

## **Intuitive Strategies for Revising Opinions:**

**Are people appropriately influenced by others?**

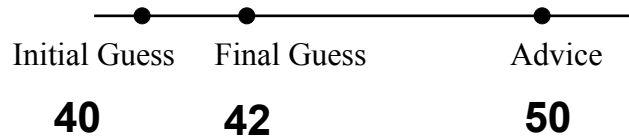
Jack Soll (INSEAD) & Richard Larrick (Duke)

### **Forshadowing the Main Results**

- What strategies do people use to revise quantity opinions?
  - lots of choosing, occasional averaging
- Empirically, people would have been better off had they averaged (on average).
- Sometimes it's better to choose. Model shows the conditions that favor averaging.

## Measuring Strategies

*How old is Madonna?*



$$\text{weight on advice} = \frac{\text{adjustment}}{\text{total difference}} = \% \text{ of adjustment toward advice}$$

$$\text{WS} = \text{weight on self} = 1 - \text{weight on advice} = .80$$

## Two Important Strategies

Averaging: WS = 0.5

Choosing: WS = 0 or 1

## Related Literature

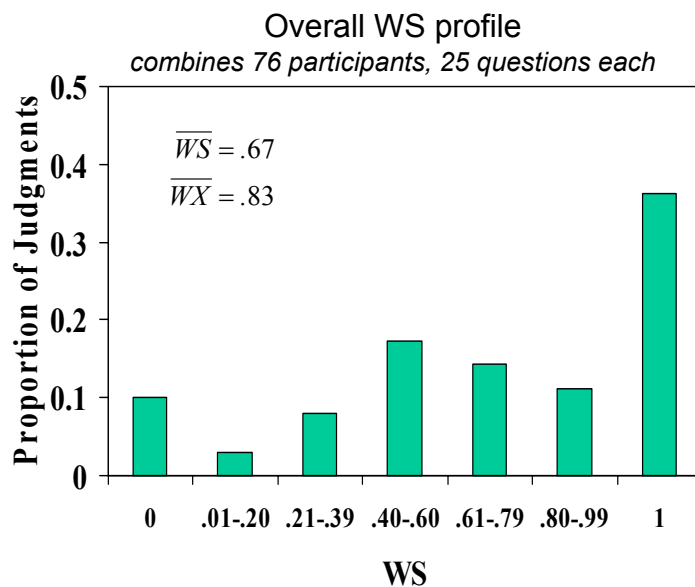
- Weighted averaging models for quantity judgments (Birnbaum & associates, 1970s)
- Recent Approach (Harvey & Fisher, 1997; Yaniv & Kleinberger, 2000; Lim & O'Connor, 1995)
  - Questions with numerical answers
  - (1) initial guess, (2) see advice, (3) revise
  - Finding: mean adjustment of 20-30% toward advice
- Two Observations
  - past research tends to focus on weighting, not accuracy
  - Means hide valuable information
    - e.g., Yaniv & Kleinberger reported median weight on self of .98

