1

(a)

False. A null hypotheses may posit one claim about our box, however our observed outcome, which may not align with the null hypotheses, *can* be explained by chance error, though this may not be the case 100% of the time. An alternative hypotheses would provide a real explanation.

(b)

False. The expectation of Y_i is the mean of the box.

(c)

False. The variance of each random variable is equal to the mean of the box.

(d)

False. Y_i represents the population mean, or the mean of the entire box universe, not the sample we choose from the box $\overline{Y}$.

(e)

False. This formula uses the variance of the sample population, the correct formula uses the variance of the box.

(f)

False. ibid. (e). Sample variance is not equal to population variance, so the denom. $n-1$ does not form an unbiased estimator as intended.

(g)

False. Significance tests inform us whether an effect is zero or non-zero but do not show impact or size of effect.

(h)

True. We can hold true that $E(Y) = (X\beta + \mu)$ given that δ_i is IID with mean 0 and variance σ^2 and $\epsilon = \mu_i + \delta_i$. Thus, ϵ_i are also independent random variables with variance σ^2.

(i)

False. Under the model in (h), $X\beta$ and μ cannot be separated from any observable data we may possess.

(j)

True. Although we may observe correlation across ε_i and ε_j, the marginal probability that subject i reads a book is still $\Phi(X_i\beta)$.

(k)

False. ϵ_i and ϵ_j are depdendent random variables, i.e. correlated. Thus, the likelihood is a product of non-independent probabilities.

(l)

False. In a probit model, the marginal effect depends on the derivative of $\phi(X_i\beta)$. We cannot multiply $\hat{\beta}$ by a change in income to get the change in probability.

(m)

No, she cannot calculate the change in predicted probability. The error lies in the pre-treatment variable X_i. This can correlate with Z_i and affect the coefficient estimate we need for any model to hold.

(n)

False. Independent of baseline characteristics, randomization is expected to eliminate concerns of confounding.

(o)

False. The treatment group would instead estimate the proportion of compliers. Further, we assume that treatment is as-if randomly assigned; the exclusion restriction is not necessary.

(p)

True. If $X_i \neq X_j$ for some i, j, then the design matrix has full rank.

(q)

False. Mechanically, the residuals in vector $e$are orthogonal to the matrix, it need not OLS assumptions to hold true.

(r)

True. In all correctly designed models with an intercept, $\sum e_i = 0$.

(s)

False. Certain conditions allow block randomization to be effective. Mainly, potential outcomes should be similar within blocks for randomization to work as intended.

(t)

False. Very few things are guaranteed in life, this is not one of them. More seriously, diff-in-diff reduces variance to a diff-in-means if and only if pre-treatment outcomes predict post-treatment outcomes. More analysis is necessary for this to be true.

(u)

True. Manipulating the running variable in an RD design violates the randomization necessary near the cutoff point that makes the design useful in the first place.

(v)

True. ATE can only be estimated consistently if SUTVA holds. However, if we can calculate the average outcome if everyone is assigned to treatment minus if everyone is not assigned to treatment (and nobody defies), then we can more clearly calculate ATE with interference.

(w)

True. CACE can also be represented by ITT divided by the compliance rate, which is a percentage between 0-1. Because of this, mechanically, CACE will always be at least as large as the ITT value.

(x)

True under certain conditions. If T_i is randomized and X_i is a pre-treatment variable, regressing Y_i on T_i and X_i gives an unbiased estimator of T_i's effect.

(y)

True. Without proper weighting, exposure to spillover is not random. Thus biasing estimates of spillover effects.

(z)

False assuming the dropout rates are large enough. This decision clearly flaws the analysis. Since the dropout rates are different between treatment and control, this clearly biases the ATE estimate. If n is very large, then maybe a difference of 1-2 between treatment and control and be mitigated by the researcher, but if rates differ by, for instance, 100 between T & C, then this is clear cause for concern.

2

I would argue that the statement the researcher posits is valid in many regards, but still requires qualification and is problematic, namely in the baseline measurement of income that is our intercept. Assuming the regression is sound, I have no problems with the relationship of an $18,000 increase, on average, with every increase of 1 education level as coded in the problem. However, we run into a concern regarding lack of observations of an individual with no income.

The researcher specifically claims that a person with "no education" would expect to earn the baseline $22,000 in income per year. However, the researcher randomly assigned education with middle school as the "lowest" amount of education one could receive – this is different from no education at all, which is that the researcher claims to be reflected in the equation. If the researcher was able to randomly assign individuals to complete no education at all, not even middle or elementary school, then the claim may be valid and we could more confidently accept the baseline intercept. However, this raises ethical and logistical concerns. Depriving someone of all education in life for an experiment and hoping they organically find work seems against the ethics of scientists.

