

Домашнее задание №2

Практическая эконометрика

21 ноября 2020 г.

1 Общие требования

Задание выполняется 1-2 студентами, сдаётся на онэкон в виде архива с 4 файлами - оформленного связного текста с ответами на вопросы, таблицами и графиками (если они необходимы по заданию) в формате pdf, и 1 файл с кодом в R (по одному для секций 2, 3). Они должны называться `HA2text.pdf` и `HA2code.R`. Архив должен называться именами студентов в формате

`kalashnov_g_suchkova_o.zip`. Файлы R должны запускаться с помощью команды `source('HA2code.R')`. В коде в первой строчке в качестве комментария напишите, пожалуйста, имена автора(ов). Часть вопросов факультативные, ответы на них оцениваются бонусными баллами (которые выходят за пределы 60 баллов на домашки).

Срок сдачи задания - 8 декабря 23:59

2 Задание

(20 баллов + 5 бонусов)

В рамках этой домашней работы мы попробуем реплицировать результаты статьи Anderson, M. L. (2012). The benefits of college athletic success: An application of the propensity score design with instrumental variables (No. w18196). National Bureau of Economic Research.

Вопросы этой работы разбиты на 4 логических блока, которые содержат 2 типа вопросов: первые 3 блока содержат письменные вопросы. 4-ый блок предлагает вам написать код.

2.1 Общие вопросы и разность разностей

(3 балла + 1 бонус)

- Какой конкретно эффект хотят оценить авторы работы (проинтерпретируйте один из результатов)? Скажите, интересует ли их эффект в среднем или эффект на воздействованных (treatment on the treated).
- Поясните, какие регрессии в таблицах с результатами являются оценками разности разностей
- Назовите два аргумента, которые авторы приводят в пользу этих оценок
- Бонус: Как еще можно получить оценку разности разностей, кроме способа, приведенного в статье?

2.2 Propensity Score

(7 баллов)

- Приведите 2 аргумента, почему авторы считают, что вероятность выиграть в соревновании не удовлетворяет предпосылке экзогенности?
- Объясните, почему предпосылка unconfoundedness может быть верна при использовании букмекерской ожидаемой разнице в очках в качестве X .
- Назовите аргумент авторов против включения наблюдаемых ими переменных вместо букмекерской ожидаемой разницы в очках
- Почему авторы используют propensity score blocking, а не propensity score weighting?
- Зачем авторы включают сам propensity score в основную регрессию? Каким свойством обладает оценка (как называется и в чем суть), если одновременно включается контроль и используется propensity score?
- Как авторы критикуют их собственный способ оценки propensity score?
- Какими способами (во всяком случае 2) авторы решают проблему отсутствия пересечения в тритментах (overlap)?

2.3 IV

(3 балла + 2 бонуса)

- Если авторы гарантируют *unconfoundedness* с помощью propensity score, какую проблему решает использование инструментальных переменных?
- Бонус: Предположим, assumption 3 не выполнено, как это меняет формулу (3). Почему, не предполагая assumption 3, мы не можем оценить интересующий нас эффект?
- Почему авторы называют технику, которой они пользуются инструментальными переменными?
- Поясните, почему first stage модель суммы propensity scores даст те же результаты, что оценка эффекта на сумму побед?
- Почему оценка суммы propensity scores будет более точной? (бонус)

2.4 Оценка эффекта

(7 баллов + 2 бонуса)

В рамках этой секции мы воспользуемся данными из статьи, чтобы получить оценку эффекта успехов команды университета на пожертвование средств университету бывшими атлетами без доверительного интервала (переменная `athletic_donors`)

В этот раз от вас не требуется точная репликация результата авторов (скорее всего цифра будет значительно отличаться)

- Оцените модель propensity score с помощью многочлена пятой степени от переменной `line` на данных `covers_data`. Добавьте в данные колонку `pscore`. Этот пункт можно пропустить и оценивать модели без propensity score.
- Создайте зависимую переменную на данных `college_data`. Зависимая переменная – это прирост пожертвований университету за 2 года к периоду $t+1$. Этот пункт можно пропустить и оценить модель не на разницах, а на абсолютных значениях
- Сцепите данные из `covers_data` и `college_data` по полям `teamname`, `year`.
- Оцените reduced form регрессию на переменную интереса, фиксированные эффекты недели и их комбинации, взвесив на обратные propensity score.

- Составьте зависимую переменную для first stage регрессии: сумма оцененных propensity score с недели $s + 1$. Оцените first stage регрессию на переменную интереса, фиксированные эффекты недели и их комбинации, взвесив на обратные propensity score.
- Оцените финальный эффект по формуле, данной в статье (2 балла)
- Оцените тот же эффект методом blocking и включите propensity score в основную регрессию (2 бонуса)