Универсиада по эконометрике 2020.

Схемы решения

Схема решения задачи 1

а) Оценка будет состоятельной. $\hat{\beta} = \frac{1}{N}\sum_{i=1}^{N}\frac{y_i}{x_i} = \frac{1}{N}\sum_{i=1}^{N}\frac{x_i\beta+e_i}{x_i} = \beta + \frac{1}{N}\sum_{i=1}^{N}\frac{e_i}{x_i}$. По закону больших чисел $\frac{1}{N}\sum_{i=1}^{N}\frac{e_i}{x_i} \xrightarrow{p} E[\frac{e}{x}]$. По свойству повторного мат. ожидания $E[\frac{e}{x}] = E[E[\frac{e}{x}|x]] = E[\frac{1}{x}E[e|x]] = 0$. Таким образом, $\hat{\beta} \xrightarrow{p} \beta$.
Критерий оценивания: 0 баллов даже в случае правильного ответа, если он получен бездоказательно или с серьезными вычислительными ошибками, полный бал только в случае использования закона Больших чисел и Повторного мат. ожидания.
б) Существует. Эффективная оценка ОМНК имеет вид:
$\widetilde{\beta} = \left(\sum_{i=1}^{N}\frac{x_i^2}{\sigma^2(x_i)}\right)^{-1}\sum_{i=1}^{N}\frac{x_iy_i}{\sigma^2(x_i)}$. Оценки идентичны, если $\sigma^2\left(x_i\right) = cx_i^2$, где c – константа.
Критерий: Правильным ответом считается верная функциональная форма.
в) Оценка останется состоятельной. Допустим: $\hat{\beta} = \left(\sum_{i=1}^{N}\frac{x_i^2}{h\left(x_i,\hat{\beta}\right)}\right)^{-1}\sum_{i=1}^{N}\frac{x_iy_i}{h\left(x_i,\hat{\beta}\right)}$ и $h(x,\hat{\beta}) \nrightarrow \sigma^2(x)$. Тогда $\hat{\beta} = \beta + \frac{1}{N}\sum_{i=1}^{N}\frac{e_i}{h(x_i,\hat{\beta})}$. По закону больших чисел $\frac{1}{N}\sum_{i=1}^{N}\frac{e_i}{h(x_i,\hat{\beta})} \xrightarrow{p} E[\frac{e}{h(x,\hat{\beta})}]$. По свойству повторного мат. ожидания $E[\frac{e}{h(x,\hat{\beta})}] = E[E[\frac{e}{h(x,\hat{\beta})}|x]] = E[\frac{1}{h(x,\hat{\beta})}E[e|x]] = 0$. Таким образом, $\hat{\beta} \xrightarrow{p} \beta$.
Критерий: 0 баллов даже в случае правильного ответа, если он получен бездоказательно или с серьезными вычислительными ошибками (например, сокращение оценки дисперсии в знаменателях является серьезной ошибкой), утверждение, что FGLS оценка состоятельна, считается бездоказательным, 3-5 балла в случае, если получен правильный ответ с незначительными вычислительными ошибками,
полный балл в случае использования ЗБЧ и свойства повторного мат. ожидания.

Схема решения задачи 2

(а) Пределы по вероятности первой, второй и третьей оценок, соответственно, составляют:

$$\widehat{\beta_1} = \frac{\sum_{i=1}^{n} y_i x_i^{(1)}}{\sum_{i=1}^{n}\left(x_i^{(1)}\right)^2} = \frac{\sum_{i=1}^{n} \beta x_i^*\left(x_i^* + \varepsilon_i^{(1)}\right)}{\sum_{i=1}^{n}\left(x_i^* + \varepsilon_i^{(1)}\right)^2}$$

$$= \frac{\beta \sum_{i=1}^{n}(x_i^*)^2 + \sum_{i=1}^{n} x_i^* \varepsilon_i^{(1)}}{\sum_{i=1}^{n}\left(x_i^* + \varepsilon_i^{(1)}\right)^2} \xrightarrow{p} \frac{E(x_i^*)^2}{E(x_i^*)^2 + \sigma_2^2} * \beta$$

$$\widehat{\beta_2} = \frac{\sum_{i=1}^{n} y_i \dfrac{x_i^{(1)} + x_i^{(2)}}{2}}{\sum_{i=1}^{n}\left(\dfrac{x_i^{(1)} + x_i^{(2)}}{2}\right)^2} = \frac{\sum_{i=1}^{n} \beta x_i^*\left(x_i^* + \dfrac{\varepsilon_i^{(1)} + \varepsilon_i^{(2)}}{2}\right)}{\sum_{i=1}^{n}\left(x_i^* + \dfrac{\varepsilon_i^{(1)} + \varepsilon_i^{(2)}}{2}\right)^2} \xrightarrow{p} \frac{E(x_i^*)^2}{E(x_i^*)^2 + \dfrac{\sigma_1^2 + \sigma_2^2}{4}} * \beta$$

$$\widehat{\beta_3} \xrightarrow{p} \frac{E(x_i^*)^2}{E(x_i^*)^2 + \sigma_2^2} * \frac{\beta}{2} + \frac{E(x_i^*)^2}{E(x_i^*)^2 + \sigma_1^2} * \frac{\beta}{2}$$

(По 6 баллов за обоснованные вычисления первых двух пределов по вероятности и еще 3 балла за третий.)

