Measuring Experimenter Demand Using Within-designs: Pre-analysis Plan

Jonathan de Quidt[*]   Johannes Haushofer[†]   Chris Roth[‡]

May 18, 2017

1   Introduction

Experimenter demand effects pose an important challenge to understand and interpret results from laboratory and field experiments (Rosenthal, 1966; Zizzo, 2010). For example, experimenter demand effects are an important obstacle to interpreting effect sizes and they constitute a threat to the external validity of experiments. Experimenter demand effects also affect the optimal design of experiments; however, little or no empirical evidence exists that studies how design features affect the relevance of experimenter demand effects. Moreover, experimental economists have a very limited understanding of the relative importance of experimenter demand effects for different outcome measures or experimental paradigms.

Here, we employ within-designs to analyze the distribution of responses to deliberately induced demand treatments which are designed to manipulate subjects' beliefs about the intentions and desires of the experimenter.

This document proceeds as follows: first, we outline the experimental design. Then in section 3, we describe the setting as well as the sample size. In section 4, we describe the analysis we would like to conduct.

---
[*]Institute for International Economic Studies (IIES); e-mail: jonathan.dequidt@iies.su.se.
[†]Department of Psychology and Public Affairs, Princeton University; e-mail: haushofer@princeton.edu.
[‡]Department of Economics, University of Oxford and CSAE; e-mail: christopher.roth@economics.ox.ac.uk.

2 Experimental Design

At the start our respondents complete a screener (Berinsky et al., 2014) that allows us to examine whether they are paying attention to our instructions. Then, we ask our participants to complete two tasks which involve real money. Participants are told that they will be paid according to the choice they will make in one of the two tasks and that the task according to which they will be paid is determined by chance. Then, our respondents are randomly assigned to play either a dictator game or an investment game.

2.1 Dictator Game

2.1.1 Task 1: No-demand condition

We let our respondents play a standard dictator game in which they receive the following instructions:

In this task you will allocate $1 between yourself and another randomly chosen participant from MTurk who will be informed that another MTurker had the chance to split money with him or her. You will simply decide how much of the $1 you want to send to the other person. You will get to keep the rest of the money for yourself.

In the first task our respondents receive no "demand treatment" and make a choice in the dictator game.

2.1.2 Task 2: Demand treatments

Then our respondents complete the same task again but, in addition, they receive one of the following two messages:

- **Positive Demand:** *You will do us a favor if you give more to the other participant than you normally would.*

- **Negative Demand:** *You will do us a favor if you give less to the other participant than you normally would.*

2.2 Investment Game

2.2.1 Task 1: No-demand condition

We let our respondents play an investment game (Gneezy and Potters, 1997) in which they receive the following instructions:

You are endowed with $1 and you can choose how much of the $1 to keep or invest in a risky project. Money that is not invested in the risky project is yours to keep. The risky project has a 40 percent chance of success (that is a 4 out of 10).

- *If the project is successful, you will receive 3 times the amount you chose to invest.*

- *If the project is unsuccessful, you will receive nothing.*

In the first task our respondents receive no "demand treatment" and make a choice in the investment game.

2.2.2 Task 2: Demand treatments

Then our respondents complete the same task again but, in addition, they receive one of the following two messages:

- **Positive Demand:** *You will do us a favor if you invest more in the project than you normally would.*

- **Negative Demand:** *You will do us a favor if you invest less in the project than you normally would.*

2.3 Demographics

Finally, we ask participants to complete a questionnaire on demographics, which includes variables on gender, age, education, income etc.

3 Setting, Sample Size and Power

We will run our experiment on Amazon Mechanical Turk, an online platform which is widely used to conduct experiments. We will only recruit participants who currently live in the United States. Moreover, workers must have completed at least 500 HITs, and they must have an overall

rating of more than 95 percent. We will also exclude any participants from our previous "demand experiments" on MTurk. We plan to recruit 1000 participants, with 500 participants randomly assigned to each of the two different games. Half of our subjects receive the "positive demand treatment" in the second task they complete, while the other half of participants receives the "negative demand treatment".

