

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 studypotanin@gmail.com **не позднее чем 06.12.2021 до 8.00 утра включительно**. Тема письма **обязательно** должна иметь следующий формат “Фамилия_Имя_Группа_Микроэконометрика_2”, например, “Потанин_Богдан_123_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.

9. Работа отправляется **только в pdf** формате и **обязательно** должна иметь название “Фамилия_Имя_Группа_Микроэконометрика_2.pdf”, например, “Потанин_Богдан_123_Микроэконометрика_2.pdf”.

10. Во всех заданиях, где требуется оценить что-либо для индивида с произвольными характеристиками, необходимо, предварительно, выписать в форме таблицы характеристики соответствующего индивида. Если характеристики произвольного индивида на протяжении работы не меняются, то можно указать их в отдельной таблице в приложении и ссылаться на нее на протяжении работы.

11. При проведении тестов необходимо в тексте работы указать нулевую гипотезу и тестовую статистику, а также её асимптотическое распределение при верной нулевой гипотезе. Исключение составляют лишь обычные тесты о значимости одного регрессионного коэффициента, по поводу которых достаточно привести лишь краткую устную интерпретацию полученного результата.

12. Формулы можно как набирать на компьютере, так и вставлять картинкой с аккуратным рукописным тестом.

II. Задания

В данной работе вы будете изучать, как различные факторы влияют на зарплату женщин с учетом неслучайного отбора в число занятых. Для этого используется встроенный датафрейм, описание которого дано в файле “Homework1.R”. Для тех, кто не использует R, прикладывается excel файл “Mroz87.xlsx” с соответствующими данными.

Часть 1. Теория и гипотезы

1.1. Выберите независимые переменные для уравнения зарплаты и уравнения занятости. Кратко теоретически обоснуйте выбор каждой из них: не обязательно со ссылками на литературу, достаточно здравого смысла. Укажите и кратко обоснуйте предполагаемые направления эффектов. Уравнение занятости должно включать по крайней мере одну переменную, которой не было в уравнении зарплаты, и одну переменную,

которая есть в уравнении зарплаты. Желательно, чтобы общая для двух уравнений переменная была непрерывной, например, возраст, а также не использовать более трех переменных в каждом из уравнений.

Часть 2. Модель Тобина

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

2.2. Опишите преимущества Тобит модели над усеченной регрессией. Объясните, в каких случаях можно использовать усеченную регрессию, но не получится использовать Тобит модель: приведите гипотетический (можно использовать фантазию) пример.

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

2.4. Для индивида с произвольными характеристиками укажите (предварительно записав используемые для расчетов формулы):

A) $E(y^*)$

Б) $E(y)$

В) Вероятности того, что индивид работает

2.5. Для индивида с произвольными характеристиками рассчитайте предельный эффект любой переменной (не дамми), входящей линейно (предварительно записав используемые для расчетов формулы) на:

A) $E(y^*)$

Б) $E(y)$

В) Вероятности того, что индивид работает

2.6*. Добавьте в модель нелинейный эффект (например, квадрат). Повторите предыдущий пункт для переменной, имеющей нелинейный эффект.

2.7**. При помощи LR теста проверьте гипотезу о гомоскедастичности в Тобит модели, предварительно формально записав предполагаемые нулевой гипотезой ограничения на параметры, асимптотическое распределение тестовой статистики (при верной нулевой гипотезе) и максимизируемую в гетероскедастичной Тобит модели функцию правдоподобия. При этом уравнение дисперсии должно включать по крайней мере одну переменную, не входящую в основное уравнение. Укажите негативные

последствия, к которым может приводить отсутствие учета гетероскедастичности при условии ее наличия в Тобит модели.

2.8***. Для индивида с произвольными характеристиками в Тобит модели с гетероскедастичной случайной ошибкой рассчитайте (предварительно записав соответствующую формулу) предельный эффект переменной (не дамми), входящей и в основное уравнение, и в уравнение дисперсии, на:

А) $E(y^*)$

Б) $E(y)$

В) Вероятности того, что индивид работает

Часть 3. Модель Хекмана

3.1. Оцените модель Хекмана с помощью метода максимального правдоподобия, предварительно записав максимизируемую функцию правдоподобия и указав независимые переменные в уравнении занятости, которое должно иметь по крайней мере одну переменную, не входящую в уравнение зарплаты. Результат представьте в форме таблицы (можно, например, использовать выдачу из stata, R или python).

3.2. Опишите отличия модели Хекмана от модели Тобина.

3.3. Воспользуйтесь методом Хекмана, основанным на двухшаговой процедуре и сравните оценки, с полученными с использованием метода Хекмана, основанном на методе максимального правдоподобия. Опишите относительные преимущества и недостатки обоих методов.

3.4. Проинтерпретируйте значимость и значение оценки корреляции между случайными ошибками в обоих оцененных моделях. Укажите, можно ли было бы обойтись оцениванием обычной МНК модели.

3.5. В любой из двух оцененных в данном разделе моделей для индивида с произвольными характеристиками рассчитайте (предварительно записав формулу):

А) $E(y^* | z = 1)$ и $E(y^* | z = 0)$

Б) предельный эффект любой переменной (не дамми), входящей линейно и в основное уравнение, и в уравнение занятости, на $E(y^* | z = 1)$ и $E(y^* | z = 0)$.

Часть 4. Модель Ньюи.

4.1*. Опишите преимущества и недостатки метода Ньюи по сравнению с методом Хекмана.

4.2***. Взяв за основу любую модель бинарного выбора (для простоты можно и параметрическую, например, логит), произвольную сглаживающую функцию и используя leave-one-out кросс-валидацию для подбора степени полинома, оцените модель Ньюи, описав осуществленные для ее построения шаги. Результат представьте в форме таблицы, содержащей оценки и бутстрапированные стандартные ошибки. Сравните оценки модели Ньюи и моделей Хекмана, основанных на двухшаговой процедуре и ММП.