EXAMINATION QUESTION PAPER - Take-home examination

DRE 70061
Panel Data/Microeconometrics

Department of Economics

Start date:	04.05.2017	Time 09.00
Finish date:	05.05.2017	Time 15.00

Weight:	100% of DRE 7006
Total no. of pages:	3 incl. front page
No. of attachments files to question paper:	0
To be answered:	Individually

Answer paper size:	10 excl. attachments
Max no. of answer paper attachment files:	1
Allowed answer paper file types:	pdf
Allowed answer paper attachment file types:	pdf

The exam paper must also include a table of contents, a summary at the beginning and a bibliography/reference list at the conclusion of the paper. These pages are counted separately from the main paper. You will find a thorough explanation of how to use quotes and references on the BI library webpage: www.bi.edu/library/training-and-support/citing-and-referencing. For information on formal requirements and template paper; see www.bi.edu/templatepaper

Question 1

Imagine that you are a policy maker interested in the relationship between a mother's welfare participation and the birth weight of her child. You have available a panel data set for the US states for two years, 1987 and 1990.[1] The key dependent variable is the percentage of births that classified as low birth weight (**lowbrth**) and the key explanatory variable is **afdcprc**, the percentage of the population in the welfare program, Aid to Families with Dependent Children (AFDC). Participation in AFDC makes poor women eligible for nutritional programs and prenatal care. The other variables, which act as controls for quality of health care and income levels, are (log) physicians per capita (**lphysic**), (log) hospital beds per capita (**lbedspc**), (log) per capita income (**lpcinc**), and (log) population (**lpopul**).

a) Create a scatterplot showing the relationship between **afdcprc** and **lowbrth** for 1990. Label the scatterpoints with the state name, and include the regression line. What is the estimated slope parameter? What is the estimated slope parameter on **afdcprc** when you control for **lphysic lbedspc lpcinc** and **lpopul**?
b) Run an auxiliary regression where you net out the effect of **lphysic**, **lbedspc**, **lpcinc** and **lpopul** on **lowbrth**. Keep the residuals and add back the mean of **lowbrth**. Similarly, run a regression of **lphysic**, **lbedspc**, **lpcinc** and **lpopul** on **afdcprc**. Keep the residuals and add back the mean of **afdcprc**. Create a scatterplot showing the relationship between the two variables you created. Label the scatterpoints with the state name, and include the regression line. What is the slope coefficient? Compare with a).
c) Now use data for both 1987 and 1990 and run a pooled OLS using traditional standard errors where, in addition to the controls mentioned in a), you also include a time fixed effect. What is the point estimate for **afdcprc**? What is the corresponding standard error? Test for serial correlation in the error term. Re-do the pooled OLS with fully robust standard errors. Comment on your findings.
d) Estimate the model using first-differences. What is the point estimate for **afdcprc**? What is the heteroscedasticity-robust standard error? Create a scatterplot as in (b).

Question 2

Using survey data from Equador, Thomas Buser (2015)[2] investigates how income affects religiousness. While cross-country evidence suggest a negative relationship, Buser comes to the opposite conclusion using a regression discontinuity (RD) design. Families that earn more go to church more often. To answer the following questions you should use Buser's data set which is available at the AEA web pages.[3]

a) Buser's sample consist of relatively poor households. Assume that they spend all their (monthly) income, so income equals expenditures. Run a regression of attendance per

[1] Please download the data set *lowbirth.dta* from it's learning.

[2] Buser, Thomas. 2015. "The Effect of Income on Religiousness." *American Economic Journal: Applied Economics*, 7(3): 178-95.

[3] https://www.aeaweb.org/aej/app/data/0703/2014-0162_data.zip

month (**attendpermonth**) on total household expenditures (**expenditures**), number of people in the household (**householdsize**), age of the respondent (**ageresponder)**, and years of schooling of the respondent (**schooling_resp**). Explain when the coefficient on **expenditures** can be given a causal interpretation. Give examples that lead to the coeffient on **expenditures** having bias.

b) According to Buser (2015, p. 181):

> *"The Ecuadorian government runs a cash transfer program called Bono Desarollo Humano (BDH), which was launched in 2003 and is aimed at the poorest 40 percent of households. … Eligibility for the BDH is determined by a households' percentile on a wealth index (called SELBEN). This index is based on a range of observable variables including household assets and housing characteristics (e.g., access to water, toilet, and shower), possession of appliances (e.g., TV, DVD, microwave), characteristics of the head of household (e.g., schooling and employment), childrens' characteristics, and household size. From these variables, the index is calculated using nonlinear principal components analysis. The variables were collected through a census of all households living in poor areas. In 2007–2008, all households in these areas were resurveyed and the definition of the index was changed (SELBEN II)[4]. The index is now composed of 59 variables covering the same categories (see Fabara 2009, for the complete list of variables). This new index was implemented in 2009, which led to changes in eligibility for many households close to the cutoff fortieth percentile. Some households who had received the transfer for over seven years suddenly lost it while others suddenly gained it."*

Explain how the cash transfer program can be used to identify how income affect religiousness using a RD design. What assumptions do you have to make? What causal parameter will you identify?

c) What is the difference between a sharp and a fuzzy RD? Explain why a fuzzy RD is appropriate in the current application. Estimate the first stage regression using a linear control function and the full bandwidth. Provide the relevant RD plot.

d) Estimate the *intention to treat* effect of BDH on attendance per month using a linear control function and the full bandwidth. Provide the relevant RD plot.

e) Using your results in (c) and (d), what is the *local average treatment effect*?

f) Estimate the *intention to treat* effect of BDH on attendance per month using a linear control function for a range of bandwidths. Compute the optimal bandwidth based on the `rdrobust`[5] package using a local linear regression and a uniform kernel. Plot the point estimates and the 95 percent confidence intervals as functions of the bandwidth chosen. Indicate the optimal bandwidth in your plot. Comment on the results.

g) Carry out standard sensitivity checks for RD analyses. Present the results from these sensitivity checks graphically. Discuss.

---

[4] In the dataset *20140162_Data.dta* this variable is called **score2_1**

[5] https://sites.google.com/site/rdpackages/rdrobust