Political Methodology Comprehensive Examination, August 2017
Department of Political Science, George Washington University

Instructions: *Read all questions before answering any of them. When you use substantive examples in your answers, we strongly prefer to see examples from political science. Answer all questions in part I. Answer 3 questions in part II. Feel free to hand write answers in a blue book, but carefully label those answers and note that you are using the blue book in your typed document. Good luck!*

Part I

1. Suppose you are on a research team testing the effect of small fluctuations in oil revenues (say, as oil prices shift) on governance in a set of countries. One of your co-authors says, "Hey, why limit ourselves to small fluctuations? We should include big shifts from oil discoveries." Another co-author disagrees and says that such events are rare and you should focus on "typical" fluctuations. What are some of the advantages and disadvantages of each approach? Which would you choose?

2. A state college admits students only if they score above 400 on a standardized achievement test. Applicants from group A have a mean of 500 and a standard deviation of 100 on this test, and applicants from group B have a mean of 450 and a standard deviation of 100. Both distributions are approximately normal, and both groups have the same size.

  (a) Find the proportion not admitted for each group.

  (b) Of the students who are not admitted, what proportion are from group B?

  (c) A state legislator proposes that the college lower the cutoff point for admission to 300, thinking that the proportion of the students who are not admitted who are from group B would decrease. If this policy is implemented, determine the effect on the answer to (b), and comment.

3. For this question, use the OLS regression output below. The data are from the 2016 American National Election Study pilot survey. The observations are Democratic identifiers (including Democratic-leaning independents). "R" denotes survey (R)espondents.

The variables used below include:

ClintonFT: Feeling thermometer for Hillary Clinton, scale of 0 to 100

Gender Discrmin: Whether or not the R feels they have personally experienced a lot of gender discrimination, 0 if no, 1 if yes

Local_terror_worry: Whether or not the R worries a lot about a terrorist event occurring in their local area, 0 if no, 1 if yes

Minwage: R?s opinion of the minimum wage on a 4 point scale (1=should be raised, 2=kept the same, 3=lowered, 4=eliminated)

Getahead= How much opportunity R sees in America today for the average person to get ahead (1=none, 2=a little, 3=a moderate amount, 4= a lot, 5=a great deal)

Femoff_issues: R?s assessment of how much female elected officials are likely to focus on issues that mainly affect women (1=a great deal more likely to focus on women, 2=moderately more likely to focus on women, 3=a little more likely to focus on women, 4=no more likely to focus on men than on women, 5=a little more likely to focus on men, 6= moderately more likely to focus on men, 7=a great deal more likely to focus on men)

Follow: how much R follows politics on a 4 point scale (1=Most of the time, 2=some of the time, 3=only now and then, 4=hardly at all)

Women: 1 for women Rs, 0 for men Rs

Black: 1 for Rs that chose black as their race, 0 otherwise

Hispanic: 1 for Rs that chose Hispanic as their race, 0 otherwise

otherRace: 1 for non-white/non-black/non-hispanic Rs (Asian, mixed, other), 0 otherwise

```
      Source |       SS           df       MS            Number of obs   =       553
-------------+----------------------------------          F(10, 542)      =      5.82
       Model |  39065.828          10   3906.5828         Prob > F        =    0.0000
    Residual | 363866.917         542  671.341175         R-squared       =    0.0970
-------------+----------------------------------          Adj R-squared   =    0.0803
       Total | 402932.745         552  729.950625         Root MSE        =     25.91
```

```
          ClintonFT |      Coef.    Std. Err.      t     P>|t|     [95% Conf. Interval]
--------------------+----------------------------------------------------------------
              women |   1.315898    2.249748     0.58    0.559    -3.103397    5.735192
   Gender Discrimin |  -.5011863    3.290931    -0.15    0.879    -6.965728    5.963355
  local_terror_worry |  2.523558     2.58036     0.98    0.329    -2.545173     7.59229
            minwage |  -3.785915    2.150771    -1.76    0.079    -8.010782     .438953
            getahead |   4.845829    1.198763     4.04    0.000     2.491038     7.20062
       femoff_issues |   .4545297    .3883526     1.17    0.242    -.3083309     1.21739
              follow |  -3.555783    1.314837    -2.70    0.007    -6.138584    -.9729816
               black |    13.9804    3.035169     4.61    0.000     8.018263    19.94254
            hispanic |  -1.475287    3.735297    -0.39    0.693     -8.81272    5.862146
            otherRace |   .9659694    4.508633     0.21    0.830    -7.890566    9.822505
               _cons |   60.70036    5.029369    12.07    0.000     50.82092     70.5798
```

(a) After running the model you produced a bunch of diagnostic plots. After careful scrutiny you decided that some of the plots of studentized residuals against the independent variables meant you needed to run a White's General Test. The result of that plus some careful thinking about race and politics made you then run a Goldfeld-Quandt Test, where you sorted your observations into groups defined by the racial categories in the analysis (black, Hispanic, white, other race). You decide that these diagnostics confirm your suspicions about a violation of the Gauss-Markov assumptions. What violation, exactly, have you decided you have on your hands and what are the

implications for your model? What could/would you do to best address the problems you have for valid inference?

