# Willingness-to-pay for Warnings: Main Tables

A. Gaduh, P. McGee and A. Ugarov

May 28, 2022

# Research Question

- How much do people value alerts (signals) about potential preventable threats?
- How do signal's probabilistic characteristics affect the willingness-to-pay for it and the welfare gains from using it?
- Applications:
  - Natural disaster warnings (tornados, floods, earthquakes)
  - Medical tests for treatable conditions
  - Investing in research on likelihood of catastrophic events (rogue AI, global warming, pandemics)
- Note: most real-life applications provide little practice with using the signal

# Overview of the Experiment

- An insurance experiment:
  - Two states of the world: bad ($\omega = 1$) and good ($\omega = 0$)
  - Probability of a bad state is $P(\omega = 1) = \pi$
  - Bad state $\implies$ loss of $\$L$
  - A perfectly protective insurance can be purchased for $\$c$
- Subject can purchase a signal $s$ before purchasing the insurance:
  - A signal is characterized by its true-positive $(P(s = 1|\omega = 1))$ and true-negative rates $(P(s = 0|\omega = 0))$

**Research objective**

How do signal characteristics affect the WTP?

- Theoretically, what should be the WTP for a signal?
- If bad states are a priori rare ($\pi L << c$) $\implies$ never protect without a signal
- The theoretical WTP $b$ for an expected utility maximizer given a signal $s$ is a solution $b^*$ to the following:

$$P(s=1)u(Y_0 - b^* - c) + \pi P(s=0|\omega=1)u(Y_0 - b^* - L) +$$

$$+(1-\pi)P(s=0|\omega=0)u(Y_0 - b^*) = (1-\pi)u(Y_0) + \pi u(Y_0 - L)$$

- A risk-neutral agent then pays:

$$b^* = \pi(1 - P(s=0|\omega=1))L - P(s=1)c$$

- The formulas become more complicated if subjects can protect without a signal (bad state are not rare enough)
- The theoretical WTP $b$ for an expected utility maximizer given a signal $s$ is a solution $b^*$ to the following:

$$P(s=1)u(Y_0-b^*-c)+\pi P(0|1)u(Y_0-b^*-L)+(1-\pi)P(0|0)u(Y_0-b^*) =$$

$$= \min[(1-\pi)u(Y_0) + \pi u(Y_0 - L), u(Y_0 - c)]$$

- A risk-neutral agent then pays:

$$b^* = \min[\pi L, c] - \pi(1 - P(s=0|\omega=1))L - P(s=1)c$$

# Hypotheses

1. Conditional on the signal's value for risk-neutral subjects, false positive and false negative rates reduce the perceived value of the signal (WTP)
    - *The opposite is true: subjects underreact to false positive and false negative rates and overpay for bad signals*
2. Conditional on the signal's value for risk-neutral subjects, false positive and false negative rates increase expected costs
    - *No: FP and FN rates have no significant effects on costs besides their predicted theoretical effect*
3. Extra: how much of these disrepancies result from belief updating issues or risk aversion?

# Risk Aversion Measurement

- Measure risk aversion based on blind protection choices:
  - Exclude obs from subjects switching back and forth
  - The lowest probability for which a subject chooses to protect is $\pi^*$
  - Calculate their coefficient of relative risk aversion $\theta$ as the solution to the following equation:

$$\pi^* u(Y_0 - L; \theta) + (1 - \pi^*) u(Y_0; \theta) = u(Y_0 - c; \theta)$$

  - Where $u()$ is the CRRA utility function:

$$u(x; \theta) = \frac{x^{1-\theta} - 1}{1 - \theta}$$

- Note: risk lovers have $\theta < 0$

# CRRA Estimates

- Most subjects are moderately risk averse:

| Probability ($\pi^*$) | $\theta$ | $N$ |
|---|---|---|
| Always protect | >2 | 1 |
| 0.1 | 2 | 2 |
| 0.15 | 1.216 | 7 |
| 0.2 | 0.573 | 17 |
| 0.25 | 0 | 7 |
| 0.3 | -0.539 | 5 |
| Never protect | <-0.539 | 8 |

- Note:
  - There are 18 subjects (out of 65) switching multiple times
  - Can use more sophisticated methods to measure risk aversion for those

# WTP for the Signal

- Theoretical value of signal for risk-neutral subject:

$$b^* = \underbrace{\min[\pi L, c]}_{\text{BP costs}} - \underbrace{\pi(1 - P(s = 0|\omega = 1))L}_{\text{False neg. costs}} - \underbrace{P(s = 1)c}_{\text{Protection costs}}$$

- Two potential approaches:

