14.127 Behavioral economics. Lecture 3

Xavier Gabaix

February 19, 2004

1 Lucas' calculation of the cost of the business cycle

- Let $\tilde{c}_t = c_t + \Delta\tilde{c}_t$ be the optimal random consumption, c_t its deterministic component and the mean of the stochastic component $E\Delta\tilde{c}_t = 0$.

- What is the welfare associated with $\Delta\tilde{c}_t$?

- In EU, the welfare equals

$$V = E\left[\int_0^T e^{-\delta t} u\left(\tilde{c}_t\right) dt\right]$$

- We measure the welfare loss by the fraction λ of consumption that people would accept to give up in order to avoid consumption variability.

- Then, λ solves

$$V = E\left[\int_0^T e^{-\delta t} u\left(c_t + \Delta\tilde{c}_t\right) dt\right] = \int_0^T e^{-\delta t} u\left(\left(1-\lambda\right) c_t\right) dt = B\left(\lambda\right)$$

- We know that $B\left(0\right) > V$. To solve the above equation let's do the Taylor expansion for small $\Delta\tilde{c}_t$

$$\begin{aligned} V &= E\int e^{-\delta t}\left(u\left(c_t\right) + u'\left(c_t\right)\Delta\tilde{c}_t + \frac{1}{2}u''\left(c_t\right)\Delta\tilde{c}_t^2 + o\left(\Delta\tilde{c}_t^2\right)\right) dt \\ &= B\left(0\right) + 0 + \int e^{-\delta t}\left(\frac{1}{2}u''\left(c_t\right) E\Delta\tilde{c}_t^2 + o\left(\Delta\tilde{c}_t^2\right)\right) dt \end{aligned}$$

- On the other hand

$$\begin{aligned} B(\lambda) &= \int e^{-\delta t} u((1-\lambda) c_t)\, dt \\ &= \int e^{-\delta t} u(c_t)\, dt + \left(-\int e^{-\delta t} u'(c_t)\, c_t dt \right) \lambda + o(\lambda) \\ &= B(0) - \lambda \int e^{-\delta t} u'(c_t)\, c_t dt + o(\lambda) \end{aligned}$$

- Thus,

$$\lambda \simeq -\frac{1}{2} \frac{\int e^{-\delta t} u''(c_t)\, E\Delta \tilde{c}_t^2 dt}{\int e^{-\delta t} u'(c_t)\, c_t dt}$$

- To calibrate u, take,

$$u(c_t) = \frac{c^{1-\gamma}}{1-\gamma}$$

- Then, $cu'' = -\gamma u'$.

- Define

$$m_t = e^{-\delta t} u'(c_t)\, c_t$$

- Thus

$$\lambda = \frac{\gamma}{2} \frac{\int m_t E\left[\left(\frac{\Delta \tilde{c}_t}{c_t}\right)^2\right] dt}{\int m_t dt}$$

- Denoting the above fraction of integrals by $\langle ., . \rangle$ (note that it is a mean with weights m_t) we write it as

$$\lambda = \frac{\gamma}{2} \left\langle E\left(\frac{\Delta \tilde{c}_t}{c_t}\right)^2 \right\rangle$$

- Suppose, $\Delta\tilde{c}_t^2 = c_t \sigma \varepsilon_t$ where ε_t are iid with variance 1. In practice $\sigma = 2\%$, thus

$$\lambda = \frac{\gamma}{2}\sigma^2 = \gamma\frac{4}{2}10^{-4} = 2\gamma \cdot 10^{-4}$$

- If $\gamma = 1$ then $\lambda \simeq 0.02\%$ for EU consumers.

- PT consumers value stability more as they are first order risk averse around their reference point.

- But their risk aversion strongly depends on horizon. Should it be yearly, monthly, daily?

- We need to have a theory of the horizon to give a proper alternative to Lucas.

2 Heuristics and the rules of thumb

- Judgment heuristic: an informal algorithm which generates an approximate answer to a problem.
- Rules of thumb are basically special cases of heuristics.
- Heuristics speed up cognition.
- Heuristics occasionally produce incorrect answers.
- The errors are known as “bias.”
- These are the unintended side effects of generally adaptive processes.

Examples:

- shade your bid in an auction for an oil parcel by 50%
- judge the distance of an object by its clarity
- judge the distance of a person by her size
- save 10% of your income for retirement
- invest (100-age)% of your wealth in stocks

- never borrow on credit cards
- leave a three second interval between you and the car in front of you

- Cognitive psychology studies the representation and processing of information by complex organisms.