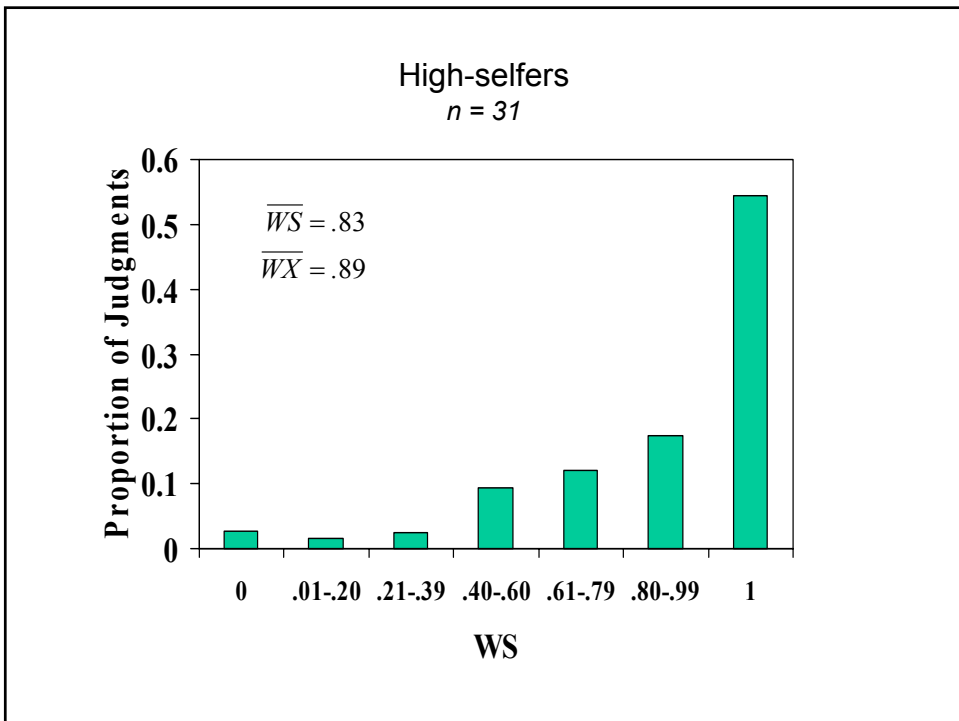
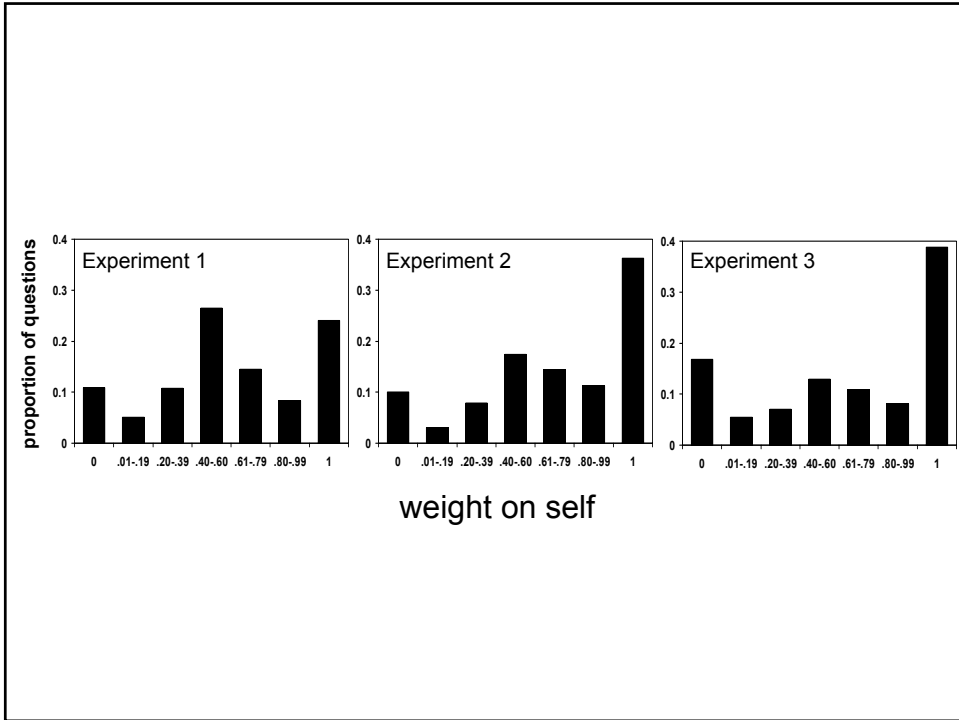
## A Paradigm for Investigating The Accuracy of Advice Use

