Supplementary Information

for

Pecunia non olet: on the self-selection into (dis)honest earning opportunities

Kai A. Konrad, Tim Lohse, Sven A. Simon

Appendix A.1

Appendix A.2

Appendix A.3

Appendix A.1

The design of our experiment faced two challenges: On one hand, our reasoning relies on the fact that any difference in the WTP between both conditions traces back to the presence of the misreporting opportunity. Hence, we minimized all other potential differences between the Honest Condition and the Untruthful Reporting Condition. However, a crucial precondition for the validity of our results was that subjects anticipated the lack/presence of the misreporting opportunity in the HC/URC when stating their WTP. This required some subtle changes in the implementation of both conditions. For example, the result of the lottery was visualized by a wheel of fortune in the URC, while there was no visualization in the HC. Moreover, the necessity to make a report per se may have affected the WTP in the URC.

In order to exclude any impact of these differences on subjects' WTP, we ran four additional sessions of a Robustness Condition (RC) and the Honest Condition.[1] This RC was identical to the URC in all aspects of the implementation, except subjects were required to make an honest report by design. Instructions were modified so that subjects were able to anticipate that they would have no opportunity to misreport. In contrast, the HC was not modified and is directly comparable to the HC in our main experiment.

Figure 7 shows the distribution of the WTP in the Robustness Condition and the Honest Condition for subjects with a consistent statement of their WTP. We focus on the condition that was first in order only (compare to section 4.1.1). Unlike for Figure 2, a visual inspection reveals no persistent 'gap' in the WTP between both conditions but rather suggests that both curves are equal. The average WTP for subjects with a WTP smaller or equal to EUR 6 in the RC is EUR 3.97 ($N = 19$) and in the HC is EUR 4.05 ($N = 20$). The difference of EUR 0.08 is not significantly different from zero (Wilcoxon-Mann-Whitney test: $p = 0.58$), and a Kolmogorov-Smirnov test does not reject the null hypothesis of equal distributions ($p = 1.00$). This result is robust to the inclusion of all subjects (EUR 5.47 vs. EUR 5.26, Wilcoxon-Mann-Whitney test: $p = 0.97$, $N = 36$ and $N = 38$).

The average within-subjects difference between both conditions (compare to section 4.1.2) is EUR 0.26 and is significantly different from zero (Wilcoxon signed-rank test: $p = 0.03$, $N = 34$). However, a majority of 88 percent of subjects have a similar WTP in both conditions (absolute difference is smaller than

[1] The sessions with a total of 74 subjects took place in January 2019 at the econlab Munich.

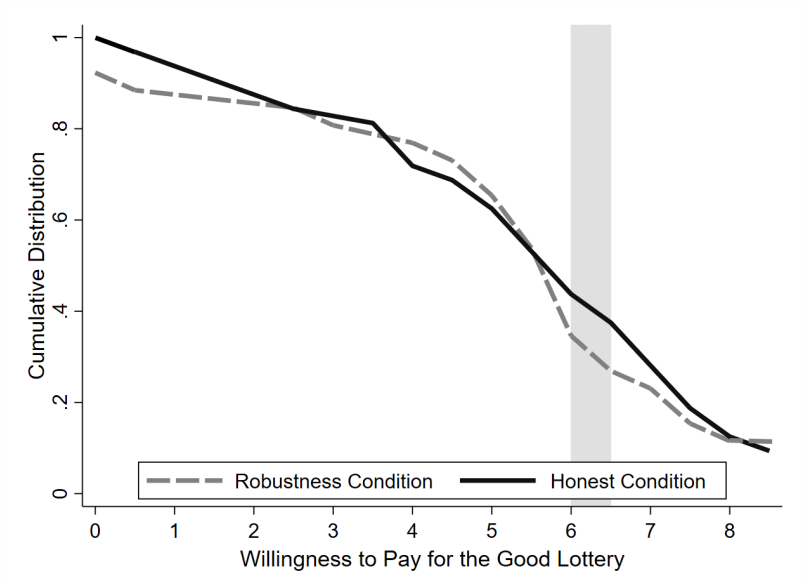

Fig. 7 Between-subjects comparison of the willingness to pay for the Good Lottery in the Robustness Condition and the Honest Condition

or equal to EUR 1), while the fraction of subjects with a negative and positive difference is 6 percent, respectively. Hence, the group of subjects with a positive difference does not outweigh the group with a negative difference, and nine out of 10 subjects have a zero difference. In the complete sample, numbers are qualitatively equivalent, but the difference is not significant (Wilcoxon signed-rank test: $p = 0.33$, $N = 74$). In summary, the comparison of the Robustness Condition and the Honest Condition reveals no evidence of a difference in the WTP between both conditions which required truthfulness by design. This suggests that the technical differences in the implementation of the URC and the HC in our main experiment did not affect the WTP in a particular direction. Instead, it is rather the opportunity to misreport which leads to treatment effects.

