Question 1

For questions 1-4, we are given the following model:

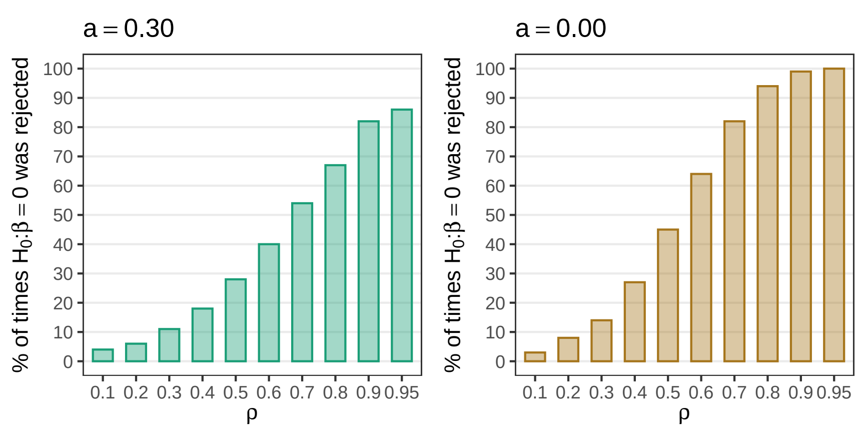
$$Y = X\beta_0 + \varepsilon \tag{1}$$
$$X = Z\Pi + V \tag{2}$$

where: Y and X are $n \times 1$ vectors which contain the endogenous variables; Z is a $n \times k$ matrix of instruments; ε and V are $n \times 1$ vectors that contain disturbances. The different rows of $\begin{pmatrix} \varepsilon \vdots V \end{pmatrix}$, are independently normally distributed, i.e.,

$$\begin{pmatrix} \varepsilon_i \\ V_i \end{pmatrix} \sim \mathcal{N}(0, \Sigma), \qquad \Sigma = \begin{pmatrix} 1 & \vdots & \rho \\ \rho & \vdots & 1 \end{pmatrix}.$$

To close the simulation scheme, we choose $n = 100$, $k = 10$ and $\Pi = a \times e_{10}$, where $e_{10} \in \mathbb{R}^{10}$ with $e_{10_1} = 1$ and all other entries are zero. Seven different values for a are proposed: $a \in \{0.3, 0.25, 0.2, 0.15, 0.10, 0.05, 0.00\}$ and ten different values for the parameter ρ: $\rho \in \{0.0, 0.1, \ldots, 0.7, 0.9, 0.95\}$. For the remaining of the question, I imposed $\beta_0 = 0$ to simulate the values. In this case, we would expect to not reject the null-hypothesis $H_0 : \beta_0 = 0$ in 95% of the cases, when using a critical value corresponding to $\alpha = 0.05$.

Figure 1: Rejection frequency of the 2SLS t-statistic for different combinations of ρ and a



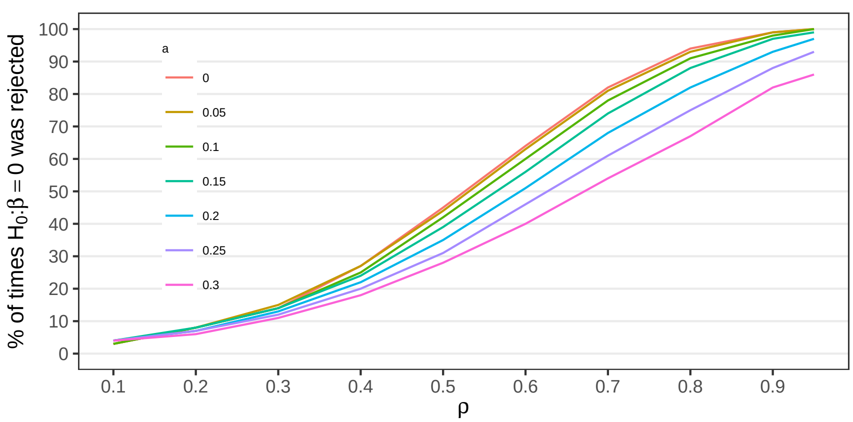
Notes: Rejection frequencies for $H_0 : \beta_0 = 0$ in the 2SLS t-statistic in 5000 simulations of the model described in (1)-(2) using $\beta_0 = 0$ and for $a = 0.30$ (left) or $a = 0.00$ (right). Frequencies in the y axis are computed as the number of relative occurrences, in 5000 replications, for each value of ρ.