(b) Using the OLS regression model, how would you test the hypothesis that "race doesn?t matter to Americans' evaluations of Hillary Clinton"? If you can test the hypothesis from this output alone, do so (set-up/report/interpret). If you cannot, explain why not and what else you?d need to know. [Go back to assuming the Gauss-Markov assumptions hold.]

(c) All else equal, what is the expected difference in Clinton ratings between black men and white women?

(d) You present the results of these models at a panel about Election 2016. One audience member asks if you can help them think through why "gender" was insignificant in your model when gendered politics seemed to be so prevalent in the campaign. You begin your reply by saying, "I think it is because the effect of gender isn't simply about women and men categorically disagreeing on Clinton. It?s buried in their attitudinal and experiential differences. Men are less likely to experience discrimination, are less supportive of the minimum wage, and are much more likely to believe Americans can still get ahead with hard work." What are you explaining to the audience and how might you present your data in a way that could help them understand your argument (e.g., are there particular figures or descriptive statistics that you would show them?)?

(e) Another audience member follows up on the previous question and your answer. "But doesn't your answer mean that the relationship between feelings about Clinton and those other measures is conditional by gender? How people use those experiences and attitudes to reflect on Clinton depends on their gender? So your model is just wrongly specified?" Take a deep breath and then explain what model the audience member wanted you to run and why they were wrong that your answer to the previous question necessarily required that model. Now, if you are right what would you expect when you ran the model this person is asking for?

(f) Another audience member wants to dismiss your models entirely because you failed to account for voters' attitudes about government policies to address gender discrimination. "That must be important to Clinton voters," he argues. You reply to the critic that you're not concerned about that, because previous research has shown unified support for such policies among Democrats. Explain the problem the audience member was claiming you had, and how your response was addressing it.

(g) One more audience member comments, "I have concerns about your dependent variable. I just don?t think that people are capable of reporting to you exactly where their feelings about Clinton fall on a 0 to 100 thermometer scale." You reply, "I agree to some extent–the evidence suggests that they really only have a general placement on the scale, and that they simply guess at exactly which number to report. In fact, I think that helps us understand why the R-squared on this model is not particularly large." You don?t seem too worried. Why not? And why did you reference R-squared?

(h) One of the reporters at this panel is actually impressed by your presentation and would like you to comment on how your work should inform the campaigns of Democratic women candidates for Congress. Would you feel comfortable commenting? Why or why not? If so, what kind of comment would you make?

Part II

1. The analysis of "clustered data" (hierarchical & multilevel data, panel and TSCS data) presents some key modeling decisions and specifications. What do you think are the two most important issues political scientists should consider when analyzing and making substantive conclusions about clustered data?

2. Surveys and experiments both face a related problem: people do not always cooperate with what the researcher hopes that they will do. Discuss this problem and its potential consequences by drawing on studies of (1) survey non-response and (2) non-compliance in field experiments. Why does this problem emerge, what impact can it have, and how can researchers seek to mitigate its impact?

3. Suppose you set up a difference-in-differences design where you look at the effect of a gun control law set up in five states in 2010. You compare the shift in crime from before to after 2010 for these five states to the shift in the remaining states that never implement this law. What are some of the threats to inferring causation from this analysis? What specific concerns does running a synthetic control analysis address? Are there other empirical tests that can address the remaining concerns?

4. A reviewer objects to your modeling choice in a journal submission. Your outcome variable is a count variable, but rather than estimating a count model, you estimated a linear regression model.

  (a) Why might the reviewer object to your model choice?

  (b) What alternative statistical models could you estimate, and what would their advantages be?

  (c) Under what circumstances can linear regression be used in a reasonable way on count data?

  (d) How could you discriminate between one of the alternative models in 2 and the linear regression model?

5. Recently a team of 72 statisticians and scientists of various stripes (including political scientists) posted a paper titled "Redefine Statistical Significance" that garnered attention in the research community. The abstract for the paper was succinct: "We propose to change the default P-value threshold for statistical significance for claims of new discoveries from 0.05 to 0.005." In the paper, the authors argued that changing the default P-value threshold would help mitigate the so-called "reproducibility crisis" in science. What do you think? What problems might this new standard solve? What problems would it create?