Will you accept without knowing what? The Yes-No game in the newspaper and in the lab

Werner Güth · Oliver Kirchkamp

Received: 21 June 2010 / Accepted: 28 February 2012 / Published online: 21 March 2012 © The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract In this paper we compare behaviour in a newspaper experiment with behaviour in the laboratory. Our workhorse is the Yes-No game. Unlike in ultimatum games responders of the Yes-No games do not know the proposal when deciding whether to accept or not. We use two different amounts that can be shared (100€ and 1000€). Unlike in other experiments with the ultimatum game we find a (small) effect of the size of the stakes. In line with findings for the ultimatum game, we find more generosity among women, older participants, and participants who submit their decision via postal mail than via Internet. By comparing our results with other studies (using executives or students), we demonstrate, at least for this type of game, the external validity of lab research.

Keywords Newspaper experiment · External validity · Yes-No game

JEL Classification C91 · C93

1 Introduction

One aim of this study is to learn more about external validity of laboratory experiments with student participants. Since students are similar in age and education such
laboratory experiments leave open the question how far results from the lab can be generalised.

To increase the variance of socio-demographic characteristics in the subject pool Roth et al. (1991) study the ultimatum game with students from different nationalities and find clear differences in behaviour between these groups. Murnighan and Saxon (1998) look at the behaviour of children and observe that generosity in the ultimatum game decreases with age. In a similar study with children Harbaugh et al. (2003) find that, once one controls for size, generosity increases with age. In a newspaper experiment with the ultimatum game (Güth et al. 2003) the medium of participation, Internet or postal mail, has an effect on generosity. Güth et al. (2007) look at a three-person ultimatum game and show that fairness and rejection rates increase with age. Köhler et al. (2007) play an ultimatum game with a heterogeneous sample of German adults illustrating that generosity increases with age and income. Integrating their experiment into an existing survey, the Dutch CentER panel, Bellemare et al. (2008) let their participants play either the ultimatum or the dictator game and confirm that generosity increases with age.

Using the trust game, Fehr and List (2004) compare the behaviour of students with that of CEOs who turn out to be more trusting, more trustworthy, and who punish less. Fehr et al. (2003) report data from a trust game with a randomly selected sample of German households. Bellemare and Kröger (2007) compare behaviour in the trust game played by students and households of the CentER panel to find a hump-shaped relation between age and trust, and a U-shaped relation between age and trustworthiness. Bornhorst et al. (2010) play a trust game with Ph.D. students of different nationalities and find significant differences in trust and trustworthiness between different regions of origin. Sutter and Kocher (2007) study a trust game with participants from different age groups and observe a hump-shaped relation between age and trust and increasing trustworthiness with age.1

To explain why non-students outside the lab and students within the lab behave differently, Pull (1999) and Selten (2000) argue that student participants in a lab environment react more clearly to subtle strategic details than non-students outside the lab. It is, hence, essential to compare behaviour of student participants with a more heterogeneous population in other games, especially those related to the previously explored ones by differing only in subtle details. This is what we want to do in this paper.2 We will use a very simple and abstract game and concentrate on analysing the effect of stakes and the effects of the subject pool on the offers and on the willingness to accept.

Our workhorse is the Yes-No game which is a game where proposers suggest how to share a given positive monetary amount and responders decide without knowing the proposal. From ultimatum games (see Camerer 2003, for a survey of ultimatum experiments), Yes-No games differ since responders in ultimatum games know what

---

1 Other games that have been studied with heterogeneous groups of participants include the beauty-contest game of Bosch-Domenech et al. (2002) or the prisoners’ dilemma in the TV show “Friend or Foe” studied by List (2006).

2 Other limitations of laboratory experiments are due to the controlled or artificial situation in the lab (see Levitt and List 2007). We concede that these limitations exist, however, we do not deal with these limitations here.
they accept or reject. Unlike to dictator experiments (e.g. Forsythe et al. 1994), the responder in Yes-No games still has full veto power in the sense that without his consent the pie of 100€ or 1000€ is lost.