  1. Regress the discrepancy between WTP $V$ and theoretical value $b^*$:
  $$V - b^* = \alpha_0 + \alpha_1 \text{FN costs} + \alpha_2 \text{Prot. costs} + \epsilon$$
  2. Regress WTP directly on its components and account for censoring at 0:
  $$V = \min[0, \beta_0 + \beta_1 \text{FN costs} + \beta_2 \text{Prot. costs} - \beta_3 \text{BP costs} + \gamma]$$
  3. Note: protection costs include costs due to false positive signals

# WTP for the Signal (Approach 1)

- Coefficient sign. different from zero is an anomaly: people overpay for bad signals

Figure: WTP for Information (Discrepancy)

| | (1) All | (2) Risk-averse | (3) Risk-loving | (4) Switchers |
|---|---|---|---|---|
| False pos. costs | .17 | .207 | .062 | .447 |
| | (1.6) | (1.5) | (0.4) | (1.0) |
| False neg. costs | .3$^{***}$ | .297$^{***}$ | .329$^{***}$ | .169 |
| | (4.8) | (3.4) | (3.3) | (0.8) |
| Constant | -.111 | -.143 | -.139 | .289 |
| | (-1.2) | (-1.0) | (-1.0) | (0.8) |
| N obs. | 744 | 336 | 354 | 54 |
| AIC | 2814 | 1243 | 1361 | 214 |
| p(coeffs=0) | 5.45e-06$^{***}$ | .00125$^{***}$ | .00434$^{***}$ | .468 |

$t$ statistics in parentheses
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

# WTP for the Signal (Approach 2, Tobit Estimation)

- Coefficient should **differ from one** in abs. value to show an anomaly (ignore stars for now)

Figure: WTP for Information (Tobit Estimation)

| | (1) All | (2) Risk-averse | (3) Risk-loving |
|---|---|---|---|
| model | | | |
| BP costs | .688*** | .625*** | .72*** |
| | (12.0) | (7.9) | (8.2) |
| Pos. signal costs | -.373*** | -.233 | -.486** |
| | (-3.0) | (-1.4) | (-2.5) |
| False neg. costs | -.587*** | -.569*** | -.578*** |
| | (-6.7) | (-4.9) | (-4.2) |
| N obs. | 744 | 336 | 354 |
| AIC | 2726 | 1213 | 1312 |
| p(coeff=1) | 6.68e-09*** | 1.04e-06*** | .00357*** |

*t* statistics in parentheses
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

# WTP for the Signal (Risk Aversion)

- Explaining the discrepancy between WTP and value with risk aversion:

Figure: WTP for Information (different risk aversion)

| | (1) $\theta = 0$ | (2) $\theta = 0.5$ | (3) $\theta = 1.0$ | (4) $\theta = 1.5$ | (5) $\theta = 2.5$ | (6) Heterogeneou |
|---|---|---|---|---|---|---|
| FP costs | .295** | .322** | .316** | .271* | .151 | .29** |
| | (0.1) | (0.1) | (0.1) | (0.1) | (0.1) | (0.1) |
| FN costs | .243*** | .346*** | .46*** | .559*** | .69*** | .343*** |
| | (0.1) | (0.1) | (0.1) | (0.1) | (0.1) | (0.1) |
| Constant | .254 | -.146 | -.664*** | -1.32*** | -1.81*** | -.411 |
| | (0.2) | (0.2) | (0.2) | (0.2) | (0.2) | (0.3) |
| Prior dummies | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 504 | 504 | 504 | 504 | 504 | 504 |
| Adjusted $R^2$ | 0.21 | 0.27 | 0.28 | 0.32 | 0.37 | 0.19 |

Standard errors in parentheses
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

# Actual Costs vs Theoretical Costs

- Calculate actual costs based on decisions made in the Informed Protection treatment and actual posterior probabilities of losses.
- Each reported participant's strategy $s$ is a tuple of numbers $(r_w, r_b)$ representing protection responses correspondingly to white and black hints
- Then the expected cost of each decision are:

$$EC(s) = \pi(P(0|1)(1 - r_w) + P(1|1)(1 - r_b))L + P(s = 1)c$$

$$+(P(s = 0)r_w + P(s = 1)r_b)c$$

- Regress expected costs on minimal theoretical costs and other signal characteristics

