GSERM 2018 – Longitudinal Data Analysis

Final Examination

Instructions

1. This examination has two questions. **You are to choose <u>one</u> question and answer it.** *Do not attempt to answer both questions.*

2. Your answer should take the form of a brief empirical analysis of the data that answers the question(s) asked. That may include plots, tables, and any other techniques you think will be useful in answering the question. Your answer may be as short or long as you feel is appropriate, but probably need be no longer than 7-8 pages, including text, tables, and graphs.

3. The exam is worth 1000 possible points.

4. **If you are opting for the "in-class" format**, you should submit your exam no later than 7:00 p.m. CET time on Friday, June 8, 2018. Please submit your exam to the instructor electronically, as a PDF file, by e-mailing it to `zorn@psu.edu`.

5. **If you are opting for the "take-home" format**, you should submit your exam no later than 5:00 p.m. CET time on Friday, June 15, 2018. Please submit your exam to the instructor electronically, as a PDF file, by e-mailing it to `zorn@psu.edu`.

Question One

The research question in this exercise is the influence of ideology on Supreme Court voting over time. In particular, the conventional wisdom states that political actors often become more politically moderate over time. To test this hypothesis, you'll examine data on the voting patterns of justices sitting on the Supreme Court during the Vinson, Warren, Burger, and Rehnquist courts (1946-1994) ($N = 32$, $T = 49$, unbalanced). The data are available on the course github repository, in the "Exam" folder, as `GSERM-2018-Exam-Q1-Data.csv`. The variables are:

- `justice` (the justice identifier variable),
- `year` (the year identifier),
- `civlib` (the percentage of liberal votes cast by that justice in civil rights and liberties decisions in that year),
- `econs` (the percentage of liberal votes cast by that justice in economics decisions in that year),

- `score` (the normed "Segal/Cover" (1989) ideology score of the justice, ranging from -1 (most right/conservative) to 1 (most left/liberal) and
- `tenure` (the number of years the justice has served on the Court, as of that year).

If the conventional wisdom is correct, one possible manifestation is that the effect of `score` on voting liberalism should be positive, but the interaction of `score` and `tenure` should be negative (as justices moderate their extremism later in their careers).

Your assignment is as follows:

1. First, examine voting liberalism in civil rights and liberties cases (`civlib`).
    - Estimate both fixed- and random-effects models for this outcome, and discuss your results, both substantively and statistically.
    - Evaluate the statistical and substantive plausibility of random effects, and compare the two sets of results in terms of model specification.
    - Lastly, fit one or more models that address the possibility of *temporal dynamics* in the data (autoregression, etc.). Discuss your findings in statistical and substantive terms.
2. Repeat the above steps for the variable on economics cases (`econs`).
3. Talk in general terms about which model(s) you prefer for these analyses, and why.

Please answer each component of this question in a short paragraph or two. In addition, please provide all computer output, and all code used to generate your findings.

Question Two

In an important article, "Mediating Civil War Settlements and the Duration of Peace" (also available here) Gurses, Rost, and McLeod (GRM) examine the durability of peace following the end of a civil war. In particular, they focus "the impact of international mediation attempts during civil war on the duration of peace once the war has ended." Put simply, their question is whether such interventions – by other countries, superpowers, the United Nations, and so forth – lead to peace settlements that are more lasting than those in which such mediation did not occur. To investigate this question, they use survival models to examine the duration of peace following more than 80 civil wars, and consider the influence of various types of mediation on the hazard of civil war recurrence.

You will be using GRM's data as they are available in .csv format on the course github repository, in the "Exam" folder. (The original data from GRM's study are available at the

paper's Dataverse). The data are survival-time data on 87 civil wars (across 55 countries) that occurred between 1946 and 1995. The response of interest is the duration of the peace (in years) following the end of a civil war; that is, the even of interest is the recurrence of civil war in a country. Details are available in the article.

In their article, GRM fit a series of models with 13 covariates/predictors, as follows:

- Mediated Agreement: Whether (=1) or not (=0) the final peace agreement was mediated.
- Mediation: Whether (=1) or not (=0) there was mediation attempt during the last year of the civil war.
- Super Power Mediation: Whether (=1) or not (=0) there was mediation attempt by a superpower.
- Total Number of Mediation Attempts: The number of such attempts during the civil war.
- U.N. Peacekeeping: An an ordinal variable reflecting the nature and degree of U.N. involvement, ranking missions from enforcement missions (4) to traditional peacekeeping missions (3), to observer missions (2), mediations (1), and no peacekeeping mission (0).
- Ethnic War: Whether (=1) or not (=0) the war was over ethnic, religious, or other ascribed characteristics.
- Log of Total Deaths: The natural logarithm of the number of deaths in the civil war.
- War Duration: The length (in years) of the civil war.
- Military Victory: Whether (=1) or not (=0) the war ended with a one-sided military victory (as opposed to a negotiated settlement).
- Democracy: The POLITY autocracy-democracy score for the country, ranging from -10 (fully autocratic) to 10 (fully democratic).
- Log of GDP: The natural logarithm of the country's GDP per capita, measured in constant 1985 U.S. dollars.
- Log of Army Size: The natural logarithm of the size of the standing army, measured as troop counts.
- War Count: The number of civil wars the country experienced during this period.

In addition, the data also contain a series of legacy variables (named `t`, `t0`, etc.) from GRM's Stata data file. Of these, `t0` can be thought of as a "start time," and `t` as a "stop time," while `d` represents the censoring/event indicator.

Your assignment is to reestimate and reevaluate GRM's analyses of post-civil war peace durations, using their data. *Note that you need not concern yourself with exactly reproducing the analyses and results reported in the article.* Instead, your assignment is to reconsider their findings with a particular focus on the four variables (the first four in the list above) that measure various aspects of mediation. More specifically:

1. Briefly describe the survival data used in GRM's analysis.
2. Specify and estimate one or more survival models on GRM's data. Discuss why you chose to fit the models you did.
3. Using whatever approaches you deem appropriate, discuss in substantive terms the associations between the four mediation variables and the hazard of civil war recurrence.
4. Conduct appropriate post-estimation diagnostics and robustness checks for your model, and discuss your findings.
5. Finally, speculate on whether and to what extent any of the various "extensions" to survival models discussed in class might be better / more appropriate for the analysis conducted, and / or provide additional insight into the question of mediation in civil wars.

Please answer each component of this question in a short paragraph or two. In addition, please provide all computer output, and all code used to generate your findings.