We find the Yes-No game an interesting game for several reasons:

- Many offers or opportunities in real life contain a certain as well as an uncertain component. E.g. a work contract might specify an explicit salary but might be silent about working hours, pensions, obligations of the worker and much more. Other examples include so-called experience goods whose quality is not known to customers or partnership proposals without knowing how reliable the partner(s) will be, as e.g. in joint ventures or spouse relationships. The ultimatum game studies as one extreme the (artificial) situation of an offer with no uncertainty at all. Everything that can be said about the offer is known to the responder. The Yes-No game looks at the other extreme: A situation where the offer is entirely unknown to the responder.

- The Yes-No game also sheds new light on motives of behaviour in situations like the ultimatum game. Do proposers make generous offers in the ultimatum game because they fear rejection of a lower offer? Since the offer is unknown to the responder low offers are not more likely to be rejected than high offers. Proposers who make high offers must have other reasons.

  Even more interestingly, responders who reject unfair offers in the ultimatum game should when they expect to receive unfair offers in the Yes-No game reject these offers, too.

  We already know from ultimatum games with private information (Güth et al. 1996), where only the proposer knows whether the pie is large or small, that most proposers who could divide the large pie offered only a fair share of the small pie which was never rejected by the responder.

  Such response behaviour can be explained either by “in dubio pro reo” (the pie could be small) or by “in dubio pro meo” (better little than nothing). In the Yes-No game accepting an unknown offer can be similarly justified (Gehrig et al. 2007). But since in the Yes-No game the proposer can be more exploitative by offering only the smallest positive amount the responder may expect less what could weaken the “in dubio pro meo” argument. Will we therefore observe more rejections (No) by responders who expect a low offer? Will these rejections be more frequent if the pie is small (100€) and punishment is cheap?

- Another reason for using the Yes-No game is, of course, that it only differs from the ultimatum game by one subtle detail, namely that one does not know the proposal when exercising one’s veto power. Is this a detail overlooked more frequently by non-students than by student participants? Furthermore, is it possible that not only students but also executives are paying more attention to subtleties like these. Do we have reasons to single out important and economically relevant subgroups of non-students?

3 The extensive form of the Yes-No game is one with imperfect but complete information whereas for the ultimatum game the extensive form has perfect and complete information. Incomplete information has been explored experimentally, for instance, by Mitzkewitz and Nagel (1993).
In this paper we want to describe properties of a heterogeneous population playing the Yes-No game. Will participants with a socio-demographic background different from students as responders also rely to the same degree on “in dubio pro reo” or “in dubio pro meo”? And will they as proposers yield to the exploitation incentive or will they feel more committed to fairness concerns, at least when the pie is small (and exploitation less profitable)?

In Sect. 2, we introduce the design of the experiment and discuss some hypotheses which, in Sect. 3, are tested with the help of the rather large data set (involving altogether 1175 participants). Section 4 concludes.

2 Experiment

In this paper we will compare five different media to participate in the experiment:

- 112 participants of a traditional lab experiment run by Gehrig et al. (2007) with low stakes,
- 64 participants of a lab experiment with high stakes where only some participants are paid,
- 303 participants of a newspaper experiment with high stakes where only some participants are paid and who chose to respond by postal mail,
- 568 participants of the same newspaper experiment with high stakes who chose to make their decision via the Internet,
- 128 business executives who play essentially the same situation with high stakes and who respond with pencil and paper.

Table 1 shows characteristics of participants for the different media of participation. Figure 1 shows boxplots of the distribution of age for the media of participation where we know the age.

For the media of participation with high stakes we used the strategy vector method, i.e. all participants submitted strategies and expectations for a pie of 100\(\text{\euro}\) and for a pie of 1000\(\text{\euro}\) and for both positions in the game (X and Y). After all participants had submitted their decisions a small number of participants were selected to be rewarded by actually playing the game. First, these participants were randomly grouped into pairs of two players. One of these two players was the proposer in the Yes-No game.

4Here we refer only to what Gehrig et al. (2007) call their “first experiment series”. The games in their “second experiment series” were embedded in a bidding mechanism rendering the data less comparable to ours.

5As usual the specifics of the company will not be revealed by us.