3

(a)

```
#add values from question
n_treat <- 510
n_control <- 448
mean_treat <- 2.34
mean_control <- 1.87
sd_treat <- 2.63
sd_control <- 2.41

# ATE
ate <- mean_treat-mean_control
ate #same as problem
```

```
## [1] 0.47
```

```
#se of ATE for CI calculation
se <- sqrt((sd_treat^2/n_treat)+(sd_control^2/n_control))
se
```

```
## [1] 0.1628713
```

```
#calculate 95% CI for ATE
lower <- ate-1.96*se
upper <- ate+1.96*se
lower
```

```
## [1] 0.1507723
```

```
upper
```

```
## [1] 0.7892277
```

Yes, I believe that the normal approximation is valid in construction this CI. Our sizes for treatment and control, 510 and 448 respectively, allow us to be confident in the central limit theorem, which supports the idea that we can use normal approximation.

(b)

The interval I constructed in part (a) is similar to the one reported by Gerber and Green, though not identical. Since G&G do not use normal approximation when reporting their CI, it is not unusual that our results are not identical. The approach they report does not rely on normal assumptions and is more explicit about the usage of random assignment in the research design, However, when we construct new intervals with normal approximation, the very similar results tell us that our new design is not immediately invalid. I believe that because our results are similar, we can believe both to be sound since there is little variation in our results. If the results from both approaches were different, then further analysis would be needed to determine the best approach.

4

(a)

```
#seed from replication code
set.seed(128)
#load data
vdl_data <- read_csv("all_samples_public.csv")
```

```
## Rows: 5247 Columns: 216
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (216): id, vietdraft, year, cohort, sex, race, vetyears, liveblks, racpu...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
#filter for correct years given section 1
vdl_data_restricted <- vdl_data %>%
  filter(cohort >= 1950 & cohort <= 1952, sex == 1)

#run regression with new outcome variable
model_liveblks <- lm(liveblks_s ~ vietdraft+factor(year), data=vdl_data_restricted)

#summary output to compare
summary(model_liveblks)
```

```
##
## Call:
## lm(formula = liveblks_s ~ vietdraft + factor(year), data = vdl_data_restricted)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.79145 -0.69334  0.05359  0.48249  2.87569
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)      -0.38676    0.15727  -2.459 0.014281 *
## vietdraft        -0.14881    0.09506  -1.565 0.118140
```

```
## factor(year)1996  1.09485    0.28943   3.783 0.000175 ***
## factor(year)1998  0.31219    0.21545   1.449 0.148001
## factor(year)2000  0.40049    0.20890   1.917 0.055822 .
## factor(year)2002  0.69187    0.26883   2.574 0.010367 *
## factor(year)2004  0.41330    0.25348   1.630 0.103668
## factor(year)2006  0.43694    0.21424   2.039 0.041961 *
## factor(year)2008  0.71370    0.23117   3.087 0.002138 **
## factor(year)2010  0.47227    0.24200   1.951 0.051591 .
## factor(year)2012  0.50286    0.24860   2.023 0.043662 *
## factor(year)2014  0.21478    0.24200   0.888 0.375259
## factor(year)2016  0.63818    0.23518   2.714 0.006900 **
## factor(year)2018  0.25569    0.25397   1.007 0.314557
## factor(year)2021  0.57771    0.19835   2.913 0.003754 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.016 on 471 degrees of freedom
##   (923 observations deleted due to missingness)
## Multiple R-squared:  0.0542, Adjusted R-squared:  0.02609
## F-statistic: 1.928 on 14 and 471 DF,  p-value: 0.02183
```

