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. USING PANEL DATA FOR AN ANALYSIS OF POVERTY

You are part of a team working on a project which aims to quantify the extent of poverty (defined as living in a household with an income lower than 60% of the national median income) among families with children over the past two decades, and to identify the factors associated with poverty at the individual and household level. The head of the research team is proposing to base the research on a data set which contains ten years of repeated cross-sections, collected at two-yearly intervals. You feel that it is important to use panel data for this project to investigate longitudinal measures of poverty, and poverty dynamics. So far, your suggestion has not been taken seriously, and you have decided to make your case in writing.

Write a memo to the head of the research team outlining the reasons why it would be a good idea to use panel data for the research, outlining in detail the insights which may be obtained from panel data research, and the reasons why these additional insights are important for an understanding of the area. You may also like to acknowledge that cross-sectional data has a number of strengths too, and to suggest ways in which cross-sectional and panel data could be combined.

In order to reinforce your arguments, you may want to refer to other authors' work in this area, and/or to present preliminary results from the British Household Panel Survey.

As well as being assessed for content, this question will also attract marks for your writing style and the persuasiveness of your argument.

This question is subject to a strict limit of 2000 words. You may also include up to two tables.

2. USING PANEL DATA METHODS TO EXPLAIN THE ONSET OF LIMITING HEALTH PROBLEMS

You are asked to write a short report on the factors influencing the onset of a health problem that limits daily activities (variable **hllt**).

Potential explanatory variables that you are asked to consider include the following list, although you may include as many others as you like, including interaction effects.

- age
- gender
- marital status
- employment status
- income
- education

Your report should include fixed and/or random effects models, and survival analysis. It would also benefit from some descriptive statistics and/or transition matrices in the initial sections.

At each stage, you should present your results clearly and comment on their implications; you should also include output from any post-estimation commands which you think appropriate. You should spend some time comparing the results from the two approaches, and commenting on these differences.

As well as presenting and commenting on your results, you should explain and motivate the purpose of your research in the initial sections, and explain how you selected your sample, recoded explanatory variables, etc.

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. The do-file should run from beginning to end without any error messages, and 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.

Be aware that there is no one “correct answer” to this question. Please provide a printout of your Stata do file containing all your syntax with your answer.

1