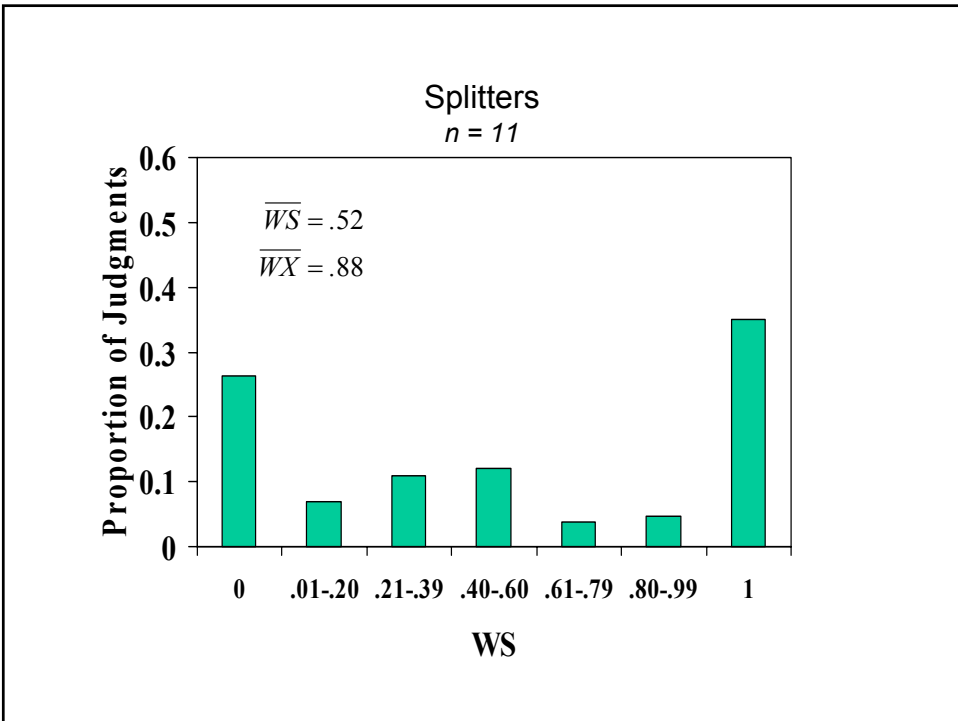
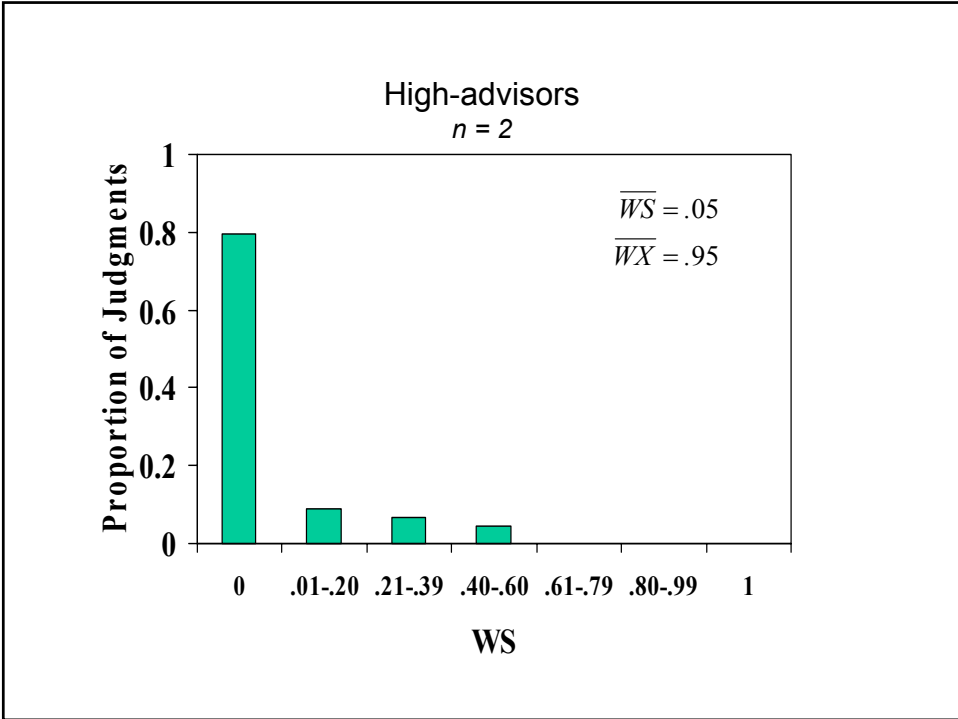
- One judge, one advisor
- Quantitative estimates with correct answer
- Responses are private
- Advice is from real person (i.e., not rigged)
- Procedure
  - Initial guesses
  - Learn someone else's guesses ("advisor")
  - Revise judgments
  - Revised judgments are paid based on accuracy

