

Семинар 15.

1. (Универсиада по эконометрике, 2016 год). Представим, что в некоторой стране уровень зарплаты работников данной специальности описывается точным соотношением:

$$Y_i = \beta_1 + \beta_2 X_i + \beta_3 Q_i,$$

где Y_i – зарплата i -го работника; X_i – бинарная переменная, которая равна 1, если i -ый работник получил высшее образование; Q_i – уровень гениальности i -го работника (ненаблюдаемая переменная).

В вашем распоряжении имеются представленные в таблице данные о 1000 работников данной специальности:

	В родном городе есть университет	В родном городе нет университета
Получил высшее образование	400 человек Средняя зарплата \$5000	100 человек Средняя зарплата \$6000
Не получал высшее образование	100 человек Средняя зарплата \$3000	400 человек Средняя зарплата \$4000

- a) Предложите способ состоятельно оценить параметр β_2 и вычислите его оценку, используя данные, которые есть в вашем распоряжении. Подробно аргументируйте выбор метода оценивания.
- b) Если доступных данных достаточно для того, чтобы осуществить какой-либо тест, характеризующий релевантность данных для применения предложенного вами подхода, то осуществите его.
2. (Универсиада по эконометрике, 2017 год.) Рассматривается модель:

$$Y_i = \beta_1 + \beta_2 X_i + \varepsilon_i,$$

где Y_i – стохастический эндогенный регрессор.

В распоряжении исследователя помимо данных о переменных X и Y есть данные еще двух переменных P и Q таких, что $cov(X_i, P_i) \neq 0$, $cov(X_i, Q_i) \neq 0$, $cov(\varepsilon_i, P_i) = 0$, $cov(\varepsilon_i, Q_i) = 0$.

- a) Докажите, что оценка двухшагового МНК для параметра β_2 , использующая переменные P и Q в качестве инструментов, будет состоятельной. Если вам требуются какие-либо дополнительные предположения, то сформулируйте их.

- b) Пусть ваша выборка состоит из 1000 наблюдений, причем вы располагаете данными о средних выборочных значениях переменных: $\bar{Y} = \bar{X} = \bar{P} = 0$, $\bar{Q} = \overline{PQ} = \overline{XQ} = \overline{P^2} = \overline{YQ} = 1$, $\overline{Q^2} = 1.5$, $\overline{XP} = \overline{YP} = 2$.

Вычислите состоятельную оценку параметра β_2 из предыдущего пункта.

3. В статье Р. Ениколопова, М. Петровой и Е. Журавской анализируется влияние телевидения на решения избирателей голосовать за ту или иную партию. В рамках моделирования на индивидуальных данных в качестве объясняющей переменной авторы используют бинарную переменную, которая равна единице, если респондент в год перед выборами смотрел телеканал НТВ, и равна нулю в противном случае. Данные по этой переменной были получены путем выборочного опроса избирателей. Зависимая переменная равна единице, если респондент голосовал за определенную партию (например, за "Единство"), и равна нулю в противном случае. Авторы включают в модель сет контрольных переменных, которые отражают индивидуальные характеристики избирателя. Поясните, с какими источниками эндогенности регрессора авторы скорее всего должны были столкнуться в своей статье. Какое решение этой проблемы в модели Вы можете предложить?

Источник: Enikolopov R., Petrova M., Zhuravskaya E. 2022. Media and Political Persuasion: Evidence from Russia // American Economic Review. Vol. 111(7). Pp. 3253-85.