Семинар «Панельные данные»

Задание выполняется на основе файла с данными mus08psidextract.dta и do-файла panel2016.do. В работе оценивается зависимость логарифма заработной платы от опыта, продолжительности рабочей недели (в часах) и длительности обучения индивида (в годах).

1. Предположите характер зависимости логарифма заработной платы от предложенных факторов.
2. Приведите описательные статистики для выбранных переменных. Проверьте, нет ли нетипичных или абсурдных значений в данных. Если необходимо все дальнейшие пункты выполняйте без учета этих индивидов.
3. Постройте гистограмму для зависимой переменной. Имеет ли смысл использование логарифма заработной платы вместо заработной платы?
4. Изучите динамику средних значений исследуемых переменных по годам.
5. Изучите диаграммы рассеяния для зависимой переменной и каждого из используемых факторов.
6. Рассчитайте коэффициенты корреляции для используемых переменных.
7. Оцените сквозную регрессию и сквозную регрессию с робастными стандартными ошибками в предположении, что дисперсия ошибок может быть разной для изучаемых индивидов (далее использовать робастные стандартные ошибки). Сравните результаты.
8. Можно ли предполагать наличие автокорреляции в остатках? Рассчитайте значения коэффициентов автокорреляции для всех лагов до шестого включительно. Подтверждается ли ваше предположение?
9. Оцените модель с фиксированными эффектами. Сравните со сквозной регрессией.
10. Найдите LSDV-оценки параметров модели. Сравните с FE-оценками.
11. Оцените модель со случайными эффектами. Сравните с FE-оценками и оценками сквозной регрессии.
12. С помощью теста Бреуша-Пагана сравните модель со случайными эффектами и сквозную регрессию.
13. С помощью теста Хаусмана сравните модели с фиксированными и случайными эффектами (обратите внимание на ограничения теста).
14. Выберите наилучшую модель на основе пунктов 12 и 13. Оцените модель с добавлением годовых дамми-переменных. Проверьте гипотезу об отсутствии временного эффекта.