- Kahneman and Tversky are two of the leaders in this field.

- They identified three important judgment heuristics in a series of path-breaking contributions in the early 1970's

- representativeness, availability, anchoring

3 People don't do Bayes' rule

- 1 in 100 people in the world have a disease.
- We have a test for it.
- If someone has the disease, she has a 99% chance of testing positive.
- If someone doesn't have the disease, she has a 99% chance of testing negative.
- Linda took the test, and tested positive.

- Assuming that Linda was drawn randomly from the population, what is the probability that she has the disease?

4 Apply Bayes rule:

- If $D =$ "has the disease" and $N =$ "doesn't have the diseases", $T+ =$"the test is positive", then,

$$\begin{aligned}
\Pr(D|T+) &= \frac{\Pr(D, T+)}{\Pr(T+)} = \frac{\Pr(T+|D)\Pr(D)}{\Pr(T+)} \\
&= \frac{\Pr(T+|D)\Pr(D)}{\Pr(T+|D)\Pr(D) + \Pr(T+|N)\Pr(N)} \\
&= \frac{(.99)(.01)}{(.99)(.01) + (.01)(.99)} = \frac{1}{2}
\end{aligned}$$

- In practice, the Bayes rate $P(D)$ is rarely used.

5 Representativeness

- decision makers use similarity or representativeness as a proxy for probabilistic thinking, e.g.,

- "Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure and a passion for detail."

- What is the probability that Steve is a farmer, salesman, airline pilot, librarian, or physician?

- How similar is Steven to a farmer, salesman, airline pilot, librarian, or physician?

- Subject rankings of probability and similarity turn out to be the same.

- OK, if similarity predicts true probability.

Why might similarity poorly predict true probability?

Consider the following example:

"Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations."

Please rank the following statements by their probability, using 1 for the most probably and 8 for the least probable.

1. Linda is a teacher in elementary school.
2. Linda works in a bookstore and takes Yoga classes.
3. Linda is active in the feminist movement.
4. Linda is a psychiatric social worker.
5. Linda is a member of the League of Women Voters.
6. Linda is a bank teller.

7. Linda is an insurance salesperson.

8. Linda is a bank teller and is active in the feminist movement.

1. (5.2) Linda is a teacher in elementary school.
2. (3.3) Linda works in a bookstore and takes Yoga classes.
3. (2.1) Linda is active in the feminist movement.
4. (3.1) Linda is a psychiatric social worker.
5. (5.4) Linda is a member of the League of Women Voters.
6. (6.2) Linda is a bank teller.
7. (6.4) Linda is an insurance salesperson.
8. (4.1) Linda is a bank teller and is active in the feminist movement.

- depending on the subject population, 80%-90% rank item 8 as more likely than item 6.

- K&T call this the conjunction effect (since the conjunctive event receives a HIGHER probability)

- done with naive subjects (undergrads from UBC and Stanford with no background in probability or statistics)

- done with intermediate subjects (graduate students in psychology, education and medicine from Stanford, who had taken several courses in probability and statistics)

- done with sophisticated subjects (graduate students in the decision science program of the Stanford Business School who had taken several advanced courses in probability and statistics)

- results are nearly identical for these three groups

- also similarity ranks perfectly coincide with probability ranks

Potential confound:

- Maybe “Linda is a bank teller,” is interpreted as “Linda is a bank teller and is NOT active in the feminist movement.”

- Response: run a between-subject design (in contrast to the within-subject design described above)

- Specifically, show some subjects (group A) the list without the conjunctive event (item 8).

- Show other subjects (group B) the list without the critical non- conjunctive events (items 3 and 6).

- Group B ranks “8” higher than Group A ranks “6”

Another experiment (conjunction effect: 68%)

Please rank the following events by their probability of occurrence in 1981.

1. (1.5) Reagan will cut federal support to local government.

2. (3.3) Reagan will provide federal support for unwed mothers.

3. (2.7) Reagan will increase the defense budget by less than 5%.

4. (2.9) Reagan will provide federal support for unwed mothers and cut federal support to local governments.

Another experiment (conjunction effect: 72%)

Suppose Bjorn Borg reaches the Wimbledon finals in 1981. Please rank order the following outcomes from most to least likely.

1. (1.7) Borg will win the match.

2. (2.7) Borg will lose the first set.

3. (3.5) Borg will win the first set but lose the match.

4. (2.2) Borg will lose the first set but win the match.

Bottom line: similarity is sometimes a poor predictor of true probability.

- Probability follows the conjunction rule: $P(A \cap B) \leq P(B)$.

