Pay For Results, Not Prestige

Robin Hanson
Economics, George Mason University
Future of Humanity Institute, Oxford

Expert Judgment Workshop
Online, June 10, 2020
Our Many Experts (= Agents)

• Decorator
• Contractor
• Priest
• Doctor
• Therapist
• Teacher
• Researcher

• C.E.O.
• Fund manager
• Reporter
• Politician
• Regulator
• Police
• Lawyer
Three Kinds of Skepticism

**Senses** – Trust my eyes, ears?

*Solution*: compare many sources

**Authority** – Trust king, scholar, priest?

*Solution*: “science” = stable metrics, shared data, math models

**Experts** – Are my experts self-serving?

*Solution*: Pick/pay by records/results
Five Ways To Pick/Pay Experts

**Loyalty** – Trust folks on your side

**Prestige** – Go with impressive elites

**Process** – Follow proper procedure

**Records** – Go with top track records

**Results** – Pay only if you get results
Heidhues & Lagerlöf 2003

2 politicians, get binary clue, pick binary policy
1 voter wants politician so policy = binary truth

\[ i, r \in E = \{1,2\}, \quad a_i, b_i, e_i \in B = \{0,1\} \quad a = (a_1, a_2) \]
\[ \omega = \{b_0(\omega), b_1(\omega), b_2(\omega)\} \in \Omega = B^3, \ \text{Cov}_p[b_0, b_i] > 0 \]
\[ \tilde{\pi}_{i\omega} = \pi_{i\omega}(0) = \Omega, \ \pi_{i\omega}(1) = \{\omega' \in \Omega : b_i(\omega') = b_i(\omega)\} \]
\[ u_i(a, r, e_i) = \mathbf{1}[r = i] - \varepsilon e_i \quad v_{0\omega}(a, r) = \mathbf{1}[b_0(\omega) = a_r] \]
1. \[ \max_{e_i \in B} \mathbb{E}_p[u_i(a, r, e_i)] \]
2. \[ \max_{a_i \in B} \mathbb{E}_p[u_i(a, r, e_i) | \pi_{i\omega}(e_i)] \]
3. \[ \max_{r \in E} \mathbb{E}_p[v_{0\omega}(a, r) | a] \]

(Use some if mix.)

HL: if \( \varepsilon < 0 \), then \( e_i = 1 \), \( a_i \) neglects \( b_i \)
Me: if \( \varepsilon > 0 \), then \( e_i = 0 \), \( a_i \) ignores \( b_i \)
N managers, get binary clue, pick binary policy
1 trader sets stock to prob. policy = binary truth

\( i \in E = \{1, 2, \ldots, N\}, \ a_i, b_i, e_i \in B = \{0, 1\} \quad a = (a_1, a_2, \ldots, a_N) \)

\( r_i \in R = [0, 1], \ \omega = \{b_0, b_1, \ldots, b_N\} \in \Omega = B^{N+1}, \ \text{Cov}_p [b_0, b_i] > 0 \)

\( \tilde{\pi}_{i\omega} = \pi_{i\omega} (0) = \Omega, \ \pi_{i\omega} (1) = \{\omega' \in \Omega : b_i (\omega') = b_i (\omega)\} \)

\( u_i (a, r, e_i) = r_i - \varepsilon e_i \quad v_{0\omega} (a, r) = -\sum_{i \in E} (r_i - \mathbf{1}[b_0 (\omega) = a_i])^2 \)

1. \( \max_{e_i \in B} \mathbb{E}_p [u_i (a, r, e_i) \mid a_i^-] \quad 3. \max_{r \in R^N} \mathbb{E}_p [v_{0\omega} (a, r) \mid a] \)

2. \( \max_{a_i \in B} \mathbb{E}_p [u_i (a, r, e_i) \mid \pi_{i\omega} (e_i), a_i^-], \ a_i^- = (a_1, a_2, \ldots, a_i) \)

BP: if \( \varepsilon < 0 \), then \( e_i = 1 \), \( a_i \) neglects \( b_i \)

Me: if \( \varepsilon > 0 \), then \( e_i = 0 \), \( a_i \) ignores \( b_i \)
Experts learn no more than clients know when reward, if
1. Info effort hidden, costly, ordered
2. Experts optimize but not directly care re topics, honesty
3. Coordinate via (a) see prior expert acts, or
   (b) 2-player 0-sum game

\[ i \in E, \; j \in C, \; a_{ij} \in A_{ij}, \; r_{ij} \in R_{ij}, \; e_{id} \in \Gamma_{id}, \; \omega \in \Omega, \; \text{all finite} \]

1. \[ \max_{(e_{id})_{d \in D}} \mathbb{E}_{p_i} [u_i(a, r, e_i) \, | \, \tilde{\pi}_{i\omega}, a_i^-] \quad \pi_{i\omega}(e_i) = \bigcap_{d \in D} \pi_{id\omega}(e_{id}) \]
2. \[ \max_{(a_{ij})_{j \in C}} \mathbb{E}_{p_i} [u_i(a, r, e_i) \, | \, \pi_{i\omega}(e_i), a_i^-] \quad a_i^- = (a_{kj})_{k \in E : k < i} \]
3. pick \[ (r_{ij})_{i \in E} \quad \mathbb{E}_{p_j} [v_j(a, r) \, | \, \pi_{j\omega}, F(a, r_j^-)] \quad r_j^- = (r_{ik})_{k \in C : k < j} \]

If A) exists focal \( \hat{\pi}_{\omega} \subseteq \pi_{j\omega} \) so \( \pi_{id\omega}(e_{id}) = , \subseteq, \text{or} \supseteq \hat{\pi}_{d\omega} \).