6To keep the experiment simple we elicited only these characteristics together with the strategies and expectations for the game. Details of the implementation are shown in the online-appendix B (see Online Supplementary Material).

7One might suspect that eliciting two offers, one for 100\(\text{\euro}\) and another for 1000\(\text{\euro}\), implies a demand effect. We fully agree that the differences in behaviour between the two amounts could perhaps be smaller had we used the game method. However, using the strategy vector method greatly simplifies the implementation of the newspaper experiment. Furthermore, here we are not interested in the absolute size of the effect but rather how this effect depends on other variables.
Table 1  Characteristics of participants

| Experiment          | Num. of participants | Known sex [%] | Females [%] | Known age [%] | Mean age [years] | Median age [years] | Known profession [%] | Blue collar [%] |
|---------------------|----------------------|---------------|-------------|---------------|------------------|--------------------|----------------------|-----------------|
| Internet            | 568                  | 91.0          | 48.9        | 82.2          | 39.0             | 40.0               | 61.4                 | 13.5            |
| Mail                | 303                  | 99.3          | 57.5        | 96.0          | 49.8             | 50.0               | 65.0                 | 27.4            |
| Lab high stakes     | 64                   | 100.0         | 56.2        | 100.0         | 23.0             | 21.0               | 100.0                | 0.0             |
| Executives          | 128                  | 0.0           | 0.0         | 100.0         | 0.0              | 0.0                | 100.0                | 0.0             |
| Lab small stakes    | 112                  | 0.0           | 0.0         | 0.0           | 0.0              | 0.0                | 0.0                  | 0.0             |
| All                 | 1175                 | 75.1          | 52.4        | 70.0          | 41.6             | 42.0               | 62.8                 | 13.7            |

Fig. 1  Boxplot of age for different media of participation with high stakes

(X-player), the other the responder (Y-player). Another random draw decided which pairs had to divide 100€ and which had to divide 1000€. For each pair the X-player chose a division. To simplify the use of the strategy vector method as well as the evaluation of the questionnaires we only allowed 10 possible divisions between 5€ and 95€ when 100€ could be divided. Similarly, 10 divisions between 50€ and 950€ were allowed when 1000€ could be divided. When the Y-player chose “yes” the amount was divided according to the proposal of the X-player. In case of “no” both players received zero. Details of the implementation are explained in the online-appendix B.

A materially opportunistic responder should accept the unknown but necessarily positive offer. Anticipating such opportunism, an equally opportunistic proposer should offer the lowest possible amount. We, however, do expect only few participants to behave in line with such common opportunism.

Whether “stakes” matter is often explored by using the same stakes in rich and poor countries, i.e., stake variation relies on large discrepancies of living conditions (see, e.g. Cameron 1999). We avoid confounding “stake” and culture in our within-subjects design: the same participants decide for a small (100€) and a much larger (1000€) pie. In view of the stake independence observed for offers in the closely related ultimatum game (Hoffman et al. 1996) we expect the relative shares, offered by proposers for both pie sizes, to be similar. Hoffman et al. (1996) also observe that rejection rates in ultimatum games are lower when stakes are higher. A reason might be that “teaching fairness to proposers” might be too expensive when stakes are high. We hence expect lower rejection rates for high stakes.

We also suspect that student participants in the laboratory might be more clearly aware of the crucial aspect of the Yes-No game and that their behaviour is closer to the equilibrium solution of the game. As in Eckel and Grossman (1998, 2001) we
expect women to be more generous than men. In line with Harbaugh et al. (2003), Güth et al. (2007), Köhler et al. (2007) and Bellemare et al. (2008) we expect that older participants offer more as proposers and reject less as responders. In line with Güth et al. (2003) we expect “more fairness in the mail than in the Internet”.

3 Results

Stakes: The left part of Fig. 2 shows boxplots for offers for the different stakes and the different media of participation. For a given medium of participation the median relative offer is always the same for 100€ and for 1000€. With small stakes of 20€ small relative offers are more frequent. That, however, might be due to several parameters that were different in the 20€ small stakes experiment by Gehrig et al. (2007) for the participants of experiments with large stakes. For an amount of 100€ the average offered share is 0.376 of the entire amount, for an amount of 1000€ the average offered share is with 0.364 slightly smaller. This difference is small, but significant.