Appendix A.2

In our main analysis, we impose a few restrictions on our data set in order to test the predictions of our theoretical framework. One of these is making a consistent statement of the WTP, i.e., not switching back and forth for different price levels or switching in the wrong direction. A second restriction is the limitation of the WTP in the Untruthful Reporting Condition, and consequently the WTP in the Honest Condition, to EUR 6.[2] This leads to a reduction in the total number of observations. In this appendix, we report (sub)section by (sub)section our findings for the unrestricted sample and show that they are in line with our results from the main analysis. It is important to note that testing our hypotheses in the unrestricted sample comes at the cost of analyzing behavior that is outside the scope of what our theoretical framework can explain.

The willingness to pay

For the WTP in the URC and the HC, no restrictions apply and we analyze the complete sample ($N = 308$).

The demand for the Good Lottery

The pseudo between-subjects comparison of the WTP in the first decision reveals that the average WTP in the URC is EUR 4.52 ($N = 154$) and the average WTP in the HC is EUR 5.38 ($N = 154$). The difference between both conditions is significantly different from zero (Wilcoxon-Mann-Whitney test: $p = 0.04$) and a Kolmogorov-Smirnov test rejects the null hypothesis that both distributions are equal ($p = 0.01$). Hence, result 1 is robust.

The within-subjects difference in the willingness to pay

The average within-subjects difference between the HC and URC is EUR 0.56 and is significantly larger than zero (Wilcoxon signed-rank test: $p < 0.01$, $N =$

[2] Several potential explanations come to mind for why the WTP might be larger than EUR 6 in the URC. First, the WTP for subjects with high lying costs ($\theta \geq 12$) could be co-determined by risk preferences as in the HC. If a subject's WTP in the HC is larger than EUR 6, the WTP in the URC should also be larger than EUR 6. Second, some subjects may disregard the misreporting opportunity, for example, due to a lack of awareness (see Fosgaard et al. 2013 and Lohse et al. 2018). These subjects will not misreport and therefore face the same decision problem as in the HC. Third, some subjects might have kleptomaniac traits, i.e., they misreport whenever there is an opportunity to do so, even if it does not pay off in the respective situation. These subjects may want to avoid the temptation to lie and may therefore have a WTP larger than EUR 6 in the URC.

308). For the decomposition into the three groups (compare to Figure 3), we find that 55.8 percent have a similar WTP in the HC and in the URC, 28.6 percent have a higher WTP in the HC as compared to the URC, and 15.6 percent have a higher WTP in the URC as compared to the HC. The third group has a slightly larger share than in the main analysis, but qualitatively the results are the same. Hence, result 2 is robust.

Reporting behavior

As subjects with an honest win outcome ($N = 131$) have no incentive to lie, the analysis of the relationship of reporting behavior and the WTP is restricted to the group of subjects with a zero outcome ($N = 177$). In this group, we distinguish between honest subjects that truthfully reported the zero outcome ($N = 98$) and dishonest subjects that lied and reported the win outcome ($N = 79$).

The willingness to pay conditional on reporting behavior

As for the main analysis, we compare the WTP conditional on the reporting behavior. Dishonest subjects have a WTP of EUR 5.15 in the HC and of EUR 3.34 in the URC, while honest subjects have a WTP of EUR 5.39 in the HC and of EUR 5.19 in the URC (compare to Figure 4). Between honest and dishonest subjects, the difference in the URC is highly significant, while we find no such difference in the WTP in the HC (Wilcoxon-Mann-Whitney test: $p < 0.01$ and $p = 0.61$, respectively). Within-subjects and between conditions, the difference in the WTP is not significant for honest subjects but is highly significant for dishonest subjects (Wilcoxon signed-rank test: $p = 0.39$ and $p < 0.01$, respectively). Hence, result 3 is robust. The interval regression analysis on the WTP in the URC (Table 1) reveals that a higher WTP in the HC leads to a significantly higher WTP in the URC, both for honest and for dishonest subjects. This finding for dishonest subjects deviates from our result in the restricted sample, potentially due to a change in the behavioral pattern for subjects with a high WTP in the URC.

