Do Control Questions Influence Behavior in Experiments?*

Catherine Roux[†] and Christian Thöni[‡]

March 2013

Abstract

Outcomes and strategies shown in control questions prior to experimental play may provide subjects with anchors or induce experimenter demand effects. In a Cournot oligopoly experiment we explore whether control questions influence subjects' choices in initial periods and over the course of a repeated game. We vary the framing of the control question to explore the cause of potential influences. We find no evidence for an influence of the control question on choices, neither in the first period nor later in the game.

JEL Classification: B41, C72, C91

Keywords: Control questions; Experimenter demand effects; Anchoring; Experimental design

*Financial support from the Richard Büchner-Stiftung is gratefully acknowledged. The usual disclaimer applies.

†University of St.Gallen. E-mail: catherine.roux@unisg.ch

‡University of Lausanne. E-mail: christian.thoeni@unil.ch; Web: http://sites.google.com/site/christianthoeni

1 Introduction

Control questions which test the subjects' understanding of instructions prior to experimental play are standard procedure in experimental economics. In these questions, one cannot avoid showing specific combinations of strategies and outcomes which may influence subjects' play, especially in the initial phase of the experiment.[1]

Although of considerable methodological interest, the influence of control questions on subjects' behavior has not yet been examined systematically. We investigate the effects of quantities shown in the control question on immediate actions and over the course of 20 periods of a repeated Cournot game. The Cournot game is interesting because, as opposed to many other games studied in experimental economics, the derivation of the Nash equilibrium is fairly complicated. We suppose that only a minority of the subjects is able to do the necessary calculations. Thus, potential influences of control questions are presumably of more importance in Cournot than in simpler games.

We see mainly two channels by which control questions may affect behavior. First, they can induce experimenter demand effects. Experimenter demand effects refer to changes of subjects' behavior due to cues about what constitutes appropriate behavior (Zizzo, 2010). Second, outcomes presented in control questions may serve as an anchor. Psychological research on how anchoring influences individuals' beliefs, value judgement and information processing is abundant and suggests that anchoring affects a variety of numerical estimates (see Furnham and Boo (2011) for a literature review).

We introduce a treatment variation to shed light on these two explanations. In half of our sample we inform the subjects of the fact that the situation in the control question is randomly generated. Subjects then know that the situation is not deliberately chosen by the experimenter and therefore contains no information about appropriate behavior. Thus, when the information about the random process is given, we should not observe any influence of the control question on behavior if the influence is driven by experimenter demand effects.[2] If, however,

[1] "[Q]uizzes [i.e., control questions] run the risk of giving unintended cues about your intentions to your subjects" (Friedman and Sunder, 1994, p.52).

[2] Zizzo (2010) distinguishes between social (wanting to please the experimenter) and cognitive

anchoring is the mechanism through which control questions influence choices the treatment variation should not matter.

2 Experimental Design and Procedures

Similar to the design of Huck et al (2004) $n \in \{2, 4, 6, 8\}$ subjects play a repeated Cournot game with linear demand and constant marginal costs of 2. We use a loaded frame explaining the game in terms of firms and production quantities. There are 20 identical periods with partner matching. The firms simultaneously and independently choose their quantity q_i from $[0, 74]$ at an increment of .1. The lower and upper limits of the choice set are not actively communicated. Only when the subjects try to enter a quantity outside the interval they are informed about the range of admissible quantities. Price is determined by the market demand $p = \max\{74 - Q, 0\}$ where $Q = \sum_i q_i$ is the total quantity produced in the market. At the end of each period players learn all individual quantities and profits in their market. Firm i's profit in the stage game is $\pi_i = \max\{74 - Q, 0\}q_i - 2q_i$. The Nash equilibrium is $q^n = 72/(n + 1)$, i.e., 24, 14.4, 10.3, and 8 for our markets with 2, 4, 6, and 8 firms.