The first item asks to compute the rejection frequency of the 2SLS t-statistic. As seen in the lectures, we cannot directly use the standard errors from the second stage to use in the test since this will not take into account the uncertainty that comes from estimating Π in the first stage. To compute the rejection frequency, then, I used the functions available in the R package `ivreg`, which automatically corrects the standard errors. In Figure 1 we observe that for a low value of

ρ ($\rho \in \{0.1, 0.2\}$), disregarding the value of a, we have that 5% of the times H_0 was rejected, as it should be. However for $\rho \geq 0.3$ the rejection frequency increases more than linearly, for both values of a displayed in the graphs. For $a = 0.00$ it seems that the error in the test is higher than for $a = 0.3$, but in both graphs we see that we cannot rely on the 2SLS procedure for our inference about β_0.

Figure 2 give a better comparison for what happens to the rejection frequency when we change the value of a: a low value $a = 0$ will have higher rejection frequencies uniformly for all values of ρ in our chosen grid. This behavior is consistent with the other lines, i.e., a lower a seems to be associated with higher rejection rates, which can be as high as virtually equal to 100% when ρ is sufficiently high, but as we can see in Figure 2, for $\rho > 0.2$ we already observe empirical rejection frequencies incompatible with what would be expected from the theoretical test. The **conclusion** is what we discussed in class: the 2SLS t-statistic is not robust under weak instruments and the inference will not be reliable. This is of great concern, given that in empirical applications the common situation is to have weak instruments.

Figure 2: Rejection frequency of the 2SLS t-statistic for different combinations of ρ and a



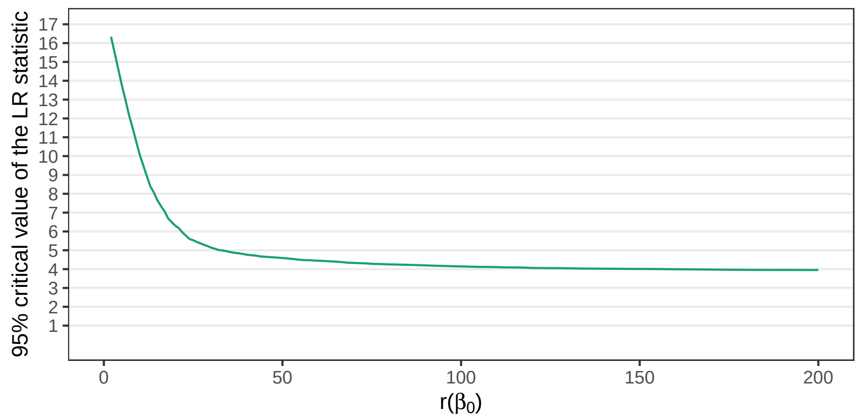
Notes: Rejection frequencies for $H_0 : \beta_0 = 0$ in the 2SLS t-statistic in 5000 simulations of the model described in (1)-(2) using $\beta_0 = 0$. Each line denotes a different value for a, ranging from 0 (top curve) to 0.3 (bottom curve). Frequencies in the y axis are computed as the number of relative occurrences, in 5000 replications, for each value of ρ and were interpolated by a simple linear approximation to facilitate the comparison between curves.

Question 2

For Question 2, we are asked to compute what are the 95% critical values of the likelihood ratio (LR) test statistic as a function of $r(\beta_0)$ for $k = 10$. Algorithm 1 (in the Appendix) has the steps implemented for this task, using s simulations and a grid of size r for $r(\beta)$.

From Figure 7 we can see that the critical value for the LR statistic is not static: we start with very high values (around 16) for values of $r(\beta_0)$ close to 0 and this has an exponential decay when we increase $r(\beta_0)$, converging to a value close to 4, for the chosen value of $k = 10$. The critical value remains fairly constant for the values in the $r(\beta_0)$ grid that are bigger than 100. We can see why this is the case by looking at the formula given in the slides for $LR(\beta_0)$:

Figure 3: Critical values of the LR statistic



Notes: Critical values for the LR statistic considering a grid $r(\beta_0) \in \{0, 1, \ldots, 200\}$ and $k = 10$. These critical values can be then compared to the actual statistic in a model and would keep the 95% confidence region.

