Решение задачи 1.

(1)

$$\hat{\theta}_{\text{МНК}} = \frac{\sum t * y_t}{\sum t^2}$$

2 балла

$$E\hat{\theta}_{\text{МНК}} = E\left(\frac{\sum t * y_t}{\sum t^2}\right) = E\left(\frac{\sum t * (\theta * t + \varepsilon_t + \varepsilon_0)}{\sum t^2}\right) = \theta$$

2 балла

$$V\hat{\theta}_{\text{МНК}} = V\left(\frac{\sum t * y_t}{\sum t^2}\right) = V\left(\frac{\sum t * (\theta * t + \varepsilon_t + \varepsilon_0)}{\sum t^2}\right) = V\left(\frac{\sum \varepsilon_t t + \varepsilon_0 \sum t}{\sum t^2}\right) =$$

$$= \left(\frac{1}{\sum t^2} + \frac{(\sum t)^2}{(\sum t^2)^2}\right)\sigma^2 \xrightarrow{T\to\infty} 0$$

4 балла

Оценка несмещена и дисперсия стремится к нулю, следовательно, оценка состоятельна.

2 балла

$Cov\left(\varepsilon_t + \varepsilon_0, \varepsilon_j + \varepsilon_0\right) = \sigma^2 > 0$, *следовательно, предпосылки КЛММР нарушены, и оценка является неэффективной.*

(2)-(3)

Следует применить обобщенный МНК.

$$\widehat{\theta_{\text{ОМНК}}} = (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}y \qquad \textit{2 балла}$$

Где $X' = (1,2,3,4), y' = (-1,4,6,8)$

$$\Omega = \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}\sigma^2 \qquad \textit{6 баллов;} \qquad \begin{pmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}^{-1} = \frac{1}{10}\begin{pmatrix} 8 & -2 & -2 & -2 \\ -2 & 8 & -2 & -2 \\ -2 & -2 & 8 & -2 \\ -2 & -2 & -2 & 8 \end{pmatrix}$$

$$\widehat{\theta_{\text{ОМНК}}} = 2{,}3 \qquad \textit{7 баллов}$$

Задача 2 (25 баллов).

Ответы:

(1) (10 баллов) $y_{11} - y_{10} - y_{01} + y_{00}$
(2) (10 баллов) $\frac{4}{m}\sigma^2$ или, что то же самое: $\frac{16}{n}\sigma^2$.
(3) (5 баллов) 4

Задача 3 (25 баллов).

Решение:

(а) Скорее всего, уровень гениальности коррелирован с решением получить высшее образование, поэтому МНК-оценка коэффициента в парной регрессии Y по X будет несостоятельной оценкой параметра θ_1. (2 балла)

В данном случае, хорошей идеей может быть использование метода инструментальных переменных. Тем более, что в нашем распоряжении есть инструмент: близость к университету. Скорее всего это хороший инструмент, так как близость к университету коррелирована с получением высшего образования, однако не коррелирована с гениальностью. (3 балла)

ММП-оценка параметра имеет вид: $\widehat{\theta_1} = \frac{COV(Y,Z)}{COV(X,Z)} = \frac{\overline{YZ}-\bar{Y}\bar{Z}}{\overline{XZ}-\bar{X}\bar{Z}}$

$$\bar{Y} = 0{,}4 * 5000 + 0{,}1 * 6000 + 0{,}4 * 3000 + 0{,}1 * 4000 = 4200$$

$\bar{X} = 0{,}5$ (половина людей в выборке получили высшее образование)

$\bar{Z} = 0{,}5$ (половина людей в выборке живут рядом с университетом)

$\overline{YZ} = 0{,}4 * 5000 + 0{,}1 * 4000 = 2400 \quad \overline{XZ} = 0{,}4$

$\widehat{\theta_1} = \frac{\overline{YZ}-\bar{Y}\bar{Z}}{\overline{XZ}-\bar{X}\bar{Z}} = \frac{2400-4200*0{,}5}{0{,}4-0{,}5*0{,}5} = 2000$ (10 баллов).

(б) В данном случае можно проверить релевантность инструментов при помощи теста на сладые инструменты. Для этого нужно вычислить F-статистику в регрессии первого шага двухшагового МНК, то есть в регрессии X по Z. Для этого нам пригодится R-квадрат в этой регрессии, который равен квадрату коэффициента корреляции между данными переменными: $R^2 = \frac{COV(X,Z)}{VAR(X)*VAR(Z)} = \frac{0{,}4-0{,}5*0{,}5}{0{,}5*0{,}5} = 0{,}3$. Соответственно, F-статистика равна: $F = \frac{0{,}3}{1-0{,}3} * \frac{1000-2}{1} = 427{,}7$

Тестовая статистика существенно больше 10, поэтому используемый инструмент является релевантным.

Пункт оценивается в 10 баллов. Если просто вычислен коэффициент корреляции между X и Z и сделан вывод о релевантности инструмента, то следует поставить 5 баллов.

Задача 4 (30 баллов)

Решение.

1) t-статистика при тестировании значимости коэффициента при переменной w равна 10/2=5.
2) Если тестировать значимость того же коэффициента с помощью двух уравнений (с помощью F статистики), то F статистика будет равна квадрату t-статистики. F=25.
3) С другой стороны эта F-статистика равна

$$F = \frac{R_{ur}^2 - R_r^2}{1 - R_{ur}^2} * \frac{n-k}{q}$$

n=103; k=3; q=1; $R_{ur}^2 = 0.8$; F=25.

Находим $R_r^2 = 0.75$

Наконец, $R_r^2 = \frac{\left(\frac{\widehat{\beta_2}}{se(\widehat{\beta_2})}\right)^2}{n-2+\left(\frac{\widehat{\beta_2}}{se(\widehat{\beta_2})}\right)^2}$. Подставляя $se(\widehat{\beta_2}) = 2$, получаем $(\widehat{\beta_2}) = 2 * 303^{0.5}$

Формула для р-квадрат может быть получена если расписать RSS через бета_2 и дисперсию икса (выборочную), а также воспользоваться формулой для ошибки бета_2 и связи ESS и s^2.

Задача 5 (40 баллов).

Решение:

Научные школы забыли бакалаврский курс эконометрики. Хуже гетероскедастичности только мультиколлинеарность. Для монополиста, максимизирующего придыль, выполняется следующее равенство (уточните у первокурсников, если забыли):

$$\frac{MP_K}{MP_L} = \frac{p_K}{p_L}.$$

В случае производственной функции Кобба-Дугласа это равенство превращается в:

$$\frac{\omega_t \beta_K K_t^{\beta_K - 1} L_t^{\beta_L}}{\omega_t K_t^{\beta_K} \beta_L L_t^{\beta_L - 1}} = \frac{p_K}{p_L}.$$

Что упрощается до:

$$\frac{\beta_K L_t}{\beta_L K_t} = \frac{p_K}{p_L}.$$

Логарифмируем $\ln L_t = const + \ln K_t$, где константа постоянна по времени. Следовательно, оба подхода подразумевают оценку моделей с коллинеарными регрессорами. Компьютер будет рушаться. Состоятельной оценки не получится. Научные школы задачу решить не смогли.