SUGGESTED ANSWERS for

Final exam, M.Sc. in Economics summer school 2021

Behavioral and Experimental Economics

Question 1: Markets

Markets can be organized in many ways and their properties depend on various characteristics. Experiments can shed light on the determinants of efficiency and "market failure".

a) Vernon Smith (JPE 1962) implemented the oral **double auction** (DA) with stationary replication to test propositions from competitive markets theory and to study properties of real markets organized as DA. What is the main proposition of standard theory concerning "efficiency" in competitive markets? How is "efficiency" measured in this experiment, and how do observations compare to the predictions of competitive markets theory? Are the experimental results surprising?

   A: Smith tests (a simplified version of) the first welfare theorem: competitive markets are efficient in the sense that they maximize the gains from trade. This prediction is driven by the price mechanism: only sellers with costs below the market price, and buyers with valuations above the market price trade (i.e., only inframarginal units are traded).
   The main finding of the experiment is that prices quickly converge to the predictions of competitive market theory and efficiency (measured as the realized sum of gains from trade relative to potential gains from trade) is very high. This finding is surprising in many ways: trading occurs at non-equilibrium prices, there is no "auctioneer" (both buyers and sellers can set prices), the parameters are such that more (in the demonstration experiment: all) units could have been traded (which would have resulted in low efficiency), and real traders are boundedly rational. Converge occurs under less restrictive conditions than assumed in the (unrealistic) competitive markets theory. Smith himself was surprised by the results (the editors of the JPE were not, but finally accepted the paper after some discussion).

b) How does a "**Pit market**" (*Hint*: think of a Bazaar) differ from a double auction (DA) market, and how do typical results from such markets compare to the results observed in the DA?
   A: A pit market is similar to a Bazaar in that market participants can walk around, physically meet each other, and haggle about prices. But because prices are publicly announced after each transaction, the pit market is more transparent than a bazaar. Experiments show that Pit markets have robust convergence to the equilibria predicted by competitive markets theory and are highly efficient (similar to the DA).

c) Smith, Suchanek and Williams (ECMA 1988) experimentally studied trading in stylized **stock markets** and find that prices do not follow fundamentals as predicted by standard theory. Explain the typical pricing patterns and discuss why they may be observed. (*Hint*: refer to how such markets differ from double auction markets and to the role of expectations)

   A: In the stock market experiment, an asset is traded that pays an uncertain dividend (while there is no uncertainty in the DA). The asset has an inner value which is known to all (rational) traders. It is determined by the expected flow of dividends (while in DA the players only know their own valuations and can therefore not predict what rational agents

would do). A main difference to the DA is that participants in the DA are *either* buyers *or* sellers whereas they can both buy and sell the stocks at any time in the stock market experiment. Hence, speculation is possible, and expectations matter: a trader might be willing to buy an asset now in order to resell it later even if s/he knows that assets currently trade above their inner value if s/he thinks prices will increase even further ("greater fool" theory). Such behavior may give rise to speculative bubbles and later market crashes. Smith et al. observe massive price bubbles and crashes towards the end of the experiment (however, bubbles are diminished and the market crashes earlier when the experiment is repeated with the same participants several times).

d) How does the number of firms competing in **oligopolistic markets** ("Cournot competition") affect consumer surplus and efficiency in standard theory? Do experiments support this view? Is it necessary that firms actually enter the market to induce the effect? (*Hint*: refer to Gächter, Thöni and Tyran, JEE 2006)
Are experimental duopoly markets more or less competitive than predicted? Why? (*Hint*: refer to Huck, Normann, Oechssler JEBO 2004)

A: Cournot markets are often remarkably well in line with theoretical predictions of standard theory according to which equilibrium quantity traded in total increases, prices fall, consumer surplus and efficiency increase with the number of competitors (e.g. when comparing duopoly and quadropoly). Gächter et al. (2006) provide evidence in support of these predictions in phase 1 of their experiment. In phase 2 they show that market outcomes are in line with predictions of standard theory when firms have the possibility to enter uncompetitive markets (i.e. they are contestable), it is not necessary that they actually do enter, the threat of entry is sufficient.
Huck et al. (2004) show that duopoly markets are somewhat less competitive than theoretically predicted because firms manage to engage in "tacit collusion" to increase their profits (hold back the quantity below the best reply even when they are not able to communicate with each other). But collusion does not succeed in their experiment with 4 firms, and markets with 5 firms are even more competitive than predicted by standard theory. Hence the title of their paper, "2 are few, 4 are many".

