

Микроэконометрика

Домашнее задание №2

I. Оформление и сроки

1. На титульном листе должны быть указаны: имя, фамилия, группа и номер домашнего задания.

2. Таблицы и графики должны быть подписаны и пронумерованы. Все оси подписываются на русском языке.

3. Все приведенные графики и таблицы должны быть проинтерпретированы со ссылкой на их номер.

4. Задание можно выполнять, используя любые программные средства (не обязательно R или Stata). К работе прикладываются файлы (R-файл, do-файл и т.д.), при помощи которых можно воспроизвести результаты вашего исследования. Из содержания приложенного файла должно быть ясно, какая часть кода или выполненных операций относится к какому заданию. Например, в .R или .do файле можно расставить комментарии по поводу того, какая часть кода относится к тому или иному заданию.

5. Работу следует отправить на почту bogdanpotanin@gmail.com до 22.12.2019 включительно. Тема письма **обязательно** должна иметь следующий формат “Фамилия_Имя_Группа_Микроэконометрика_2”. В противном случае отправленное вами задание не будет получено.

6. Все пункты работы должны быть выполнены **последовательно**. То есть оформление должно быть следующим:

Задание №1. Текст задания.

Ваш текст.

Задание №2. Текст задания.

Ваш текст.

...

Задание №N. Текст задания.

Ваш текст.

Где N – номер последнего задания.

7. Все задания выполняются индивидуально.

8. Все переменные в таблицах и на графиках должны быть понятным образом обозначены. Например, название переменной `ufj15` — не подходит, даже если в тексте

указано, что именно данная переменная означает. Если эта переменная относится к, например, полу, то её следует назвать male, gender или sex, а также дать интерпретацию кодам, например, 0 – женщина, 1 – мужчина.

II. Задания

В данной работе вы будете изучать, как различные факторы влияют на предложение труда. В данных РМЭЗ эта переменная закодирована как “uj6.2”.

Скачайте базу данных РМЭЗ за 2016 год (25 волна) по индивидам. Это те же самые данные, которые использовались нами на семинаре.

Обратите внимание, что некоторые задания имеют от одной до трех звездочек: в зависимости от уровня сложности. Выполнение заданий со звездочками не обязательно для получения оценки до восьми включительно. Для получения оценки девять необходимо выполнить по крайней мере одно задание со звездочкой. Получить оценку десять можно при условии выполнения заданий, суммарное количество звездочек в которых равняется, по крайней мере, трем: например, можно выполнить три задания с одной звездочкой либо одно задание с тремя звездочками. Верное выполнение заданий со звездочками также позволит получить дополнительные баллы, тем самым компенсируя возможные потери баллов в других заданиях вследствие допущенных неточностей.

Часть 1. Теория и гипотезы.

1.1. Выберите независимые переменные. Теоретически обоснуйте выбор каждой из них. Укажите предполагаемые направления эффектов. При этом вам понадобится как минимум одна непрерывная переменная и одна дамми переменная (не рекомендуется брать больше трех различных переменных, не считая их нелинейных преобразований: квадрат, логарифм, перемножение с целью получения переменной взаимодействия и т.д.).

1.2. Сформулируйте по крайней мере одну гипотезу о наличии эффекта взаимодействия и нелинейного эффекта (например, квадратичного). Теоретически обоснуйте выдвигаемые вами гипотезы. Включите нелинейные переменные и переменные взаимодействия в вашу модель.

1.3. Определитесь с тем, будете ли вы изучать влияние на предложение труда среди мужчин и женщин в рамках единой модели, либо остановитесь лишь на одном из полов. Обоснуйте свой выбор теоретически.

1.4. Определите границы усечения вашей зависимой переменной и теоретически обоснуйте их.

Часть 2. Обработка данных.

2.1. Приведите ваши данные в порядок, удалив пропуски и очистив используемые переменные от неточностей (осуществляется в коде, в работе это описывать не обязательно).

Часть 3. Тобит модель

3.1. Оцените тобит модель, предварительно записав максимизируемую функцию правдоподобия. Результат представьте в форме таблицы (можно, например, использовать выдачу из stata, R или python).

3.2. Проинтерпретируйте полученные значения оценок для каждой независимой переменной. Поясните, как полученные результаты соотносятся с высказанными вами ранее предположениями.

3.3. Запишите формулы, по которым можно рассчитать предельные эффекты в тобит модели для переменной, входящей линейно, в отношении:

А) $E(y^*)$

Б) $E(y)$

В) Вероятности того, что индивид работает

3.4. Проинтерпретируйте предельный эффекты для всех независимых переменных входящих нелинейно или имеющих взаимодействие на $E(y^*)$.

3.5. Проинтерпретируйте предельный эффекты для всех независимых переменных входящих линейно на $E(y)$.

3.6. Проинтерпретируйте предельный эффекты для всех независимых переменных на вероятность того, что индивид работает.

3.7. Предскажите значения $E(y^*)$, $E(y)$ и вероятности занятости для индивида с вашими характеристиками, выписав формулу, по которой осуществлялся расчет.

3.8*. Выведите, описав процесс вывода в тексте работы, формулу для расчета предельного эффекта переменной, входящей нелинейно, на $E(y)$.

Часть 4. Тестирование спецификации тобит модели

4.1. Проверьте, можно ли исключить из модели эффект взаимодействия, предварительно записав нулевую гипотезу и статистику теста, а также указав ее распределение.

4.2 Проверьте, можно ли оценивать совместную модель для тех, кто состоит в браке и тех, кто нет, предварительно записав нулевую гипотезу и статистику теста, а также указав ее распределение.

4.3. Опишите, к чему могут приводить, во-первых, гетероскедастичность, а во-вторых, нарушение допущения о распределении случайной ошибки в тобит модели.

4.4*. Запишите функцию правдоподобия для тобит модели с гетероскедастичной случайной ошибкой и формально опишите тестирование гипотезы о наличии гетероскедастичности. Для тех, кто работает в R, дополнительно следует оценить параметры данной модели используя функцию с семинара и представить результаты в форме таблицы, а также сделать вывод о наличии или отсутствии гетероскедастичности.

Часть 5. Сравнение моделей.

5.1. Оцените две дополнительные модели: с помощью МНК и усеченной регрессии, представив результаты оценивания в форме таблицы, а также выписав функцию правдоподобия для усеченной регрессии.

5.2. Какие оценки являются более эффективными: тобит модели или у усеченной регрессии. Ответ обоснуйте.

5.3. Выберите лучшую модель на основании критерия AIC.

5.4*. Выберите лучшую модель по критерию MSE, посчитанному на тестовой выборке.

5.5***. Оцените модель при допущении о том, что случайные ошибки имеют распределение Лапласа (https://en.wikipedia.org/wiki/Laplace_distribution), используя функцию <http://ugrad.stat.ubc.ca/R/library/rmutl/html/Laplace.html> (для R) или <https://www.stata.com/manuals/fnfunctionsbyname.pdf> (для stata см. описание функции где фигурирует laplace), предварительно выписав функцию правдоподобия. Предварительно запишите функцию правдоподобия. Представьте результаты оценивания в форме таблицы с указанием реализаций оценок коэффициентов, стандартных ошибок и p-value. Для тех, кто осуществляет задание в stata, вместо реальных данных можно использовать симулированные с кратким описанием процесса их генерации в тексте работы.