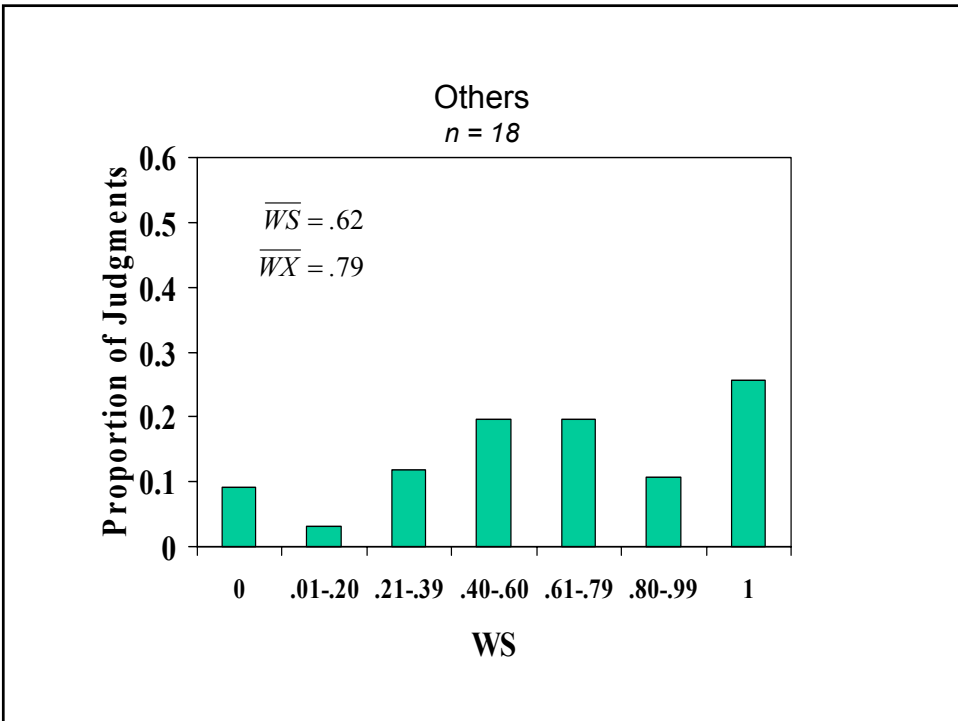
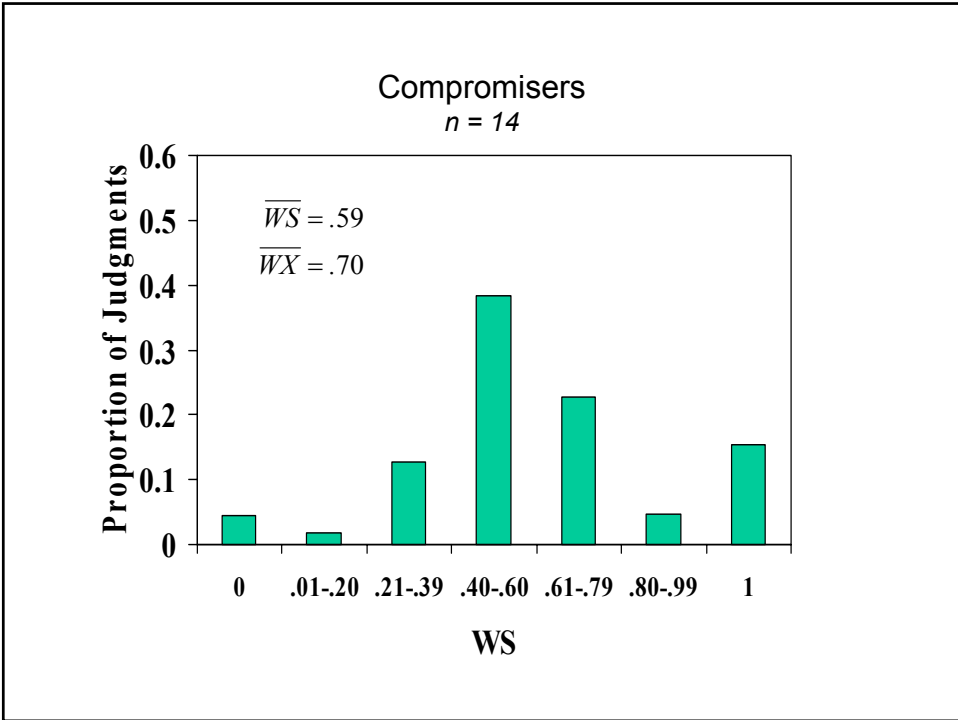
## Feedback Study *Predicting MBA Salaries*

- Everyone took a 5-question pre-test
- 25-Item Prediction Test (1 week later)
  - Initial guesses for 25 new schools
  - Received advice
    - Some saw pre-test MAD scores, own & advisor's
  - Revised guesses
  - Paid for accuracy of revised guesses









## Frequency of Strategies by Feedback

*percentages in red*

| Feedback on pretest | High self                                   | High advisor                              | Splitter                                   | Comp                                       | Other                                      | $\overline{WS}$ | $\overline{WX}$ |
|---------------------|---|---|--|--|--|-----------------|-----------------|
| Better              | 15<br><span style="color: red;">71.4</span> | 0<br><span style="color: red;">--</span>  | 0<br><span style="color: red;">--</span>   | 1<br><span style="color: red;">4.8</span>  | 5<br><span style="color: red;">23.8</span> | .77             | .84             |
| Worse               | 5<br><span style="color: red;">21.7</span>  | 2<br><span style="color: red;">8.7</span> | 7<br><span style="color: red;">30.4</span> | 4<br><span style="color: red;">17.7</span> | 5<br><span style="color: red;">21.7</span> | .54             | .83             |

