Jean-Robert Tyran
Department of Economics,
University of Copenhagen

Suggested Answers in Final exam for MA course "Behavioral and Experimental Economics"

August, 2012 (2 hours, closed book)

Question 1: Causal effects

Consider an outcome of interest Y that depends on a list of determinants, i.e. $Y = f(X_1, \ldots, X_N)$. A causal effect of X_1 on Y is the effect of varying X_1 holding fixed all other determinants $Z = (X_2,\ldots,X_N)$.
Answer questions a) to d) with reference to the notation introduced in Falk and Heckman (*Science* 2008).

a) Does an observed causal effect depend on the level of Z?
(Hint: refer to "separability")
A: Unless f is separable in X_1, so that $Y = \phi(X_1) + g(Z)$, the level of the Y response to change in X_1 will depend on the level of Z.

b) Does an observed causal effect depend on the level of X_1?
(Hint: refer to "separability")
A: Even in the separable case, unless $\phi(X_1)$ is a linear function of X_1, the causal effect of X_1 depends on the level of X_1 and the size of the variation of X_1.

c) Suppose a laboratory experiment identifies a strong causal effect of X_1 on Y (given Z) and a field experiment identifies a weaker effect of X_1 on Y (given Z'). What can be concluded for the ability of results from experiments to "generalize" to other environments Z''?
A: It is not clear that one causal effect has more predictive power than the other. It often depends "how similar" Z'' is to either Z or Z'. In general, a theory of how alternative environments Z relate to the effect of a change in X_1 to Y is needed to be able to say whether the result of one or the other experiment "generalizes better" to other environments.

d) Explain how "randomization" serves to neutralize the effect of uncontrolled determinants (X_u) on Y.
A: Suppose factor X_U cannot be controlled (or even be observed) by the experimenter. Suppose factor X_1 is varied in a controlled way, and all other factors Z are held constant. Randomized allocation of participants to treatments makes sure that the level of X_U is the same in both treatments. Randomization plus control of X_1 and Z allows to isolate $\Delta Y/\Delta X_1$ at a given (perhaps unknown) level of X_U. In contrast if participants select into treatments according to X_U, some effect $\Delta Y/(\Delta X_1 \Delta X_U)$ is measured.

e) Provide an example (of research using non-experimental data) in which an observed difference between two groups is not causal but is likely to result at least in part from selection.
A: Examples discussed during lecture: Married men live longer and are healthier than unmarried, Unemployed volunteering for active labor market policy (e.g. taking a computer course) are more likely to find a job.

Question 2: Competitive Markets

Consider the Double Auction institution (Smith JPE 1962). Suppose there are 5 buyers (B) and 5 sellers (S) with 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 the prediction (according to standard theory) for equilibrium

- Prices and quantities (hint: note that units are discrete)
  A: price between 70 and 75, quantity is 5
- Surplus in equilibrium
  A: 165

b) Suppose that a tax of 25 is levied on all sellers in this market. Calculate equilibrium

- Prices and quantities (hint: note that units are discrete)

  A: Price between 85 and 90, quantity 3

- Surplus in equilibrium

  A: 65

c) What is the typical observation in such markets? (e.g. Smith JPE 1962)

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

Question 3: Loss aversion

a) Name the three key assumptions of Kahneman and Tversky (ECMA 1979) about how people evaluate risky prospects (in particular if they involve the possibility of losses).
A: 1. Risky prospects are not evaluated in terms of outcomes (final wealth, as in EU) but in terms of changes w.r.t. to reference point r (initial wealth). 2. Losses weigh more heavily than gains of equal size (kink at the reference point): loss aversion, 3. Concave in gains but convex in losses → risk aversion in gains but risk loving in losses.

b) What is the endowment effect? Provide an example.
A: Endowment effect says that people have a higher willingness to accept to give up a good currently in their possession than a willingness to pay for the same good that is not in their possession. They value a good more of it's theirs than if it's not. Example: In the "decorated mug experiment" (Kahneman, Knetsch and Thaler JPE 1990), students asked about twice as much to sell their mug than they were willing to pay to obtain the mug.

c) How does the endowment effect relate to the assumptions of Prospect theory?
A: While prospect theory was originally developed to explain choice under risk, it has later also been applied to choices that do not involve risk. The endowment effect is related to the assumption of "loss aversion" (a tendency to strongly prefer avoiding losses to acquiring gains) as people seem to mentally code giving up on something in their possession as a loss.

d) What does the existence of an endowment effect imply for the Coase Theorem? (Hint: refer to property rights)
A: The Coase theorem claims that (under a number of strong assumptions) externalities can be internalized by negotiations (voluntary agreements) and do not require direct government intervention. This claim holds (under standard assumptions) independent of whom the property rights are allocated to, i.e. efficiency is invariant with respect to that allocation. However, if agents are loss averse, WTA > WTP and, in general, different agreements will result depending on whom the property right is assigned to.

