

Pre-analysis Plan Outline for The Economic Consequences of Depression

JONATHAN DE QUIDT

Institute for International Economic Studies

JOHANNES HAUSHOFER

Princeton University

GAUTAM RAO

Harvard University

FRANK SCHILBACH

MIT

PIERRE-LUC VAUTREY

MIT

January 25, 2019

CONTENTS

1	Introduction	3
2	Research Design	3
3	Econometric Models	4
3.1	Data	4
3.2	Reduced-Form ITT Approach	4
3.2.1	Imbalances	5
3.3	IV Approach	5
3.4	Heterogenous Effect Approach	5
3.5	Correlation Analysis	5

3.6	Dropping Binary Outcomes with Bunching	6
4	List of Outcomes	7
4.1	Analysis of Beliefs Outcomes	7
4.1.1	Overview	7
4.1.2	Initial Confidence	8
4.1.3	Belief Updating	8
4.1.4	Hiring Schemes Decisions - Reservation Wage	8
4.1.5	Hiring Schemes Decisions - Belief-Dependent Option	9
4.2	"Saving a Note" Task	9
4.3	Defaults	9
4.4	Risk and Social Preferences (Revealed Preference Approach)	10
4.5	Risk, Time and Social Preferences (Global Preferences Survey modules)	10
4.6	Other Variables	11
5	Multiple Hypothesis Testing	11
5.1	Classes of Outcomes	11
5.1.1	First Stage: Current Depression Level	11
5.1.2	Beliefs: Initial Level	11
5.1.3	Beliefs: updating	11
5.1.4	Hiring Scheme Decisions	11
5.1.5	Decision-Making	12
5.1.6	Other Outcomes	12

1. INTRODUCTION

The goal of this study is to estimate the effects of treating Major Depressive Disorder ("Depression") on economic behavior. The study follows up on two recently-completed randomized control trials of low-cost psychotherapy to treat depression in Goa, India, conducted by the NGO Sangath and led by Professor Vikram Patel. We will study the impact of the treatments (and the resulting reduced levels of depression) on beliefs, preferences and decision-making, and other economic outcomes.

The families of primary outcomes that we are interested in are: (i) levels of beliefs about one's own ability relative to others, (ii) differences in belief updating upon receiving feedback, (iii) decision to apply for a short-term job offer with an uncertain prospect that depends on one's ability, (iv) revealed preferences and decision-making (time preferences, risk preferences, social preferences, default effects). We will adjust for multiple hypothesis testing within families of primary outcomes when appropriate.

Secondary outcomes include: self-reported consumption levels, human capital investments, locus of control, sleep quality, loneliness, intimate partner violence, female empowerment, and mood. This part is intentionally exploratory. We have measures for many outcomes of interest as well as sub-components of the belief measures, decisions to apply for jobs, etc. We consider any results found on these outcomes as suggestive, exploratory findings intended to guide further research focusing on specific outcomes. We will not adjust for multiple hypothesis testing on these outcomes.

Two distinct approaches will be used: an assessment of the effect of the treatment based on the original assignment to treatment and control in the two RCTs (intention to treat) on outcomes, and an instrumental variable strategy in order to quantify the causal effect of current diagnosed depression on outcomes.

2. RESEARCH DESIGN

This study follows up on two randomized controlled trials of low-cost psychotherapy run by the NGO Sangath in Goa, India. Results from the first trial, PREMIUM, have been published (Patel et al. (2017) and Weobong et al. (2017)), while the second trial, SHARE, is being analyzed. Please see the original papers as well as the two trial protocols for more details (Patel et al. (2014) and Sikander et al. (2015)).

The present follow-up study consists of two days of face-to-face interactions between participants

of the aforementioned trials and surveyors administering scripted surveys and economic tasks. We will attempt to contact all participants from both the treatment and control arms of SHARE and PREMIUM via phone and/or home visits. Therefore our target sample is composed of the 495 participants of PREMIUM and 280 participants of SHARE, for a total of 775 participants. Our final sample will be all the participants we manage to re-contact and who consent to participate.

3. ECONOMETRIC MODELS

In all regressions standard errors will be clustered at the participant level when there is more than one observation per participant.

3.1. Data

Our preferred sample is composed of PREMIUM and SHARE trial samples pooled together. However, if the analysis reveals that the “first stage” effect on current levels of depression is substantially weaker in one trial sample than in the other (this could be both due to different populations and intervention details, and due to a different amount of time since intervention), we will focus on the trial which shows impacts.

