Applied Statistical Analysis II

Replication Exercise by Claire Mooney

Based on “*Personalising Moral Reframing in Interpersonal Conversation*”

Joshua L. Kalla, Adam Seth Levine and David E. Broockman

Sourced from the Harvard Dataverse

1. **Background**

One of the authors of this study, Joshua L Kalla, described[[1]](#footnote-1) how this field study was designed to provide a basis for how canvassers can calvinize participants to take action, and possibly change their views, on the issue of abortion, which could possibly be applied for other policy issues.

1. **Details of the Survey**

The survey “*Personalising Moral Reframing in Interpersonal Conversation*” was carried out in the US state of Maine and was designed to engage voters on the issue of abortion.

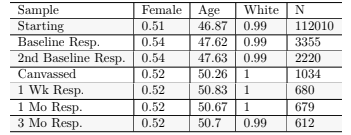
Agreement to take part in the survey was secured in advance, with over 112,000 individuals on the register of voters, representing some 71,000 households, contacted by post and invited to participate. All were invited to take the baseline survey, online. To improve precision, a second baseline survey was conducted.

In total 3,355 voters completed at least one baseline survey. Seven were subsequently removed at their own request. No inducement was provided for participation.

Canvassers, then called on households within a specified geographical area.

Subsequently follow-up surveys were issued at 1 week, 1 month and 3 months, after the canvass.

The following table sets out the numbers engaged and the representativeness of the experiment at each stage.



1. **Placebo Procedure**

In the door-to-door canvassing stage and at the subsequent 3 follow-up surveys, the voters were randomly assigned, half to the treatment and half to a placebo group.

The purpose of the placebo was to identify people who were at home and who would otherwise have been in a position to undertake the survey, having already expressed their interest to do so. The placebo provides a baseline for comparison with the group who engage with the treatment survey.

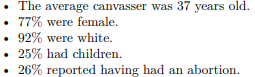
At each of the stages in the canvass, all conditions in the placebo and treatment groups were identical, up until the line of questioning.

For those in the placebo group the line of questioning was shorter and focussed on issues not related to abortion.

The average time for the treatment group interviews was 11.8 minutes (standard deviation of 8.39 minutes), as compared with an average time of under 1.99 minutes (standard deviation of 1.9 minutes), for the placebo engagements.

1. **Canvass**

The demographic profile of the canvassers, who were all volunteers, was as follows:



In addition to receiving extensive advance training on abortion issues, they were trained on how to identify moral values underpinning opinions given and to connect with the voter on shared values and on non-judgemental exchange narratives.

While canvassers engaging with the placebo group had a short-scripted line of questioning, those undertaking treatment interviews were provided with a carefully planned script for their interaction with voters involving:

* How to establish initial contact;
* Collection of initial responses;
* Engagement with follow up question - to understand voters’ experience which shaped their views;
* Identification with the values shared between the canvasser and voter;
* Further engagement on the voter’s view of what the experience a woman seeking an abortion should have;
* The initial question of how permissible abortion should be, was then asked again and, any deviation from their initial response, was discussed.

1. **Follow-up Surveys**

The surveys covered a number of political, social and cultural issues, only some of which were related to abortion.

Follow-up surveys were subsequently carried out at 1 week/1 month/3 month intervals were sent by email to any of the voters who answered their door, whether for the baseline survey or the placebo.

The below items appeared on multiple surveys; the # sign below will be replaced with the survey number in the analysis:

• The baseline survey is survey 0;

• The second baseline survey is survey 1;

• the 1-week survey is survey 2;

• the 1-month survey is survey 3, and

• the 3 month survey is survey 4.

Again, half of those sent postal surveys were sent either a baseline survey or a placebo.

1. **Recording of Responses**

A number of approaches were taken to recording responses on policy, social and stigma issues including:

* A 5 or 7 point scale was given for questions asking how much do respondents agrees or disagrees with specific statements on abortion or abortion action taking, with a neutral midpoint of “neither agree or disagree” as appropriate;
* Respondents were asked a 1-100 feeling thermometer (or therm) question for other issues such as planned parenthood or attitudes on women who have had abortions, where it was felt that scaled responses were not appropriate.

1. **Analysis of Survey**

At the pre-analysis stage, multiples indices were combined to test hypotheses. This was done to increase precision and limit the potential for multiple hypothesis testing.

Indices were created using factor analysis, with rescaling of the factors to have a mean of 0 and standard deviation of 1.

Analysis was carried out using Ordinary Least Squares regression or OLS.

1. **Reported Outputs**

The dataset had 112010 observations. The replication was originally in an R markdown file which I then transferred to an R file.

Whilst there was a large amount of data, I found the test of the design assumptions to be the most interesting. The aspect of the survey was the covariate balance among All Subjects, Compliers and Reporters.

In the paper the tables were quoted to:

“demonstrate that balance on pre-treatment observable attributes is maintained among the original universe of pre-survey respondents randomized to each group, the sub-sample that was canvassed, and the sub-sample that was both canvassed and successfully re-interviewed. Each table shows the mean value for the covariate (measured in the baseline survey before treatment) under each condition as well as the p-value from a one-way ANOVA test. The first table considers all voters who were randomly assigned after having taken the pre-survey (all subjects); the second table considers all voters who were successfully contacted (compliers); the remaining tables consider all voters who responded to the first through third post-surveys (reporters).”