Question 4: Biases in probability estimates

a) Consider the following scenario: For a woman at age 40 who participates in routine screening, the probability of breast cancer is 0.01. If a woman has breast cancer, the probability is 0.9 that she will have a positive mammogram. If a woman does not have breast cancer, the probability is 0.1 that she will still have a positive mammogram. Now imagine a randomly drawn woman from this age group with a positive mammogram. What is the probability that she actually has breast cancer? (Hint: use Bayes' rule)
A: prob (cancer | test positive) = (0.9 * 0.01) / (0.9 * 0.01 + 0.1 * 0.99) = about 8% (0.0833, to be exact)

b) What probability would a person prone to the Base-rate fallacy estimate for the woman at age 40 to have breast cancer in the example above? (explain why)
A: The "base rate" in the example above is 0.01, i.e. how often breast cancer occurs in the sample. The base-rate fallacy claims that people tend to ignore the base rate. Since it is low in the example above, such people would tend to overestimate the probability. Typically, they would think it is is 90% (i.e. the reliability of the test, given that someone has cancer).

c) Kagel, Ganguly and Moser (JRU 2000) translate a similar scenario into an asset market context in which an "analyst" provides a report (signal) about success or failure of some company. The authors find that individual estimates are systematically biased, but prices in the market reflect the correct (Bayesian) probability of success in some cases and were far off in other cases. Explain under what conditions (and why) either outcome prevailed.
A: The authors used different scenarios in which the prediction according to Bayes' rule has is higher or lower than the guess a person with Base-rate fallacy would make. Market prices are better predicted by Bayes' Rule if Baysian agents have higher values and *unlimited buying* power (were able to buy all assets). Thus, a single trader with correct (Bayesian) beliefs is sufficient to drive market price to the rational prediction. But BRF is better predictor of prices if Bayesian agents have lower values (and limited selling power). BRF now think the asset has high value, and they drive the prices up to irrationally high levels.

Question 5: Voting and redistribution

The questions below refer to Höchtl, Sausgruber and Tyran (EER forthcoming).

a) Explain what the numbers in line/column (3;2) and (2;3), i.e. 0.51 and 0.99 show. (Hint: the calculations assume that the share of inequality-averse voters is ρ = 20%)

	1	2	3	4	5	6	7	8
1	-	0.96		0.97		0.98		0.99
2	0.64	-	0.99		0.99		1.00	
3		0.51	-	1.00		1.00		1.00
4	0.82		0.41	-	1.00		1.00	
5		0.66		0.33	-	1.00		1.00
6	0.90		0.66		0.26	-	1.00	
7		0.97		0.85		0.21	-	1.00
8	0.94		0.80		0.50		0.17	-

A: The numbers in the table show the probability of obtaining voting outcomes in line with standard theory in a simulation assuming that the share of self-interested voters is 80% and the share of inequality averse-voters is ρ = 20% (i.e. if there are two income classes). Simulation results as a function of the number of rich voters (the lines) and the number of poor voters (the columns).
The table shows, for example, that the probability is much lower when rich voters slightly outnumber poor voters (e.g. 51% in case 3;2) than in the reverse case (e.g. 99% in case 2;3), and this asymmetry grows with the size of the electorate.

b) In Höchtl et al., voters decide on redistribution from "rich" to "poor". In all treatments, the prediction for voting of self-interested poor voters is to opt for $t = 8$, the prediction for the rich voters is $t = 1$.
Consider treatment PMV (when the poor are in majority). How do average votes by the poor and the rich compare to predictions? How does the aggregate outcome compare to the prediction? Explain.

A: Average votes of the rich are biased towards "excessive" redistribution (avg. is 3.9), and so are the average votes of the poor (average is 8.9). So, on average both groups support higher redistribution than predicted by standard theory (= 8 in the aggregate). But the aggregate outcome is driven by the median voter. Since the median voter is poor and the median voter does not seem to be inequality averse, the prediction for the aggregate is in line with standard predictions despite most voters deviating from standard predictions.

Question 6: Cooperation and punishment

a) What is the standard game-theoretic prediction in the Public Good Game (or, voluntary contribution mechanism) if played once? (Hint: $0 < a < 1 < an$)

(A: free riding ($g_i = 0$) is dominant, and zero contributions by all participants is the unique (inefficient) equilibrium.)

b) Explain how the "strategy method" can be used to elicit cooperator "types" (e.g. in Thöni, Tyran and Wengström JPubE 2011). Describe the profile (slope) for a free rider and of a conditional cooperator. What distribution of "types" do the authors find in the Danish population?