3.2. Reduced-Form ITT Approach

This is our primary approach. It analyzes the effect of the treatment (inherited from the initial studies we are following up on and randomly assigned) on a range of outcomes.

We will use the "double machine learning" approach from Chernozhukov et al. (2016) (algorithm **DML1**) to select control variables. Our primary implementation will use the Random Forest algorithm to predict outcome and treatment based on controls with a 2-folds sample splitting procedure. The sample splitting will stratify participants from SHARE and PREMIUM in the pooled approach, so that there will be a balanced number of SHARE and PREMIUM participants in the auxiliary and the main samples. The set of nuisance parameters (controls) will be composed of all the baseline variables collected before the intervention (listed in the original codebooks) as well as surveyors fixed effects (answers to multiple choice questions will be converted in a set of dummies for all possible choices).

3.2.1 Imbalances

The variation in treatment status that we exploit is inherited from two randomized control trials. However, conditional on participation in our study there may be imbalances on baseline observables (those collected for the trials SHARE and PREMIUM). We attempt to address such imbalances through the Chernozhukov et al. (2016) approach.

3.3. IV Approach

Using the current measure of depression (PHQ-9 score) as the endogenous variable and past assignment to treatment as the instrumental variable, we will provide IV estimates of the effect of depression on outcomes. We will only do so conditional on the first-stage effect being sufficiently strong.

As a secondary specification, we will also run the IV regression using past (post-treatment) values of depression (PHQ-9 score) as the endogenous variable, to understand if effects on outcomes are driven by past effect on depression. Again, we will do so conditional on the first-stage effect being sufficiently strong.

We will primarily focus on the ITT because the exclusion restriction is not entirely satisfied in our view, PHQ-9 may not perfectly capture current depression, and treating depression may have lagged effects even if differences in depression have waned.

3.4. Heterogenous Effect Approach

We will follow the approach from Chernozhukov et al. (2018) to test for heterogenous effects. When heterogenous effects are detected on an outcome we will use the method to determine the baseline characteristics of individuals that are the most and the least affected by the treatment. This specification will be considered secondary.

3.5. Correlation Analysis

We also plan to show descriptive statistics and correlations between raw outcomes and levels of depression.

3.6. Dropping Binary Outcomes with Bunching

We will drop outcomes that are binary for which we find that 90% of the subjects or more have the same outcome. Some outcome measures may turn out to be poorly calibrated ex-post and fail to generate an interior choice for most respondents. We will drop such outcomes in order to increase power on other outcomes.

4. LIST OF OUTCOMES

4.1. Analysis of Beliefs Outcomes

4.1.1 Overview

A section of our study consists of a laboratory experiment with a timed task (making as many bead bracelets as possible) and a sequence of self-assessments of one's performance relative to other participants. Participants are asked their degree of confidence to have scored in the top half of a group of participants similar to them (in terms of number of bracelets made and beads strung in the unfinished one). They report this confidence by dividing a liter of water between a beaker labeled "top half" and a beaker labeled "bottom half". The surveyor weighs the beakers to obtain a continuous belief measure.

The confidence is measured right after doing the task. The participant is then given a signal that is the truth with probability $2/3$ and a lie with probability $1/3$, and report their beliefs again. Each participant receives 5 signals.

Subjects then make a series of decisions about whether or not to accept a job related to the same task. The first decision is a BDM elicitation of the minimal wage they would accept in order to make 1000 bracelets in a month at home: on a list of wages, they have to say whether they want to accept or reject the job for each wage. A wage is then drawn at random and the relevant decision is implemented.¹

The second decision pertains to a hiring scheme that presents participants with a wage contract that takes into account their performance on the bracelet making task. The participants can choose from two options. The participant can either choose a sure option where they do not get a job but instead get a small amount of money with certainty. On the other hand they can choose a performance based option, wherein they will get a job opportunity to make 1,000 bracelets for a large wage if they are in the top half, but will get nothing if they are in the bottom half relative to the other 9 bracelet makers they are being compared to.

Individual Variables

Some of these variables are not outcomes of interest but are needed to analyze the outcomes.

¹More precisely, with a large probability nothing happens, and with a small probability a wage is drawn and the relevant decision implemented. This is made clear to participants.

