SUGGESTED ANSWERS for

Final exam, M.Sc. in Economics summer school 2020

Behavioral and Experimental Economics

Question 1: Methodological aspects

Following the notation in Falk and Heckman (Science, 2008), consider an outcome of interest Y that depends on a list of determinants, i.e., $Y = f(X_1, \ldots, X_N)$, where $Z = (X_2, \ldots, X_N)$. Some of these determinants (X_U) may be difficult to observe or control. Answer the questions below with reference to this notation.

a) Provide an example to explain Y, X_1 and Z in a specific experiment. Provide examples of determinants that are easy and at least one that is difficult to observe or control (X_U). (*Hint*: refer to a simple experiment, e.g., in an individual work context)

A: An example (discussed in Falk and Heckman 2008) is a "real effort" experiment in which subjects are asked to perform a simple task (e.g., calculating cross-sums) and are paid a piece rate for each correct answer. The researcher may want to study how output (Y) responds to the piece rate (X_1). Output will depend on other factors Z like the nature of the task (how boring or difficult it is), the subject's skill (whether the subject is able to work with numbers), the technology at disposition (e.g., whether it is a computerized or paper-and-pen experiment), the time available to work on the task and other factors (e.g., room temperature, whether the subject slept well the day before). The wage rate, technology and room temperature are easy to observe and control while X_U refers to variables like skill or whether the subject slept well.

b) How can the causal effect $\Delta Y/\Delta X_I$ be measured in your example? How does the researcher deal with X_U? Will the observed effect still (despite the researcher following best practice) depend on X_U?

A: The causal effect is isolated by ceteris paribus variation of X_I (e.g., by doubling the piece rate) in treatment t and by holding it constant in the other treatment (c, the control) and by observing the difference $\Delta Y_t - \Delta Y_c$. If subjects are randomized across treatments and if the sample is sufficiently large, the distributions of X_U will be the same in the two treatments. The researcher will observe $\Delta Y/\Delta X_I$ at a given (perhaps unknown) level of X_U. That is, randomization will "neutralize" the effect of the unobserved variable in the sense that it has the same effect in both treatments.

c) How do experimental economists make sure the experiment can be "replicated" by other researchers in practice? (*Hint*: explain how "control" can be exerted in your example above and refer to the notation above).

A: Replication is aided by exerting as much control as possible and reporting how this was done in detail. In particular, by reporting the parameters (e.g., the piece rate), procedures (e.g., subjects were not allowed to talk to each other), the software and experimental instructions, and by providing an exact protocol (e.g., how exactly the experiment was conducted, what the subjects were told when). (This is a simple experiment. In more complex environments, other factors like institutions ("the rules of the game") and endowments also need to be controlled).

d) In what ways will the presence of X_U limit ideal replication?

A: A researcher trying to replicate an earlier experiment will run his or her experiment at a different time and location, recruiting different subjects. Hence, some of the levels of the unobserved determinants are likely to be different (be at level X_U' rather than X_U). For example, the recruited subjects in the replication may have better training in mathematics and are therefore more productive.

e) Discuss how the fact that subjects volunteer to participate in economic experiments may bias findings on the relevance of other-regarding preferences. (Hint: refer to "selection".)

A: Subjects in standard lab experiments are often students who are recruited with the promise of earning money during the experiment. Thus, people with (unrepresentatively) strong preferences for earning money may select into lab experiments. One may thus expect them to be less other regarding than subjects from the general population (however, the general population may differ also in many other ways from students, e.g., is richer, older, and less educated on average). A robust finding from comparing findings in standard lab experiments with the same experiments conducted on samples representative of the general population seems to be that standard subjects are not more other-regarding (and if a difference is found, less other-regarding) than "normal" people. We briefly discussed studies by Slonim et al. JEBO 2013, Falk, Meier, Zehnder JEEA 2013, Hoffman and Morgan JEBO 2015.

Question 2: Markets

Consider a double auction market experiment with 5 buyers and 5 sellers and the following induced values.

ID number	Value of first unit	Value of second unit
B1	**75**	**65**
B2	**85**	**60**
B3	**95**	**55**
B4	**105**	**50**
B5	**110**	**45**

ID number	Value of first unit	Value of second unit
S1	**55**	**90**
S2	**60**	**85**
S3	**65**	**80**
S4	**70**	**75**
S5	**55**	**95**

a) Calculate equilibrium price, quantity and surplus (i.e., the sum of consumer and producer rent) predicted by competitive market theory in the example above.