Reporting behavior conditional on the willingness to pay

Finally, we investigate the effect of self-selection on reporting behavior. Around 70 percent of subjects in the group 'Bad Lottery–self-selected' are dishonest ($N = 60$), 31 percent of subjects in the group 'Bad Lottery–assigned' are dishonest ($N = 67$), and 32 percent of subjects in the group 'Good Lottery–self-selected'

are dishonest ($N = 50$)(compare to Figure 5). The difference between the first and the second group is highly significant, as is the difference between the first and the third group (χ^2-test: $p < 0.01$, respectively). The difference between subjects with a high WTP that were involuntarily assigned to the Bad Lottery and those that were assigned to the preferred Good Lottery is small and not significant (χ^2-test: $p = 0.94$). The probit regression analysis on the reporting behavior (Table 2) reveals that a higher WTP in the URC is related to a higher propensity to misreport. Hence, result 5 and result 6 are robust.

Estimation of lying costs

As in our main analysis, we focus on subjects with the zero outcome in the URC ($N = 177$). The estimation of the lying costs reveals a pattern similar to the restricted sample: 19 percent of subjects have lying costs of zero, while 55 percent have prohibitively high lying costs and abstain from misreporting. A fraction of 12 percent of subjects have an intermediate range of lying costs between $\theta \in (0, 12)$, while 14 percent of subjects have lying costs $\theta > 12$ but are dishonest. In sum, result 7 is robust.

Appendix A.3

In our 'Robustness discussion' (section 5), we address the potential consequences of the observability of the true lottery outcome, and individual lying behavior, in our data. As a robustness test, we condition on whether a subject felt observed or not. Based on our post-experimental control question regarding on the Untruthful Reporting Condition (exact wording 'Did you feel observed at any point in time?'), we compare the behavior between subjects that did not feel observed (main analysis: 78 percent, $N = 99$; complete sample: 75 percent, $N = 230$) and those who did (22 percent, $N = 28$; 25 percent, $N = 78$), and repeat our main analysis in both subgroups. Note that this subdivision leads to a small sample size for some of the comparisons. Nevertheless, we confirm all our findings for subjects who did not feel observed, and obtain qualitatively similar results for subjects that felt observed. As this appendix only states the results, please refer to the Robustness discussion for an interpretation of the results.

Between-subjects comparison of both subgroups

For the between-subjects comparison of subjects who felt observed and those who did not, we find no significant differences in the WTP on the condition in order first (Wilcoxon-Mann-Whitney test: $p = 0.53$ and $p = 0.34$, $N = 86$ and $N = 74$, URC and HC) as well as for differences in the within-subjects comparison of the difference in both conditions (Wilcoxon-Mann-Whitney test: $p = 0.98$, $N = 121$). However, subjects that felt observed are marginally less likely to cheat (35 percent vs. 53 percent, χ^2-test: $p = 0.10$, $N = 114$) and the estimated lying costs are significantly larger (EUR 10.29 vs. EUR 7.16, Wilcoxon-Mann-Whitney test: $p = 0.02$, $N = 114$).

Replication of the analysis in both subgroups (section 4)

For the subgroup analysis, the same restrictions as in the main analysis (section 4) apply, and we further divide the sample into subjects that did not feel observed and those who did. Results for the complete sample as analyzed in Appendix A.2 are qualitatively similar.

The demand for the Good Lottery

For the pseudo between-subjects comparison of the WTP in the first condition (Figure 8), subjects who did not feel observed have a WTP of EUR 2.36 in the

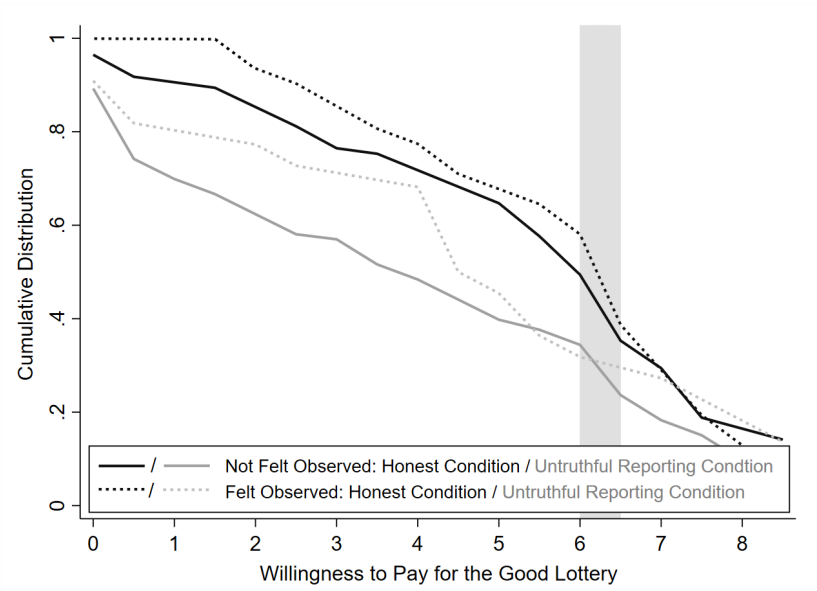

Fig. 8 Between-subjects comparison of the willingness to pay for the Good Lottery in the Untruthful Reporting Condition and the Honest Condition. Solid lines indicate subjects who did not feel observed, and dotted lines subjects who felt observed.