##(b)

```
#create vetstatus
vdl_data <- vdl_data %>%
  mutate(vetstatus=ifelse(vetyears > 0, 1, 0))

#subset cohorts
cohort_70 <- filter(vdl_data, cohort >= 1944 & cohort <= 1950)
cohort_71 <- filter(vdl_data, cohort == 1951)
cohort_72 <- filter(vdl_data, cohort == 1952)

#models for seperate cohorts
model_70 <- lm(vetstatus ~ vietdraft+factor(year), data = cohort_70)
model_71 <- lm(vetstatus ~ vietdraft+factor(year), data = cohort_71)
model_72 <- lm(vetstatus ~ vietdraft+factor(year), data = cohort_72)

#compare models
summary(model_70)
```

```
##
## Call:
## lm(formula = vetstatus ~ vietdraft + factor(year), data = cohort_70)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5047 -0.4143 -0.3617  0.5592  0.6705
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       0.414291   0.046917   8.830   <2e-16 ***
## vietdraft         0.039142   0.025431   1.539    0.124
## factor(year)1982 -0.062416   0.069319  -0.900    0.368
```

```
## factor(year)1983 -0.084759   0.068669  -1.234    0.217
## factor(year)1984 -0.055047   0.068362  -0.805    0.421
## factor(year)1985 -0.052549   0.068291  -0.769    0.442
## factor(year)1988  0.020938   0.072335   0.289    0.772
## factor(year)1989  0.050246   0.079175   0.635    0.526
## factor(year)1990 -0.047017   0.078781  -0.597    0.551
## factor(year)1991 -0.044543   0.074861  -0.595    0.552
## factor(year)1993 -0.002189   0.076264  -0.029    0.977
## factor(year)1994  0.051244   0.105336   0.486    0.627
## factor(year)2010 -0.079695   0.067869  -1.174    0.240
## factor(year)2012 -0.007226   0.068920  -0.105    0.917
## factor(year)2014 -0.012677   0.066764  -0.190    0.849
## factor(year)2016  0.012133   0.064869   0.187    0.852
## factor(year)2018 -0.037064   0.065512  -0.566    0.572
## factor(year)2021  0.051097   0.059109   0.864    0.387
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4937 on 1509 degrees of freedom
##   (1465 observations deleted due to missingness)
## Multiple R-squared:  0.009244,   Adjusted R-squared:  -0.001918
## F-statistic: 0.8282 on 17 and 1509 DF,  p-value: 0.6612
```

```
summary(model_71)
```

```
##
## Call:
## lm(formula = vetstatus ~ vietdraft + factor(year), data = cohort_71)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4505 -0.2031 -0.1345 -0.0748  1.0003
##
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)       0.20314    0.07091   2.865  0.00437 **
## vietdraft         0.10080    0.03657   2.756  0.00609 **
## factor(year)1982  0.03910    0.09135   0.428  0.66888
## factor(year)1983 -0.11943    0.10220  -1.169  0.24319
## factor(year)1984 -0.20339    0.09604  -2.118  0.03475 *
## factor(year)1985 -0.05748    0.10352  -0.555  0.57903
## factor(year)1988 -0.06868    0.10456  -0.657  0.51164
## factor(year)1989 -0.12834    0.10894  -1.178  0.23940
## factor(year)1990  0.08539    0.14290   0.598  0.55046
## factor(year)1991 -0.15093    0.11065  -1.364  0.17326
## factor(year)1993  0.14654    0.11236   1.304  0.19285
## factor(year)1994 -0.17638    0.13293  -1.327  0.18525
## factor(year)2010 -0.06897    0.09834  -0.701  0.48348
## factor(year)2012 -0.10345    0.09834  -1.052  0.29340
## factor(year)2014 -0.05903    0.10015  -0.589  0.55590
## factor(year)2016 -0.12630    0.09403  -1.343  0.17991
## factor(year)2018 -0.10829    0.10735  -1.009  0.31368
## factor(year)2021 -0.04812    0.08366  -0.575  0.56547
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3745 on 438 degrees of freedom
##   (460 observations deleted due to missingness)
## Multiple R-squared:  0.05966,    Adjusted R-squared:  0.02316
## F-statistic: 1.635 on 17 and 438 DF,  p-value: 0.05248
```

```
summary(model_72)
```

```
##
## Call:
## lm(formula = vetstatus ~ vietdraft + factor(year), data = cohort_72)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.22280 -0.14106 -0.12348 -0.08943  0.93042
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        0.111603   0.051369   2.173   0.0303 *
## vietdraft         -0.002214   0.034238  -0.065   0.9485
## factor(year)1982   0.028024   0.076299   0.367   0.7136
## factor(year)1983   0.071020   0.078255   0.908   0.3646
## factor(year)1984  -0.042027   0.081174  -0.518   0.6049
## factor(year)1985  -0.003986   0.082030  -0.049   0.9613
## factor(year)1988   0.014781   0.100149   0.148   0.8827
## factor(year)1989   0.066170   0.097825   0.676   0.4991
## factor(year)1990   0.025465   0.088718   0.287   0.7742
## factor(year)1991   0.039172   0.091690   0.427   0.6694
## factor(year)1993  -0.033999   0.107347  -0.317   0.7516
## factor(year)1994  -0.110773   0.130866  -0.846   0.3977
## factor(year)2010   0.014089   0.078866   0.179   0.8583
## factor(year)2012   0.031676   0.090070   0.352   0.7252
## factor(year)2014   0.089209   0.080478   1.108   0.2682
## factor(year)2016   0.111193   0.082970   1.340   0.1809
## factor(year)2018   0.018072   0.079579   0.227   0.8205
## factor(year)2021  -0.019956   0.065993  -0.302   0.7625
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3408 on 456 degrees of freedom
##   (466 observations deleted due to missingness)
## Multiple R-squared:  0.01726,    Adjusted R-squared:  -0.01938
## F-statistic: 0.471 on 17 and 456 DF,  p-value: 0.9651
```

(c)

No, the results are not consistent with randomly scrambled treatment labels. Draft eligibility strongly predicts military service among the key 1951 birth cohort, demonstrating that the lottery assignment meaningfully influenced life outcomes. Although compliance is weaker among other cohorts, especially 1952, the presence of a strong first stage in the main eligible group supports the interpretation that treatment labels were consequential, not scrambled.