Prior to the first period subjects are confronted with a control question in which they have to calculate their payoff for a combination of the own quantity (q_i^{cq}) and the average quantity of the other players ($\bar{q}_{-i}^{cq}$). The wording is:

Please consider the following situation:	
Your production quantity	10
Average production quantity of the 3 other firms	8.2
What is your profit in this situation (rounded to integers)?	___

Both quantities displayed are generated randomly from the uniform distribution in

(using all information received from the experimenter as cues about what constitutes optimal behavior) experimenter demand effects. In principle, our treatment variation shuts down both channels because a random number conveys no information about either optimal behavior or the experimenter's preferred outcome. If, however, subjects, rightly so, guessed that the purpose of the experiment is to find out how they react to random numbers, a positive effect could well be explained by social experimenter demand effects. We deem this very unlikely because the instructions were focused on the Cournot game while the control question was only a minor preparatory step.

$[0, 74/n]$.[3] We adjust the interval for group size to avoid situations with very high quantities that result in negative payoffs.[4] To facilitate the calculation subjects have access to a profit calculator which allows them to determine the payoff consequences of hypothetical combinations of their quantity and the average quantity of the other firms. The control question is the only occasion when the subjects observe a specific combination of quantities provided by the experimenters.

Our empirical strategy is to test whether the situation presented in the control question has predictive power for the quantity decisions, formally, we look for the slope of the function $q_i = f(q_i^{cq})$. To explore potential explanations for an influence of the control question on behavior we introduce a treatment variation: in *Standard* we present the subjects with the control question as shown above. In *Random* we add the following information between title and situation: "(The situation in this control question was generated purely **randomly**)". The allocation to treatments was randomized within sessions such that half of the subjects were in *Standard* and half were in *Random*.

What might be the size of the effect? While we are not aware of studies providing a quantitative estimate for the strength of experimenter demand effects, we rely on the results reported by Ariely et al (2003) to form a prior about the size of potential anchoring effects. According to their results f' should be roughly .24, i.e., whenever the anchor increases by one unit, the average willingness to pay increases by about .24 dollars.[5]

We have observations from 214 subjects. The numbers of observations are 30 (17), 64 (35), 72 (32), and 48 (23) for the markets with 2, 4, 6, and 8 firms respectively; the numbers of observations in *Random* are shown in parantheses. The sessions were run in the lab at the University of St.Gallen during February and October 2012 and were programmed in z-Tree (Fischbacher, 2007).[6] Recruitment

[3] q_i^{cq} is rounded to an integer, and $\bar{q}_{-i}^{cq}$ is rounded to a multiple of .1.

[4] Negative payoffs are still possible if both draws are close to the upper bound. In our experiment this situation did not occur.

[5] Table 1 in Ariely et al (2003, p.76) provides only quintiles. We estimate the size of the anchoring effect by regressing the average price of a quintile on the presumed mean value of the last two digits of the social security number in the respective quintile, namely $\{10, 30, 50, 70, 90\}$. A separate estimate for each of the six goods gives us $\{.22, .43, .20, .27, .19, .12\}$. The average size of the anchoring effect is thus .24.

[6] After the subjects had played the repeated Cournot game, the experiment continued with a

was done with ORSEE (Greiner, 2004). Subjects were randomly allocated to computer terminals such that they could not infer with whom they would interact. During the entire experiment communication was not allowed. We provided written instructions which informed the subjects of all the market features, and we read out loud the most important points. Profits were calculated in 'Guilders', and 1000 Guilders were worth 1.10 CHF for the duopolies. We adapted the exchange rate for the four, six and eight-firm markets such that Nash earnings were equalized across markets. Payments consisted of a show-up fee of 10 CHF plus the sum of the profits over the course of the experiment. Sessions lasted for about 105 minutes, and the average earning was about 37 CHF ($\sim$ 39 USD). The subjects were undergraduate students in economics, business, international studies and law at the University of St.Gallen, and had not participated in oligopoly experiments before.

3 Results

We start by examining the effect of control-question quantities on output choices in the first period. We pool the data from the two treatments. Figure 1 plots the relation between q_i^{cq} and q_i in the first period separately for the different markets. For orientation we add a dashed 45 degree line, the location of observations where subjects choose exactly their control-question quantity. The solid line is an OLS regression line, and we indicate the 95 percent confidence interval for the slope (shaded area). The plots show no systematic relationship. All slope coefficients are insignificant at $p > .16$.

Cournot game with punishment. The results are discussed in Roux and Thöni (2013).