-
- Number of bracelets made *Bracelets*
 - Actual rank in the randomly selected pool of comparison
 - Initial probabilistic belief of being in the top half p_0
 - Sequence of 5 signals s_1, s_2, s_3, s_4, s_5 . These are not outcomes, but rather generated randomly by the experimental software but their realizations are used in the analysis.
 - Sequence of 5 updated beliefs of being in the top half, following signals: p_1, p_2, p_3, p_4, p_5
 - Decisions made on the two hiring tasks

4.1.2 Initial Confidence

The first outcome of interest is the initial belief of the participant to have scored in the top half p_0 . Specifically, we will compute the degree of overconfidence as $y_i = p_0 - \pi_i$ where π_i is the true probability that they are in the top half given our entire participant pool and the sampling process. The average treatment effect on y_i will be estimated as described above with double machine learning to select controls.

4.1.3 Belief Updating

Belief updating will be analyzed as in Mobius et al. (2011) but allowing for distinct updating weights in treatment and control. We will drop participants with extreme reported beliefs ($\hat{\mu}_{it} = 0$ or 1). Our primary outcomes will be:

- The estimated weights on good and bad news respectively,
- the difference in these weights (asymmetric updating), and
- conservatism in updating (estimated reaction to signals when pooling positive and negative signals, relative to the bayesian benchmark).

4.1.4 Hiring Schemes Decisions - Reservation Wage

Participants first report their reservation wage for a job involving crafting 1000 bracelets over a month at home, through a BDM scheme (see the overview above). The outcome variable will be the arithmetic mean of the two wages around the switching point (the highest wage for which the job is refused and the lowest wage for which the job is accepted). This outcome is considered secondary.

4.1.5 Hiring Schemes Decisions - Belief-Dependent Option

Participants then choose between getting a lump sum of Rs. 300 for no work, or entering a hiring scheme which delivers a job to make 1,000 bracelets at home over a month for Rs. 3,000 if one is in the top half and nothing (no job, no lump sum) if one is in the bottom half. Our outcome variable here will be the binary decision to enter the risky, performance-dependent hiring scheme.

First, it will be analyzed as most other outcomes using the double machine learning approach with potential baseline controls. **Second**, we will seek to measure how much of any treatment effect can be attributed to differences in beliefs, risk preferences, and cost of effort.

4.2. "Saving a Note" Task

In this task we give participants a Rs. 100 note at the end of the first session (out of two sessions in total). The serial number of the note is recorded. The participant is told that they will receive that amount plus an additional amount upon their second visit if they show us the same Rs. 100 note that we gave them initially. We will ensure that the note they show us on the second visit is the same as the one we gave them initially by verifying that the serial number on the note is the same as the one we had recorded on their first session. The outcome variable y_i will then be a dummy capturing whether they returned with the original Rs. 100 note or not.

4.3. Defaults

During the recruitment call or visit, participants are told that they will be offered a gift on their first session: either a cheap amount of rice or Rs. 100, randomly assigned, with more people assigned to rice (80%). They are also told that by calling in before their first session, they have the opportunity to switch this gift to another one (rice or Rs. 100).

We will check that almost no participant assigned with Rs. 100 calls to switch to rice, to confirm that Rs. 100 has more value to most participants. We will then focus our attention on participants who received rice. The outcome variable y_i will then be a dummy capturing whether they called to switch to the (more valuable) Rs. 100 or not.

4.4. Risk and Social Preferences (Revealed Preference Approach)

Participants complete the following risk and social preference activities .

1. **Dictator Game:** A standard dictator game in which participants choose how much of a cash endowment to give to an unknown recipient. The outcome variable y_{it} is the amount of money sent.
2. **Risk Preference Task:** Here participants are presented with a series of choices in which they indicate whether they prefer a risky or a safe option. Participants who choose the safe option A get a fixed amount of cash for sure, while those who choose the risky option B are faced with a 50-50 coin toss that dictates their earnings. The choices the participants face vary in terms of the lower bound that they might earn after choosing the risky option B. We are interested in the point at which the participant switches from choosing the risky option to the safe option. Our outcome variable will be the arithmetic mean of the two lower bounds encircling the switching point.
3. **Loss Aversion Task:** This task is identical to Risk Preference except that the low amount in the risky option B is a loss of cash (taken from a sum that is presented as acquired). Again, we are interested in the point at which the participant switches from choosing the risky option to the safe option, and the outcome variable will be the arithmetic mean of the two lower losses encircling the switching point.