If we compare behavior in the positive and negative demand condition we have a power of .8 to detect effect sizes of .25 of a standard deviation at α=.05 if we consider the risk and the dictator games separately. If we pool results across games we have a power of .8 to detect effect sizes of .18 of a standard deviation at α=.05. In the regressions in which we exploit within-individual variation our power for any given effect size will be higher.

4   Analysis

4.1   Baseline Balance

We will compare the mean covariates for respondents in the positive demand conditions with those of respondents in the negative demand condition. We consider the following covariates:

- gender

- age

- log income (income is the midpoint of the interval specified by the respondent)

- household size

- ethnicity (dummies for White, Black, Hispanic, and Asian)

- employment status (dummies for unemployed, part-time employed, and employed full-time)

- education (dummy for person with at least bachelor degree)

- experience on MTurk (number of HITs completed on MTurk)

- political orientation (taking value one for Republicans and zero otherwise)

We will regress each of these variables on an indicator variable taking value 1 if they are in the positive demand condition and value zero if they are in the negative demand condition.

4.2 Figures

We will separately consider the within variation in our four treatment arms:

- Investment game with positive demand

- Investment game with negative demand

- Dictator game with positive demand

- Dictator game with negative demand

We will plot the distribution of the difference in behavior between task 2 and task 1 separately for each of our main treatment arms of interest. In addition, we show scatter plots of the data with task 2 on the y-axis and task 1 on the x-axis.

4.3 Main Specification

4.3.1 Within-variation

For each of the four different treatment arms, we separately compare mean behavior in task 2 to mean behavior in task 1. In other words, we have two choices, j, for each of our respondents, i. We estimate the following equation:

$$y_{ij} = \alpha_0 + \alpha_1 \text{demand}_{ij} + \eta_i + \varepsilon_{ij} \tag{1}$$

y_{ij} are people's raw choices in the two tasks. demand_{ij} takes value 1 for the choice in which we induce a demand treatment, i.e. in task 2. $demand_{ij}$ takes value zero for choices in task 1, i.e. choices in which no demand treatment was administered. η_i are individual-level fixed effects. We will cluster standard errors at the individual level.

4.3.2 Defiers

We will also calculate the fraction of *strict defiers*, i.e. respondents who are in the positive (negative) demand treatment and whose choice in task 2 is strictly lower (higher) than their choice in task 1.

We will estimate equation 1 using the subsample of strict defiers separately for our four different treatment arms. This will provide us with the average response of defiers.

5

4.3.3 Compliers

We will also calculate the fraction of *strict compliers*, i.e. respondents who are in the positive (negative) demand treatment and whose choice in task 2 is strictly higher (lower) than their choice in task 1.

We will estimate equation 1 using the subsample of strict compliers separately for our four different treatment arms. This will provide us with the average response of compliers.

4.3.4 Sensitivity

To calculate the sensitivity of behavior, we compare behavior for participants in task 2 that receive the positive demand treatment with the behavior of participants that receive the negative demand treatment in task 2. To do so, we estimate the following equation:

$$y_i = \alpha_0 + \alpha_1 \text{positivedemand}_i + \varepsilon_i \tag{2}$$

y_i is people's choice normalized at the game level using the mean and standard deviation of behavior in task 1.[1] positivedemand_i takes value one for respondents that receive the positive demand treatment in task 2, while it takes value zero for participants who receive the negative demand treatment. We will run this regression for (i) all respondents from the dictator game, and (ii) all respondents from the investment game. In these regressions we will apply robust standard errors.

4.3.5 Comparison to previous estimates

We will compare our estimates of sensitivity (difference behavior in the positive and negative demand treatment arm) between the within-design as well as the between design. Specifically, we will compare choices normalized using the mean and standard deviation of behavior in task 1 from the within-design (outlined in this PAP) to our normalized choices[2] from the previous between-design separately for both the dictator game and the investment game. For the between-design estimates, we only consider data from experiment 1 using the strong demand treatment[3] in which choices were incentivized.