- If better than advisor, people dismiss advice
- If worse than advisor, people use a variety of strategies
  - WS changes
  - WX does not

## Frequency of Strategies by Feedback

| Feedback on pretest | High self                                   | High advisor                              | Splitter                                   | Comp                                       | Other                                      | $\overline{WS}$ | $\overline{WX}$ |
|---------------------|---|---|--|--|--|-----------------|-----------------|
| Better              | 15<br><span style="color: red;">71.4</span> | 0<br><span style="color: red;">--</span>  | 0<br><span style="color: red;">--</span>   | 1<br><span style="color: red;">4.8</span>  | 5<br><span style="color: red;">23.8</span> | .77             | .84             |
| Worse               | 5<br><span style="color: red;">21.7</span>  | 2<br><span style="color: red;">8.7</span> | 7<br><span style="color: red;">30.4</span> | 4<br><span style="color: red;">17.7</span> | 5<br><span style="color: red;">21.7</span> | .54             | .83             |
| No Feedback         | 11<br><span style="color: red;">34.4</span> | 0<br><span style="color: red;">--</span>  | 4<br><span style="color: red;">12.5</span> | 9<br><span style="color: red;">28.1</span> | 8<br><span style="color: red;">25.0</span> | .70             | .82             |

## Improvement from Initial MAD

| Strategy     | n  | Intuition | Averaging |
|--------------|----|-----------|-----------|
| High Self    | 31 | 8.9%      | 13.3%     |
| High Advisor | 2  | -3.8%     | 15.4%     |
| Splitter     | 11 | 8.1%      | 19.0%     |
| Compromise   | 14 | 14.2%     | 12.8%     |
| Other        | 18 | 14.2%     | 20.4%     |

So far,

- People tend to choose or average
- In this study, averaging beats intuition
- Our model compares averaging and choosing
- But, why not consider optimal weights?
  - now, a brief detour

- Case for averaging
  - Groups (Zajonc, 1962; Einhorn et al., 1977)
  - Improper Linear Models (Dawes & Corrigan, 1974)
  - Statistics/Forecasting (Winkler, 1984; Clemen, 1989)
- Case for choosing
  - One-reason decision making (Todd & Gigerenzer, 2000)
  - Take the best (Gigerenzer & Goldstein, 1996)
- Our analysis
  - We concur with both camps that fine-tuning the weights won't help, and could hurt.
  - Whether averaging or choosing works best depends on the environment.