B) if \( \pi_{i\omega}(e_i') \subset \pi_{i\omega}(e_i) = \hat{\pi}_{\omega} \), then \( u(a, r, e_i') < u(a, r, e_i) \).

Then \( \pi_{i\omega}(e_i) \supseteq \hat{\pi}_{\omega} \)
Obstacles To Paying For Results

*Individuals:*
- Wait to get paid
- Pay more than have
- Noise makes pay risk
- Evaluation stress

*Organizations:*
- Market power
- Bad act bundles
- Unmeasured results

How to Measure Anything, Hubbard 2010
Buy Health, Not Health Care

• Buy life, pain, disable, & health insure from LMO
• They pick/pay medicine
  • Internalize tradeoffs!
• Advise re other acts
• Pick min premium LMO

• Set life, etc. insure to own value of life, etc.
• Extra insure payout to Anti-LMO, who pays part of LMO premium
• Anti-LMO can’t act:
  • Few employees
  • Transparent
  • Sits far away
Who Vouches For You?

(Not change how pick what is crime, fines, court rulings.)

1. Require all get **VOUCHER** to cover *all* legal liability
   - voucher = shows can pay (& reinsured), on hook until replace

2. All crimes punished officially via **FINES**
   - Voucher pays fine F re client act, except crime of no voucher
   - Fine F pays for bounty B, victim help C, govt. revenue R
   - Voucher-client contract specifies: *premium, punish, co-liable groups, limits on movement, monitor/privacy*
   - Fines F can vary with freedom limits via $\frac{1}{p(\text{catch})}$ **

3. All enforcement via **BOUNTY** hunters
   - First to convince court paid, T.B.D. rights re evidence collected
   - Blue-wall-of-silence only if *all* hunters forgo bounties
Buy Low, Sell High

“Pays $1 if Trump wins”

Will price rise or fall?

E[ price change | ?? ]

Lots of ?? get tried, price includes all!
Advantages

• Numerically precise
• Consistent across many issues
• Frequently updated
• Hard to manipulate
• Need not say who how expert when
  • Issue is not experts vs. amateurs
• At least as accurate as alternatives
• Easy to subsidize so $ only to info adders

(Tournaments: org picks Qs, players, scoring fn., consensus fn. Similar gains, but for them must trust org who manages)
Estimates from Prices

$1 \text{ if } A \quad \leftrightarrow \quad p(A) \quad $1

$1 \text{ if } A \& B \quad \leftrightarrow \quad p(A \& B) \quad $1

$1 \text{ if } A \& B \quad \leftrightarrow \quad ? \quad $1 \text{ if } A

\[
p(A \& B) \quad \quad \frac{p(A \& B)}{p(B \mid A)} \quad \quad p(A)
\]
Estimates from Prices

$1 \text{ if A} \leftrightarrow p(A) \quad $1$

$1 \text{ if A&B} \leftrightarrow p(A&B) \quad $1$

$1 \text{ if A&B} \leftrightarrow p(B|A) \quad $1 \text{ if A}$

$\$ x \leftrightarrow E[x] \quad $1$

$\$ x \text{ if A} \leftrightarrow E[x|A]*p(A) \quad $1$

$\$ x \text{ if A} \leftrightarrow E[x|A] \quad $1 \text{ if A}$
SCICAST.ORG
18 months
1200 questions
130K forecasts
240 forecasts/day
Single Payer Decision Markets

$1 if Lifespan Up & Single Payer

$1 if Lifespan Up & Not Single Payer

$1 if Not Single Payer

$1 if Single Payer

P(S)

P(not S)

P(L | S)

P(L | not S)

Compare!
If I Had A Million

• Two markets for each Fortune 500 firm:
  \[ E[ \text{Stock price} \mid \text{CEO leave this quarter} ] \]
  \[ E[ \text{Stock price} \mid \text{CEO stay this quarter} ] \]

• Lots press, CEOs study, try to influence
  • Let them pay to delist firm?

