

PS 150C/350C
 Spring 2003
 Final Exam

This is an open book, take-home exam. You have 24 hours to complete and return your exam to Matt Levendusky and myself by e-mail. Unlike computing assignments throughout the quarter, your answers must be your own work.

1. Each year, experts are asked to score Latin American countries on the level of democracy of each country, by assigning a score between zero (most undemocratic) to one hundred (most democratic). Let $i = 1, \dots, n$ index the countries, let $j = 1, \dots, m$ index the experts, and let $t = 1, \dots, T$, such that y_{itj} is the rating of country i at time t by expert j . In the questions below, unless specified otherwise, assume the ε_{itj} have expectation zero and are iid.

- (a) (5 points) Interpret the parameters α_i in the model

$$y_{itj} = \alpha_i + \varepsilon_{itj}^{(a)},$$

- (b) (5 points) Interpret the parameters α_{it} in the model

$$y_{itj} = \alpha_{it} + \varepsilon_{itj}^{(b)}$$

- (c) (5 points) Interpret the parameters α_{it} in the model

$$y_{itj} = \alpha_{it} + \delta_j + \varepsilon_{itj}^{(c)}$$

- (d) (7 points) Under what conditions would we obtain the same estimates of α_{it} from the models in the last two questions?

- (e) (7 points) Interpret the following model for these data:

$$\begin{aligned} y_{itj} &= \alpha_{it} + \delta_j + \varepsilon_{itj}^{(e)} \\ \text{var}(\varepsilon_{itj}^{(e)}) &= \sigma_j^2 \end{aligned}$$

How would we test this specification against the model in part (c)?

2. Consider a model for panel data $y_{it} = \mathbf{x}_{it}\boldsymbol{\beta} + \alpha_i + \varepsilon_{it}$ where $i = 1, \dots, n$ indexes units, $t = 1, \dots, T$ indexes time and $\boldsymbol{\varepsilon}_t$ is a n -by-1 vector $(\varepsilon_{1t}, \dots, \varepsilon_{nt})'$ with

$$\begin{aligned} E(\boldsymbol{\varepsilon}_t) &= \mathbf{0}, \forall t = 1, \dots, T \\ \text{var-cov}(\boldsymbol{\varepsilon}_t) &= \boldsymbol{\Sigma}_{(n \times n)}, \forall t = 1, \dots, T \\ E(\mathbf{x}_{it}\boldsymbol{\varepsilon}_{it}) &= \mathbf{0}, \forall i, t \\ E(\mathbf{x}_{it}\alpha_i) &= \mathbf{0}, \forall i, t. \end{aligned}$$

- (a) (5 points) Explain the role of Σ .
- (b) (5 points) What additional assumptions are required as to make OLS an optimal estimation strategy for this model?
- (c) (5 points) When OLS is not appropriate for these data, how else might we estimate β and the α_j ?
3. (40 points) Sean Theriault and Jen Lawless have **data** on congressional careers summarized below, and available from my web site. You can read these data into R by downloading and saving the data file, and then using the `dget` command to read it into R:

```
dat <- dget(file="JenSean.dpt")
```

The data cover the 98th through the 107th U.S. House of Representatives, and are at the level of the individual legislator. The dependent variable of interest is whether the legislator decides to retire; i.e., $y_{it} = 1$ if legislator i retires at election t and zero otherwise, although Theriault and Lawless coded their outcome variable `retire3` by the actual nature of the decision, which I cross-tabulate here by Congress:

Congress	Did Not Retire	Retired, Left Politics	Sought Higher Office	NA
98	413	9	13	8
99	405	17	14	4
100	414	11	11	7
101	410	11	16	9
102	368	52	15	6
103	387	27	24	4
104	385	35	16	5
105	410	18	13	3
106	403	20	10	5
107	397	18	19	8

Some of the NAs are from legislators dying in office, resigning midterm or appointed to executive office. Covariates available for analysis include:

- `yrsh` years served in Congress
- `child` dummy variable for whether legislator has children
- `female` dummy variable
- `age` age of legislator in years
- `married` dummy variable
- `redist` dummy variable, is legislator's district subject to redistricting (only takes on "yes" values for the 102nd and 107th Congresses).
- `majority` dummy variable, is legislator in party with a majority of the House
- `democrat` dummy variable

- `winpct` proportion of vote won at previous election

Use these data to write a short research note on the determinants of retiring from Congress. Be sure to justify your choice of statistical model, the validity of underlying assumptions, handling missing data, etc. You should also carefully and precisely elaborate the effects of the covariates. Use whatever tables and graphs you need to make your conclusions and/or to assess model fit, etc. Don't hand in raw output from your statistical software.

4. (40 points) A large literature investigates the impact of partisan composition of government on economic growth. In an article that appeared in the *American Political Science Review* in 1991, Alvarez, Garrett and Lange use a data set on 16 OECD countries over 14 years (1971-1984, inclusive). These data have been repeatedly re-analyzed (e.g., Western's 1998 *AJPS* article). The **data** are available from my web site. As for the previous question, after you download the data file, you can read it into R using the `dget` command.

The outcome of interest is called `y` in the data file, a measure of economic growth (percentage change in real GDP). The covariates available for analysis are:

- `growth.lag` lagged `y`
- `imports` price movements of OECD imports
- `exports` price movements of OECD exports
- `demand` vulnerability to demand in the OECD (the annual OECD economic growth rate weighted by the "openness")
- `left` percent of cabinet seats held by parties of the left
- `lo` a scale measure of labor organization (combining information about size and centralization of labor unions).

Details on the measures and their sources appear in the 1991 AGL article. Of primary substantive importance is the impact of left governments, and a possible interaction between left party government and union movement strength.

Analyze these data in a brief report that addresses both the statistical and substantive issues. You should begin with some exploratory analyses, seeing how much of the variation in the dependent variable is cross-sectional or longitudinal, and check for outliers by country, year, and both. Your report should contain a set of models, checking for the robustness of the effects of substantive interest to changes in specification. You should report some specification tests: are the residuals "well-behaved" (e.g., homoskedastic and serially uncorrelated by country, by time, are fixed effects warranted over one or more dimensions of the data). Finally, you should also be alert to the nature of the interaction between left party government and union movement strength: for your preferred model, interpret these effects carefully and in terms that a generalist read will understand.

Hint: if you are using R, be sure to see the code examples I have up on my web site.