# Actual Costs vs Theoretical Costs

Figure: Actual Exp. Costs vs Theoretical Costs

| | (1) OLS | (2) OLS | (3) FE | (4) FE | (5) FE |
|---|---|---|---|---|---|
| Optimal exp. costs | .979*** | .549*** | .987*** | .733*** | 1.06*** |
| | (13.1) | (2.9) | (11.5) | (6.0) | (10.2) |
| Prior prob. | -.689 | -3.3** | -.607 | -2.15** | -.18 |
| | (-0.9) | (-2.5) | (-0.8) | (-2.5) | (-0.2) |
| False neg. rate | | -2.48*** | | -1.88*** | |
| | | (-3.4) | | (-3.1) | |
| False pos. rate | | -1.04 | | | .71 |
| | | (-1.4) | | | (1.0) |
| Constant | -.707*** | -.542*** | -.711*** | -.637*** | -.754*** |
| | (-6.2) | (-4.5) | (-7.4) | (-6.6) | (-6.8) |
| Observations | 743 | 743 | 743 | 743 | 743 |
| Adjusted $R^2$ | 0.38 | 0.39 | 0.43 | 0.44 | 0.43 |

*t* statistics in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

# Additional Complementary Tables

1. Factors affecting informed protection responses
2. The effect of beliefs on informed protection
3. How accurate are their beliefs?
4. Decomposition of belief updating: priors vs signals

# Informed Protection: Determinants

Figure: Informed Protection

| | (1) All | (2) All | (3) Good quiz | (4) Good quiz |
|---|---|---|---|---|
| Informed protection | | | | |
| Posterior prob. | 2.15*** | .662*** | 2.26*** | .638*** |
| | (19.1) | (3.3) | (17.7) | (3.0) |
| Prior prob. | | 1.13*** | | 1.17*** |
| | | (4.1) | | (3.8) |
| Gremlin says Black | | 1.34*** | | 1.46*** |
| | | (8.8) | | (8.8) |
| Constant | -.662*** | -1.03*** | -.717*** | -1.1*** |
| | (-14.2) | (-11.2) | (-14.2) | (-10.9) |
| Observations | 1487 | 1487 | 1259 | 1259 |
| *AIC* | 1467.25 | 1394.01 | 1211.48 | 1137.59 |

*t* statistics in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

# Informed Protection: Do Subject's Beliefs Matter?

Figure: Informed Protection: Response to Reported Beliefs

| | (1) All | (2) All | (3) Good quiz |
|---|---|---|---|
| Informed protection | | | |
| Belief | 2.18*** | 2.62*** | 2.8*** |
| | (18.5) | (18.2) | (17.0) |
| Belief error | | 1.52*** | 1.41*** |
| | | (11.5) | (9.3) |
| Constant | -.762*** | -.881*** | -.963*** |
| | (-14.3) | (-15.7) | (-15.9) |
| Observations | 1487 | 1487 | 1259 |
| *AIC* | 1566.82 | 1413.23 | 1146.78 |

*t* statistics in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

# Belief Updating: Correlation

Figure: Belief Elicitation: Belief vs Posterior

| | (1) All | (2) Good quiz | (3) Dishonest greml |
|---|---|---|---|
| Posterior prob. | .644*** | .693*** | .524*** |
| | (37.5) | (39.2) | (21.8) |
| Constant | .175*** | .15*** | .236*** |
| | (21.7) | (19.8) | (23.4) |
| Observations | 1488 | 1260 | 992 |
| Adjusted $R^2$ | 0.53 | 0.60 | 0.38 |

*t* statistics in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

# Belief Updating: Decomposition

- Posterior probability $\mu = P(B|S = x)$ that the ball is black conditional on a hint $S = x$ can be written as:

$$\ln\left(\frac{\mu}{1-\mu}\right) = \lambda_0 + S_B + S_W$$

- With $\lambda_0 \equiv \ln(p/(1-p))$ representing (transformed) prior beliefs
- And $S_B$, $S_W$ describing the effect of new evidence:

$$S_B \equiv I(S = B)\ln(P(s = B|B)/P(s = B|W))$$

$$S_W \equiv I(S = W)\ln((1 - P(s = B|B))/(1 - P(s = B|W))$$

# Belief Updating: Decomposition

Figure: Belief Elicitation: Decomposition

| | (1) OLS | (2) FE | (3) Good quiz, FE |
|---|---|---|---|
| lt_prior | .237*** | .182*** | .187*** |
| | (3.9) | (4.0) | (4.0) |
| signalB | .426*** | .865*** | .992*** |
| | (5.1) | (6.4) | (6.7) |
| signalW | .439*** | 0 | 0 |
| | (5.7) | (.) | (.) |
| Constant | | -.54*** | -.632*** |
| | | (-6.0) | (-6.6) |
| Observations | 332 | 332 | 288 |
| Adjusted $R^2$ | 0.29 | 0.29 | 0.34 |

$t$ statistics in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$