---

[1] We will also present these results based on raw choices.

[2] We normalize these choices using the mean and standard deviation from incentivized behavior in the no-demand condition.

[3] This is the same demand treatment that was applied in this experiment.

We estimate the following equation:

$$y_i = \pi_0 + \pi_1 \text{positivedemand}_i + \pi_2 \text{positivedemand}_i \times \text{between}_i + \pi_3 \text{between}_i + \varepsilon_i \qquad (3)$$

positivedemand_i takes value 1 for people in the positive demand condition, while it takes value zero for people in the negative demand condition. between_i takes value 1 for people in the between design (experiment 1) and value zero for people in the within-design (the experiment outlined in this PAP). Our key object of interest is whether $\pi_2 \neq 0$. In these regressions we will apply robust standard errors.

We will also re-estimate equation (3) from above using the raw choice data (i.e. not normalized) rather than the z-scored data.

References

Berinsky, A. J., Margolis, M. F., and Sances, M. W. (2014). Separating the shirkers from the workers? making sure respondents pay attention on self-administered surveys. *American Journal of Political Science*, 58(3):739–753.

Gneezy, U. and Potters, J. (1997). An experiment on risk taking and evaluation periods. *The Quarterly Journal of Economics*, pages 631–645.

Rosenthal, R. (1966). Experimenter Effects in Behavioral Research.

Zizzo, D. J. (2010). Experimenter Demand Effects in Economic Experiments. *Experimental Economics*, 13(1):75–98.

Experimental Instructions

You are being invited to take part in this research study. Before you decide to participate in this study, it is important that you understand why the research is being done and what it will involve. Please take the time to read the following information carefully. Please ask the researcher if there is anything that is not clear or if you need more information.

Purpose of the research:
You are asked to participate in a research study conducted by Johannes Haushofer, PhD.

Your participation in this research is completely VOLUNTARY. If you choose to participate you may subsequently withdraw from the study at any time without penalty or consequences of any kind. If you choose not to participate, that will not be penalized in any way or incur any losses.

The aim of the study is to examine social, and economic behavior. In particular, the study will test your performance and preferences on simple behavioral tasks. These tasks will be explained to you separately.

To be eligible to participate you have to be 18 years and above.

Study Procedures:
If you volunteer to participate in this study, you may be asked to make simple economic decisions, about which you will have been precisely instructed beforehand. Depending on your decisions, you may earn additional money in the experiment.

The experiment will last approximately 2 minutes.

Benefit and Risk:
There are no direct benefits to you from participating in this research study, other than the specified compensation. However, if you wish, you can send an email message to Johannes Haushofer (haushofer@princeton.edu) and we will send you a copy of any manuscripts based on the research (or summaries of our results). There are no foreseeable risks to participating in this study. If you feel uncomfortable, you may stop participating at any time.

Confidentiality:
Your participation in this study will remain confidential, and your responses will be kept private, and no identifying information such as your name, phone number etc. will be collected at any point during
this research. Please note that research data will remain confidential, accessible only to the principal
investigator and the research staff in charge of this project. To preserve anonymity, any link between
your survey response and your mTurk account will be deleted upon payment.

When the results of the research are published or discussed in conferences, no information will be included that would reveal your identity.

Compensation:

You will be compensated for participating in the study. You will receive a minimum payment of 25 cents for the 2-minute study. You may also receive additional payment for your choices in the behavioral tasks.

Who to contact with questions:

PRINCIPAL INVESTIGATOR:
    Johannes Haushofer
    Woodrow Wilson School of Public Policy and Department of Psychology
    Princeton University
    Phone: (617) 360-1605
    Email: haushofer@princeton.edu

2. If you have questions regarding your rights as a research subject, or if problems arise which you do not feel you can discuss with the Investigator, please contact the Institutional Review Board at:

Office of Research Integrity and Assurance
Human Research Protection Program
Assistant Director
Phone: (609) 258-0865
Email: irb@princeton.edu

I understand the information that was presented and that:

    A. My participation is voluntary, and I may withdraw my consent and discontinue participation in the project at any time. My refusal to participate will not result in any penalty.