(б) Для всех трех оценок предел по вероятности по модулю меньше β, так что все три оценки несостоятельны.

Модули отклонения пределов по вероятности от истинного значения для первой, второй и третьей оценок, соответственно, составляют:

$$\left|\beta - \widehat{\beta_1}\right| \xrightarrow{p} = \frac{\sigma_2^2}{E(x_i^*)^2 + \sigma_2^2} |\beta|$$

$$\left|\beta - \widehat{\beta_2}\right| \xrightarrow{p} \frac{\dfrac{\sigma_1^2 + \sigma_2^2}{4}}{E(x_i^*)^2 + \dfrac{\sigma_1^2 + \sigma_2^2}{4}} |\beta|$$

$$\left|\beta - \widehat{\beta_3}\right| \xrightarrow{p} \left(1 - \frac{1}{2}\frac{E(x_i^*)^2}{E(x_i^*)^2 + \sigma_2^2} - \frac{1}{2}\frac{E(x_i^*)^2}{E(x_i^*)^2 + \sigma_1^2}\right) |\beta|$$

Третья оценка нам не подойдет, так как из того, что $\sigma_2^2 < \sigma_1^2$, следует, что

$$\left|\beta - \widehat{\beta_3}\right| > \left|\beta - \widehat{\beta_1}\right|.$$

(2 балла за эту идею.)

Выбор между первой второй оценками зависит от соотношения между дисперсиями ошибок измерения.

Первую оценку следует выбирать в том случае, когда для нее отклонение от истинного значения меньше:

$$\frac{\sigma_2^2}{E(x_i^*)^2 + \sigma_2^2} < \frac{\frac{\sigma_1^2 + \sigma_2^2}{4}}{E(x_i^*)^2 + \frac{\sigma_1^2 + \sigma_2^2}{4}}.$$

Или, что эквивалентно:

$$\sigma_2^2 < \frac{\sigma_1^2}{3}.$$

Соответственно, вторую оценку следует выбирать тогда, когда

$$\sigma_2^2 > \frac{\sigma_1^2}{3}.$$

Наконец, первая и вторая оценки равнозначны тогда, когда

$$\sigma_2^2 = \frac{\sigma_1^2}{3}.$$

Отметим, что по условию исследователю указаны используемые дисперсии, так что он всегда сможет сделать выбор на основе указанных рекомендаций.

(8 баллов за вычисление пограничного соотношения между дисперсиями и итоговый вывод на основе этого сравнения.)

Схема решения задачи 3

Problem 3

(a) If returns are a random walk, then $r_t = r_{t-1} + u_t$, and u_t is white noise with a variance equal to σ^2.

$$V(j) = \frac{var\left(r_t - r_{t-j}\right)}{j * var(r_t - r_{t-1})} =$$
$$\frac{var(r_{t-j} + u_t + u_{t-1} + \cdots + u_{t-j+1} - r_{t-j})}{j * var(r_{t-1} + u_t - r_{t-1})} =$$
$$\frac{var(u_t + u_{t-1} + \cdots + u_{t-j+1})}{j * var(u_t)} = \frac{j * \sigma^2}{j * \sigma^2} = 1$$

Variance ratio should be equal 1 in that case. So null hypothesis of no predictability in this case is similar to testing if variance ratio is exactly 1. One of the possible disadvantages is an alternative hypothesis. If we get that variance ratio is not one, we will only know that returns are not random walk, but we will know nothing about their predictability.

(b) According to the table, 3 and 5 horizon returns are predictable (dividend growths are not predictable). We use the usual significance test to check.

The main problem here (probably because of which 1-year returns are unforecastable) is endogeneity. Price dividend ratio drives returns but returns can also alert prices due to demand on financial markets. Possible way to overcome the problem is to use instrumental variables or GMM (latter is widely used in finance), but for these we need a solid theoretical background and really good instrument. Easier way to solve the problem is to use VAR model to account for a simultaneous causality.

(c) Here all coefficients are highly significant. It may mean that OLS specification gave biased estimates. We can infer from VAR that both returns and price dividend ratio are predictable and one Granger causes another (null hypothesis is no causality). Potential trouble is that other variables (such that risk-free rate or dividend growth can potentially affect returns and price dividend ratio and thus should be included in the model).