A: price is between 70 and 75, quantity is 5 (to find the answer, sort bids from high to low, asks from low to high, find the intersection point which determines which are the infra-marginal units), surplus is 165.

b) What is the typical observation for these variables in double auction markets? (e.g., Smith JPE 1962)

A: These markets converge well to standard predictions, i.e. prices and quantities are in (or close to) equilibrium values, and efficiency tends to be high (typically 95% or more).

c) What are the sufficient conditions according to the Induced Value Theory (Smith, AER 1982) to induce the buyer and seller values in such an auction? (*Hint*: consider preferences $V(m_0 + \Delta m, z)$). Are these conditions also necessary?

A: Monotonicity: V_m exists and is positive for each combination of (m, z) ; Salience: the payment Δm depends on the actions of the subject; Dominance: The changes in utilities of a subject during the experiment depend importantly on Δm. The influence of z can be neglected. The conditions are sufficient but not necessary. Markets usually also converge to equilibrium (as in our demonstration experiment) when no money is paid.

d) Consider the following strategic market game in which two firms compete over quantities What do Huck et al. (JEBO 2004) find for collusion when 2 such firms compete vs. when 5 such firms compete?

A: The authors find no collusion with 5 firms. Such markets tend to be even more competitive than predicted: quantities are larger (and prices are below) Nash equilibrium predictions. But with 2 players there is tacit collusion, i.e., quantities are smaller than predicted.

Question 3: Biases in probability estimates

a) Consider the following scenario: The prevalence of the Covid19 virus in the population is 1%. The test accuracy is 90%. That is, if a person is infected by the virus, the probability is 0.9 that s/he will have a (correct) "positive" test result (and conversely, the probability that the test provides a (correct) "negative" result, i.e. shows s/he has not been infected when s/he is indeed not infected is also 90%). However, if a person is not infected, the probability is 0.1 that the test result is (falsely) positive. Now imagine mass testing is implemented in a country (i.e. everyone is tested), and a randomly drawn person receives a positive test result.

a1) What is the probability that the person is actually infected with Covid19? (*Hint*: use Bayes' rule)

A: prob (infected $\mid$ test positive) = (0.9 * 0.01) / (0.9 * 0.01 + 0.1 * 0.99) = about 8% (0.0833, to be exact)

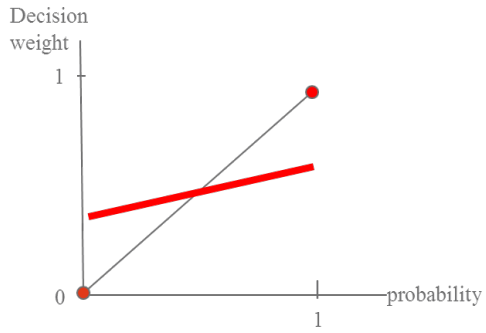
a2) What probability would a person prone to the Base-rate fallacy estimate for the person in the example above? (explain why).

A: The "base rate" in the example above is 0.01. It shows how often Covid19 occurs in the sample. The base-rate fallacy claims that people tend to ignore the base rate. Typically, they would think it is 90% (i.e., the accuracy of the test).

b) "Probability weighting" is one of the assumptions of prospect theory (Kahneman and Tversky, ECMA 1979).

What does the assumption imply for estimating the prevalence of Covid19 under current conditions (i.e. with low prevalence)? How will perception of the risk to be infected differ when the prevalence goes from 0 to 2% vs. from 10% to 12% (*Hint*: draw a diagram).

A: "Probability weighting" assumes that people do not proportionally perceive differences in probabilities («weight») but think in terms of «does not happen», «may happen», «will happen» (people tend to overreact to small probability events but under-react to medium and large probabilities. Overweighting occurs when weighting function is above 45 degree line).



c) Croson and Sundali (JRU 2005) provide evidence for the "Gambler's fallacy". Explain the basic intuition of the fallacy (Hint: refer to the "Law of small numbers"). What is the main finding of the study?
A: The GF results from the "law of small numbers" (Tversky and Kahneman 1971) and is a belief in quick reversal if a deviation from expected proportions is observed in a series of random events (false belief in negative autocorrelation).
The authors study data (in particular the "outside bets", i.e. "Red" vs. "Black") from cameras mounted above roulette tables in a Casino. They find that if a sufficiently long streak of "Red" occurs, gamblers increasingly bet on "Black", consistent with GF.

