"Thinking"
&
"Superforecasting"

Munir A. Sheikh

December 2016

Two impressive books on rationality and forecasting

"Thinking, Fast and Slow" by *Daniel Kahneman*

And

"Superforecasting: The Art and Science of Prediction" by *Philip Tetlock and Dan Gardner*

Some Experiment Results

1: Treatment for lung cancer

- **Participants**: Professors at Harvard Medical School, split into two groups, 1 and 2
- **Experiment**: one-month survival rates for surgery
- **Evidence provided**:
  - Group 1: "one-month survival rate is 90%"
  - Group 2: "10% mortality rate in the first month"
- **Results:** 84% of Group 1, 50% of Group 2 chose surgery

2. True value of a house on sale

- **Participants**: Professional real estate agents, split into two groups, 1 and 2
- **Experiment:** Put a price on the value of a house already listed for sale
- **Evidence provided:**
  - **Group 1:** all information in the listing but asking price artificially increased significantly
  - **Group 2:** all information in the listing but asking price reduced significantly
- **Result:** The first group's average price was 41% higher than the other group: priming effect

3. Lottery

- **Participants:** Economics Nobel laureates Arrow, Friedman, Samuelson and renowned statistician Savage

- **Experimenter:** Maurice Allais (another Economics Nobel Laureate)

- **Experiment:** Which lottery would you prefer?

- **Information provided:** Choice within two options
  - Option A: "61% chance to win $520,000 (A1: value $317,200) OR 63% chance to win $500,000 (A2: value $315,000)"
  - Option B: "98% chance to win $520,000 (B1: value $509,000) OR 100% chance to win $500,000 (B2:value $500,000)"

- **Result:** Group selected A1 and B2
  - Obvious preference for certainty: *Contradicts utility theory and rational expectations*

4. Cab driver behaviour

- **Participants:** New York cab drivers
- **Experiment:** Hours driving taxi on pleasant days and rainy days
- **Information:** fares are higher on rainy days and lower on pleasant days in response to changing demand
- **Result:** Cab drivers drive more hours on pleasant days and less on rainy days, contradicting basic economic result of positive supply response to prices

5. Undergoing colonoscopy

- **Participants:** colonoscopy patients
- **Experiment:** which procedure do they prefer? Two groups, 1 and 2.
- **Options:**
  - Group 1: Option 1 which takes longer, has variable pain, with cumulative pain of X1, and ends with a pain level of Y1
  - Group 2: Option 2 which is shorter, has variable pain, with cumulative pain of X2, and ends with a pain level Y2
  - X1 is significantly more than X2 but Y1 is significantly less than Y2
- **Result**: Option 1 is preferred by a wide margin; only the peak is remembered, not cumulative pain—problem with life satisfaction questions

6. Another lottery

- **Participant:** A friend of Paul Samuelson
- **Experiment:** Reaction to a biased lottery
- **Experimenter:** Paul Samuelson
- **Options:**
  - 1. A single coin flip, with a head call prize of $200 and a tail call loss of $100
  - 2. 100 similar coin flips
- **Result:** Option 1 was was rejected, Option 2 accepted:
  - Option 1 response irrational; Option 2 response inconsistent with Option 1 response, and irrational

7. Accident eyewitness

- **Participants:** Participants reviewing a court case following an accident
- **Experiment:** How credible is the lone witness to an accident that involved a taxi?
- **Situation:** The only witness to a taxi involved in an accident says the taxi was Blue, with two types of taxis in town, with the other being Green
- **Information provided:** 85% of the cabs in the city are Green and 15% are Blue; Got a psychologist to test the witness' credibility. Result: can trust him 80% of the time.
- **Response:** The Blue taxi driver is likely 80% guilty.
  - Probability of it being a Blue taxi is only 41%-- completely ignored.

8. Shopping

- **Participants**: shoppers wanting to buy a dinnerware set; split into 3 groups 1, 2 and 3
- **Experiment:** how much would they pay for a dinnerware set?
- **Options:**
  - Group 1 shown a set with 40 pieces, 9 broken
  - Group B shown a set with 24 pieces, none broken, including all those from Option 1
  - Group 3 shown both sets
- **Average Prices Offered:**
  - Group 1: $23 (because of broken pieces)
  - Group 2: $33 (all pieces in good shape)
  - Group 3: $32 for Set A and $30 for set B (joint evaluation).