- The probability that Linda is a feminist bank teller (feminist $\cap$ bank teller $= A \cap B$) is less than the probability that Linda is a bank teller (bank teller $= B$).

- Similarity relations do not follow the conjunction rule.

- E.g., similarity between a blue square and a blue circle (blue $\cap$ circle $= A \cap B$) is *greater* than the similarity between a blue square and a circle (circle $= B$).

6 Applications of representativeness:

- insensitivity to prior probabilities of outcomes
- insensitivity to sample size
- misconceptions of chance
- insensitivity to predictability
- the illusion of validity?
- misconceptions of regression

6.1 Insensitivity to base rates

- Problem 1:
  - Jack's been drawn from a population which is 30% engineers and 70% lawyers.
  - Jack wears a pocket protector.
  - What is the probability Jack is an engineer?

- Problem 2:
  - Jack's been drawn from a population which is 30% lawyers and 70% engineers.
  - Jack wears a pocket protector.
  - What is the probability Jack is an engineer?

- We will denote Problem 1 probability by p_1 and Problem 2 probability by p_2.

- If $E =$ "Engineer," $W =$ "wears a pocket protector", and $Gi =$ "Problem i", then, the Bayes law says

$$p_i = \Pr(E|W \& Gi) = \frac{\Pr(W, E|Gi)}{\Pr(W|Gi)} = \frac{\Pr(E|Gi)\Pr(W|E \& Gi)}{\Pr(W|Gi)}$$

and

$$\begin{aligned}&\Pr(W|Gi)\\&= \Pr(E|Gi)\Pr(W|E \& Gi) + (1 - \Pr(E|Gi))\Pr(W|(\text{not}E) \& Gi)\end{aligned}$$

for $i = 1, 2$.

- Using the above mentioned Bayes laws

$$
\begin{aligned}
\frac{p_i}{1-p_i} &= \frac{\frac{\mathsf{Pr}(E|Gi)\,\mathsf{Pr}(W|E\&Gi)}{\mathsf{Pr}(W|Gi)}}{1-\frac{\mathsf{Pr}(E|Gi)\,\mathsf{Pr}(W|E\&Gi)}{\mathsf{Pr}(W|Gi)}} \\
&= \frac{\mathsf{Pr}\,(E|Gi)\,\mathsf{Pr}\,(W|E\&Gi)}{\mathsf{Pr}\,(W|Gi)-\mathsf{Pr}\,(E|G1)\,\mathsf{Pr}\,(W|E\&Gi)} \\
&= \frac{\mathsf{Pr}\,(E|Gi)\,\mathsf{Pr}\,(W|E\&Gi)}{(1-\mathsf{Pr}\,(E|Gi))\,\mathsf{Pr}\,(W|\,(\mathsf{not}E)\,\&Gi)}.
\end{aligned}
$$

- We implictly assume that conditional probabilities of wearing a pocket protector are the same in both problems. In an actual experiment we need to make this explicit, and embed it in the story we tell.

- The above assumption means that

$$\text{Pr}\left(W|E\&G1\right) = \text{Pr}\left(W|E\&G2\right) = \text{Pr}\left(W|E\right)$$

  and

$$\text{Pr}\left(W|\left(\text{not}E\right)\&G1\right) = \text{Pr}\left(W|\left(\text{not}E\right)\&G2\right) = \text{Pr}\left(W|\text{not}E\right).$$

- This allows us to compute simplify the ratio we computed and arrive at

$$\frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}} = \frac{\frac{\Pr(E|G1)\Pr(w|E)}{(1-\Pr(E|G1))\Pr(W|\text{not}E)}}{\frac{\Pr(E|G2)\Pr(W|E)}{(1-\Pr(E|G2))\Pr(W|\text{not}E)}}$$

$$= \frac{\frac{\Pr(E|G1)}{1-\Pr(E|G1)}}{\frac{\Pr(E|G2)}{(1-\Pr(E|G2))}} = \frac{\frac{.3}{1-.3}}{\frac{.7}{1-.7}} = \left(\frac{3}{7}\right)^2$$

- But, in the lab:

$$\frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}} \approx 1$$

- What happens when we give the subjects no information other than base rates?

- What happens when we change the description to something uninformative like, “Jack went to college.”

6.2 Insensitivity to sample size

- subjects assess the likelihood of a sample result by asking how similar that sample result is to the properties of the population from which the sample was drawn

- A certain town is served by two hospitals. In larger hospital, 45 babies born per day. In smaller hospital, 15 babies born per day. 50% of babies are boys, but the exact percentage varies from day to day. For a period of 1 year, each hospital recorded the days on which more than 60 percent of the babies born were boys. Which hospital do you think recorded more such days?