(d)

I believe that non-interference assumption is likely to hold, but cannot be absolutely guaranteed. There is no clear and present medium in which I can envision the draft eligibility affects the outcome variable of interest of another individuals. In this context, the draft eligibility of one person only affects that person. Unless there were some channel in which this data was widely available during the time, which I find hard to believe, I believe that SUTVA holds.

(e)

The exclusion restriction would be violated if draft eligibility influenced attitudes through channels that are not military service, which may be unobserved in our context. Beyond the scope of the model, it may be possible that eligibility itself changed personal outcomes and life trajectories, but I do not see that as plausible. However, if alternative pathways of life based on eligibility are significant, then we have greater reason to believe that the exclusion restriction is violated.

(f)

When draft eligibility had little or no effect on military service (as in early birth cohorts before 1950 and late cohorts after 1952), we observe no significant effects of eligibility on racial attitudes. This null result in placebo cohorts supports the argument that the observed effects among the primary eligible cohorts (1950–1952) are indeed operating through military service, not some confounding factor correlated with birth cohort or survey year. If eligibility were correlated with racial views for reasons unrelated to military service, we would expect to see effects across all cohorts, not just the ones where eligibility altered life experiences. Thus, the placebo tests strengthen the internal validity of the ITT analysis.

(g)

Draft eligibility does predict military service among the main cohorts, and the placebo tests offer some reassurance that the variation mattered where it should. Still, potential violations of the exclusion restriction like changes in education, employment, etc. make it harder to be certain that military service is the only channel linking draft eligibility to racial attitudes. Because of these concerns, the ITT estimates should probably be interpreted as the effect of being draft-eligible, rather than purely the effect of serving. The analysis moves us closer to understanding the relationship, but it does not completely resolve all identification challenges.

5

(a)

Given the models above, the direct effect of the treatment can be identified through β_3, which represents being assigned treatment with the likelihood of signing the petition, with a constant level of concern across all individuals. Indirect effects can be identified through β_1 and β_4, where the former captures the effect of treatment on concern for women's violence and the latter captures the effect of such concern on signing the petition.

(b)

The two main assumptions necessary in this context are the monotonicity and exclusion restrictions assumptions. For monotonicity, the researcher must ensure that there is total compliance and no defiers exist. If this is in a laboratory setting where the researcher can manipulate the experiment, I think this assumption is plausibly true and holds. For the exclusion restriction, we must ensure that the instrument affects the outcome only through treatment. More specifically, assignment to the video needs to influence petition signing only through changes in attitudes of women's safety from the video.

(c)

I would agree with the researchers assessment given the results presented. Given the experiment design is large in observations and the analysis is well-powered, I am confident in the conclusion that there is little support for an indirect effect through concern given β_4 is small and statistically insignificant. Since we are discussing indirect effects here, any value of direct effects, even if large or significant, are not applicable here. A direct effect can be observed to support a number of alternative hyptheses, but here, a specific indirect effect cannot be supported.

(d)

The pattern suggests that concern for women's safety might mediate the effect of the information treatment on petition-signing. Concern and signing rates are highest when participants see both the harassment statistics and a testimonial, lower for stats-only and testimonial-only groups, and lowest in the control group. This pattern fits the idea that increased concern is an important part of why participants decide to sign.

However, this does not necessarily prove mediation. For that inference to be valid, you would need sequential ignorability to hold, accurate measurement of the mediator, no treatment–mediator interaction unless explicitly modeled, and the exclusion restriction, meaning the treatment affects signing only through concern, not some other path. If any of those assumptions are wrong, then it becomes less clear to make a clean causal claim.

(e)

One ethical consideration is exposing potentially unsuspecting participants to violent or disturbing content without knowing. Without knowing the complete history of each individual participating, it is impossible to know how one will react. One way researchers could mitigate this without violating any necessary of our assumptions, researchers could warn that the experiment does involve sensitive subjects, but not necessarily violence against women. This could help "screen" individuals who may know that they are more sensitive to such disturbing content without exposing what the treatment is.