FORECASTING SUPERFORECASTERS

Natalie Herrmann, Jessica Ortiz, Qizun Xie, Kristina Lerman, Fred Morstatter, Andres Abeliuk

1. Problem Statement

- Knowing future probabilities, *forecasts*, is useful for making decisions.
- Research competitions for forecasting have created a profile for what type of people are consistently adept at forecasting—superforecasters.
- Goal: improve superforecaster profile.
- Method: identify possible instances of Simpson's paradox.

2a. Superforecaster Profile

Based on previous work, superforecasters have:
- certain cognitive abilities and approaches: higher intelligence, actively open-minded thinking style
- certain mindsets: more analytical, more probabilistic
- certain work ethics: more likely to update forecasts, greater motivation to be high performing (Tetlock & Gardener, 2016; Mellers, Stone, Murray, et al., 2015)

2b. Simpson's Paradox
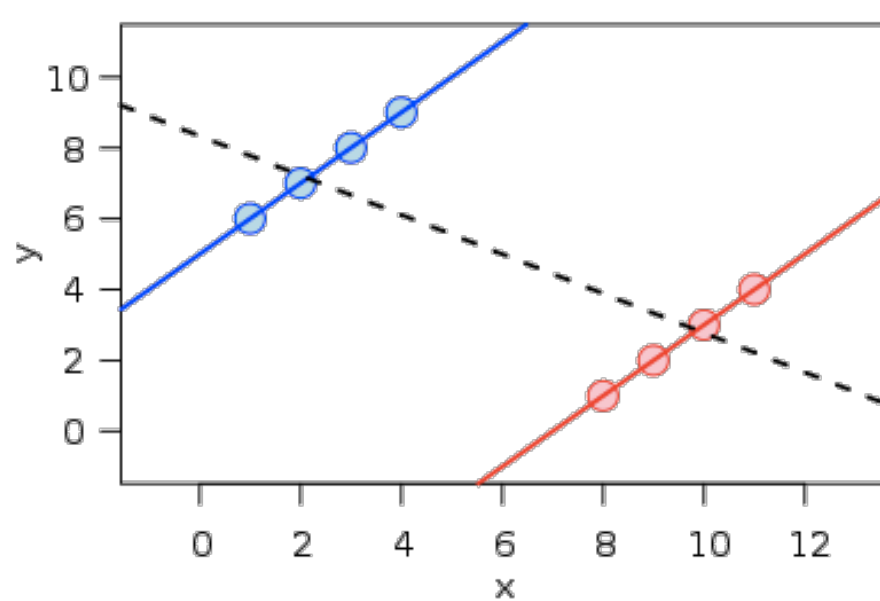
A trend for an aggregated population (e.g. dotted line) is different, perhaps opposite, of trends for disaggregated subgroups (e.g. blue line, red line).

3. Data & Data Cleaning

- Data: forecasting and psychological data from 4-year Aggregative Contingent Estimation (ACE) forecasting competition.
- Data Cleaning: Simpson's paradox program cannot run with incomplete data, so we imputed average scores into missing data.
  - Imputed = calculate average of all completed scores for each test, fill in average score into empty spaces.

4. Brier Score Calculation

$$BS = \sum_{t=1}^{N}(f_t - o_t)^2$$

- N = number of outcomes
- t = particular outcome
- f_t = probability forecasted for t
- o_t = actual outcome
  - (1 = true, 0 = false)

- Brier scores measure forecast accuracy.
  - Lower scores = more accurate.
  - Higher scores = less accurate.
- Brier score for each participant is the average of each question's Brier score.
  - Overall Brier score for year is the average of the Brier score for first and last forecast.
- Problem: participants choose questions they answer and questions vary in difficulty.
  - Solution: standardize (find z-score) Brier scores by question.