$$LR(\beta_0) = \frac{1}{2}\left[kAR(\beta_0) - r(\beta_0) + \sqrt{(kAR(\beta_0) + r(\beta_0))^2 - 4r(\beta_0)\left[kAR(\beta_0) - LM(\beta_0)\right]}\right] \quad (3)$$

We can see that when $r(\beta_0) \to 0$, Equation (3) reduces to

$$LR(\beta_0) = \frac{1}{2}\left[kAR(\beta_0) + \sqrt{(kAR(\beta_0))^2}\right] \approx kAR(\beta_0),$$

which should be compared to the $\chi^2_{(10)}$ 95% critical value of 18.31. On the other hand, when $r(\beta_0) \to \infty$, we have that $LR(\beta_0) \to LM(\beta_0)$, which has a 95% critical value from a $\chi^2_{(1)}$ and is approximately 3.84. Note that these values are in line with what is being showed in the graph from Figure (7).

Question 3

When we repeat the exercise from Question 1 to the AR, LM and LR statistics, we found a complete different picture than the one in Figure 2. As we can see in Figure 4, all three rejection frequencies are close to the 5% threshold. There is a little bit of noise in the LR (middle graph) for higher values of ρ, but still these are very small variations. The AR statistic by definition does not change with the value of a, hence all lines overlap. The conclusion is that all three tests are robust to weak instruments and constitute a good alternative to the 2SLS t-statistic.