    B. By signing this agreement, I do not waive any legal rights or release Princeton University, its agents, or you from liability for negligence.

By clicking "Yes, I agree" below, I give my consent to participate in this study:

○ Yes, I agree to participate in this survey.

○ No, I do not agree to participate in this survey.

[Next >>]

When a big news story breaks people often go online to get up-to-the-minute details on what is going on in the world. We want to know which websites people trust to get this information. We also want to know if people are paying attention to the question. To show that you've read this much, please ignore the question and select ABC News and Huffington Post as your two answers.

When there is a big news story, which is the one news website would you visit first? (Please only choose one)

☐ New York Times website   ☐ The Drudge Report   ☐ The Associated Press (AP) website

☐ Huffington Post   ☐ Google News   ☐ Reuters website National Public Radio (NPR) website

☐ Washington Post website   ☐ ABC News website   ☐ USA Today website

☐ CNN.com   ☐ CBS News website   ☐ None of these websites

☐ FoxNews.com   ☐ NBC News website   ☐ New York Post Online

☐ MSNBC.com   ☐ Yahoo! News

[Next >>]

Risk: Positive Demand

We will now ask you to complete two tasks which involve real money. You will be paid according to the choice you will make in one of the two tasks. The task according to which you will be paid is determined by chance.

Next >>

We will now ask you to complete task 1.

Next >>

You are endowed with $1 and you can choose how much of the $1 to keep or invest in a risky project.

Money that is not invested in the risky project is yours to keep.

The risky project has a 40 percent chance of success (that is a 4 out of 10).

- If the project is successful, you will receive 3 times the amount you chose to invest.
- If the project is unsuccessful, you will receive nothing.

What amount would you like to invest in the project?

Amount invested:

$0 _____ $1

Next >>

We will now ask you to complete task 2.

Next >>

You are endowed with $1 and you can choose how much of the $1 to keep or invest in a risky project.

Money that is not invested in the risky project is yours to keep.

The risky project has a 40 percent chance of success (that is a 4 out of 10).

- If the project is successful, you will receive 3 times the amount you chose to invest.
- If the project is unsuccessful, you will receive nothing.

You will do us a favor if you invest more in the project than you normally would.

What amount would you like to invest in the project?

Amount invested:

$0   ▭──────────────────────────────── $1

Next >>

Risk: Negative Demand

We will now ask you to complete two tasks which involve real money. You will be paid according to the choice you will make in one of the two tasks. The task according to which you will be paid is determined by chance.

Next >>

We will now ask you to complete task 1.

Next >>

You are endowed with $1 and you can choose how much of the $1 to keep or invest in a risky project.

Money that is not invested in the risky project is yours to keep.

The risky project has a 40 percent chance of success (that is a 4 out of 10).

- If the project is successful, you will receive 3 times the amount you chose to invest.
- If the project is unsuccessful, you will receive nothing.

What amount would you like to invest in the project?

Amount invested:

$0 ▢————————————————————— $1

Next >>

We will now ask you to complete task 2.

Next >>

You are endowed with $1 and you can choose how much of the $1 to keep or invest in a risky project.

Money that is not invested in the risky project is yours to keep.

The risky project has a 40 percent chance of success (that is a 4 out of 10).

- If the project is successful, you will receive 3 times the amount you chose to invest.
- If the project is unsuccessful, you will receive nothing.

You will do us a favor if you invest less in the project than you normally would.

What amount would you like to invest in the project?

Amount invested:

$0 | $1

Next >>

DG: Positive Demand

We will now ask you to complete two tasks which involve real money. You will be paid according to the choice you will make in one of the two tasks. The task according to which you will be paid is determined by chance.

Next >>

We will now ask you to complete task 1.

Next >>

In this task you will allocate $1 between yourself and another randomly chosen participant from MTurk who will be informed that another MTurker had the chance to split money with him or her.

You will simply decide how much of the $1 you want to send to the other person. You will get to keep the rest of the money for yourself.

What amount would you like to give to the other MTurk participant?