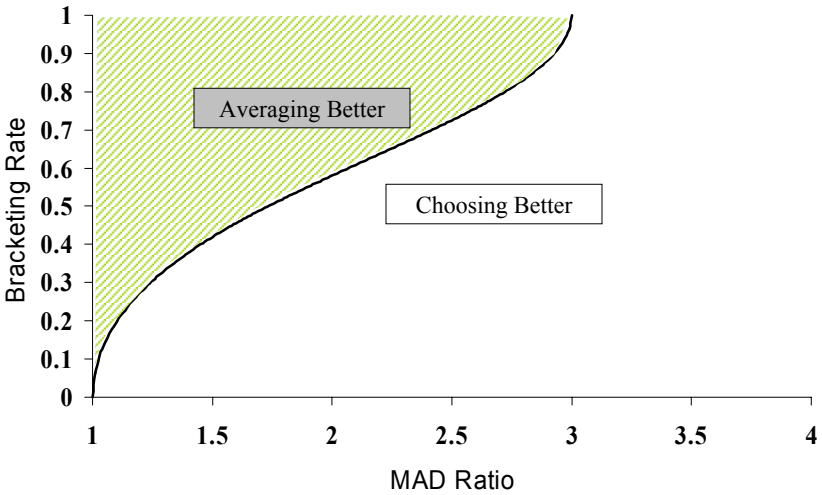
## The Model

*Compares averaging & choosing*

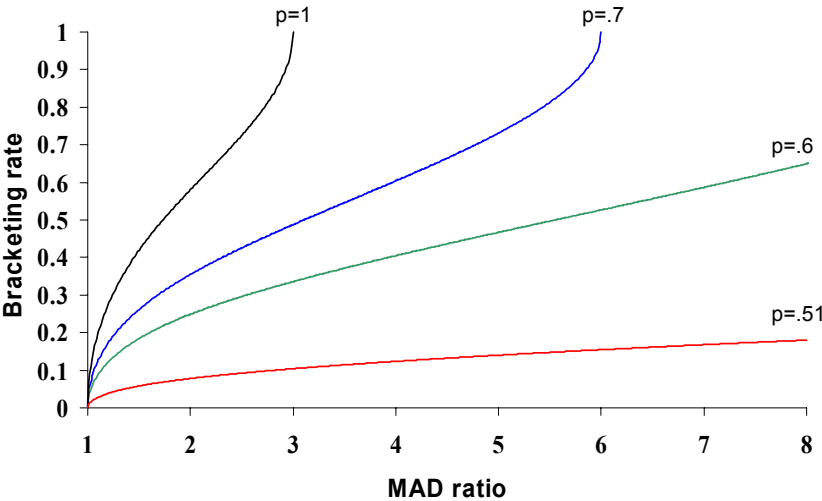
### Three Parameters

- MAD ratio, highest over lowest
  - *Measures relative expertise*
- Bracketing rate
  - *Combines bias and correlation into one measure*
- Probability of Detecting the Better Judge
  - *the one with the lower MAD over set of q's*

### Expert Detection Probability = 1.0



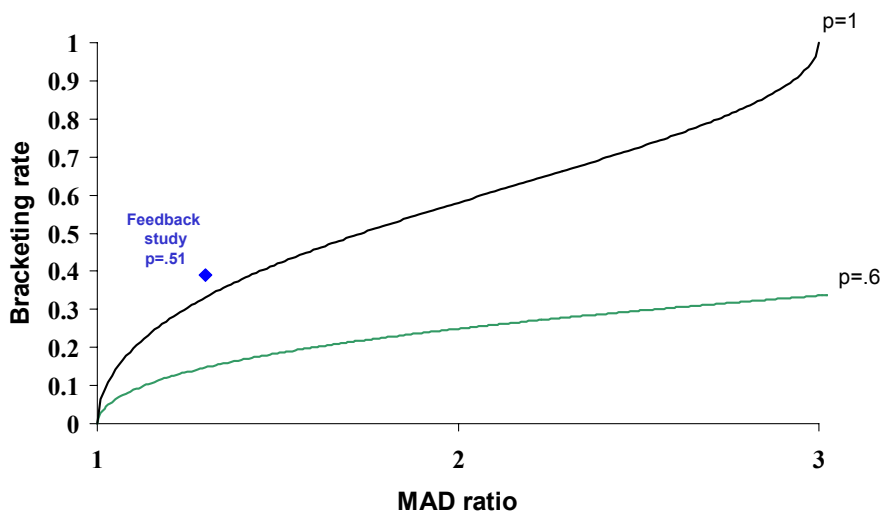
### Averaging vs. Choosing as a Function of Expert Detection



## Feedback Study Aggregate Accuracy Statistics

|                   | MAD Ratio | Bracketing | Means |      | Mean % Improvement over initial |                    | Expert detect'n |
|-------------------|-----------|------------|-------|------|---------------------------------|--------------------|-----------------|
|                   |           |            | WS    | WX   | Avg'ing                         | Intuitive revision |                 |
| No Feedback       | 1.35      | 38%        | 0.70  | 0.82 | 15%                             | 10%                | 50%             |
| Positive Feedback | 1.37      | 39%        | 0.77  | 0.84 | 13%                             | 11%                | 52%             |
| Negative Feedback | 1.24      | 39%        | 0.54  | 0.83 | 19%                             | 10%                | 52%             |

### Averaging vs. Choosing as a Function of Expert Detection



## Conclusions

- W-shaped WS distribution suggests that people mainly average and choose.
  - It's not anchoring and adjustment.
  - Ecologically, this is the 'right' portfolio of strategies.
- People favor choosing.
  - overconfidence in identifying expert (Soll & Larrick, 2002)
  - misappreciation of averaging (Larrick & Soll, 2002)
  - intuitive loss function may justify choosing
- Strategy is sensitive to feedback. There is a motivational effect.
- Model explains why averaging out- or underperforms intuition in individual cases.

## Prescriptions?

- There are many situations when averaging is not appropriate
  - It is very costly to average if there is a large expertise difference (and the expert can be identified!)
  - Future research should describe environments. E.g., what types of contexts favor averaging? ...favor choosing?
- Our rule of thumb:  
If you are highly certain that one judge is more than twice as accurate than the other, you should choose; otherwise, you should average.