e) Some markets are beset by "**quality uncertainty**" in the sense that buyers are uncertain whether they will receive high or low quality for a given price from sellers (for whom it is more costly to produce high quality goods).
Huck, Lünser and Tyran (GEB 2012) experimentally test the efficiency effects of two conditions to reduce inefficiency. Explain the design, hypotheses and main results.

A: The authors use a binary trust game in which the subgame-perfect equilibrium (if played once) is that sellers provide low quality and buyers stay out of the market. There are 4 buyers and 4 sellers in each market. They implement a 2-by-2 design with the possibility of reputation formation (buyers know the full or only partial history of sellers) and competition (buyers can choose where to shop vs. random assignment), plus a control without possibility of reputation formation (no information). The authors hypothesize that both factors should increase efficiency as they both serve to "discipline" sellers (buyers can reward good behavior of sellers by choosing them, and punish sellers by shunning them). The results are that efficiency is very low in the baseline condition (about 8%), intermediate without competition (at about 35%, and the amount of information does not play a big role in that case), while efficiency is very high with competition (about 80%). Interestingly, there are very pronounced end-game effects in all conditions, showing that strategic considerations (rather than fairness concerns alone) made the difference. Hence the title of the paper: competition fosters trust (which, in turn, promotes efficiency).

Question 2: Money illusion and the indirect effects of bounded rationality

a) What role does "**money**" play in standard microeconomics? How are prices defined in the standard model of consumer choice? What does standard theory assume about the consequences of simultaneously doubling nominal prices and incomes?
(*Hint*: $f(\lambda \mathbf{x}) = \lambda^r f(\mathbf{x})$)

A: Money plays no role in the standard model of consumer choice which is a "real" theory. Prices are defined as relative (not absolute or nominal) prices. The question studied in this model is what bundles of goods consumers choose with a given budget and given relative prices (which determine the opportunity set), assuming utility maximization (trying to reach the highest indifference curve). The "homogeneity postulate" is that a doubling of nominal incomes and prices has no effect on chosen bundles, i.e., (net) demand functions are assumed to homogeneous of degree zero (i.e. $r = 0$ above).

b) Why do people use money and why is **thinking in nominal terms** about economic transactions common?

A: Money has many advantages, it serves as a medium exchange and accounting, and as a store of value. Thinking in nominal terms about transactions is intuitive and common because doing so is salient, simple, natural, and it is often a good heuristic (esp. when there is no inflation).

c) Can people **avoid falling prey** to "money illusion"?
(*Hint*: refer to system 1 vs. 2)

A: Thinking in terms of money (rather than in terms of relative prices) about economic transactions is intuitive, it is an example of "system 1" thinking (Kahneman´s "intuitive mind" which is fast, automatic, effortless), but by activating "system 2" (the "reflective mind" which is slow, conscious, effortful) we can overcome it (e.g. make an effort to correctly translate nominal into real values by deflationing.

d) Agell and Bennmarker (2003) ran a survey among Swedish human resource managers on the **acceptability of wage cuts**.
Describe the variation between the two scenarios that were randomly presented to respondents. What is the main result of comparing the two scenarios? Provide a psychological explanation for why the effect was observed.

A: The authors present one of two scenarios (which are equivalent in real terms) to respondents (both involve a real wage cut of 5%). In one treatment, inflation is low and the real cut comes with a nominal wage cut, in the other treatment, inflation is high and the real wage cut comes with a nominal wage increase. The main finding is that human resource managers think that workers find the real wage cut more acceptable when it comes with nominal wage increase than when it comes with a nominal wage cut. The explanation for this observation is that the effect of loss aversion is thought to be more pronounced when real losses are more salient. This is the case when they come with nominal losses (and real losses are "obscured", i.e. require more activation of "system 2" when they come with a nominal wage increase). Hence, the psychological mechanism is money illusion plus loss aversion (or, "nominal loss aversion").

e) A limitation of the survey evidence cited in the previous question is that it cannot easily distinguish between **direct and indirect effects** of money illusion. Explain what is meant by these effects and why this is so in the scenario in question d).