$$z = \frac{x - \mu}{\sigma}$$

$\mu =$ Mean
$\sigma =$ Standard Deviation

- Superforecaster = performed better than one standard deviation and answered more than 5 questions.
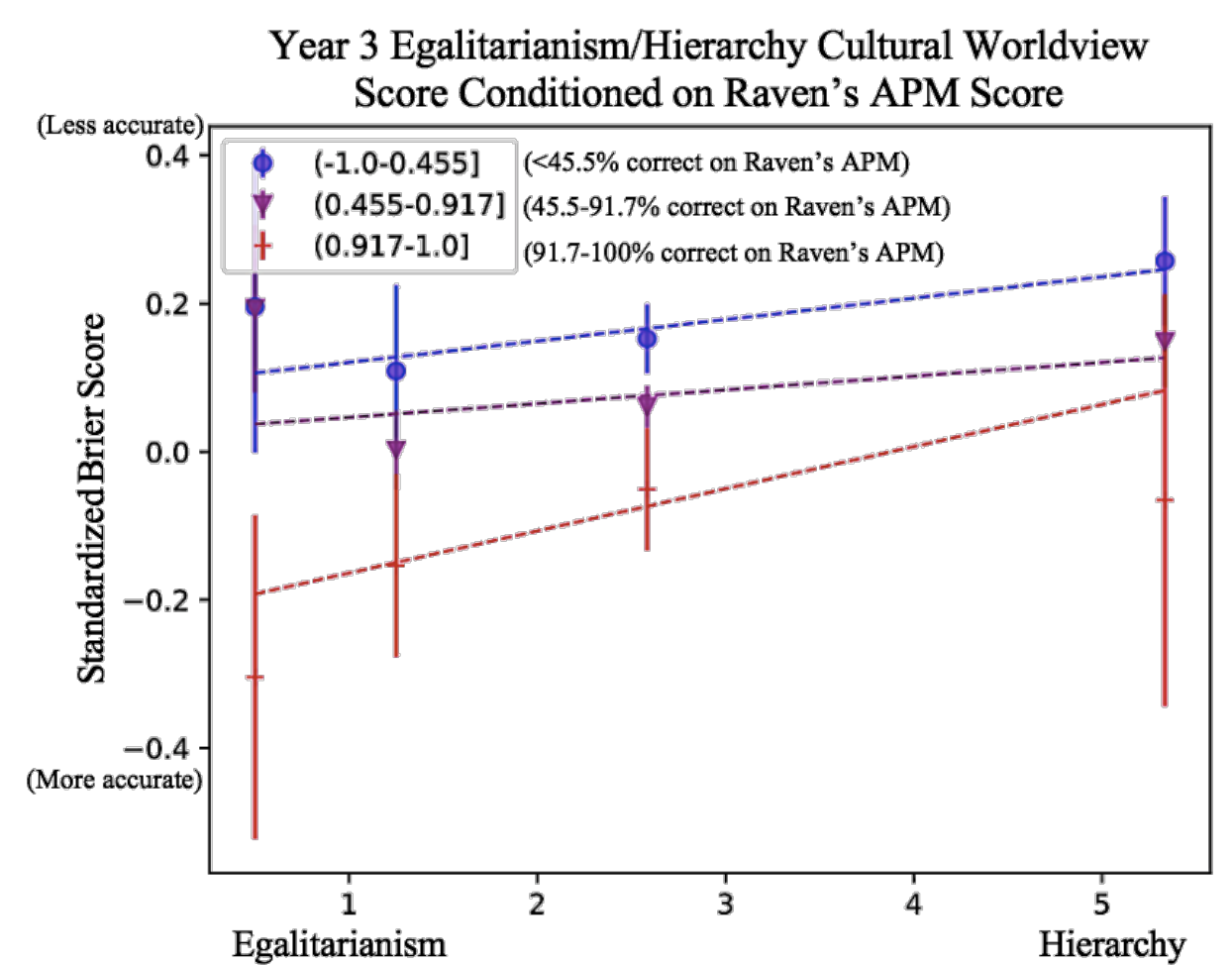
5. Results

Cultural Worldview Scale:
- Egalitarianism: prioritizing equality for all people (~liberal)
- Hierarchy: maintain roles/ranks in society (~conservative)

