

Quasi-Replication of “Alcohol and Self-Control: A Field Experiment in India”

Buddy Anderson

12/3/2020

Replication Paper

The paper partially replicated in what follows is Frank Schilbach’s *Alcohol and Self-Control: A Field Experiment in India*. As I am interested in experimental economics, particularly field experiments, this was an excellent learning opportunity for a multitude of reasons. For one, I learned a little about STATA. My experience with coding in general has been limited up to this point so exposure to a new language is always valuable, particularly a language that is widely used in academia. Secondly, I learned a little more about R than I previously knew. And arguably most importantly, I dissected an experiment and attempted to follow the reasoning of a practiced experimental economist so as to better understand what it is I might be doing for my dissertation.

An experience with STATA

- After finding an interesting paper I needed to make sure I could actually run their code
- I went through the process of running their code, which took a great deal longer than anticipated as I had to repeatedly alter certain aspects or install packages before the code would actually run without error
- For example:

```
*****
*****
*Figure 4 lower panel:
*The Impact of Incentives on Day Drinking and Overall Drinking
*Time of first drink
*****

*Buddy had to install distplot 11/02/2020
ssc install distplot, replace

* BIG ISSUE HERE

*distplot line time_first_drink if day_in_study > 4 & day_in_study < 20, ///
*mc(gsl) by(tx_group) xtitle("Time of day (24h)") ytitle("Fraction of individuals who started drinking") ///
*lc(gsl forest_green maroon) lwidth(medthick) ///
*lpattern(solid longdash dot solid) legend(label(1 "Incentives") label(2 "Choice") label(3 "Control") rows(1)) ///
*xline(18, lcolor(navy)) xsc(r(6 24)) xlabel(6(2)24) text(0.8 15.4 "Study office opens {&arr}", color(navy)) ///
*graphregion(color(white)) bgcolor(white) title("Time of First Drink") recast(line)

*graph export "$figures/3d_Time_drinks_figure_FINAL.eps", replace
```

Replicating STATA

- After successfully running the STATA code and generating the same figures and tables the authors used in their paper I shifted to replicating some of their results
- Of primary interest was graphical replication as those are the things that normally stick in your mind from a study
- The following few slides compare and contrast figures that I was able to reproduce with some level of success

Now with R

First, let's source the code from a seprate script

- The following graph was called from a script that is part of this project, but not part of the Rmarkdown document
- This is an extremely useful reference tool for future projects requiring real time updating with data!!!

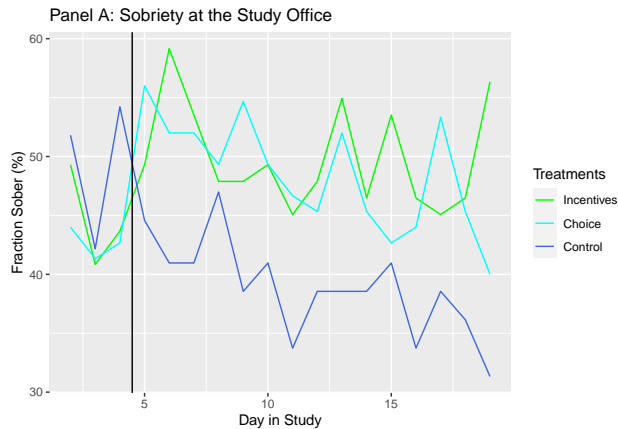


Figure 1: Graph Generated in Seprate Script

Now compare to original

- These graphs are fairly similar, but certainly more could be done to make them closer to identical