URC and of EUR 3.74 in the HC, and subjects who felt observed have a WTP of EUR 2.87 in the URC and of EUR 4.32 in the HC. In both subgroups, the difference between the conditions is significant (Wilcoxon-Mann-Whitney test: $p < 0.01$ and $p = 0.05$, $N = 126$ and $N = 34$). A Kolmogorov-Smirnov test also rejects the null hypothesis of equal distributions in the subgroup that did not feel observed ($p < 0.01$).

The within-subjects difference in the willingness to pay

In both the subgroup of subjects who felt observed and those who did not, the average within-subjects difference between the HC and URC is significantly larger than zero (EUR 0.94 and EUR 1.12, Wilcoxon signed-rank test: $p = 0.02$ and $p < 0.01$, $N = 27$ and $N = 94$). For the decomposition into the three groups of subjects (compare to Figure 3), we find an almost identical distribution of subjects who felt observed and those who did not: 63 and 61 percent have a similar WTP in the HC and in the URC, 33 and 34 percent have a higher WTP in the HC as compared to the URC, and 4 and 5 percent have a higher WTP in the URC as compared to the HC.

The willingness to pay conditional on reporting behavior

For clarity, we state results for each subgroup separately (compare to Figure 4). In the subgroup that did not feel observed, dishonest subjects have a WTP of EUR 3.45 in the HC and of EUR 0.73 in the URC, while honest subjects have a WTP of EUR 3.00 in the HC and of EUR 2.80 in the URC. Between honest and dishonest subjects, the difference in the URC is highly significant, while we find no such difference in the WTP in the HC (Wilcoxon-Mann-Whitney test: $p < 0.01$ and $p = 0.47$, respectively $N = 53$). Within-subjects and between conditions, the difference in the WTP is not significant for honest subjects but is highly significant for dishonest subjects (Wilcoxon signed-rank test: $p = 0.89$ and $p < 0.01$, $N = 23$ and $N = 30$). In the subgroup that felt observed, dishonest subjects have a WTP of EUR 4.67 in the HC and of EUR 2.58 in the URC, while honest subjects have a WTP of EUR 4.58 in the HC and of EUR 3.87 in the URC. Between honest and dishonest subjects, neither the difference for the URC nor for the HC is significant (Wilcoxon-Mann-Whitney test: $p = 0.30$ and $p = 0.99$, respectively $N = 18$). Nevertheless, the gap in the WTP for the URC is much larger than for the HC, and we qualitatively confirm our results in this subgroup. Within-subjects and between conditions, the difference in the WTP is not significant neither honest subjects nor dishonest subjects (Wilcoxon signed-rank test: $p = 0.33$ and $p = 0.13$, $N = 12$ and $N = 6$).

Reporting behavior conditional on the willingness to pay

For clarity, we state results for each subgroup separately (compare to Figure 5). For subjects who did not feel observed, 76 percent in the group 'Bad Lottery–self-selected' ($N = 46$) are dishonest, 22 percent in the group 'Bad Lottery–assigned' ($N = 18$) are dishonest, and 26 percent in the group 'Good Lottery–self-selected' are dishonest ($N = 19$). The difference between the first and the second lottery allocation is highly significant, as is the difference between the first and the third allocation (χ^2-test: respectively $p < 0.01$). In contrast, there is no significant difference between 'Bad Lottery–assigned' and 'Good Lottery–self-selected.' For subjects who felt observed, 50 percent in the group 'Bad Lottery–self-selected' ($N = 6$) are dishonest, 41 percent in the group 'Bad Lottery–assigned' ($N = 17$) are dishonest, and 13 percent in the group 'Good Lottery–self-selected' are dishonest ($N = 8$). None of the differences in dishonest reporting behavior reaches significance due to the low number of observations in this subgroup, but results are broadly in line with our main analysis.

Estimation of lying costs

The estimation of the lying costs reveals a diverging pattern for the subjects who did not feel observed and subjects who felt observed: 36 percent and 10 percent of subjects have lying costs of zero, while 47 and 65 percent have prohibitively high lying costs and abstain from misreporting. A fraction of 17 and 25 percent of subjects have an intermediate range of lying costs between $\theta \in (0, 12)$.