Coursework Assignment 2011-2012 SC968-AS-MI

**SC968 Panel Data Methods for Sociologists**

Candidates must answer **BOTH** questions

The questions are **NOT** of equal weight. Question 1 is worth 40%, and Question 2 is worth 60%.

1. **WRITING A PROPOSAL FOR RESEARCH USING PANEL DATA MODELS**

The Department of Health has announced that it intends to fund a research project investigating the effects of shift-working on mental health; in particular, on whether people’s mental health is adversely affected by (a) working at night, and by (b) irregular working hours.

Your institution has decided to apply for this research funding. Your colleague has made a first draft of a proposal, but was obviously in a great hurry when he wrote it. His proposal as it stands reads:

“This is an extremely important and timely research question. For this analysis we propose to use the British Household Panel Survey (BHPS). The BHPS contains two variables which measure mental health: the Likert scale, which ranges from 0 to 36, and the Caseness scale, which ranges from 0 to 12. It also contains information on whether people work at night, and a question asking whether people work irregular shifts. This means we will be able to investigate the relationship between shift working and mental health.”

This proposal clearly needs a great deal more work before it has any chance whatsoever of attracting funding. It is terrible in so many ways: it hasn’t mentioned any hypotheses for why a relationship might exist between shift working and mental health; it hasn’t mentioned what other variables you might control for in regressions; and worst of all, it hasn’t said anything about the analytical techniques you might employ, or what insights a panel data analysis might offer above a cross-sectional analysis.

Your colleague has asked you to redraft the proposal. When preparing the proposal, you might want to consider including the following elements which address some of the shortcomings of the first draft.

1. An outline of why the proposed area of research is important, and what previous studies have found. For this assignment, you are not required to present a comprehensive literature review, but you may want to refer to a couple of studies in the area.
2. Details of the research questions that you propose to address, together with associated hypotheses.
3. A section making the case for using panel data in the research, and explaining how panel data can inform this question in ways which cross-sectional data cannot.
4. A description of the methods which you propose to use for the analysis (which will include the methods taught in SC968, but which need not be restricted to these methods), and the particular aspects of the research question which will be addressed by each of these methods. The potential strengths and limitations of each of the methods should be pointed out.
5. A brief outline of the data you will use (here, the BHPS). This section should also include brief descriptions of how you will define the variable(s) describing shift work and irregular working hours, and of the explanatory variables which you will be seeking to use in multivariate analysis (these variables should not be restricted to the variables present in the SC968 teaching data set, but should be present in the BHPS).

Information on the contents of the BHPS is available at <http://www.iser.essex.ac.uk/bhps/documentation>

It is not expected that the research proposal will contain a large volume of analytical results. However, you can include a couple of tables showing descriptive statistics, cross-tabulations, transition matrices, etc. if they serve to illustrate the importance of the topic, or to underline the reasons why a particular approach is to be taken.

Proposals of this type are often subject to strict limits on length. For this assignment, you are limited to 2000 words (excluding tables) – although an excellent proposal could be somewhat shorter than this.

**2. USING PANEL DATA MODELS TO ESTIMATE THE EFFECTS ON PSYCHOLOGICAL WELLBEING OF WORKING MORE OR FEWER HOURS THAN ONE WOULD IDEALLY LIKE.**

The BHPS contains a question on people’s preferred working hours, which asks:

Thinking about the hours you work, assuming that you would be paid the same amount per hour, would you prefer to...

1. Work fewer hours
2. Work more hours
3. Continue working the same hours

This variable is called JBHRLK and is contained in the data set ASSESSMENT2012, which may be found on the SC968 website <http://www.iser.essex.ac.uk/resources/module-sc968/course-materials>

The data set is similar to the training file you have been working with during the course, but note that it contains 18 waves of data and a subset of variables appropriate for the assignment.

Your task is to analyse the relationship between people’s psychological wellbeing, as measured by variables HLGHQ1 and/or HLGHQ2, and their preferences over their hours of work. When thinking about this task, you should consider the other variables that might explain the relationship between people’s psychological wellbeing and their preferences over working hours, and the variables that might modify or interact with preferences over working hours.

This is an open-ended task: you may use any or all of the panel data models covered in SC968 which you consider to be appropriate. We won’t tell you which ones to use, but obviously it would be a good idea to showcase a range of models.

You should discuss how you interpret the problem; how you selected the models you used; your choice of additional explanatory variables and interaction terms; how you selected your sample (for example, whether you restricted your sample by age, sex, labour market status etc), and how you coded or recoded the variables in your regressions. You should present your findings in tables and figures, as appropriate, and discuss these findings in detail, including a discussion of whether findings differ between models, and what you conclude from them.

Please also supply with this assignment an ELECTRONIC copy of your Stata do-file. One third of the marks for this question will be awarded on the basis of this do-file. It should:

1. Allow the people marking your assignment to replicate your results exactly, starting from the same data set which was supplied to you: thus, it should contain commands for cleaning and re-coding variables, sample selection, etc, as well as for regressions.
2. Be well documented. It is not necessary to insert comments explaining the purpose of every single command in the do-file, but you should document the purpose of each section of code.
3. Contain all the commands necessary to generate not only the results you report, but also any results which you discuss but don’t report in tables (for examples, any specifications which you reject in favour of better ones, such as those including interaction terms which turn out to be insignificant).
4. It is good practice to divide your do-file into two sections, one where you create and re-code variables, and another where you perform the analysis. It is good practice to list the commands in the same order that they appear in your written text, and to add comments showing which STATA commands produce the output for which tables in your text.

As well as the do-file, marks will be given for evidence that you have thought about the problem thoroughly; an appropriate and accurate use of analytical tools; clear presentation of your results; correct inferences from your results; good organisation of your material, and clear writing.

There is a strict page limit for this question: answers, including tables and references, must be a maximum of 10 sides of A4 paper. The font must be at least 12 point for the text, and 10 point for the tables.