Conclusion

- Let us be clear about what rationality means in economics:
  - "The only test of rationality is not whether a person's beliefs and preferences are reasonable, but **whether they are internally consistent**." (*Kahneman)*
- Richard Thaler's two species: Econs and Humans. Econs believe in rationality.
- Examples show Humans are "irrational", including those noble laureats (unknowingly), who specialize in and are the strongest proponents of rationality
  - "Intuitive statistics are hard for Humans"

An Irony

- Paul Romer (*New York University)* wrote the following about some big names in the rational expectations camp:

  "Bob Lucas, Ed Prescott, and Tom Sargent…. Prior to 1980, they made important scientific contributions to macroeconomic theory… when these contributions provoked return fire that could be sarcastic, dismissive, and wrong- headed….they developed a bond of loyalty that would be admirable and productive in many social contexts. Two examples illustrate the bias that loyalty can introduce into science (comment: System 2 totally at the mercy of System 1). "

Why are Humans not rational (in the Econ sense)?

The Two-system Framework

- Modern day psychologists use a framework of human mind that has two distinct parts, which are labeled System 1 and System 2 (*Following Stanovich and West, and used by Kaheneman and Tetlock*).

- *System 1* operates automatically and quickly, with little or no effort and no sense of voluntary control.

- *System 2* allocates attention to the effortful mental activities that demand it, including complex computations.

System1 and System 2 Working Together

- System 1 works continuously and quickly produces impressions, intuitions, intentions and feelings and shares them with System 2.

- System 2 reviews these and, if it agrees, results in voluntary actions.

- The two systems together generally work fine.

- Most of what we do originates in System 1 but System 2 takes over when things get difficult.

- The division of labour between the two is highly efficient and minimizes effort and optimizes performance.

System 1

- System 1 "jumps to conclusions" and is affected by many things such as:
  - The desire to be *coherent* (easier to do this the less information there is)
  - *Priming* (e.g. picture of staring eyes hanging over a coffee machine, speed monitors)
  - *Cognitive Ease* (e.g. size of printed letters, stock symbols)
  - *Mood* affects System 1 solutions
  - *The halo effect* (e.g. word sequencing)
  - *"What you see is all there is (WYSIATI)":* limited evidence as the basis for quick decisions
  - *Framing (*e.g. 90/10 % chances of surviving surgery)
  - *Question substitution* (unknowingly substitute a difficult question with an easy one and answer it): life satisfaction determined by current mood
  - *Anchoring (*e.g. value of a house on sale)
  - *Endowment effect(* e.g. sale price much higher than what you would pay for a commodity you own)
  - *Mental accounts (e.g. lost movie tickets, ignorance of sunk costs)*
  - *Talent and skill (*these experiences and abilities reside in System 1 memories and can produce better quick decisions, e.g. sixth sense)
  - *Ignores duration effects (e.g. pain of colonoscopy) and has biased memories*
  - *Ignores base rate information (e.g. taxi accident)*
  - *Suffers from planning fallacy (i.e. best case scenarios)*

System 2

- System 2 is the thoughtful and logical part and has the following characteristics:
  - It is lazy
  - Gets information (e.g. memories) stored in System 1 to act
  - As a default, wants to endorse or rationalize ideas and feelings generated by System 1
  - But can act independently when it finds System 1 has reached a foolish conclusion, but that requires an investment of attention from System 2, which reduces the body energy level (like any physical activity), putting a constraint on how much it can do
  - It is not a paragon of rationality as its abilities and access to knowledge are limited
  - It has no ability to distinguish between a skilled or a heuristic response of System 1.

Systems 1 and 2 interactions

- System 2 often endorses System 1 views. Finds excuses to justify initial System 1 views.

- Sometimes it overrides these. But, you feel internal discomfort because of the inherent conflict between the two systems.

- Chances of System 1 going off the rails higher when System 2 is busy.

Some Consequences

Example: Wrong Behaviour

- When someone is especially good to you, you are nice to him.

  – Next time, he is not so good.

- When someone is bad to you, you are nasty.

  – Next time he is not so bad.

- Conclusion: Being nasty works, being nice does not.

- This is System 1 at work. Ignored "mean regression".

Wrong Choice for School Support

- Gates Foundation: $1.7 billion support for smaller school size.

- It was based on a survey of 1,662 schools in Pennsylvania that showed small schools were better by a factor of 4.

- Easy to construct a credible story, which System 1 does.

- Same survey showed, when examined further, that worst schools were also small schools. Problem of small samples and small areas.

Costly Outcomes

- Behaviour, based on System 1's aversion to risk, can be very costly.
- Examples include:
  - Insurance policies with low or no deductibles;
  - Expensive warranties.
- Exaggerated optimism, again a System 1 bias, has costs as well.
  - Examples include: 90% of drivers claiming they are better-than-average; sum of probabilities of basketball teams winning the championship at 240%.
  - Worse, new small businesses' expectation of their survival is near 100%, when only 30% survive.