Question 4: Information aggregation in voting

It has been argued that a "dividend of democracy" can be obtained through positive information aggregation in majority voting (e.g. by Condorcet 1785).

a) Explain the information aggregation effect using an example of a common interest situation with 3 voters where each voter has a probability of $p_i = 0.6$ to make the correct choice. Assume that abstention is not allowed and voters cast their votes independently and sincerely. *Hint*: Condorcet Jury Theorem.

A: The groups makes the correct choice when a majority of votes is cast for the correct option. This is the case if all three voters vote for the correct alternative, which happens with probability $0.6 \times 0.6 \times 0.6 = 21.6\%$. The group also makes the correct choice when two out of three voters make the correct choice, which happens with probability $0.6 \times 0.6 \times 0.4 = 14.4\%$, and there are three possibilities for one of the three voters to cast the wrong vote, hence the total probability to make the correct choice for the committee is: $21.6\% + 3*14.4\% = 64.8\%$ which is more than the individual probability to make the correct choice $p_i = 60\%$. The difference is called information aggregation effect.

b) How is information aggregation related to the "wisdom of the crowds" and the claim that there might be a "dividend of democracy" in this case?

A: Information aggregation is stronger in larger committees. As the size of the committee goes to infinity, the probability of the committee to make the correct choice (quickly) goes to 1 under the assumptions in this example (especially $p > 0.5$). In this sense, "larger crowds" are "wiser" and there might be a "dividend of (inclusive or direct) democracy".

c) Morton, Piovesan and Tyran (GEB 2019) experimentally investigate information aggregation in voting by letting people vote on the correct answers to quiz questions, some of which are easy, some are "hard".

The authors compare outcomes in an "opinions" treatment (OT) vs. a baseline treatment (BT). How does this comparison speak to what the authors call "the dark side of the vote"? (*Hint*: describe the treatment difference first) Why?

A: In the Opinions treatment (OT) all voters learn what answers other (previous) voters thought are right but not whether these are the "right" answers, similar to an opinion poll. Specifically, voters are told the %age of subjects who answered A and B in baseline treatment (BT, which was run before the OT). In the BT voters do not get any information about how others decided in previous votes.

The "dark side of the vote" refers to negative information aggregation predicted by the Condorcet Jury Theorem when $p_i < 0.5$ (i.e. when voters are biased as in the "hard" questions). The authors observe positive information aggregation with easy questions but negative information aggregation with hard questions in the baseline treatment. In the Opinions treatment both positive and negative information aggregation becomes more pronounced. The reason is that voters tend to ignore their own signals and follow the public information. Doing so can be rational under some assumptions.

d) Mechtenberg and Tyran (GEB 2019) study information aggregation in a setting in which subjects can delegate the choice to an expert or can demand to make the choice themselves by majority voting.

The authors investigate the extent of "rational ignorance". What does it mean in the context of their experiment? *Hint*: Refer to the cost of information acquisition.

A: The authors use a design with a common interest situation (as in Morton et al. above), but with costly information acquisition. That is, voters can buy information (in which case they obtain an informative signal $p > 0.5$ about whether A or B is the right choice). If they do not, they get an uninformative signal ($p = 0.5$). The rational decision to buy information is complex in this setting in which abstention is possible. Buying the information has a cost but also a benefit, through information aggregation. This benefit is decreasing with the number of informed voters. Buying information is like providing a public good, it improves the outcome for all, hence free-riding incentives prevail. Because of these incentives the cost of information acquisition may dominate the benefit for a self-interested voter and he rationally may decide to remain uninformed (i.e. to be "rationally ignorant").

Question 5: Cooperation, Honesty, and Institutions

a) Gächter, Herrmann and Thöni (Science 2008 and 2010) investigate cultural and institutional determinants of cooperation. In the figures below, phase 1 (periods 1-10) is a standard linear public goods game, phase 2 is a peer punishment game.