A: The direct effect of money illusion refers to the extent to which nominal thinking distorts choices of a decision maker because he himself is confused (is not individually rational). The indirect effect results from *beliefs* of a decision maker about how other people are confused by money illusion, and the strategic (optimal) response to this belief. The question asked in the survey mentioned in question d) above was whether human resource managers think that workers find a real wage cut acceptable. If managers say workers are likely to find a real cut more acceptable, it may be because the human resource managers are themselves prone to money illusion (= direct effect) or they are not but think that workers are. (in either case, managers may be hesitant to cut nominal wages for fear of negative consequences like unions calling for a strike or demotivated workers providing low effort).

f) How does the design of the experimental study of **Fehr and Tyran (GEB 2007)** correct for the limitation discussed in e)? Briefly describe the game and its three (pure-strategy) Nash equilibria.
(*Hint*: refer to pareto-rankability in nominal vs. real terms). What do the authors find?

A: The authors implement a coordination game in which subjects are in the role of firms who repeatedly choose nominal prices (in groups of 5 or 6 players). Real payoffs are determined by relative prices. The game has 3 pure-strategy Nash equilibria which are pareto-rankable (2 are stable, one is unstable). However, the pareto-dominant equilibrium in real terms (A) has the lowest payoffs in nominal terms, and the inferior stable equilibrium in real terms (C) has the highest nominal payoffs. Hence, players who are themselves prone to money illusion or players who think that other players are prone to money illusion are drawn to C. The authors implement 4 treatments which vary whether subjects have to form expectations (when they play against other humans (H) vs. against computers who choose best replies to the choice of the subject, in NC and RC) and the nominal vs. real representation of payoffs (N vs. R). In NH and RH, the game has strategic uncertainty, while in NC and RC it is an individual optimization task (only direct effects can matter here). Thus, the authors identify the direct effect by allowing (in NH) or precluding (in NC) an indirect effect (while money illusion itself is identified by comparison of N vs. R treatments).
The authors find that coordination fails in the sense that in the NH all groups converge to equilibrium C. But in in NC many make optimal choices (pick A). This shows that the coordination failure was to some extent due to direct effects of money illusion (even after many repetitions 18% of subjects choose equilibrium C in NC), but is mostly due to indirect effects (84% choose C in NH).

g) Consider a **guessing game** in which subjects choose numbers between 0 and 100 with $n > 2$ players and $1 > p > 0$. How is p related to indirect effects of bounded rationality?
(*Hint*: compare the case above to $-1 < p < 0$).

A: p shapes strategic incentives and the effect of indirect effects. With $1 > p > 0$, strategic complementarity prevails and rational players have an incentive "to follow the crowd". Players who think that other players choose high numbers have an incentive to also choose high numbers in this case, and indirect effects of bounded rationality add to the direct effects, and overall effects are multiplied. With $-1 < p < 0$, strategic substitutes prevail, rational players who think that others choose high prices now have an incentive to choose low numbers, and indirect effects compensate the direct effects, such that the overall effects are mitigated.

Question 3: Fairness and institutions

Much evidence suggests that people are heterogeneous with respect to social preferences and fairness ideals. Experiments can shed light on the nature and consequences of this heterogeneity.

a) Evidence from the dictator game has often been cited as showing that people have heterogeneous prosocial preferences, and that few dictators choose an equal split when sharing money with a stranger. In contrast, evidence from **Ultimatum games** shows that the equal split is common. Why do we see a difference in sharing behavior in these games? (*Hint*: refer to differences in strategic incentives)