Wrong Conclusion: Unhappy Marriage (before and after the marriage)

Life Satisfaction

Some Policy Implications

An Additional Reason for Public Policy

- Economics suggests two reasons for a policy intervention:
  - When markets fail;
  - Equity objectives.
- Behavioural economics, based on "irrational" Human beviour suggests another reason for intervention:
  - "Humans need help (Kahneman)".
  - But policy makers are Humans too!

Policy Examples

- Gallons per mile rather than miles per gallon
- Regulations to make contracts simple and use the "right" language
- Opt-out clause for organ donation:
  - Organ donation rates with opt-out: 100% in Austria; 86% in Sweden. Only 12% in Germany and 4% in Denmark, with opt-in.
- Opt-out policies for savings under US plan "Save More Tomorrow", a proposal accepted by Congress from Thaler and Benartzi.
- Nudge Unit in the UK headed by Thaler
- Thaler's coauthor, Sunstein, appointed by President Obama in the Office of Information and Regulatory Affairs
- Starting wars, high degree of litigation because of over-optimism, small business support

Forecasting and Superforecasting

Kahneman's View

- "Intelligence and knowledge would improve forecasting but the benefits would taper off fast. People armed with Ph.Ds and decades of experience may be only a tad more accurate than the attentive readers of *New York Times.*"

- This view is grounded in his description of System 1 with all its biases, the laziness of System 2, its tendency to endorse System 1 views, and the difficulty to forecast beyond a very short term because of uncertainty.

Klein's Research

- Research on decision-making by professionals
- Example of a commander of firefighters, who ordered his men out of a burning house, not knowing why, and the house collapsed soon after
- He just knew—"the hallmark of an intuitive judgment"
- Klein-Kahneman collaboratory research

Background to "Superforecasting"

- Tetlock conducted a forecasting tournament from 1984-2004 that concluded that the average expert is "roughly as accurate as a dart-throwing chimpanzee" in forecasting.

- He dug deeper into the issue in the Good Judgments Project sponsored by US intelligence agencies, with thousands of volunteers—all ordinary people-- submitting many thousands of forecasts.

- He began as a skeptic but became an optimistic skeptic, given his research results and in the context of views expressed by Kaheneman and Klein.

- His conclusions flow from Kahneman's two-system model, adjusted for the insight he got from his two projects.

What Tetlock found

- "How predictable something is depends on what we are trying to predict, how far into the future, and under what circumstances".

- While forecasting is critical to all of us, nobody keeps scores of how good or bad forecasters are. As Bill Gates commented: "You can achieve incredible progress if you set a clear goal and find a measure that would drive progress towards this goal….This may seem basic, but it is amazing how often it is not done…".

Superforcasters

- Some people are much better than others in forecasting.
- It is not a gift from heaven but a skill that can be learnt, practiced and improved.
- It is easier to forecast events that are relatively short term, less uncertain and linked to the past with available data.
- This result is not inconsistent with Kahneman's model, but rather fits in with his System 1 augmented with talent and skill and System 2 trained to get out of its laziness and do its job.

Requirements for Superforecasting

- Forecasts must be:
  - Precise, not vague, so they can be tested
  - One must keep scores, so we know how good or bad they are; this is critical to produce the right incentives
  - They must have a clear time frame; generalities cannot be evaluated
  - Considered as probabilistic, not binary outcomes
  - Good forecasting requires humility, use of statistics and other information, ability to synthesize, and a desire for continuous improvement.
  - Teams, properly organized, can improve forecasting.

Characteristics of Superforecasters

- Tetlock list ten commandments for aspiring superforecasters:
  - Triage
  - Break seemingly intractable problems into tractable sub-problems
  - Strike the right balance between inside and outside views
  - Strike the right balance between under- and overreacting to evidence
  - Look for the clashing causal forces at work in each problem
  - Strive to distinguish as many degrees of doubt as the problem permits but no more
  - Strike the right balance between under- and overconfidence, between prudence and decisiveness
  - Look for the errors behind your mistakes but beware of rearview-mirror hindsight biases
  - Bring out the best in others and let others bring out the best in you
  - Master the error-balancing cycle

Tetlock vs Kahneman

- While Tetlock is right that his model is consistent with Kahneman's, there is tension between their views.

- Tetlock calls himself an "optimistic skeptic" who thinks that superforecasting can be taught.

- Kahneman concludes:

  – " The way to block errors…ask for reinforcements from System2. Unfortunately this sensible procedure is least likely…when it is needed most."

- Could the superforecasters be like Wayne Gretzky and Sidney Crosbie, which would make Tetlock/Kahneman views consistent?