Raven's Advanced Progressive Metrics (APM): measures fluid intelligence (ability to solve novel problems)
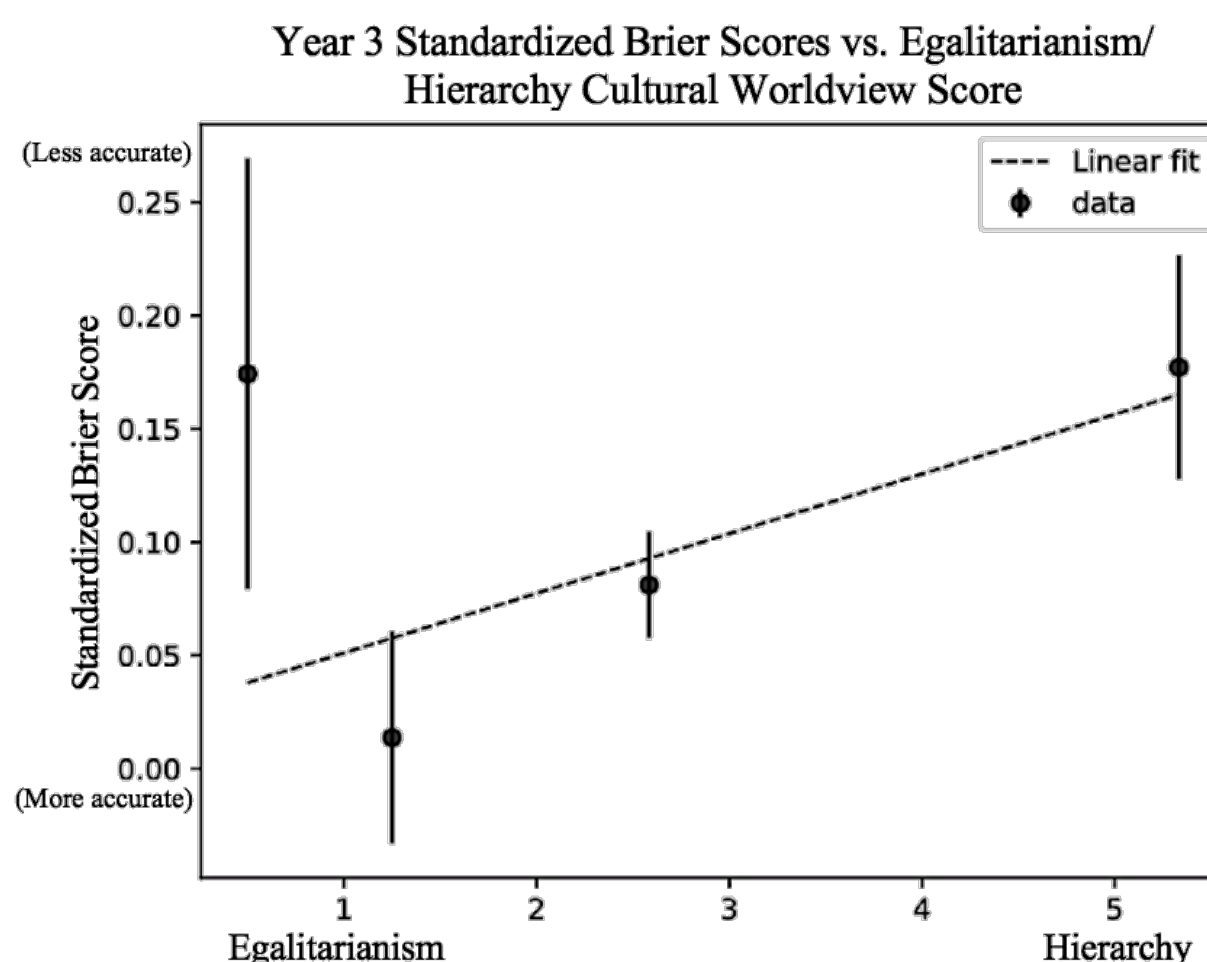
Aggregate: a greater endorsement of a hierarchical worldview correlates to a worse Brier score

Disaggregate: Brier scores decrease by intelligence and by hierarchical worldview

6. Discussion

- Simpson's paradox is present in forecasting.
  - Results suggest that cultural worldview impacts forecasting performance.
- Future work:
  - Focus on cultural worldviews and forecasting.
  - Improve Simpson's paradox program to run data with missing responses—data imputation may add bias to results.

USC Viterbi
School of Engineering
Information Sciences Institute