The output from the survey which I replicated was the covariate balance from all responders ,on all of the principal aspects of the survey, in the third follow-up survey.

You should be aware that the output reflects the views of females **only**.

What I found interesting in this output was the higher average age of the responders.

Table

Description automatically generated

As abortion more directly affects women in age groups which we would expect them have children. I was interested in exploring **their** views and **not** the views of all.

I found it striking that in the reported survey results the average age of both the treatment and placebo respondents was over 50 years.

This is generally accepted to be above child-bearing age.

I looked to explore the dynamics of those who are of child-bearing age.

1. **Code**

Below is the code that they used to set up and produce the tables.

**Text, letter

Description automatically generated**

A screenshot of a computer

Description automatically generated**Text

Description automatically generatedText

Description automatically generated**

This is the code I ran to replicate the data for age under 50. I had already previously run all the covariate balance among all the subjects, compliers and reporters.

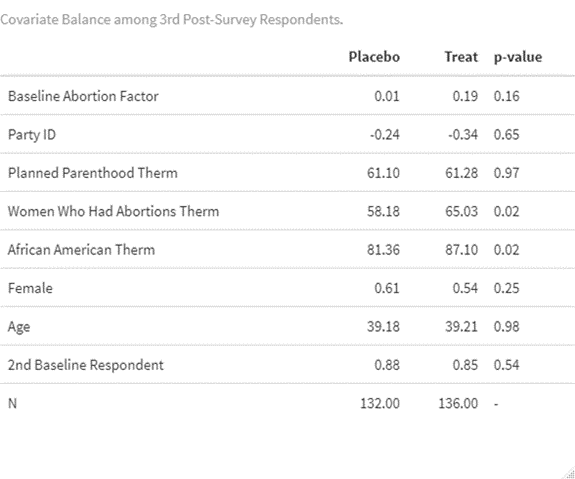
**Text

Description automatically generated**

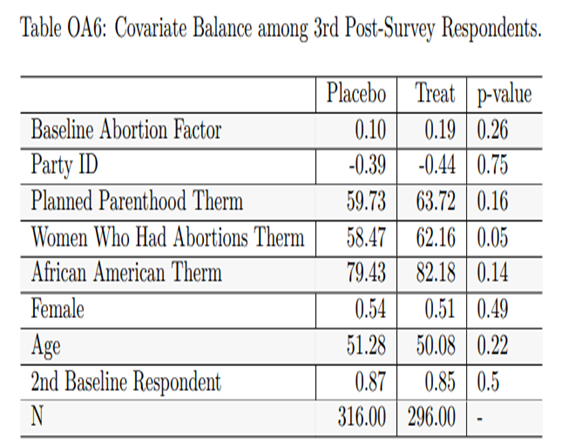
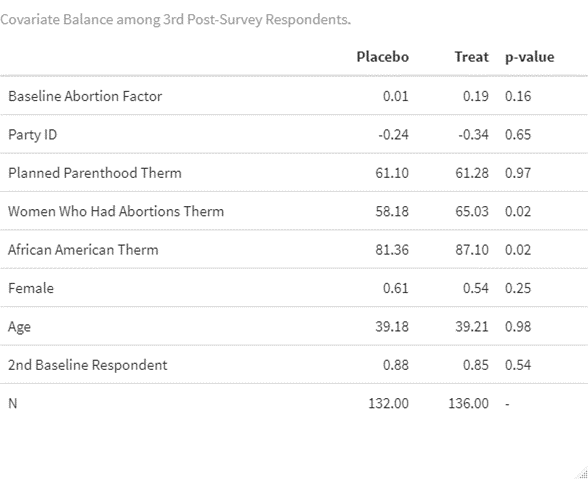
**Text

Description automatically generated**

The Output



1. **Analysis**

The Covariate Balance among 3rd Post-Survey Respondents and then my replicated results for The Covariate Balance among 3rd Post-Survey for Respondents under the age of 50.

This shows the two outputs, the replicated and then the results for when age is set at under 50.

As you can see the average age – excluding those over 50 years - is now just over 39.18 years.

What is interesting is that the P values for two of the results have changed.

For these 2 results the P value is now below 0.05.

In the reported results they were at or above this critical threshold.

Looking closer at these results -

For women who had previously had an abortion the P value has fallen from 0.05 to 0.02.

Similarly, the African American therm is now 0.02 down from the previously reported level of 0.14 based on the inclusion of all age groups.

If I was to conduct further analysis, I may look at a different population. This study was conducted in Maine and therefore the population was almost 100% white. If I was to continue my study and replication, I may choose a more diverse population.

However, my replication demonstrates that a change in the age profile to reflect those regarded as more directly engaged with the issue of abortion is more likely to give rise to question these two important outputs from the study.

1. https://politicalscience.yale.edu/news/joshua-kalla-personalizing-moral-reframing-interpersonal-conversation-field-experiment [↑](#footnote-ref-1)