

Универсиада «Ломоносов» по эконометрике 2024
40 баллов (100%)
Комментарии по критериям проверки отборочного этапа

Задание 1 (10 баллов)

А) (6 баллов)

По 1 баллу за каждую из оценок коэффициентов α (изменится на $100 * (1 - \hat{\beta})$) и β (оценка не изменится) с пояснениями (например, с выкладками через свойства ковариации и дисперсии).

2 балла за ответ и пояснения о коэффициенте детерминации R^2 . R^2 остаётся неизменным, в качестве обоснования, например, можно привести выкладки через равенство R^2 квадрату коэффициента корреляции между зависимой переменной и регрессором и свойства ковариации и дисперсии.

2 балла за ответ и пояснения о статистической значимости оценки коэффициента β . Статистическая значимость не изменится. В качестве обоснования можно привести формулу расчётной t-статистики¹, в знаменателе которой стоит стандартная ошибка оценки коэффициента β , рассчитываемая через сумму квадратов остатков (не изменяется) и выборочную дисперсию регрессора (не изменяется).

Б) (4 балла)

По 1 баллу за каждую из оценок коэффициентов α (не изменится) и β (уменьшится в 100 раз) с пояснениями (например, с выкладками через свойства ковариации и дисперсии).

1 балл за ответ и пояснения о коэффициенте детерминации R^2 .

1 балл за ответ и пояснения о статистической значимости оценки коэффициента β .

Задание 2. (7 баллов) Гарри Поттер и строгая мультиколлинеарность

А) (3 балла)

1 балл за ответ «нет» и любое пояснение про строгую мультиколлинеарность/линейную зависимость, которая не позволяет вообще получить какие-либо оценки коэффициентов из уравнения (1):

$$y_i = \alpha + \beta_1 * G_i + \beta_2 * S_i + \beta_3 * K_i + \beta_4 * H_i + \beta_5 * d_i + \varepsilon_i \quad (1)$$

1 балл за пояснения, как проверить гипотезу: убрать из (1) одну из бинарных переменных, кодирующих факультет Хогвартса, чтобы избавиться от строгой мультиколлинеарности и оценить уравнение вида:

$$y_i = \alpha + \beta_1 * G_i + \beta_2 * S_i + \beta_3 * K_i + \beta_5 * d_i + \varepsilon_i \quad (1^*)$$

Нулевая гипотеза тогда формулируется как $\beta_1 = \beta_2 = \beta_3 = 0$.

Альтернативная гипотеза:
$$\begin{cases} \beta_1 \neq 0 \\ \beta_2 \neq 0 \\ \beta_3 \neq 0 \end{cases}$$

¹ См., например <https://books.econ.msu.ru/Introduction-to-Econometrics/chap02/2.4/>

Обратите внимание, что запись $\beta_1 + \beta_2 + \beta_3 = 0$ неверна, т.к. это другая гипотеза!

1 балл за схему проверки сформулированной гипотезы через F-тест: как выглядит расчётная статистика, табличная, критерий их сравнения.

Б) (4 балла)

2 балла за ответ с пояснениями: оценки из уравнения (3)

$$y_i = \alpha + \beta_1 * G_i * d_i + \beta_2 * S_i * d_i + \beta_3 * K_i * d_i + \beta_4 * H_i * d_i + \varepsilon_i \quad (3)$$

можно получить. Уравнение не нужно изменять.

Приведём пример данных:

| i | d_i | G_i | S_i | K_i | H_i |
|---|-------|-------|-------|-------|-------|
| 1 | 1 | 1 | 0 | 0 | 0 |
| 2 | 0 | 1 | 0 | 0 | 0 |
| 3 | 1 | 0 | 1 | 0 | 0 |
| 4 | 0 | 0 | 1 | 0 | 0 |
| 5 | 1 | 0 | 0 | 1 | 0 |
| 6 | 0 | 0 | 0 | 1 | 0 |
| 7 | 1 | 0 | 0 | 0 | 1 |
| 8 | 0 | 0 | 0 | 0 | 1 |

(У Гарри всего 200 наблюдений, т.е. пусть такую таблицу можно повторить $200/8=25$ раз).

Тогда:

| i | d_i | G_i | S_i | K_i | H_i | $G_i * d_i$ | $S_i * d_i$ | $K_i * d_i$ | $H_i * d_i$ |
|---|-------|-------|-------|-------|-------|-------------|-------------|-------------|-------------|
| 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 2 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| 4 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 6 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 7 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |
| 8 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

Столбцы с произведениями $G_i * d_i, S_i * d_i, K_i * d_i, H_i * d_i$ не являются линейно независимыми.

1 балл за гипотезу

1 балл за схему проведения F-теста.