(d) We now propose a way that can be implemented here:

First, run VAR(1) of dividend growth and price-dividend ratio. Note that we cannot include returns because otherwise we would obtain a linearly dependent system. Let us denote VAR coefficient as follows:

	Δd_t	$p_t - d_t$
Δd_{t-1}	ϕ_d	μ_d
$p_{t-1} - d_{t-1}$	b_d	ϕ

Using Campbell-Shiller decomposition, we get:

$$
\begin{aligned}
r_{t+1} \approx \kappa_1(p_{t+1} - d_{t+1}) + \Delta d_{t+1} - (p_t - d_t) & \\
&= \kappa_1 \phi(p_t - d_t) + \kappa_1 \mu_d \Delta d_t + \kappa_1 \varepsilon_{t+1}^{pd} + \phi_d \Delta d_t \\
&+ b_d(p_t - d_t) + \varepsilon_{t+1}^{d} - (p_t - d_t) \\
&= \beta_r(p_t - d_t) + b_r \Delta d_t + \varepsilon_{t+1}^{r}
\end{aligned}
$$

where the last expression is a predictive regression. From there we can derive coefficients:

$$
\beta_r = \kappa_1 \phi + b_d - 1
$$

$$
b_r = \kappa_1 \mu_d + \phi_d
$$

Variance of the predictive coefficient is $\kappa_1^2 Var(\phi) + Var(b_d)$. Thus, we obtain t-statistics: $\frac{\beta_r}{\sqrt{(\kappa_1^2 Var(\phi) + Var(b_d)})}$.

Getting IRFs is challenging. We cannot simply plot them as our shocks are not orthogonal. To see that, notice that from Campbell-Shiller decomposition we get $\varepsilon_{t+1}^{r} = \kappa_1 \varepsilon_{t+1}^{pd} + \varepsilon_{t+1}^{d}$ which means that shocks are correlated. To obtain orthogonal shocks we should make a Cholesky decomposition.

Схема решения задачи 4

Задание 4

(a) Потенциальный источник несостоятельности оценки интересующего нас эффекта — это эндогенность принятия решения о службе в армии: по условию, индивиды сами решают, служить ли им. Возможно, есть какие-то специфические характеристики, которые побуждают индивида служить в армии и одновременно влияют на его будущую зарплату.

Набор оцененных моделей позволяет нам преодолевать эту эндогенность двумя путями: при помощи включения контрольных переменных (прочих факторов, влияющих на зарплату) и при помощи инструментальных переменных.

Модели из третьей таблицы используют оба эти пути (там есть и контрольные переменные, и для оценивания используется двухшаговый МНК (2МНК). Кроме того, в обоих случаях тест Хаусмана позволяет при уровне значимости 5% отвергнуть гипотезу о состоятельности оценок обычного МНК, так как в обоих случаях соответствующее P-значение меньше пяти сотых. Поэтому нам придется отвергнуть модели из первой таблицы и сконцентрироваться на моделях из третьей таблицы.

Для состоятельности 2МНК-оценок требуется релевантность и экзогенность используемого инструмента.

Для проверки релевантности осуществим F-тест на слабые инструменты. Его описание можно найти, например в учебниках (Сок, Ватсон, 2015) или (Картаев, 2019). Вычислив соответствующую F-статистику, обнаруживаем, что для модели №1 из таблицы №3 она существенно больше 10 (чего уж там, она даже больше 500), а для модели №2 она существенно меньше 10. Таким образом в первой модели инструмент релевантен, а во второй — нет. Так что модель №2 из таблицы №3 нам тоже не подойдет.

Из содержательных соображений можно заключить, что инструмент в модели №1 из таблицы №3 экзогенен. Действительно, вряд ли военный опыт родителей будет непосредственно влиять на зарплаты их детей.

Таким образом, остановимся на модели №1 из таблицы №3. Проверив гипотезу о равенстве нулю коэффициента при переменной SWORN_SWORD, заключаем, что эта переменная значима. Так что опыт военной службы влияет на зарплату индивида.

Коэффициент при этой переменной равен 0,239. Следовательно, при прочих равных (то есть при одинаковых образовании и опыте работы) индивид, который служил в армии, получает зарплату на

$$(e^{0,239} - 1) = 27\%$$

выше по сравнению с жителем Вестероса, который в армии не служил.

Для получения полного балла за этот пункт требуется выбрать верную модель и привести необходимые обоснования. В частности:

- осуществить тест Хаусмана,
- вычислить нужные F-статистики и осуществить тест на слабые инструменты,
- тестировать значимость коэффициента при переменной SWORN_SWORD и интерпретировать его.

(б) В этом пункте засчитывались любые разумные идеи по преодолению потенциальной эндогенности (при наличии должного обоснования. Например, предложения, связанные с рассмотрением альтернативных функциональных форм или альтернативных наборов контрольных переменных. А также предложение по использованию двух инструментов одновременно вместо одного (что позволило бы тестировать их экзогенность при помощи теста Саргана или его аналога).