The right part of Fig. 2 shows frequencies of rejections. For three out of four media of participation (postal mail, Internet, executives) rejection rates are lower for stakes of 1000€. In the lab situation we have exactly one rejection and this for 1000€. Taken together, for 100€ altogether 6.6% of all participants reject, whereas for 1000€ only 4.5% of all participants reject. The difference is significant.

8 The data in this paper was analysed with R version 2.14.1 (2011-12-22) (R Development Core Team 2011).

9 An exact paired Wilcoxon test based on the Shift Algorithm by Streitberg and Röhmel (1986) yields a $p$-value of 0.00004, a paired $t$-test yields a $p$-value of 0.00008.

10 A one-sided Fisher’s exact test for independence of the amount yields a $p$-value of 0.0234.
Fig. 3  Average relative offers and rejection rates for males and females. Graphs show own choices (relative offer and reject) and expected choices (expOffer and expReject) of the other player. The lines are lowess-splines (Cleveland 1981) based on the defaults of R’s \texttt{plsmo} function (see R Development Core Team 2011. \texttt{plsmo} is a function that plots smoothed estimates based on a lowess smoother). The gray area is a 95%-confidence band for the average behaviour of males and females for all media of participation where we know the age.

\textbf{Lab versus field:}  While there is a small effect of stakes on rejection rates Fig. 2 reveals a large effect of the subject pool. Students in laboratory experiments are much closer to the game theoretic solution and reject rarely, regardless whether stakes are 20\(€\), 100\(€\), or 1000\(€\). Rejection rates outside the lab are significantly higher than inside the lab.\(^{11}\) There is no significant difference between rejection rates for high stakes and the experiment with small stakes done by Gehrig et al. (2007).\(^{12}\)

\textbf{Age and gender:}  The left graph in Fig. 3 shows how offers and expected offers depend on age and the gender of the decision makers. Offers clearly increase with age. Furthermore, females are more generous. This is in line with Eckel and Grossman (1998, 2001).\(^{13}\) Expected offers do not seem to depend on gender.

The right graph in Fig. 3 shows how rejection rates depend on age and gender. Also rejection rates increase with age. There does not seem to be a systematic difference between males and females in rejection rates. Expected offers are lower than actual offers and expected rejection rates are higher than actual rates. In particular female expectations of rejection rates are too pessimistic.

\textbf{Age and medium of participation:}  The left part of Fig. 4 shows how average offers depend on the age group and on the medium of participation. For media of participation where we have a large number of young and old participants (postal mail,

\(^{11}\)For high stakes we find rejection rates in the lab 0.8\% and rejection rates elsewhere 5.9\%. A Fisher’s exact test for independence yields a \(p\)-value of 0.0086.

\(^{12}\)For high stakes we find a rejection rate of 0.8\% while Gehrig et al. (2007) find for small stakes 0\%. A Fisher’s exact test for independence of lab and field yields a \(p\)-value of 1.

\(^{13}\)A mixed effects regression where we control for stakes, age, the medium of participation, gender, and the subject pool finds age and gender to be significant (see specification 2 in Table 2 in the online appendix).
Internet) offers clearly increase with age. For the remaining media of participation (lab small stakes, lab large stakes, executives) the offers are approximately consistent with the age group. This holds for actual offers as well as for expected offers and is in line with what Harbaugh et al. (2003), Güth et al. (2007), Köhler et al. (2007) and Bellemare et al. (2008) find for ultimatum games.

The right part of Fig. 4 shows the relation between rejection rates and age. Although actual rejection rates hardly depend on the stakes, expected rejection rates do so very much.14

Correlation between own choice and expectations: Do participants who make generous offers expect similar offers? And are participants who expect frequent rejections more likely to reject themselves? The answer to both questions is “yes”. The left graph in Fig. 5 shows a bubble-plot of expectations over offers. We see that on the individual level own relative offers and expected offers are clearly correlated. Participants who make small offers expect small offers from others. Participants who are generous expect generous offers from other participants. The right graph in Fig. 5 shows a mosaicplot of actual and expected rejections. Again, we find that expectations are in line with choices. Participants who expect low rejection rates seldom reject. There are not many participants who expect a high rejection rate but those who do will reject rather frequently.