Задание 3 (12 баллов)

А)

1 балл — за правдоподобие, 1 балл — за правильно найденные $\Lambda(\alpha)$ и $\Lambda(\gamma)$, где $\gamma = \alpha + \beta$ (например, а можно было и без замены переменных). Остальные 2 балла — за правильный ответ (итого 4 балла).

Если в решении не написано, почему нашли максимум, а не седло или минимум, то за это не снимались баллы. Выпуклость логгосса считалась общим знанием.

Б)

Всё хорошо — 2 балла, есть неаккуратности — 1 балл, остальные случаи — 0 (итого 2 балла)

В)

По 0.5 за каждую правильно посчитанную ошибку, остальной балл — за правильный вывод о качестве модели (итого 2 балла).

Г)

формально обосновано отсутствие экстремума (ссылкой на монотонность или неразрешимость FOC) — 2 балла, просто дан правильный ответ, но нет строгого обоснования — 1 балл, остальные случаи — 0 (итого 2 балла)

Д.

Если написано что-то разумное о любой модели прогноза вероятности банкротства, то 1 балл, иначе 0 (итого 1 балл)

Топ самых частых ошибок в задаче 3:

1. В пункте в) участник вычисляет вероятности ошибок I/II рода, деля на все 2000 наблюдений, а не только на те, где $y = 1$ или 0
2. Там же либо участник ничего не говорит о качестве модели, либо говорит неоднозначные фразы / апеллирует к высокой ассигасу или низкой вероятности одной из ошибок
3. В пункте г) участник просто подставляет нуль в вычисленные в п. а) оценки и заключает, что оптимума нет (за это 1 балл из 2)
4. В пункте а) участник либо бросает задачу, выписал FOC, либо получает неправильные оценки (и, соответственно, автоматически теряет пункт б))
5. В пункте б) участник выписывает условные вероятности $P(y = 1 | x = 1)$ и $P(y = 1 | x = 0)$ и либо бросает пункт на этом, либо неправильно собирает из них $P(y = 1)$

Задание 4 (11 баллов)

По 2 балла за проверку каждой гипотезы: определение спецификации уравнений для проверки двух гипотез из условия; запись трёх гипотез в терминах модели (через коэффициенты уравнения).

Для первой гипотезы «В развитых странах наблюдается эффект декаплинга, а в странах с развивающимися рынками — не наблюдается» достаточно было линейного уравнения вида:

$$y_{it} = \alpha + \beta_1 * x_{it} + \beta_2 * x_{it} * d_i + \varepsilon_{it} \quad (*), \text{ где:}$$

y_{it} — показатель загрязнений или использования природных ресурсов (необходимо конкретизировать в пункте про данные) в стране i в году t ;

x_{it} — показатель производства товаров и услуг (или темпов роста ВВП – необходимо уточнить показатель и источник в пункте про данные) в стране i в году t ;

d_i — бинарная переменная, равная 1, если страна i относится к развитым, и 0, если к развивающимся;

ε_{it} —случайный шок (шум)

Тогда гипотезе о том, что в развитых странах наблюдается эффект декаплинга, сумма коэффициентов $\beta_1 + \beta_2$ должна быть значимой и отрицательной.

А «а в странах с развивающимися рынками не наблюдается эффект декаплинга» соответствует β_1 незначимый либо значимый положительный.

Также в решениях встречались различия абсолютного и относительного декаплинга и сравнение коэффициентов с 1 – тоже верно.

Альтернативно, опираясь на слово «тренды» в условии задания, можно было бы проверить гипотезу 1 аналогично с работой Cohen et al. (2017) «Emissions and Growth: Trends and Cycles in a Globalized World» // IMF WP 17/191. <https://www.imf.org/en/Publications/WP/Issues/2017/08/30/Emissions-and-Growth-Trends-and-Cycles-in-a-Globalized-World-45202>. Выделить (например, фильтром Ходрика-Прескотта) для каждой страны трендовую составляющую в темпах роста реального ВВП и выбросах парниковых газов, а затем оценить по очищенным данным для каждой страны долгосрочную эластичность из регрессии трендовой части выбросов парниковых газов в зависимости от трендовой части темпов роста реального ВВП.