- The large hospital?
- The small hospital?
- About the same (within 5% of each other)

6.3 Misconceptions of chance (the law of small numbers)

- people expect that a sequence of events generated by a random process will represent the essential characteristics of that process even when the sequence is short

- so if a coin is fair, subjects expect HHH to be followed by a T (gambler's fallacy)

- if girls are as likely as boys, subjects expect GGG to be followed by B

- so $BGGBBG$ is viewed as a much more likely sequence than $BBBBBB$

- people expect that the essential characteristics of the process will be represented, not only globally in the entire sequence, but also locally in each of its parts

- even scientists make this mistake, overpredicting the likelihood that small sample results will replicate on larger samples

- All families of six children in a city were surveyed. In 72 families the exact order of births of boys and girls was $GBGBBG$.

- What is your estimate of the number of families surveyed in which the exact order of births was $BGBBBB$?

In standard subject pools $\approx$ 20% get it right and the median estimate is 30.

6.4 Insensitivity to predictability

- predictions are often made by representativeness

- if a company is described favorably (e.g., lots of profitable new products) we predict outcomes that are similar (e.g., high future stock returns)

- the predictions are unaffected by the reliability of the information and the predictability of the outcomes

Subjects presented with several paragraphs describing the performance of a student teacher during a single practice lesson. Subjects were asked to evaluate the quality of lesson (in percentile scores). Other subjects were asked to predict, also in percentile scores, the standing of each student teacher 5 years after the practice session. The judgments were identical.

6.5 Misconceptions of regression to the mean

- extreme outliers tend to regress toward the mean in subsequent trials (e.g., best performers on the midterm, fighter pilots with the best landings, tall fathers)

- but intuitively, we expect subsequent trials to be representative of the previous trial, so we fail to anticipate regression to the mean

7 Availability

People assess the frequency of a class or the probability of an event by the ease with which instances or occurrences can be brought to mind

- What percentage of commercial flights crash per year?

- What percentage of American households have less than $1,000 in net financial assets, including savings accounts, checking accounts, CD's, stocks, bonds, etc... (but not counting their most recent paycheck or their defined benefit and defined contribution pension assets)?

- What is the population of greater Boston (USA)?

- What is the population of greater Osaka (Japan)?

Examples:

a class whose instances are easily retrieved will appear more numerous than a class of equal frequency whose instances are less retrievable

- “Does this list contain more names of men or women?”

- When the list contains male names that are slightly more famous than the female names, subjects conclude that the list is disproportionately male.

- When the list contains female names that are slightly more famous than the male names, subjects conclude that the list is disproportionately female.

- Subjects erroneously conclude that the class (sex) that had the more famous personalities was the more numerous.

What makes something salient, and hence retrievable?

- familiar (Harrison Ford vs. Geraldine Page)
- important (death of a parent in a car accident vs. car accident reported on evening news)
- personal (uncle Bob's story about his Volvo vs. statistical report on Volvos)
- recent (yesterday's "close call" vs. stale "close call")

7.1 Biases due to the effectiveness of a search set:

Suppose one samples words (3+ letters) at random from English texts.

- Is it more likely that the words will begin with an r or have an r as the third letter?

- What is the probability that a seven-letter word would end in "ing"?

- What is the probability that a seven-letter word would have "n" as its sixth letter?

7.2 Biases of imaginability:

- Suppose you had 10 people who you had to organize into committees of k members.

- How many different committees of k members can you form?

- What if $k = 2$?

- What if $k = 8$?

Claim: Suppose you have N objects in a set and you want to choose subsets of size k. The number of subsets of size k is equal to the number of subsets of size $N-k$.

Proof: For every subset of size k you can form a subset of size $N-k$ made up of the objects excluded from the original k-subset.

Suggestive evidence for availability effects in the real world

- people with older siblings prepare more for retirement
- advertising

How should you overcome availability effects (when making important decisions)?

- enumerate extensive lists of possible outcomes
- simulate to identify outcomes that you hadn't imagined (when the door is opened it hits the rear-view mirror)
- understand the ways in which your memory base is biased; these biases will effect your probability judgments (e.g., you under-rehearse unpleasant memories, leading to biased availability-based inferences)

8 Anchoring

Anchors seem to matter:

E.g., starting points, frames, defaults, etc....

- Is the Mississippi River more or less than 70 miles long? How long is it?

- Is the Mississippi River more or less than 2000 miles long? How long is it?