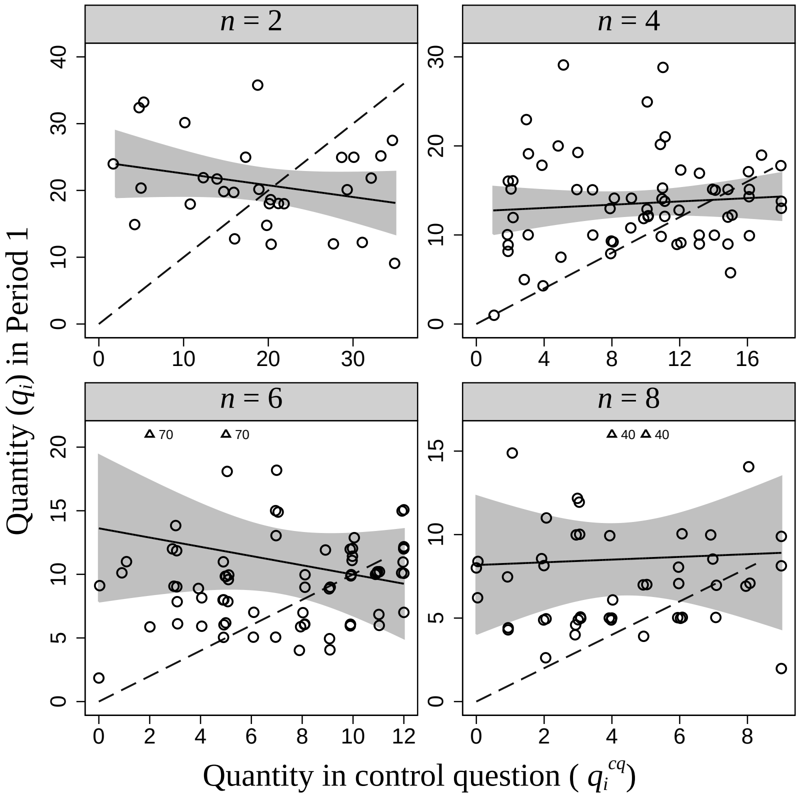

Figure 1: Quantity chosen in the first period of the Cournot game in relation to the quantity in the control question, separately for the number of firms in the market. Dashed line is the 45 degree line, solid line shows predicted quantities from an OLS regression, shaded area shows the 95% confidence interval for the slope of the regression line. Dots are jittered. Outliers with very high quantities are indicated as triangles and not drawn to scale; the small numbers denote the q_i of the outliers.

Result 1. *The quantities shown in the control question do not seem to have an effect on subjects' quantity choices in the first period of the experiment.*

Insignificant results must, of course, be interpreted with great caution. If the empirical setup does not offer sufficient power to detect the effect the lack of significant results cannot be interpreted in a meaningful way. To address this concern, we use a Monte-Carlo simulation and answer the following question: how likely is it to observe our parameter estimates if the true effect of the control-question quantity on output decisions was .24, as the results by Ariely et al (2003) suggest (see footnote 5)? We randomly draw quantities q_i^r from a normal distribution with mean and standard deviation as observed in our experiment. For each n we draw the same number of observations as in our experiment. We use our 214 control-question quantities from the experiment and calculate hypothetical quantities $q_i^h = q_i^r + .24(q_i^{cq} - q_i^r)$.[7] We ran four times 2000 regressions with q_i^h as the dependent and q_i^{cq} as the independent variable. We compare the distribution of coefficients for q_i^{cq} with the slope coefficients of the four regression lines in Figure 1. For example, for $n = 2$ we estimate a slope of $-.176$ in our sample. In the 2000 regressions on simulated data only a single estimate resulted in a coefficient lower than $-.176$, suggesting a p-value of .0005. For $n = 4$ we observe a slope of .092. This is obviously much more likely to happen if the true slope is .24, resulting in a p-value of .073. The estimated p-values for six and eight firm markets are $p = .011$ and $p = .317$. Using Fisher's method to combine the p-values from the four independent tests results in $p = .0001$, i.e., our estimates are very unlikely to be observed if the true effect of the control question on quantities is .24.[8] Even if we consider an effect only half as strong (slope of .12) we find a joint p-value of .010.[9]

Still it is premature to conclude that experimenter demand or anchoring effects are unimportant. The analysis so far does not take into account that subjects are presented with two quantities in the control question, their own and the others' average quantity. It might be that the effect runs via the latter, e.g. if subjects tend to play best response to the others' quantities shown in the control question. We use OLS estimates explaining q_i by both q_i^{cq} and $\bar{q}_{-i}^{cq}$ for each number of firms to address this issue. All eight coefficient estimates are insignificant with $p > .2$.