---

14A mixed effects regression where we explain expected rejection rates as a function of stakes, age, the medium of participation, gender, and the subject pool finds stakes to be highly significant (see Table 5 in the online appendix).
Fig. 5  Own choices and expectations in the newspaper experiment. The area of the circles (in the left diagram) and the area of the rectangles (in the right diagram) are proportional to the number of observations. The solid line in the left diagram shows an OLS regression of expectations on offers.

4 Conclusion

Lab research is often questioned by arguing that the stakes are minor, student participants are not representative, and experimental games are far too abstract. Here we concentrate on the first two issues, i.e. whether the size of the pie matters and whether student participants are representative. The Yes-No-game is simple enough to be understood by reasonably educated newspaper readers and executives. It also captures some important aspect of life, namely the need to accept or reject some deal whose profitability has already been determined or manipulated but is not known to the responder.

With respect to the first issue, size of stakes, we have explored stake dependence by quite high pie sizes of 100\(\text{€}\) and 1000\(\text{€}\). Of course, the random selection of only 40 participants questions the stake size. Still, there is little evidence for random payment effects (see, for instance Cubitt et al. 1998). While Hoffman et al. (1996) reports no significant effect of the stake size for the ultimatum game we find for the Yes-No game a significant effect of stakes on the aggregate level. Of course, the effect is small and is only significant due to the large number of 1175 participants. Still, it is possible that also for other games the impact of stakes could be small, but different from zero. With a limited number of participants in the lab small effects might be harder to detect than with a large number of newspaper readers.

With respect to the second point, representativeness of the student population, we agree that students belong to a rather narrow age bracket. We have found three important socio-demographic variables: age, gender, and the medium of participation (postal mail vs. Internet). Here our observations confirm most of our expectations. In line with Eckel and Grossman (1998, 2001) women are more generous. Similar to Güth et al. (2003) there is more “fairness in the mail and more material opportunism in the Internet”. Consistent with Harbaugh et al. (2003), Güth et al. (2007), Köhler et al. (2007) and Bellemare et al. (2008) we find older people to be more generous.

© Springer
In particular—once we control for age—results from the field are surprisingly consistent with behaviour in the lab. Newspaper readers with the age of laboratory participants behave very similar to laboratory participants. Older participants are more generous and reject more frequently. The newspaper study can, thus, be seen as a natural extension of the laboratory. It allows us to easily access a more heterogeneous subject pool.

There are many games which so far have only been researched in the lab. We have seen in this paper that one has to be careful when extrapolating from these laboratory results to the field. We have also seen that newspaper experiments can help us to learn more about external validity since the group of participants is more heterogeneous than those in the laboratory. Newspaper experiments also include typically many more participants. This allows us to detect small, but perhaps interesting, effects.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

Bellemare, C., & Kröger, S. (2007). On representative social capital. *European Economic Review, 51*(1), 183–202.

Bellemare, C., Kröger, S., & van Soest, A. (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica, 76*(4), 815–839.

Bornhorst, F., Ichino, A., Kirchkamp, O., Schlag, K., & Winter, E. (2010). Similarities and differences when building trust: The role of cultures. *Experimental Economics, 13*(3), 260–283.

Bosch-Domenech, A., Montalvo, J. G., Nagel, R., & Satorra, A. (2002). One, two, (three), infinity, … and lab beauty-contest experiments. *American Economic Review, 92*(5), 1687–1701.

Burnham, T. C. (2007). High-testosterone men reject low ultimatum game offers. *Proceedings of the Royal Society B, 274*, 2327–2330.

Camerer, C. F. (2003). *Behavioral game theory*. New York: Princeton University Press.

Cameron, L. A. (1999). Raising the stakes in the ultimatum game: Experimental evidence from Indonesia. *Economic Inquiry, 37*(1), 47–59.

Cleveland, W. S. (1981). Lowess: A program for smoothing scatterplots by robust locally weighted regression. *The American Statistician, 35*, 54.

