

SØK3514 H2021 Assessment guidelines

Q1.

a) The students can build on the standard economic growth model for example as presented in section 2 in the article by Miguel and Roland (2011) to introduce a constraint on saving that may lead to a poverty trap after armed conflicts. Answers need not include a lot of math but should be to the point and explain the **temporary** and **long term** effects from armed conflicts (huge reduction in capital stock) with and without savings constraints. A simple regression model may be formulated to discuss the main econometric challenges. Pure cross section data for the postconflict period and panel/pooled cross section time series with pre and post conflict data information are relevant data types. Should be able to discuss long and short run mechanisms according to the theoretical considerations. Main challenges to be discussed is omitted variables and endogeneous selection of areas of armed conflict, regardless of whether cross section or panel data are used.

b) The student should be able to illustrate the effect of bombing by multiplying the coefficient $-0.0004 \times 32.2 = -0.0129$ which is the effect on poverty share (in percentages) in region with average bombing compared with a region not bombed at all. The effect is negative and tiny relative to the average poverty rate at 41.1 percentage points. Further, it is opposite in sign to what would be expected under the poverty trap hypothesis.

c) The problems with omitted variables and endogeneous selection of areas exposed to bombing should be discussed as a source of endogenous bombing intensity and the implied bias in OLS estimator. The best candidates may explain possible direction of bias due to endogeneity.

d) Should explain how the IV-2SLS method may remove the OLS bias, and explain the requirements the instrumental variable should satisfy, i.e. relevance and exclusion restriction. Whether these requirements are fulfilled for the instrument used in col (6), the distance from 17-latitude) should be discussed. The students should explain how the first stage equation can be used to test the relevance requirement, and refer to and discuss the significance of the estimated coefficient in front of the distance variable as reported in table 3 in the paper to evaluate the relevance. Students should note that the exclusion restriction is untestable. Students should describe how the endogeneity of the bombing intensity variable can be tested following the approach in Woolridge 15-5a, and what assumptions this test is based on. The best students would note and explain that introduction of an additional instrumental variable makes it possible to test for overidentification restrictions (Woolridge 15-5b) and how that relates to the requirements.

Q2.a)

Students should formulate a simple regression model with applications to higher education as dependent variable and regional unemployment rate as key explanatory variable. The dependent variable could be measured as a regional level variable (the number of applicants) or as a dummy measured at the individual x region level. Both representations will do, but in the following presentation the former assumptions are used. Students could discuss potential challenges when having i) cross section data from one period available and ii) panel data from

several periods. In case i), the usual problems with endogeneity with cross section data is relevant to discuss and in case ii) students should discuss to what extent the inclusion of region fixed and period fixed effects enables the researcher to control for some (time-constant region and pure macro level observables and unobservables. Should also discuss that sufficient within region variation in the dependent (application) and variable of interest (unemployment) is required for this approach to identify the causal effect of unemployment.

Finally, in the panel case (ii), students should consider the situation where for example omitted variables varying over time within periods and correlated with the outcome is a source of bias. This may suggest an IV approach that also naturally lead to question Q2b)

Q2.b)

This is an open question, which can be dealt with in different ways. The proposal from the commentator can be discussed under several different assumptions about the nature of the effect of the corona virus or the restrictions imposed because of the corona virus. To the extent that there is a common corona virus effect across regions, the effect is captured by the inclusion of time period fixed effects. If the corona virus or restrictions varies across regions, some may argue that these may be used as instrumental variables for the unemployment rate. But good students should provide a critical discussion of whether the exclusion requirements are fulfilled-i.e. to what extent it is plausible that corona virus or their implied regional restrictions affect applications only through the unemployment channel.

Q3

a) Standard discussion of RDD and distinguishing between sharp and fuzzy RDD as done in Lee and Lemieux (2010) on the reading list and may exemplify the latter by the applications in Angrist and Lavy (1999) study of class size effects or both by the Petterson-Lidbom (2012) study of council size effects in Finland and Sweden, respectively.

b) Should relate the discussion of the different regressions done in Petterson-Lidbom (2008) in the text to the answer on a). Issues are full sample regressions with parametric controls for nonlinear effects of the running variable (here: share of votes), typically polynomials versus regressions using the discontinuity samples and including control variables or not.

c) Answer here can naturally build on the answer in b) and discuss the specification tests in Petterson-Lidbom section 4.1. Actually it is possible answer b) and c) together.