Review of the thesis “Experiments in Corruptibility” by Iryna Momotenko

The thesis consists, after a brief introductory chapter, of three substantial research papers, each self-contained with its own references list and appendices. All appear to be single-authored (the author uses the pronoun ‘I’ when referring to herself as author). Each reports an experimental study addressing issues of corruption and anti-social behaviour. All experiments are properly designed and conducted, the data analysis is substantial. There is no doubt in my mind that the thesis meets the standards required for a doctoral thesis. Below are some comments that may improve the thesis. All can be dealt with editorially.

Chapter 1

The first paper studies how information on peer behaviour influences subjects’ propensity to engage in corrupt acts. A corrupt act is here modelled as extracting resources from a common pool. To emphasise the corrupt nature of the act, there is a (tiny) probability that the perpetrator is caught and stripped of all payment. The money remaining in the pool is then multiplied and evenly redistributed. Thus, the game is a reframed voluntary contribution mechanism.

There are three main treatments (and two minor ones, which I comment on below). They differ in the information given to subjects about behaviour in previous sessions. In one treatment, the control, no information is given. In the positive information treatment (PosInfo) subjects were told the six lowest extractions in previous sessions (which happened to be all zero). In the negative information treatment (NegInfo) they were told the six highest extractions (which happened to be all the maximum of 250). This is clever: Given the strong tendency of subjects towards the edges, it is likely that the bottom six observations out of 24 are all zero, and the top six are all the maximum. In this sense, the actual information content is rather limited. Yet the phrase ‘The six biggest withdrawals were 250,250,250,250,250,250’ makes it look like the subject pool is thoroughly corrupt, while the phrase ‘The six smallest withdrawals were 0,0,0,0,0,0’ makes it look like withdrawals were very rare. In fact, these are two true statements about the very same data set.

The main results are very interesting. Indeed, the manipulation has a strong effect in the expected direction. Mean extractions are 43% lower in PosInfo than in NegInfo. The effect is mainly driven by the positive information. While extractions are indeed higher in NegInfo than in the Control, the difference is small and not significant, while the difference between Control and PosInfo is large and significant. The author also elicits positive and normative expectations about actual and appropriate withdrawals and finds that the treatment manipulation works in the same way there.

I was a bit confused about the discussion (and implementation) of the ‘strategy method’. Strategy method is generally understood as the elicitation of complete strategies for later movers in multi-stage games. This allows the experimenter to collect second-mover data also on nodes that have not been reached because the first mover has not chosen the corresponding action. The present experiment involves a simultaneous-move one-shot game, so strategy method is not applicable (or, rather, it becomes trivial as the one decision is the complete strategy). Apparently, the author means something different. Subjects are asked to ‘imagine’ that the six highest withdrawals were 250, and to ‘imagine’ that the six lowest were 0 in two previous sessions. Thus, it seems like a hypothetical scenario in some way (the concrete instructions for these treatments were not provided in the appendix), but for real in another way (since both statements are simultaneously true). Anyway, it is very different from how strategy method is generally understood (though admittedly, there is no legal definition of the term and the now common use was even disapproved
of by the originator of the term, Reinhard Selten). The paragraphs about the possible use of strategy method should be rewritten and a different term should be used.

Minor points.

1. The detection probability is 0.003 and thus ridiculously low. The author states historical reasons for it. But the cited precedent involved a repeated game, in which the probability was the probability of getting caught in each repetition. Hence the overall probability for repeat offenders to get caught was substantially higher. It would be interesting to see whether this tiny probability would also matter in the one-shot game. It shouldn’t really, but it could since the very existence of the lottery also carries a moral message. But this is a matter for future research.

2. The 2018 JLEO paper ‘The Effect of Social Norms on Bribe Offers’ seems to do similar things. Would it be worth mentioning?

Chapter 2

The second paper studies whether higher wages decrease (or possibly even increase?) the corruptibility of public officials. This is a very important question still awaiting empirical proof. The author designs an experiment based on an existing bribery game. The game models the reciprocal relationship between a firm and a public official. The firm offers a transfer to the official, which the official can accept or reject. At the second stage the official can choose an action that favours the firm but damages the public (here represented by a charity donations to which are withheld). As in the previous experiments on this kind of bribery game, there is at each stage a probability that a corrupt exchange is detected and the game ends for both players without payment.

The experimental variable is the wage that is paid to the public official. In the baseline earnings potentials for both players are pretty even. In three other treatments the wage level for the official is increased. It is hypothesised that the corruptibility of officials either decreases with rising wage (because officials have more to lose in case of detection) or that the relationship is U-shaped (because very wealthy people tend to be less risk-averse). The author finds some indication of a U-shaped relationship.