• Track ROI of firms follow advice vs. not
  • Few years to see significant difference
  • Shame/sue boards to follow advice
## 2016 US President Race

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Nominate?</th>
<th>Win?</th>
<th>Win if Nom.?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton</td>
<td>78%</td>
<td>46%</td>
<td>58%</td>
</tr>
<tr>
<td>Warren</td>
<td>4.8%</td>
<td>2.9%</td>
<td>62%</td>
</tr>
<tr>
<td>Biden</td>
<td>4.8%</td>
<td>2.9%</td>
<td>62%</td>
</tr>
<tr>
<td>Omalley</td>
<td>3.8%</td>
<td>2.0%</td>
<td>51%</td>
</tr>
<tr>
<td>Rubio</td>
<td>22%</td>
<td>11%</td>
<td>50%</td>
</tr>
<tr>
<td>Paul</td>
<td>9.1%</td>
<td>3.8%</td>
<td>42%</td>
</tr>
<tr>
<td>Bush</td>
<td>31%</td>
<td>18%</td>
<td>59%</td>
</tr>
<tr>
<td>Walker</td>
<td>20%</td>
<td>7.7%</td>
<td>39%</td>
</tr>
</tbody>
</table>

oddschecker.com in Feb 2015
## Block Size Changes

<table>
<thead>
<tr>
<th>Metric</th>
<th>No change</th>
<th>Block size increase (82.4% likelihood)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>$280.33</td>
<td>$555.57</td>
</tr>
<tr>
<td>Network hash rate</td>
<td>3,255.17 PH/s</td>
<td>5,032.64 PH/s</td>
</tr>
<tr>
<td>Daily transaction volume</td>
<td>168,438.22 tx/day</td>
<td>193,773.08 tx/day</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>5,222.57 nodes</td>
<td>1,265.37 nodes</td>
</tr>
</tbody>
</table>

## Replace-by-fee Patch

<table>
<thead>
<tr>
<th>Metric</th>
<th>No change</th>
<th>Replace-by-fee patch committed (13.6% likelihood)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price</td>
<td>$533.00</td>
<td>$342.54</td>
</tr>
<tr>
<td>Network hash rate</td>
<td>4,969.68 PH/s</td>
<td>3,132.09 PH/s</td>
</tr>
<tr>
<td>Daily transaction volume</td>
<td>192,603.80 tx/day</td>
<td>168,406.73 tx/day</td>
</tr>
<tr>
<td>Number of nodes</td>
<td>1,720.41 nodes</td>
<td>3,500.76 nodes</td>
</tr>
</tbody>
</table>
Academia

Today:
• Main function: credential people as *impressive*
  • status via association to: patrons, students, media.
• Each discipline self-evaluates, defends turf
• Consensus, intellectual progress are side effects

Proposal:
• Patrons fund markets on *Questions That Matter™*
  • Settle: directly measurable or long term consensus
• Research labs: hire people, invest, win via trade
• *Can Also:* study replicate, pubs. if hire, cites if publish, GPA if admit
Info Accounting Proposal – Buy Side

**Charge Question Posers By**

- Estimate $ value of info
  - E.g., demand variance
- Match $ subsidy level
  - E.g., liquidity in LMSR
- Can vary by price, time
- Charge subsidy used
  - Plus admin. overhead
Info Accounting Proposal – Sell Side

Credit Competitors By

• Use a new “color of money”
• Track actual subsidy gained
• If stats show consistent gain,
  • Minus time taken to compete
• Separate accounts to track
  • Bias correction procedures
  • Official stat use, trade first?
• Score budget gatekeepers too
Three Premises and a Conclusion

1. It is not too hard to tell rich happy nations from poor miserable ones long after the fact.
2. Governments largely fail by not aggregating available information.
3. Speculative markets are our best known institution for aggregating information.

∴ Try to vote on values, but bet on beliefs.
Our Democracy

Passed Bills => Laws
Vote On Values  \hspace{20mm} But \hspace{20mm} Bet On Beliefs

\[ E[\text{National Welfare} \mid \text{Alternative}] > ? \]
\[ E[\text{National Welfare} \mid \text{Status Quo}] \]
Futarchy’s Main Rule

*If valid market estimates National Welfare higher given proposed policy, it becomes law.*

- Unless market on future-defined NW vetoes
- Start with existing policies, fee/reward for proposals
- Each proposal says:
  - short pre-set time window to approve
  - how handle conflicts with other laws
- High standards at base, recurse to relax standards
- Treaties could agree to include other nation’s NW.

Concerns

*Intuitive, Common*
- Rich get more “votes”
- Buy policy via market
- Market bubbles
- More lies, secrecy
- Do harm to win bets
- Forgo popular illusions

*Regarding Natl. Welfare*
- Policy in NW defn.
- Max NW meas. errors
- Slow catch NW bugs
- Corrupt NW measure
  - More when more care
More Concerns

**Same As Democracy**
- Guarantee rights
- Time-inconsistency
- Induce policy cycling
- Expressive voting
- Too many policies
- Military needs secrecy
- Losers fight transition

**Economics, Technical**
- Markets too thin
- Most policy too small
- When “clear” diff.
- What if destroy Earth
- Infinity never comes
- Risk premia distortion
- Decision selection bias