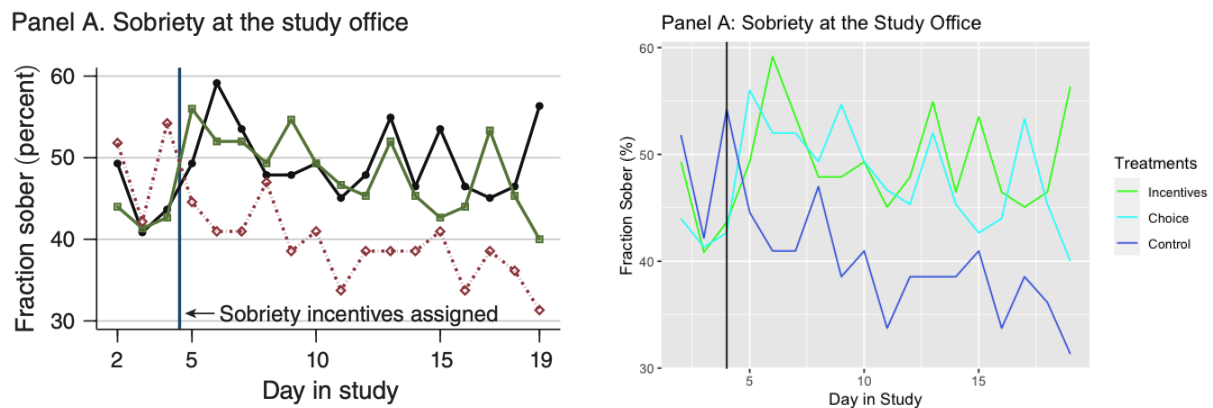


Figure 2: Comparrison of Figure 3, Panel A

Let's do the same type of thing with Table 3

Just kidding, this was advised against. We'll circle back if we have time.

Instead, lets consider an extension

- Maybe the results are being driven by heavy drinkers
- For instance, if you look at a simple histogram of the typical amount of drinks had by subjects...

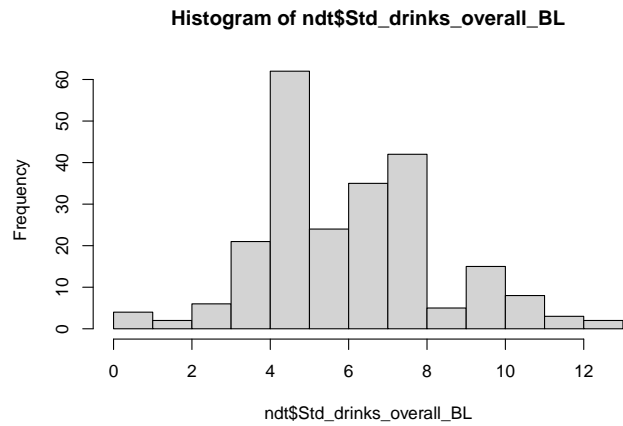


Figure 3: Histogram of Daily Standard Drinks

- ... you can see that a substantial portion of the subjects drink 7 or more drinks daily
- So what happens if we look at heterogeneous effects of “heavy” drinkers as oppsed to all others?

Heavy Drinkers vs. Others

- We define heavy drinkers as individuals that reported drinking more than 6 drinks each day
- In what follows we report the results of the same regression performed with data containing only heavy drinkers and data containing only non-heavy drinkers

Model 1

- The results of Table 1 come from the following model

Insert Model Here

- The data for this model contain only observations for heavy drinking individuals
- There are no controls here as we simply wanted to develop a working code and then make improvements to the model as to better approximate the effects of each predictor

Table 1: Heavy Drinkers

	Sober at Office	BAC	Number of Drinks
	(1)	(2)	(3)
Incentives	0.124*** (0.026)	-0.041*** (0.007)	-1.044*** (0.202)
Choice	0.069*** (0.025)	-0.010 (0.006)	-0.412** (0.190)
Control			
Constant	0.292*** (0.016)	0.117*** (0.004)	3.692*** (0.123)
Observations	1,980	1,744	1,742
R ²	0.012	0.022	0.015
Adjusted R ²	0.011	0.021	0.014
Residual Std. Error	0.473 (df = 1977)	0.112 (df = 1741)	3.376 (df = 1739)
F Statistic	11.638*** (df = 2; 1977)	19.501*** (df = 2; 1741)	13.414*** (df = 2; 1739)

Note:

*p<0.1; **p<0.05; ***p<0.01

Model 2

- The results of Table 2 come from a second model

Insert Model Here

- The data for this model contain only observations for non-heavy drinkers

Table 2: Non-heavy Drinkers

	Sober at Office	BAC	Number of Drinks
	(1)	(2)	(3)
Incentives	−0.021 (0.027)	−0.005 (0.005)	−0.149 (0.147)
Choice	0.003 (0.027)	0.008 (0.005)	−0.280* (0.145)
Control			
Constant	0.563*** (0.020)	0.050*** (0.004)	1.767*** (0.106)
Observations	2,142	1,874	1,872
R ²	0.0005	0.004	0.002
Adjusted R ²	−0.0005	0.003	0.001
Residual Std. Error	0.497 (df = 2139)	0.090 (df = 1871)	2.549 (df = 1869)
F Statistic	0.511 (df = 2; 2139)	3.764** (df = 2; 1871)	1.870 (df = 2; 1869)

Note:

*p<0.1; **p<0.05; ***p<0.01