4.5. Risk, Time and Social Preferences (Global Preferences Survey modules)

The Global Preferences Survey module delivers scale-based, self-reported measures of these preferences. The numbers on these scales will be used as robustness checks, while our primary outcome measures are the revealed-preferences outcomes described above. Respondents indicate, on a 1 to 10 scale:

1. In general, how willing you are to choose uncertain outcomes in real life?
2. How willing are you to give to good causes without expecting anything in return?
3. How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?
4. How willing you are complete tasks at the earliest, and not leave them for later/postpone them?

4.6. Other Variables

In addition, we will collect information via surveys on consumption, labor supply and earnings, education and human capital investments, intimate partner violence, women's empowerment, mood, loneliness, sleep quality and locus of control. Summary scores for each of these outcomes will be computed. These scores will constitute our outcome variables. We will also collect beliefs about the effects of treatment on depression levels right at the intervention and in the follow up, and on other outcomes such as productivity.

5. MULTIPLE HYPOTHESIS TESTING

5.1. Classes of Outcomes

5.1.1 First Stage: Current Depression Level

Our first stage reflects whether treatment has long-lasting effects on Depression as measured by the currently-administered PHQ9.

5.1.2 Beliefs: Initial Level

The first primary outcome is initial overconfidence as defined above.

5.1.3 Beliefs: updating

Primary outcomes are updating coefficients in reaction to good and bad news, their difference (asymmetric updating), and the updating coefficient when pooling good and bad news (conservatism).

5.1.4 Hiring Scheme Decisions

This class of outcomes comprises one primary outcome: the decision to enter the hiring scheme that depends on performance.

5.1.5 Decision-Making

This class of outcomes includes risk, time and social preferences and susceptibility to defaults effects. Primary measures are the revealed-preference decisions. Secondary measures are the answers to the Global Preference Survey questions.

5.1.6 Other Outcomes

This class of outcomes (all secondary) is composed of consumption, education and human capital investments, intimate partner violence, women's empowerment, mood, loneliness, sleep quality and locus of control, labor supply and earnings, the reservation wage measured in our experiment (see 4.1.4) and the performance at bracelets making.

REFERENCES

- Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins**, "Double/debiased machine learning for treatment and causal parameters," *arXiv preprint arXiv:1608.00060*, 2016.
- , **Mert Demirer, Esther Duflo, and Ivan Fernandez-Val**, "Generic machine learning inference on heterogenous treatment effects in randomized experiments," Technical Report, National Bureau of Economic Research 2018.
- Mobius, Markus M, Muriel Niederle, Paul Niehaus, and Tanya S Rosenblat**, "Managing self-confidence: Theory and experimental evidence," Technical Report, National Bureau of Economic Research 2011.
- Patel, Vikram, Benedict Weobong, Abhijit Nadkarni, Helen A Weiss, Arpita Anand, Smita Naik, Bhargav Bhat, Jesina Pereira, Ricardo Araya, Sona Dimidjian et al.**, "The effectiveness and cost-effectiveness of lay counsellor-delivered psychological treatments for harmful and dependent drinking and moderate to severe depression in primary care in India: PREMIUM study protocol for randomized controlled trials," *Trials*, 2014, 15 (1), 101.
- , —, **Helen A Weiss, Arpita Anand, Bhargav Bhat, Basavraj Katti, Sona Dimidjian, Ricardo Araya, Steve D Hollon, Michael King et al.**, "The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: a randomised controlled trial," *The Lancet*, 2017, 389 (10065), 176–185.
- Sikander, Siham, Anisha Lazarus, Omer Bangash, Daniela C Fuhr, Benedict Weobong, Revathi N Krishna, Ikhlaz Ahmad, Helen A Weiss, LeShawndra Price, Atif Rahman et al.**, "The effectiveness and cost-effectiveness of the peer-delivered Thinking Healthy Programme for perinatal depression in Pakistan and India: the SHARE study protocol for randomised controlled trials," *Trials*, 2015, 16 (1), 534.
- Weobong, Benedict, Helen A Weiss, David McDaid, Daisy R Singla, Steven D Hollon, Abhijit Nadkarni, A-La Park, Bhargav Bhat, Basavraj Katti, Arpita Anand et al.**, "Sustained effectiveness and cost-effectiveness of the Healthy Activity Programme, a brief psychological treatment for depression delivered by lay counsellors in primary care: 12-month follow-up of a randomised controlled trial," *PLoS medicine*, 2017, 14 (9), e1002385.