Cubitt, R. P., Starmer, C., & Sugden, R. (1998). On the validity of the random lottery incentive system. *Experimental Economics, 1*(2), 115–131.

Eckel, C. C., & Grossman, P. J. (1998). Are women less selfish than men?: Evidence from dictator experiments. *Economic Journal, 108*, 726–735.

Eckel, C. C., & Grossman, P. J. (2001). Chivalry and solidarity in ultimatum games. *Economic Inquiry, 39*(2), 171–188.

Fehr, E., Fischbacher, U., von Rosenbladt, B., Schupp, J., & Wagner, G. G. (2003). A nation-wide laboratory. Examining trust and trustworthiness by integrating behavioral experiments into representative surveys (CESifo Working Paper 866). CESifo GmbH.

Fehr, E., & List, J. A. (2004). The hidden costs and returns of incentives—Trust and trustworthiness among CEOs. *Journal of the European Economic Association, 2*(5), 743–771.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). The statistical analysis of experiments with simple bargaining games. *Games and Economic Behavior, 347–369.*

Gehrig, T., Güth, W., Levati, V., Levinský, R., Ockenfels, A., Uske, T., & Weiland, T. (2007). Buying a pig in a poke: An experimental study of unconditional veto power. *Journal of Economic Psychology, 28*(6), 692–703.

Greiner, B. (2004). An online recruitment system for economic experiments. In K. Kremer & V. Macho (Eds.), *Forschung und wissenschaftliches Rechnen. GWDG Bericht* (Vol. 63, pp. 79–93). Göttingen: Ges. für Wiss. Datenverarbeitung.
Güth, W., Schmidt, C., & Sutter, M. (2003). Fairness in the mail and opportunism in the Internet: A newspaper experiment on ultimatum bargaining. *German Economic Review, 4*(2), 243–265.

Güth, W., Schmidt, C., & Sutter, M. (2007). Bargaining outside the lab—A newspaper experiment of a three-person ultimatum game. *Economic Journal, 117*, 449–469.

Güth, W., Huck, S., & Ockenfels, P. (1996). Two-level ultimatum bargaining with incomplete information: An experimental study. *Economic Journal, 106*(436), 593–604.

Harbaugh, W. T., Krause, K., & Liday, S. G. (2003). *Bargaining by children* (Working paper). Economics Department, University of Oregon.

Hoffman, E., McCabe, K. A., & Smith, V. L. (1996). On expectations and the monetary stakes in ultimatum games. *International Journal of Game Theory, 25*, 289–301.

Köhler, C., Kuklys, W., Struck, O., & Fischer, S. (2007). *Ultimatum behavior, individual characteristics, and moral attitudes on the labor market* (Arbeitspapier 15, SFB 580). University of Jena.

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives, 21*(2), 153–174.

List, J. A. (2006). Friend or foe? A natural experiment of the prisoner’s dilemma. *Review of Economics and Statistics, 88*(3), 463–471.

Mitzkewitz, M., & Nagel, R. (1993). Experimental results on ultimatum games with incomplete information. *International Journal of Game Theory, 22*(2), 171–198.

Murnighan, J., & Saxon, M. S. (1998). Ultimatum bargaining by children and adults. *Journal of Economic Psychology, 19*, 415–445.

Pull, K. (1999). *What is the fair wage?—A model of as-if-co-operation* (Discussion paper 58). Quint-Essenzen, Trier.

R Development Core Team (2011). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. ISBN 3-900051-07-0

Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991). Bargaining and market power in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An experimental study. *American Economic Review, 81*, 1068–1095.

Selten, R. (2000). Thünen-Vorlesung: Eingeschränkte Rationalität und ökonomische Motivation. *Zeitschrift Für Wirtschafts- und Sozialwissenschaften, 9*, 129–157 (special issue).

Streitberg, B., & Röhmel, J. (1986). Exact distributions for permutations and rank tests: An introduction to some recently published algorithms. *Statistical Software Newsletter, 12*(1), 10–17.

Sutter, M., & Kocher, M. G. (2007). Trust and trustworthiness across different age groups. *Games and Economic Behavior, 59*, 364–382.