Three hypotheses:

- People make estimates by starting from an initial value that is adjusted to yield the final answer; typically, these adjustments are insufficient potentially because the adjustment is stopped when the answer becomes sufficiently close to the correct answer (Slovic and Lichtenstein; Kahneman and Tversky; Quattrone et al; Wilson et al; see too the literature on *satisficing*).

- Subjects take the question as a hint from the experimenter (Kahneman and Tversky; Schwarz).

- Subjects subconsciously recruit memories consistent with the anchor (Strack; Wilson and Brekke; Gilbert).

Kahneman and Tversky's first anchoring experiment:

- subjects were asked to estimate the percentage of African countries in the UN
- first spin Wheel of Fortune $\rightarrow$ random number
- guess whether % African $>$ random number
- then guess % African
- spin $= 10 \rightarrow$ % African $= 25$
- spin $= 60 \rightarrow$ % African $= 45$

- Where does anchoring bias matter?

- E.g. in conducting surveys

Similar experiment run in *last* year's class. I generated a "random number" by transforming subject birthdates into a number between 0 and 100.

- What day of the month were you born?
- Call this number x.
- $y = 3x$.
- Let $z =$ % African countries in UN
- Is $y > z$? Yes or no?
- What is the value of z?

- Two groups: 2 minute group and 6 minute group

- For each group run regression:

$$z = \alpha + \beta y + \varepsilon$$

- If subjects were well-calibrated:

$$\alpha = \frac{53}{189} = 28$$

$$\beta = 0$$

- 2 min.: $\alpha = 17$ ($t = 3.9$), $\beta = .131$ ($t = 1.7$)

- 6 min.: $\alpha = 27$ ($t = 5.6$), $\beta = -.011$ ($t = 0.1$)

- Why does the 6 minute group show no effect?

But other experiments of this type do yield strong effects:

- wheel of chance (Tversky and Kahneman)
- randomly chosen card (Cervone and Peake)
- experiment number (Switzer and Sniezek)
- social security number (Wilson, Houston, Etling, and Brekke)

Other types of anchoring experiments have been run by Jacowitz and Kahneman:

- Is the Mississippi River more or less than 70 miles long? How long is it?

- Is the Mississippi River more or less than 2000 miles long? How long is it?

Mississippi (mi)	70	2000	300	1500
Everest (ft)	2000	45500	8000	42550
Meat (lbs/year)	50	1000	100	500
SF to NY (mi)	1500	6000	2600	4000
Tallest Redwood (ft)	65	550	100	400
UN Members	14	127	26	100
Female Berkeley Profs	25	130	50	95
Chicago Population (mil.)	0.2	5.0	0.6	5.05
Telephone Invented	1850	1920	1870	1900
US Babies Born (per day)	100	50000	1000	40000

My feelings: Anchoring effects are strongest when anchors have implicit information value and when subjects don't have much time to think about the problem.

8.1 Biases due to insufficient adjustment

- Within 5 seconds, estimate the product:

- $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$

- $1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$

- First sequence median guess: 2250.

- Second sequence median guess: 512.

- Correct answer: 40,320.

8.2 Biases in the evaluation of conjunctive and disjunctive events

- People tend to overestimate the probability of conjunctive events (e.g., all 15 parts of my thesis will proceed as planned, enabling me to hand it in on time).

- People tend to underestimate the probability of disjunctive events (at least one part of my thesis will go horribly wrong, leading to a terrible sequence of all-nighters).

8.3 Anchoring in the assessment of subjective probability distributions (overconfidence)

How many people live in...

- Ankara, Guiyang, Changsha, Rochester, Sheffield, Frankfurt, Sao Paulo, Mogadischo, Greensboro, Ottawa

- these cities were picked randomly from a list of the world's 380 largest cities

Write down your 98% confidence intervals for each city

X such that there is a 99% chance that the true value is greater than X

Y such that there is a 99% chance that the true value is less than Y

Answers:

Ankara (3.4), Guiyang (1.5), Changsha (1.8), Rochester (1.1), Sheffield (1.3), Frankfurt (2.0), Sao Paulo (18.4), Mogadischo (1.2), Greensboro (1.2), Ottawa (1.1)

Subjects generate 99% confidence intervals which only contain 80% of the true answers.

What should you do to avoid (exploit) anchoring bias?

- Initiate negotiations with your own anchor.

- When trying to elicit information, don't anchor respondents answers if you want unbiased responses.

- When trying to elicit information, anchor respondents answers if you want biased responses.

- Almost always blow up your estimates of uncertainty.