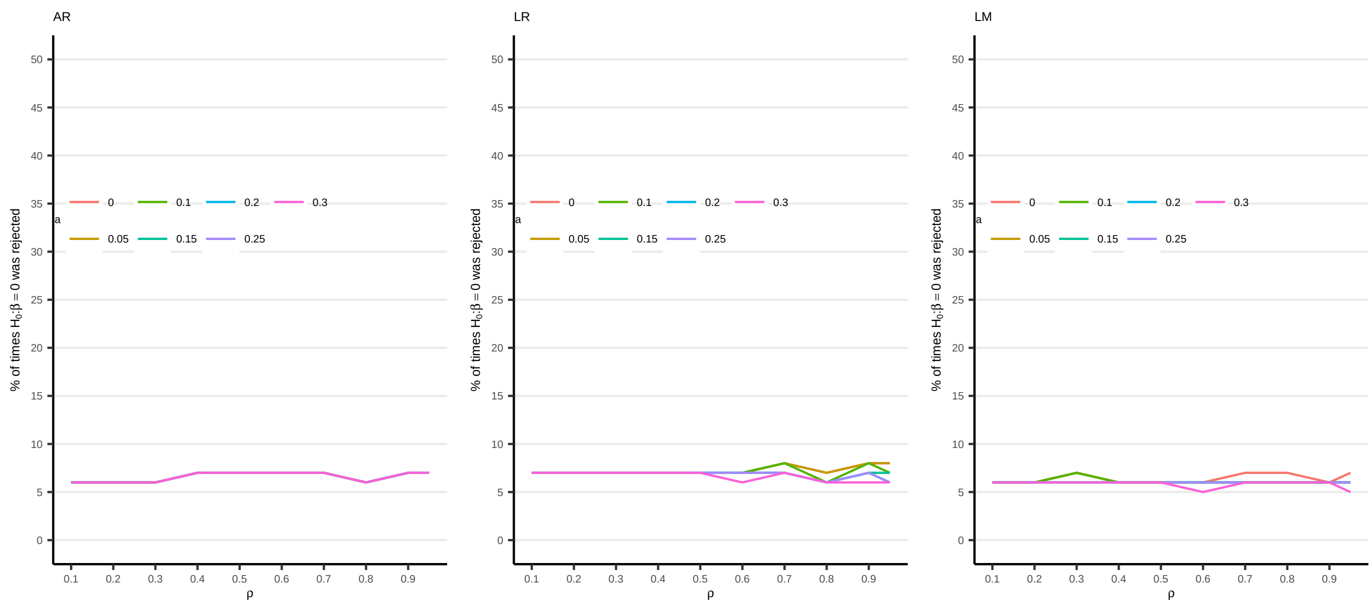
Question 4 - see appendix

Question 5

Items a-d

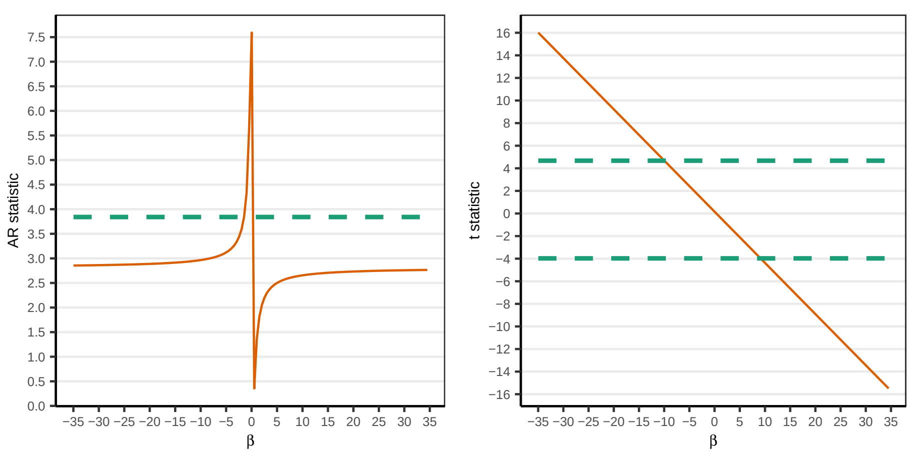
(Item a) See the notes in Figure (5).

Figure 4: Rejection frequency of the AR, LR and LM test statistics for different combinations of ρ and a



Notes: Rejection frequencies for $H_0 : \beta_0 = 0$ in the AR (left), LR (middle) and LM (right) statistics in 5000 simulations of the model described in (1)-(2) using $\beta_0 = 0$. Each line denotes a different value for a, ranging from 0 (top curve) to 0.3 (bottom curve). Frequencies in the y axis are computed as the number of relative occurrences, in 5000 replications, for each value of ρ and were interpolated by a simple linear approximation to facilitate the comparison between curves. The AR statistic does not depend on the value of a, hence the curves overlap.

Figure 5: 95% confidence set for the AR and the 2SLS t-statistic



Notes: The confidence set for the 2SLS t-statistic is represented in the right graph and corresponds to the interval $[-3.97, 4.67]$. The AR statistic is computed using the formula given in the lectures and it is compared with the 95-th quantile of a $\chi^2_{(1)}$ distribution (since only one instrument is used). We note in the left panel that the confidence set is unbounded and disjoint, i.e., $CI_{AR} = (-\infty, -1.48] \cup [0.02, +\infty)$, which suggests that our instrument is weak. Critical values are indicated with dashed green lines.

(Item b) Yes, the confidence sets from Figure (5) are quite different, due to the presence of a weak instrument. As seen in Question 1, the size of the test using the 2SLS t-statistic is not going to be the correct one, hence it is not advisable for inference in this model.

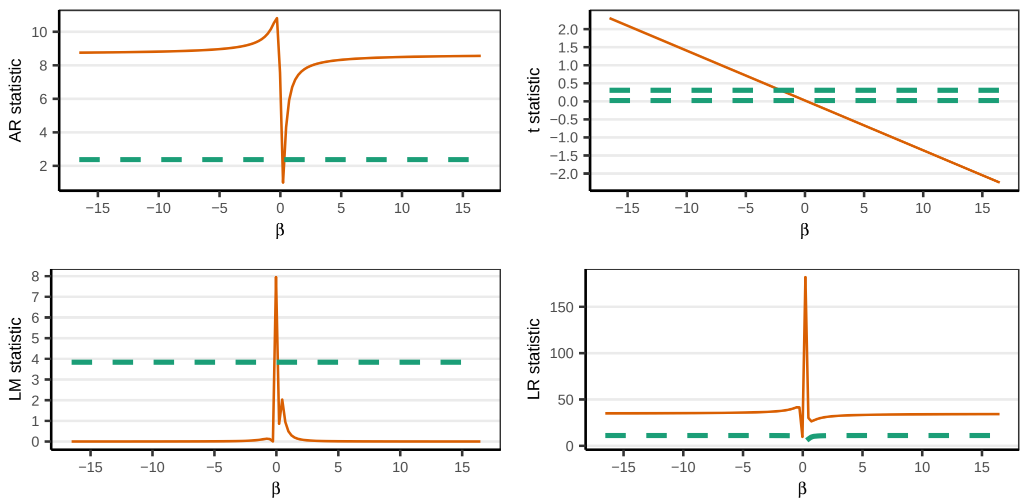
(Item c) The value of the F-Statistic in the first stage is given by 2.81, which is below the suggested threshold of 10 from Stock and Yogo (2005) to obtain the correct limiting distribution for the 2SLS statistic. As expected, the AR statistic converges to the value of the F-statistic as $|\beta|$ increases.