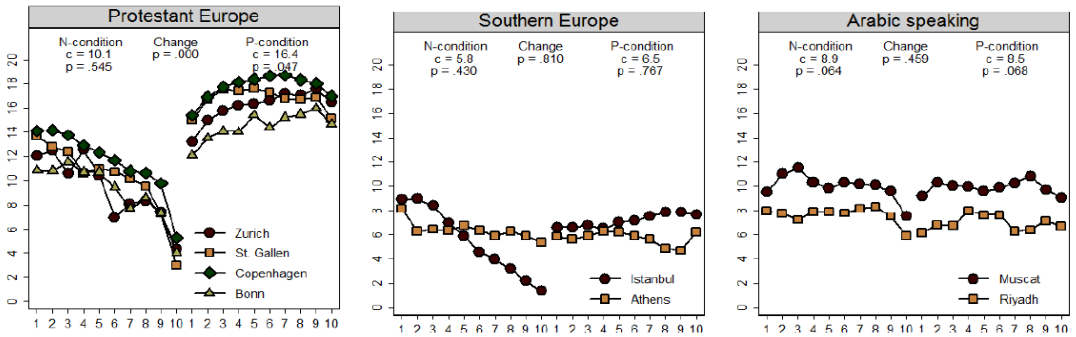
a1) What is the prediction of standard theory for phase 2? Explain.

A: No punishment and no cooperation. The reason is that punishment is costly to the punisher, hence nobody wants to punish in the last period. Given no punishment, nobody cooperates (free rider problem).

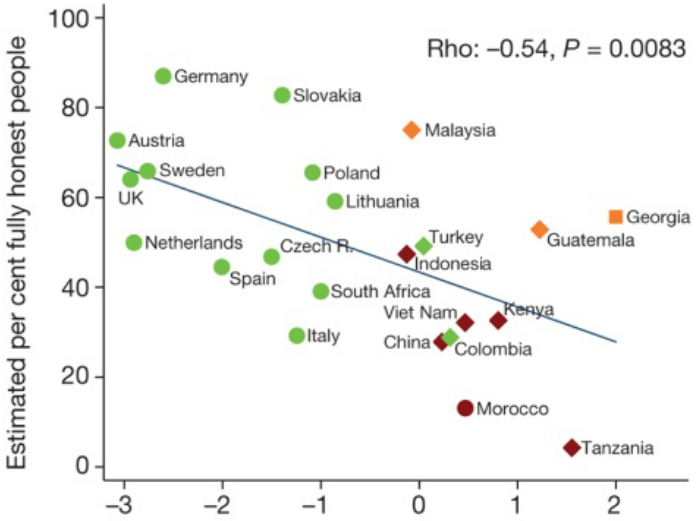
a2) How do the authors explain variation in efficiency across countries in phase 2?

A: They find that "antisocial punishment" (i.e., low cooperators punishing high cooperators) is high in countries in which "*norms of civic cooperation*" are low (e.g. attitudes to tax evasion, abuse of the welfare state, or dodging fares on public transport is perceived to be ok) and where the "*rule of law*" is weak (people do not perceive law enforcement institutions as fair, impartial etc.). It seems to be that punishment by peers in

not accepted in such countries because players do not think that cooperation is the appropriate course of action. As a consequence, they (blindly) exert counter-punishment which depresses efficiency.



b) Gächter and Schulz (Nature 2016) relate and index of prevalence of rule violations (PRV indicated on the *x*-axis) to a measure of "intrinsic honesty" in 23 countries as shown below.



b1) What does the PRV reflect? (*Hint*: the index is composed of three elements)
A: political fraud (measured by an indicator of political rights by Freedom House that measures the democratic quality of a country's political practices); tax evasion (proxied by the size of a country's shadow economy) and corruption (measured by the World Bank's Control of Corruption Index)

b2) How do the authors measure "intrinsic honesty"? (for 2'568 students in 23 countries)

b3) How is the "per cent of fully honest people" (*y*-axis) estimated?

b4) What does the color coding in the figure above reflect?

A: (b2) intrinsic honesty is measured by a die-rolling experiment. Participants were asked to roll a six-sided die placed in a cup twice, but to report the first roll only. Die rolling was unobservable by anyone except the subject. Participants were paid according to the number they reported. Reporting a 1 earned the participant 1 money unit, claiming a 2 earned 2 money units etc., except that reporting a 6 earned nothing. (b3) The estimated percentage of fully honest people is estimated form those reporting 6 (i.e., zero earnings). (b4) The quality of institutions in each country, proxied by the constraints on executives. Green, yellow, red means high, medium, low quality, respectively.

c) Cohn et al. (Science 2019) (deliberately) “lost” wallets in public places in 39 countries. How do the authors interpret the rates of reporting? These rates vary widely across countries. Name two correlates of the reporting rate.

A: they interpret the reporting rate as civic honesty. The find positive significant effect for cultural proxies of moral norms (e.g. more protestantism, weaker family ties) and state institutions (inclusive political systems proxied by plurality rule, longer democratic history, more constraints on the executive, longer history of primary schooling)