Our results so far suggest that the specific situation shown in the control question had no systematic influence on behavior in the first period of the Cournot game. Although *Standard* and *Random* were meant to discriminate between two causes for an effect we have just shown to be inexistent, the treatment variation is

[7] We also restrict the hypothetical quantities to feasible quantities, $q_i^h \in [0, 74]$.

[8] According to Fisher's method the p-values from k independent tests can be aggregated to a joint test, where the test statistic $-2\sum_{i=1}^{k} \ln(p_i)$ is χ^2 distributed with $2k$ degrees of freedom. See Fisher (1958, p.99f).

[9] Alternatively, we may identify the effect in the entire sample using dummies to control for the number of firms. In this model, we measure a coefficient of $-.115$ for q_i^{cq}. Assuming an effect size of .24 we find zero out of 2000 simulated estimates to result in a lower coefficient, with an effect size of .12 we find four estimates with a lower coefficient, suggesting a p-value of .002.

Table 1: OLS estimates for quantity in period 1.

	$n = 2$	$n = 4$	$n = 6$	$n = 8$
q_i^{cq}	−0.146	−0.070	−0.465	0.144
	(0.196)	(0.203)	(0.613)	(0.292)
Random × q_i^{cq}	0.060	0.287	0.164	−0.310
	(0.236)	(0.273)	(0.871)	(0.606)
$\bar{q}_{-i}^{cq}$	0.420	−0.079	−0.539	−0.027
	(0.247)	(0.175)	(0.586)	(0.175)
Random × $\bar{q}_{-i}^{cq}$	−0.625**	0.165	0.844	−1.252
	(0.264)	(0.255)	(0.684)	(1.096)
Random (D)	7.107	−5.076	−7.082	9.796
	(7.989)	(4.152)	(10.666)	(6.964)
Constant	17.486**	15.240***	18.146*	6.451***
	(7.237)	(2.922)	(9.373)	(1.657)
F-test	2.0	0.4	0.4	0.5
Prob > F	0.114	0.861	0.869	0.761
R^2 adjusted	0.222	−0.041	−0.032	0.035
N	30	64	72	48

Notes: OLS estimates for individual quantity (q_i) in period 1 with robust standard errors. Independent variables are the quantities shown in the control question (q_i^{cq} and $\bar{q}_{-i}^{cq}$), a dummy for the treatment variation *Random*, and interactions. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

still interesting. The two effects could potentially work in opposite directions and cancel each other out in the whole sample. To investigate this possibility we ran OLS estimates where we interact both control-question quantities with a dummy for the treatment variation. Table 1 shows the results. With one exception, all interaction effects as well as the dummy for *Random* are far from significance, the exception being the interaction between the treatment dummy and the other firms' quantity for $n = 2$. Given that the suggested effect is absent in the other three regressions, we believe that this simply reflects spurious correlation and should not be interpreted as an effect. The insignificant F-tests and the low adjusted R^2 also suggest that the explanatory power of the five variables is basically zero.

Result 2. *We find no difference between the treatments* Standard *and* Random. *Informing the subjects about the fact that the situation in the control question is randomly generated does not seem to influence the effect of the control question on* *behavior in the game.*

So far, we only considered data from the first period of the game. Given that we observe no systematic effect in the first period we do not expect to find one in later periods either. We run the same estimates as reported in Table 1 for all 20 periods with robust standard errors clustered on group level and controlling for period. With one exception all coefficients of control-question quantities and interactions are insignificant ($p > .1$). The significant coefficient refers to q_i^{cq} in the estimate for $n = 6$ (coefficient: $-.105$, $p = .020$). Alternatively, we model the dynamics of the quantity choice by adding lagged q_i and $\bar{q}_{-i}$ as explanatory variables. All control-question quantities are insignificant ($p > .1$), but in the estimate for $n = 2$ we observe a weakly significant effect for q_i^{cq} ($-.046$, $p = .094$). As the coefficients are insignificant or negative we find no support for a positive relation between the control-question quantity and the quantity choices of the subjects over the course of the 20 periods.