(Item d) For the case with only one instrument ($k = 1$), the AR, LM and LR statistics are the same and all three follow a chi-square distribution with one degree of freedom. To see that the result holds for the LR statistic, we just plug in the corresponding values in Equation (3) to simplify the expression.

Items e-f

(Item e) See the notes in Figure (6).

Figure 6: 95% confidence set for the AR, LM, LR and the 2SLS t-statistic using 4 instruments



Notes: The formats of the confidence sets for the AR and the 2SLS t-statistic in comparison to item a remain unchanged, but the scale slightly change. We still have the suggestion of weak instruments. The LM conficence region is obtained by using the quantiles of a qui-square with 1 degree of freedom, while the LR critical values are obtained iteratively for each value of β from the grid in the simulation. Critical values are indicated with dashed green lines.

(Item f) When we increase the number of instruments, it would be expected that the confidence set for the AR and the 2SLS t-statistic to be more similar (which leads me to be suspicious of my code for Figure 6, but I couldn't find the problem to correct it). In any case, since the problem of the 2SLS t-statistic is related to weak instruments, it would be expected that its performance improves when adding more instruments. We can confirm this by observing that we also obtain narrower intervals for the LM and the LR statistic.

Appendix

Question 2

Algorithm 1: Pseudo-algorithm to compute the 95% critical value for the LR statistic

Input: Number of simulations S, k, a grid $rgrid$ of size R, an empty $S \times R$ matrix vLR.
Output: $LRcv$, an array $R \times 1$ of critical values for the LR test.
begin
1. Compute S random values from a χ^2_1 and from χ^2_{k-1} and store in $\Psi_{(1)}$ and $\Psi_{(k-1)}$

for $r \in \{1, \ldots, R\}$ **do**
for $s \in \{1, \ldots, S\}$ **do**

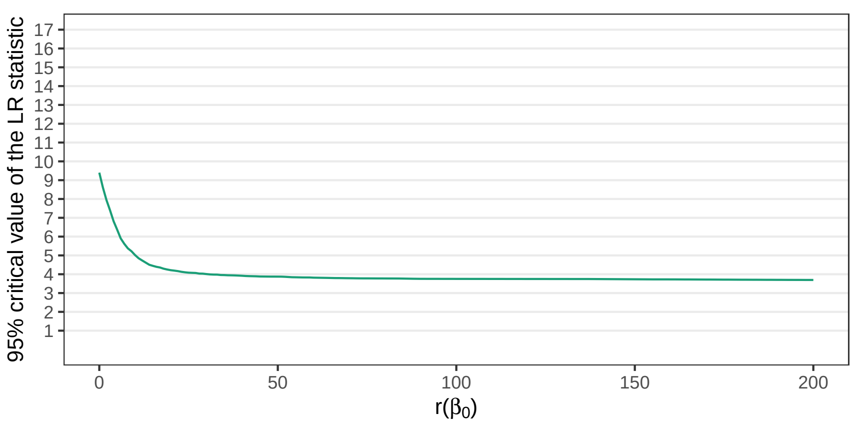
$$vLR(s,g) = 0.5 \cdot \left(\Psi_{(k-1)}(s) + \Psi_{(1)}(s) - \text{rgrid}(r) + \sqrt{\left(\Psi_{(k-1)}(s) + \Psi_{(1)}(s) + \text{rgrid}(r)\right)^2} \right.$$
$$\left. \sqrt{-4 \cdot \text{rgrid}(g) \cdot \Psi_{(k-1)}(s)} \right)$$

end
end
4. For each column of vLR, find the 95-th quantile.

end

Question 4

Figure 7: Critical values of the LR statistic for $k = 4$



Notes: Critical values for the LR statistic considering a grid $r(\beta_0) \in \{0, 1, \ldots, 200\}$ and $k = 4$. These critical values can be then compared to the actual statistic in a model and would keep the 95% confidence region. In comparison with Question 2, we see that the critical value when $r(\beta_0) \to 0$ is smaller, which is expected since now we have less instruments.