Sensorveiledning:

Q1.

a)DiD estimator of the policy effect: 890-800-(1130-1100)=60

b) Econometric model where the outcome y denotes the number of hours worked

(1) $y = \beta_0 + \delta_0 d1993 + \beta_1 Tiltakssonen + \delta_1 d1993 \cdot Tiltakssonen + u$

where $d1993$ is a dummy equal to 1 for observations from 1993, and equal to 0 if not. Tiltakssonen is a dummy equal to 1 if the individual lives in a municipality in "Tiltakssonen" and 0 if not, and u is the usual error term. It should be explained that the OLS estimate of the δ_1 coefficient in (1) equals the simple DiD estimate of the policy effect obtained in a), see also the discussion of the parameters in Table 13.3 in Woolridge. Within the regression framework, confidence interval around the parameter can be obtained in the usual way, see textbook.

c)The model in (1) can be extended with variables Age and Comp to account for the suggestions made by the Ministry.

(2) $y = \beta_0 + \delta_0 d1993 + \beta_1 Tiltakssonen + \delta_1 d1993 \cdot Tiltakssonen + \gamma_1 Age + \gamma_1 Complete + u$

δ_1 will still denote the policy effect, although not in the simple form as in a) and b), but the interpretation is similar.

Students should comment on the fact that the effects estimated in a), b) and c) can be interpreted as causal policy effect under the parallel trend assumption. That is the assumption that the development in outcomes in the treated areas ("Tiltakssonen") in the absence of treatment (the unobserved counterfactual) would be similar to that in the control group (the other areas)

d) Exploiting the information in d), the model in b) can be reformulated in the following way and similar to the discussion in chapter 13-2a in Woolridge:

High=1 if the individual has higher education, 0 otherwise

(3)$y = \beta_0 + \beta_1 High + \beta_2 Tiltakssonen + \delta_0 d1993 + \delta_1 d1993 \cdot High +$

$+\delta_2 d1993 \cdot Tiltakssonen + \delta_3 d1993 \cdot High \cdot Tiltakssonen + u$

In (3) the term $\delta_1 d1993 \cdot High$ allow the trends in outcome for the high educated individuals to differ from that of low-educated, while the term $\delta_1 d1993 \cdot Tiltakssonen$ allows the trend in outcome to differ between Tiltakssonen and other areas. Thus, the formulation I (3) allow us to control for these different trends in the regression.

δ_3 is the policy effect and can be interpreted as the causal effect under the assumption that the difference in trends between high and low educated people is the same across areas in the absence of treatment.

Q2.

a)Interpretation of coefficients in this log-lin regression model:

(*) $lfaminc = \beta_0 + \beta_1 hedu + \beta_2 wedu + u$

The estimated coefficient in front of ***hedu***, 0.0439 means that an additional year of husbands education increase family income by approximately 4.4%. Similarly the coefficient of 0.039 in front of ***wedu*** implies that one additional year of wife's education increase family income by approximately 3.9%.

b)In column (2) wife's education, ***wedu***, is omitted from the model. Since the correlation matrix in Table 2 indicates that ***wedu*** and ***hedu*** are highly positively correlated (correlation coefficient of 0.59), the model in column (2) has an omitted variable problem that leads to a positive bias in the ***hedu***-coefficient in column (2). The coefficient partly accounts for the effect of the omitted variable (***wedu***). The increase in the ***hedu*** coefficient from col (1) to col (2) confirms the omitted variable bias derived in eq (3.46) on p. 85 in the textbook.

c)The students should explain that the coefficient in front of ***hedu*** in column (3) is the difference in the impact of ***hedu*** and ***wedu*** on family income, i.e $\beta_1 - \beta_2$, while the coefficient in front of ***edutot*** is β_2. This implies that column (1) and (3) is basically the same regression, but identifies different parameters as is also confirmed by the fact that the constant term and the R-square is the same in col (1) and (3).

d)A test of the hypothesis that husbands and wifes education has equal effect on family income is then simply a test of whether the coefficient in front of ***hedu*** in the model in column (3) is different from zero which can be tested with a simple t-test. The t-statistic is 0.26 which is far below critical values in the t-distribution. So we cannot reject the hypothesis that the impacts of husband and wife education are equal.

e)Adding ***kl6***: Interpretation: One more kid under 6 reduce family income by approximately 17%., all else equal. Holding the number of kids under 6 constant, column (4) also indicates that the estimated returns to education for both husband and wife increase, with the greater increase going to the wife (from 0.039 to 0.042) whose working hours would be more likely to be affected by the presence of young children.

f) Table 4 is the results from a RESET test of functional form misspecification discussed in ch 9-1a in textbook. Adding polynomials of predicted ***lfaminc*** is a way to test if the specification in (*) above is a valid simplification of a more general nonlinear relationship between ***lfaminc*** and ***hedu*** and ***wedu*.** The regression in Table 4 corresponds to

$$lfaminc = \beta_0 + \beta_1 hedu + \beta_2 wedu + \gamma_1 \widehat{lfaminc}^2 + \gamma_2 \widehat{lfaminc}^3 + u$$

Adding the two variables to the structural model allow the effect of ***hedu*** and ***wedu*** to be highly nonlinear. The RESET test can then be done by testing the hypothesis $\gamma_1 = \gamma_2 = 0$ by a standard F-test. Using the R^2-version of the F-test ($\frac{(R^2_{UR} - R^2_R)/2}{(1-R^2_{UR})/423}$, see textbook p.145) implies an F-statistic equal to approximately 2.06 which is clearly below the critical value in the F-distribution with 2 and 423 degrees of freedom. Thus, based on the RESET test we do not reject the specification in column (1) against general nonlinear alternatives.

g) The short and to the point answer is that the statement does not make sense since panel data methods (First difference or within-group estimators) requires within unit variation in the variable of interest in order to reduce omitted variable problems. Since education for these couples is likely to be completed, there is likely no or very little variation in wedu and hedu over time.