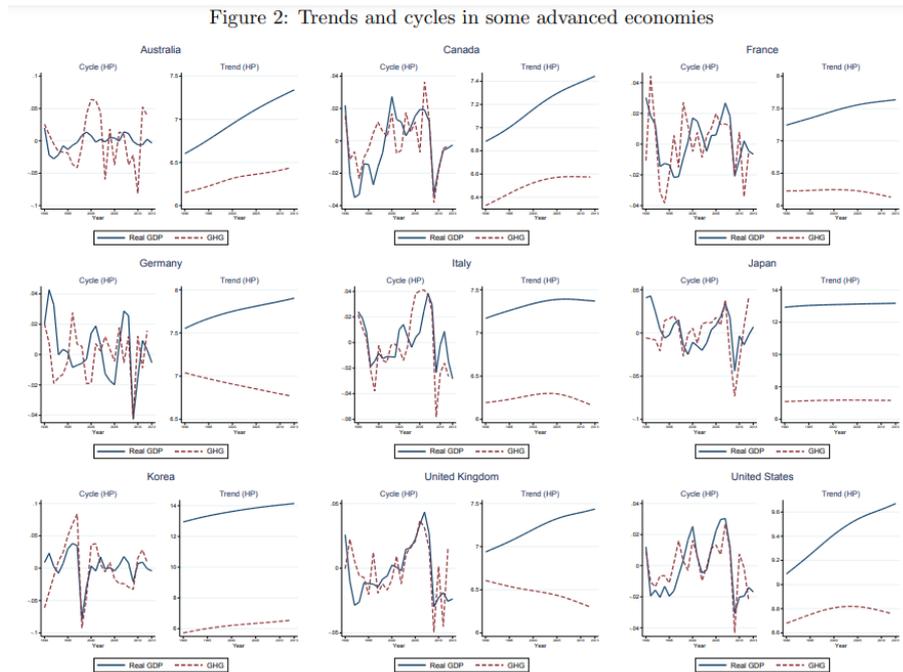


Рис.1 Источник – Cohen et al. (2017)

Гипотеза 2 «Имеет место так называемая «экологическая кривая Кузнеца», то есть закономерность: «при росте дохода на душу населения уровень деградации окружающей среды сначала растет, а затем — по мере достижения определенного уровня благосостояния — он начинает снижаться»»

Достаточно было оценить уравнения с квадратичной зависимостью:

$$y_{it} = \alpha + \beta_1 * x_{it} + \beta_2 * x_{it}^2 + \varepsilon_{it} \quad (**), \text{ где:}$$

y_{it} — показатель загрязнений или использования природных ресурсов (необходимо конкретизировать в пункте про данные) в стране i в году t ;

x_{it} — доход на душу населения (необходимо уточнить показатель и источник в пункте про данные) в стране i в году t ;

x_{it}^2 — он же во второй степени;

ε_{it} —случайный шок (шум)

Тогда для подтверждения кривой Кузнецца необходимо, чтобы оценки соответствовали параболе ветвями вниз и вершиной правее нуля по оси дохода, то есть $\beta_1 > 0$ и $\beta_2 < 0$.

Тогда вершина параболы $\frac{-\beta_1}{2\beta_2}$.

Альтернативный ответ с пороговой регрессией Хансена тоже верен.

Гипотеза 3 «Чем больше доля углеводородов в экспорте страны, тем выше тот уровень благосостояния, после достижения которого начинается нисходящая часть экологической кривой Кузнецца.»

Для её проверки достаточно немного изменить уравнение с квадратичной зависимостью:

$$y_{it} = \alpha + \beta_1 * x_{it} + \beta_2 * x_{it}^2 + \beta_3 * x_{it} * oil_{it} + \varepsilon_{it} \quad (***) , где:$$

y_{it} — показатель загрязнений или использования природных ресурсов (необходимо конкретизировать в пункте про данные) в стране i в году t ;

x_{it} — доход на душу населения (необходимо уточнить показатель и источник в пункте про данные) в стране i в году t ;

x_{it}^2 — он же во второй степени;

oil_{it} — доля углеводородов в экспорте (необходимо уточнить показатель и источник в пункте про данные) в стране i в году t ;

ε_{it} —случайный шок (шум)

Тогда вершина параболы рассчитывается как $\frac{-(\beta_1 + \beta_3 oil_{it})}{2\beta_2}$. И при соблюдении условий $\beta_1 > 0$, $\beta_3 > 0$ и $\beta_2 < 0$ она положительно зависит от доли углеводородов в экспорте страны.

Альтернативно можно было

*либо 1) оценить уравнение (***) для каждой стран, чтобы получить для каждой страны свою величину вершины параболы и оценить новое уравнение зависимости вершины от доли углеводородов в экспорте.*

*либо 2) оценить уравнение (***) по двум группам стран – экспортёрам углеводородов и остальным странам и сравнить вершины параболы. Тогда нужно предварительно оговорить критерий разделения стран.*

Описание данных: какая выборка (страны, период) и какие статистические показатели используются – 2 балла

Указание метода оценки, необходимых тестов и критериев оценки качества эконометрического уравнения – 2 балла.

За аргументацию со ссылками на теорию и/или опубликованные эмпирические работы и/или отчёты ведомств, международных организаций – 1 балл. Если просто дана ссылка на работу, но нет объяснения, чем она пригодилась в решении, то 0.5 балла.