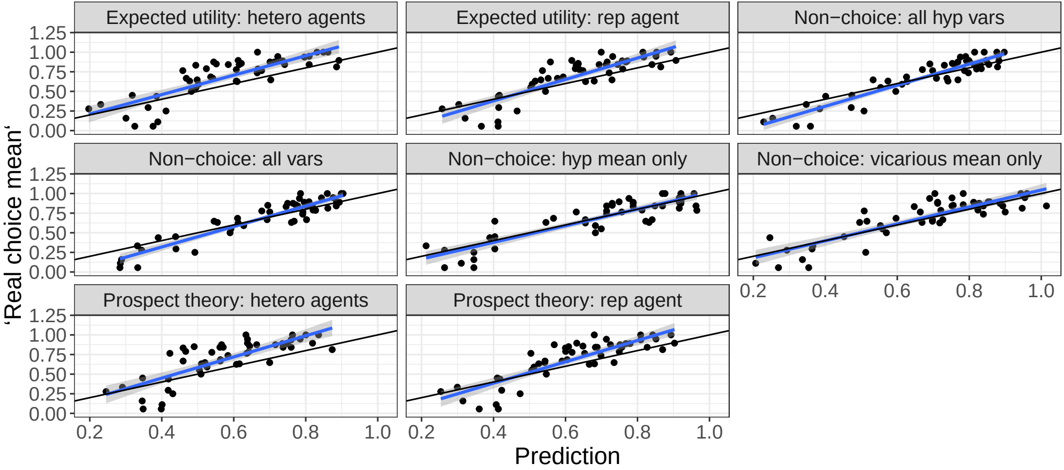
Back

In-Sample Performance



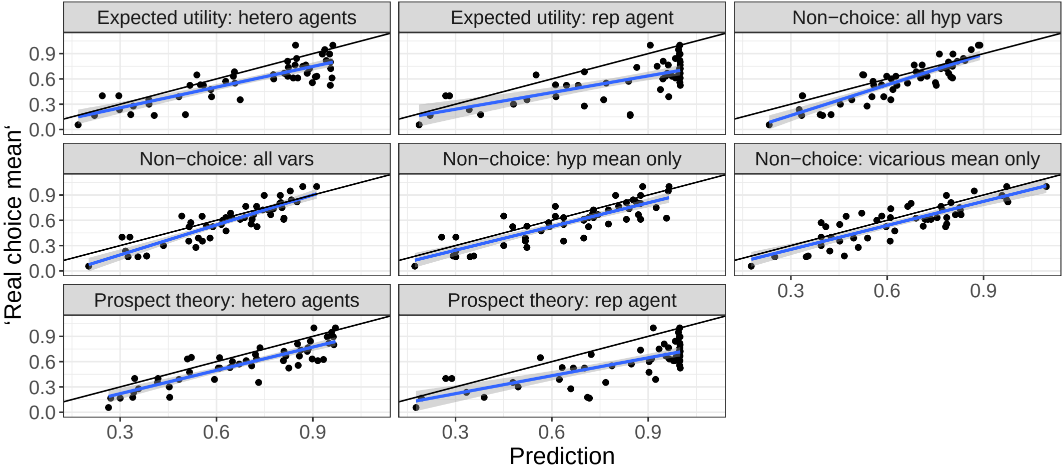
Back

Out-of-Sample Performance: Interior



Back

Out-of-Sample Performance: Triangle 2



▸ Back