4 Conclusion

Control questions which test subjects' understanding prior to experimental play are standard in laboratory experiments. Control questions usually show specific combinations of strategies and outcomes, and thus, may influence behavior, either because subjects perceive them as cues about desired behavior or because of unconscious anchoring effects.

We investigate the influence of control questions on behavior in a repeated Cournot game. We also vary the framing of the control question to explore the likely causes of the potential effect. We find no evidence for an influence of the quantities presented in the control question on subjects' output choices, neither in the first period nor later in the game. Moreover, our treatment variation does not affect this finding.

A word of caution should be added here. Our results should not be misread as being evidence against the importance of experimenter demand effects per se. Such effects may well be introduced by factors other than control questions such as the lack of experimenter-subject anonymity. Demand effects related to perceived social pressure from the experimenter have been shown to play a role in games where fairness concerns are important (Zizzo and Flemming, 2011).[10] While these sources of experimenter demand effects may still be relevant, our results suggest that potential concerns on experimenter demand effects induced by the control questions are unwarranted.

Finally, since we investigated one particular game it remains an open question as to what extent our results carry over to other experimental games or to designs which make use of a number of control questions as opposed to a single one. In the introduction we argued that, due to its complexity, the Cournot game provides an environment which is particularly vulnerable to influences of control questions. In view of our results, we conjecture that play should not be affected by the situations shown in control questions in games simpler than the Cournot game. While it seems natural to assume that in simpler environments anchoring effects would be of minor importance, it is less clear whether the same is true for the experimenter demand channel. We leave it to future research to decide whether our conjecture holds or not.

References

Ariely D, Loewenstein G, Prelec D (2003) Coherent arbitrariness: stable demand curves without stable preferences. Quarterly Journal of Economics 118(1):73–105

Bardsley N (2008) Dictator game giving: altruism or artefact? Experimental Economics 11(2):122–133

Barmettler F, Fehr E, Zehnder C (2012) Big experimenter is watching you! Anonymity and prosocial behavior in the laboratory. Games and Economic Behavior 75(1):17–34

Fischbacher U (2007) Z-tree - zurich toolbox for readymade economic experiments. Experimental Economics 10(2):171–178

Fisher RA (1958) Statistical Methods for Research Workers, thirteenth edn. Oliver and Boyd, Edinburgh

Friedman D, Sunder S (1994) Experimental Methods: A Primer for Economists. Cambridge University Press

[10]Research on this issue is inconclusive. There is evidence that giving in dictator games is affected by experimenter demand effects via the choice set (Bardsley, 2008) and the variation of experimenter-subject anonymity (Hoffman et al, 1994), while Barmettler et al (2012) argue that the latter has no influence on measured social preferences.

Furnham A, Boo HC (2011) A literature review of the anchoring effect. Journal of Socio-Economics 40(11):35–42

Greiner B (2004) An online recruitment system for economic experiments. In: Kremer K, Macho V (eds) Forschung und wissenschaftliches Rechnen, gwdg beric edn, Göttingen, pp 1–15

Hoffman E, McCabe K, Shachat K, Smith V (1994) Preferences, property rights and anonymity in bargaining games. Games and Economic Behavior 7(3):346–380

Huck S, Normann HT, Oechssler J (2004) Two are few and four are many: number effects in experimental oligopolies. Journal of Economic Behavior & Organization 53(4):435–446

Roux C, Thöni C (2013) Collusion among many firms: the disciplinary power of targeted punishment. DEEP Working Paper 1302

Zizzo DJ (2010) Experimenter demand effects in economic experiments. Experimental Economics 13(1):75–98

Zizzo DJ, Flemming P (2011) Can experimental measures of sensitivity to social pressure predict public good contribution? Economics Letters 111(3):239–242