A: The strategic situation is very different in the two games, and incentives transform the extent to which a given deviation from the standard assumption manifests itself in behavior. The dictator game (DG) is not really a game because the receiver is totally passive. In contrast, in the Ultimatum game (UG), the proposer suggests a share of the "pie" (s) to be passed to the responder who can accept or reject that proposal (in which case both get nothing). A proposer can, thus, be "punished" by a rejection, and the cost of the punishment to the responder increases in s. A self-interested proposer with incomplete information, i.e., who is uncertain about the social preference of the responder (i.e., what s/he considers to be the fair share) may want to reduce the probability of rejection by giving $s >> 0$ which may (indirectly, through an increase of the acceptance probability) increase own expected earnings. Empirically, acceptance rates increase with s, and are very low for $s = \varepsilon$ but close to 100% for 50:50 proposals. Thus, the difference between DG and UG is not explained by differences in generosity of the proposer but by the strategic nature of the game and a (presumed) willingness to punish unfair behavior by responders which depends on the perceived "fairness" (= how close it is to 50:50) and the cost of punishment.

b) Prasnikar and Roth (QJE 1992) study the **multi-proposer Ultimatum Game** to show that competition (between proposers) can importantly shape how given (pro-social) preferences are transformed by incentives. Explain the design and discuss their results.

A: Design: 9 proposers simultaneously submit offers from 0$ to 10$ to a single responder. The responder may exclusively accept the best offer s_h. If s/he rejects, all 10 players get a payoff of 0. If he accepts, the responder gets s_h, the one proposer who submitted the best (accepted) offer gets 10 - s_h. All others get 0. Results: Highly unequal outcomes prevail immediately and near-perfect equilibration prevails after a few periods, as predicted by standard game theory ($9.95 and $10 are equilibrium offers which are accepted by the responder). Proposers indeed engage in fierce competition and the entire rent goes to the responder. Thus, even though a considerable share of players is prosocial, such preferences do not translate into observed behavior in this (market) setting which generates very uneven outcomes (1 person gets everything, 9 people get nothing).

c) **Almas, Cappelen and Tungodden (WP 2019)** use an online redistribution experiment to shed light on the question why there is less (net) income inequality in the U.S. than in Scandinavia. (*Hint*: The design involves two "workers" and a "spectator" who make choices in 3 main treatments called Luck, Efficiency, and Merit).

c1) How do the three treatments differ?

A: Workers each do 3 simple sentence unscrambling tasks and are paid 0 USD or 6 USD, depending on the treatment.
- In Luck it is randomly determined which workers get 0 USD or 6 USD, and redistribution of earnings by the spectator is costless

- Efficiency is the same as Luck except that there is an efficiency cost (of 100%) of redistribution
- In Merit, the more productive worker gets 6 USD, the less productive gets 0 USD, there is no cost of redistribution

c2) The authors argue that differences in "**fairness views**" are a potential explanation for differences observed in income inequality in the US and in Scandinavia. Explain the three views considered by the authors and what they imply for spectator choices in the respective treatments.
(*Hint*: Think of which aspect of "fairness" these views are supposed to capture)

A: Spectators who decide how to redistribute the money between the two workers are supposed to hold one of the following views:

- "Libertarian": inequality due to luck and productivity is considered fair (no redistribution predicted)
- "Meritocratic": inequality due productivity differences is considered fair, but not differences due to luck (prediction: the more productive worker gets more in treatment Merit, but earnings are equally shared when income is determined randomly)
- "Egalitarian": neither source of inequality is considered fair (prediction: each worker gets half independent of the source of inequality)

c3) What are the **main results** of the study? Name 3 results. (*Hint*: refer to similarities and differences in responses across the countries. The authors also elicit individual characteristics of respondents)

- If inequality is due merit instead of luck: inequality acceptance increases in both countries → source of inequality (i.e., fairness views) matter in both US and NOR
- Cost of redistribution: no effect in US and NOR
- US spectators accept more inequality than NOR spectators in all treats
- Libertarians are more common in US, egalitarians more common in Norway
- Conservatives and men implement more inequality in both countries, highly educated only in US
- Spectator choices correlate with attitudes on redistribution in society in both countries (suggesting some external validity of the experiment)

c4) Almas et al. (2019) is an example of an **online experiment**. Name one advantage and one limitation of doing such an experiment online vs. in the laboratory.