Again, this is a very solid effort. I have no objections against the design. Maybe it is slightly less original than the other two studies, but this is not really a flaw. The results are interesting. The author indeed finds a U-shaped relationship between wage and corruptibility. This is all the more impressive as this was indeed a hypothesis formulated ex ante. Having said that, the evidence in the data isn’t all that strong, and I believe the author is a bit too confident in her claim that she has found that the relationship is U-shaped. Yes, figure 2.5 shows such a relationship, but none of the pairwise comparisons is significant (most do not even come close). The regressions are not that conclusive either (in general, I believe that if the straightforward nonparametric tests fail, there is not much point in digging deep into regressions to find something, because that something is likely to be a false significance). I think the results are interesting enough without these in my view unwarranted claims. It is not a failure if an experiment does not produce the exciting and conclusive evidence the author might have hoped for. It could be that the experiment was just underpowered to detect such a subtle effect with statistical significance. It should be noted that the repeated bribery game can produce messy data because everything is interdependent in a multi-round relationship. Let me stress that my objection does not require substantial changes. Just tone down the language a little bit. I feel that ‘The results suggest that the relationship may be U-shaped’ is more in line with the evidence than ‘The results show that the relationship is U-shaped’.
There was one minor design issue that appeared somewhat weird to me. I very much like the choice of a charity as the target for negative externalities, as compared to the various methods used in the literature. The problem is that there might be disagreement about the worthiness of the charity. For instance, politically right-leaning persons might be reluctant to donate to Greenpeace, or radical atheists might object to church-run charities. There are two ways to address this issue. One is to try to find a charity that appears uncontroversial and hope that it really is. Another is to provide a list of diverse charities and let the subject choose. The author does let each subject choose a charity, but then eventually the computer picks randomly among those charities that are chosen (I reckon this means chosen by at least one subject, so it might not exclude many candidates). Thus it might well happen that some subjects end up donating to a charity they despise. Why was this done that way? I am sure the author had good reasons for her choice, but a brief discussion would be in order.

Chapter 3

The third paper uses the same paradigm as the second. This time it is used to figure out whether the subjective socio-economic status (SSES) affects a person’s corruptibility. This theory challenges the conventional wisdom that higher salaries make officials less corrupt. It argues that people with high social status tend to see themselves as above the law and make them more willing to engage in corrupt acts. Thus increasing salaries might backfire.

The author sets up a lab experiment using the paradigm of the second paper, but this time all different wage levels for the officials are implemented in the same session. The assignment to the wage groups were through a tournament. Subjects had to answer quiz questions, and were then ranked according to their success. The lower half were assigned the role of firms (here the disadvantaged role what earnings potential is concerned), the upper half were assigned officials with wage group rising with quiz success. To reinforce the status induced by this method, the author primes the subjects for a sense of social status using methods known from psychology. This is a very clever design that brings a new methodology to the experimental study of corruption.

The results are somewhat mixed. The propensity of officials to engage in bribery generally decreases with their wage level, contrary to the SSES hypothesis. However, this can partly be explained by the lower bribes that richer officials are offered (perhaps understandable that the much poorer firms are reluctant to transfer big money to the much richer officials). When looking only at offers above a certain threshold, very rich officials are indeed more corrupt. Again, I wouldn’t count this as strong or even conclusive evidence. I always become sceptical when presented with conclusions that are based on too many qualifiers, like they only hold for a subset of subjects, only for one set of decisions etc. If the hypothesis is not supported in its original formulation, then mining the data for support by introducing many ifs can be misleading. There is no need for it anyway. The data are interesting the way they are. It’s a pity if they are not perfectly conclusive, but that’s just the way it is. Perhaps, and this is noted by the author at the end, there was just not enough power in the experiment. Resources are limited, especially for graduate students. In this paper it is the design that shines, not the conclusive results.

There is one issue which I personally have no strong feelings about, but I know many people have, and the author should be prepared to defend herself. For the quiz that determines the roles, the author uses sets of questions that have different levels of difficulty. This is to avoid confounding effects between skills and corruptibility. This is certainly clever, but I am sure not everybody agrees that there is no deception about not telling subjects about the different levels of difficulty. The design is built upon the subjects’ perception of themselves as low or high status, induced by the tournament outcome. Thus the experiment relies on the subjects believing that the game wasn’t
rigged, while in fact it was. Could subjects have been suspicious? Well, it can’t be entirely ruled out. If a subject knows that he is quite good at this kind of quizzes, but finds himself at the bottom of the list, would he feel that there is something wrong and the ranks in the tournament do not represent the true skills? I personally don’t think it’s likely, but it is not impossible.

Summary

This is a very good thesis. It definitely meets the standards for a doctoral dissertation. For the final submission only minor editing is required. The chapters could easily be converted into research articles publishable in highly ranked journals.