(A: In the strategy method, players indicate their contributions conditional on all possible contribution levels by others. The set of all conditional choices constitutes a cooperation profile. If a player indicates a flat profile at zero, he is classified as a free rider. If a player indicates a monotonically increasing profile, he is classified as a conditional cooperator. Thöni et al. find that conditional cooperators are the most common group in a (close to) representative Danish sample.)

c) Gächter, Herrmann and Thöni (Science, 2008) observe substantial variation across countries in the punishment game (e.g. Fehr and Gächter AER, 2000). How do the authors explain this variation?

(A: depends strongly on the tendency to engage in antisocial ("perverse") punishment of cooperators which in turn is shaped by "norms of civic cooperation" and "rule of law" as measured in surveys)

d) Markussen, Putterman and Tyran (2011, WP) implement a game with voting on formal sanctions. What is the prediction of standard theory for voting and contributions in treatment DC, i.e. when $s = 0.8$ and $c = 2$, if the alternative is no sanctions? How do these predictions change in treatment DE, i.e. when $s = 0.8$ and $c = 8$? How do experimental results compare for voting in DC and DE?

Hint:

$$\begin{aligned}\pi_i^{FS} &= (1-s)(20-C_i)+0.4\sum_{j\in g}C_j - c \\ &= 20(1-s)+(0.4+s-1)C_i+0.4\sum_{j\neq i}C_j - c\end{aligned}$$

(A: The predictions for DE and DC are the same: $s = 0.8$ means formal sanctions are deterrent, i.e. rational and self-interested agents do not free ride: $C_i = 20$. Therefore, both DE and DC will be accepted in voting. The results show that DC is mostly accepted – in phase 6 almost 90% vote for DC – but support is much lower for DE – only about 30% vote for DE over no sanctions.)

Question 7: Cooperation and competition

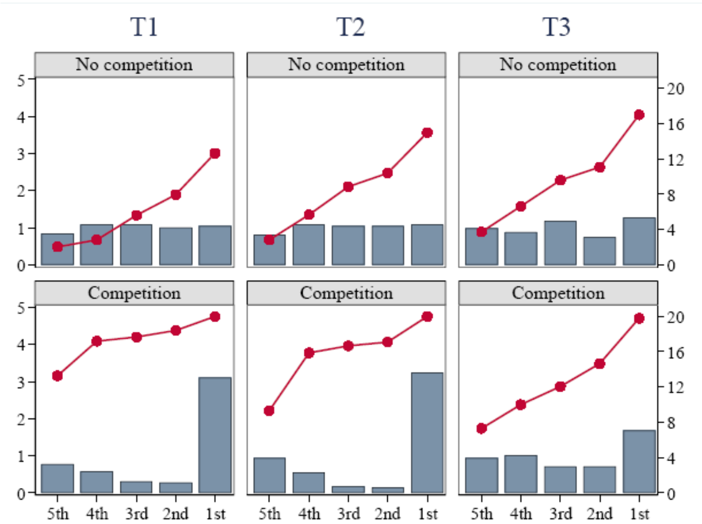
In Reuben and Tyran (EJPE 2010) k = 5 groups of n = 3 players compete as follows

$$\pi_{ik} = \pi_i \times f(r_k),\ \pi_i = y - c_i + \alpha \sum_j c_j,$$

a) What is the profit-maximizing contribution choice for player i in the following situation?
Suppose E = 20, α = 0.4. The average contribution in each of the other k-1 groups is 12, and the other n-1 members of player i's group contribute 20 on average.

A: zero (this holds in all treatments), earning player i π_i = 20 – 0 + 0.4(40) = 36. Since player i's group is first anyway (i.e. with a positive contribution by i), i's contribution does not affect the ranking (i.e. r_k), not contributing is dominant (α < 1)

b) Consider the following figure from Reuben and Tyran (EJPE 2010) and explain the following.



- How do treatments T1, T2, T3 differ?
A: they differ by the incentives in rank competition, i.e $f(r_k)$. They are steep in T1 (1.0, 0.8, 0.6, 0.4, 0.2), punish not winning strongly (1.0, 0.5, 0.5, 0.5, 0.5) or flat (1.0, 0.95, 0.9, 0.85, 0.8). The aim of the treatment variation is to see whether the scheme works (better) when incentives are steep, and whether "graded" incentives are more effective.
- What do the red lines show in the diagram above? (briefly comment on the findings)
A: Red line shows the avg. group contribution to the PG (right scale) by group rank. The line is at higher values (further up) with competition than without (i.e. when $f(r_k)$ = 1 for all).
- What do the bars show? (briefly comment on the findings)
A: Height of bar shows the number of groups in rank r (left scale). Without competition: groups are equally likely to be in any rank (and it does not matter for payoffs what rank one's group is). With competition, in T1 and T2: about 3 groups are ranked first on avg. Hence, the title of the paper: Everyone can be a winner: several groups can be ranked first, and because such groups fully cooperate, most (70% in T1 and 56% in T2) is better off (i.e "is a winner) with than without competition. But the effect is weaker when incentives to compete are flat (T3)