A: Advantages: Low cost of recruiting large sample (6000 subjects in total), ability to generate a heterogeneous sample of the adult population (e.g. w.r.t. ideological preferences, income, gender, age). Limitations: only simple experiments can be implemented online, limited control (e.g. attentiveness, understanding of instructions), selection bias (e.g. rich people participate less)

c5) Good research answers at least one question, but other questions remain unanswered and the results often raise new questions. Discuss two **open issues** in Almas et al. (2019).
(*Hint*: name reasons other than fairness preferences for why the US may have less redistribution than Norway, or refer to proximate vs. ultimate causes)

A: Unclear where fairness views come from (fairness views are a proximate cause). What is the role of institutions in shaping such views (school, media, competitiveness of society)

Alternative explanations: Supply-side factors: maybe rich people have more influence on the political system in US than in Norway, differences in history/ethnic heterogeneity (some evidence suggest that people are more willing to share with people who are similar)

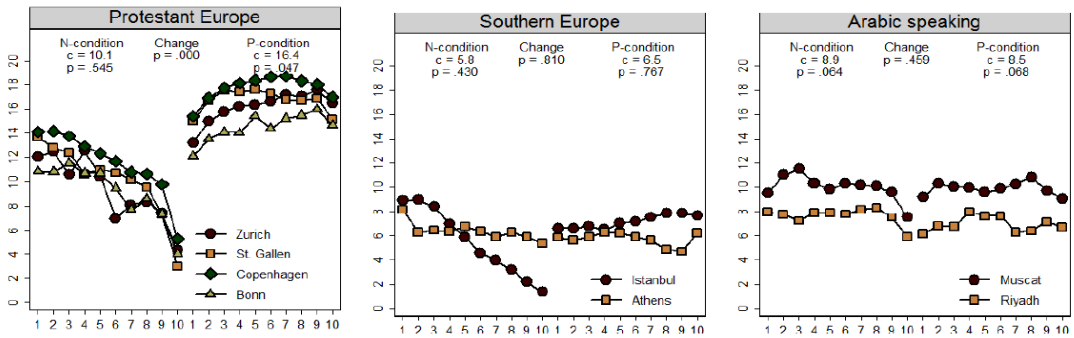
d) **Free-riding in cooperation games** is considered unfair by many. The extent to which free-riding is observed depends on institutions and the incentives they generate. Gächter, Herrmann and Thöni (Science 2008 and 2010) investigate cultural and institutional determinants of cooperation. In the figures below, phase 1 (periods 1-10) is a standard linear public goods game, phase 2 is a peer punishment game.

d1) What is the **prediction** of standard theory for phase 2? Explain.

A: Prediction assuming common knowledge of rationality and self-interest in a one-shot game: No punishment and no cooperation. The reason is that punishment is costly to the punisher, hence nobody wants to punish which also holds in the last period of the game. By backwards induction, this holds for the entire phase. Given no punishment, nobody cooperates (an extreme free rider problem prevails).

d2) How do the authors explain **variation in efficiency across countries** in phase 2? (*Hint*: refer to norms of civic cooperation and rule of law)

A: They find that "antisocial punishment" (i.e., low cooperators punishing high cooperators) is common in countries in which "*norms of civic cooperation*" are low (measured e.g. by 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 the case 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.



Question 4: Democracy

It has been claimed that a "dividend of democracy" can be obtained through positive information aggregation in majority voting.

a) Explain the **information aggregation** effect using an example of a common interest situation with $n = 3$ 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 0.6×0.6

× 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 "inclusive democracy"? (Hint: refer to variation in n)
A: A "crowd" is a large group of people. 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). Explain how does this treatment comparison speaks to what the authors call "the dark side of the vote"?
(*Hint*: describe the treatment difference first and then refer to p_i)
A: In the Opinions treatment (OT) all voters learn what answers other (previous) voters thought to be correct but not whether these are indeed correct, 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. That is, voting yields worse choices than subjects would have made individually with hard questions. The reason is that voters tend to ignore their own signals and follow the public information. Doing so can be rational under some assumptions discussed in the paper.

d) **Mechtenberg and Tyran (GEB 2019)** study information aggregation in a setting in which subjects can delegate a 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 this expression 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 in 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 s/he may rationally decide to remain uninformed (i.e., to be "rationally ignorant").