Amount sent to the other MTurker:

$0 [   ] $1

Next >>

We will now ask you to complete task 2.

Next >>

In this task you will allocate $1 between yourself and another randomly chosen participant from MTurk who will be informed that another MTurker had the chance to split money with him or her.

You will simply decide how much of the $1 you want to send to the other person. You will get to keep the rest of the money for yourself.

You will do us a favor if you give more to the other participant than you normally would.

What amount would you like to give to the other MTurk participant?

Amount sent to the other MTurker:

$0   [———————————————————————————] $1

Next >>

DG: Negative Demand

We will now ask you to complete two tasks which involve real money. You will be paid according to the choice you will make in one of the two tasks. The task according to which you will be paid is determined by chance.

Next >>

We will now ask you to complete task 1.

Next >>

In this task you will allocate $1 between yourself and another randomly chosen participant from MTurk who will be informed that another MTurker had the chance to split money with him or her.

You will simply decide how much of the $1 you want to send to the other person. You will get to keep the rest of the money for yourself.

What amount would you like to give to the other MTurk participant?

Amount sent to the other MTurker:

$0 $1

Next >>

We will now ask you to complete task 2.

Next >>

In this task you will allocate $1 between yourself and another randomly chosen participant from MTurk who will be informed that another MTurker had the chance to split money with him or her.

You will simply decide how much of the $1 you want to send to the other person. You will get to keep the rest of the money for yourself.

You will do us a favor if you give less to the other participant than you normally would.

What amount would you like to give to the other MTurk participant?

Amount sent to the other MTurker:

$0 [   ] $1

Next >>

Demographics

What was your combined annual household income (before taxes) in 2015?

- ○ Less than $10,000
- ○ $10,000 – $14,999
- ○ $15,000 – $19,999
- ○ $20,000 – $29,999
- ○ $30,000 – $39,999
- ○ $40,000 – $49,999
- ○ $50,000 – $59,999
- ○ $60,000 – $69,999
- ○ $70,000 – $79,999
- ○ $80,000 – $89,999
- ○ $90,000 – $99,999
- ○ $100,000 or more

Which of these describes your current situation most accurately?

- ○ Employed full-time
- ○ Employed part-time
- ○ Self-Employed
- ○ Unemployed and looking for a job
- ○ Unemployed but not looking for a job
- ○ Retired
- ○ Student
- ○ Other

Which ethnicity do you belong to?

- ○ White
- ○ African American
- ○ Hispanic
- ○ Asian
- ○ Native American
- ○ Arab
- ○ Other

What is your religion?

- ○ Christianity
- ○ Judaism
- ○ Islam
- ○ Hinduism
- ○ None
- ○ Other

[ ]

Next >>

Pick the category that describes you best:

- ○ Mechanical Turk is my main source of income.
- ○ I work on Mechanical Turk to supplement my income.
- ○ I work on Mechanical Turk as a hobby.
- ○ Other:

[ ]

How many people are there in your households?

[▾]

Which of these describes you most accurately?

- ○ Male
- ○ Female

Next >>

What year were you born?

[▾]

What is the highest level of education you have completed?

○ 12th grade or less
○ Graduated high school or equivalent
○ Some college, no degree
○ Associate degree
○ Bachelor's degree
○ Post-graduate degree

What category would best describe your political orientation?

○ Extremely liberal
○ Very liberal
○ Slightly liberal
○ Neutral
○ Slightly conservative
○ Very conservative
○ Extremely conservative
○ Other:
[ ]

How Many HITs have you previously completed on MTurk?

○ Between 500 and 1,000
○ Between 1,000 and 2,000
○ Between 2,000 and 3,000
○ Between 3,000 and 4,000
○ Between 4,000 and 5,000
○ Between 5,000 and 10,000
○ Between 10,000 and 15,000
○ Between 15,000 and 20,000
○ Between 20,000 and 30,000
○ Between 30,000 and 50,000
○ Between 50,000 and 100,000
○ More than 100,000

In which state do you currently reside?

[▾]

Next >>