Question 1. Suppose that annual earnings and marjiuana usage are determined jointly by

$$log(earnings) = \beta_0 + \beta_1 marijuana + \beta_2 educ + u_1$$
$$marijuana = \gamma_0 + \gamma_1 log(earnings) + \gamma_2 educ + \gamma_3 fine + \gamma_4 prison + u_2$$

where fine is the typical $fine$ assessed for people possessing small amounts of marijuana and $prison$ is a dummy variable equal to one if a person can serve prison time for being in possession of marijuana for personal use. Assume $fine$ and $prison$ can vary with the country of residence.

(a) If $educ$, $fine$, and $prison$ are exogenous, what do you need to assume about the parameters in the system to consistently estimate the β_j.

There is a simultaneity bias, since the $marijuana$ and $educ$ are jointly determined with the dependent variable $log(earnings)$. If $educ$, $fine$, and $prison$ are exogenous, they are uncorrelated with u_1 and correlated with $marijuana$.

We need to assume about the $\gamma_3 = \gamma_4 = 0$ to consistently estimate the β_j

The hypothesis $\begin{cases} H_0 : \gamma_3 = \gamma_4 = 0 \\ H_1 : \text{at least one of } \gamma_i \neq 0 \end{cases}$

$$F = \frac{(RRSS - URSS)/2}{URSS/(n-5)} \quad \sim F_{(2,n-5)}$$

If we do not reject H_0 then we can't consistently estimate the β_j

(b) Explain in detail how you would estimate the β_j assuming the parameters are identified.
Using 2SLS
1st stage : $\widehat{marijuana} = \hat{\alpha_0} + \hat{\alpha_1} educ + \hat{\alpha_2} fine + \hat{\alpha_3} prison$
2nd stage : $log(\widehat{earning}) = \hat{\beta}_0 + \hat{\beta}_1 \widehat{marijuana} + \hat{\beta}_2 educ$

(c) Do you have overidentification?
The number of IV = 2 ($fine$, $prison$)
First equation is overidentified and the second equation is underidentified.

Question 2. Let us consider the demand for fish. Using 97 daily price ($avgprc$) and quantity ($totqty$) observations on fish prices at the Fulton Fish Market in Manhattan, the following results were obtained by OLS.

$$log(\widehat{totqty_t}) = \underset{(.163)}{8.244} - \underset{(.176)}{.425}\, log(avgprc_t) - \underset{(.226)}{.311}\, mon_t - \underset{(.223)}{.683}\, tues_t - \underset{(.220)}{.533}\, wed_t + \underset{(.220)}{.067}\, thurs_t$$

The equation allows demand to differ across the days of the week, and Friday is the excluded dummy variable. The standard errors are in parentheses.

(a) Interpret the coefficient of log($avgprc_t$) and discuss whether it is significant.

The coefficient of log($avgprc_t$) is the price elasticity of $totqty_t$. It means the effect of average fish prices on demand of fish. 1% increase in fish prices reduces fish demand by 0.425% on average.

$$\begin{cases} H_0 : \beta_1 = 0 \\ H_1 : \beta_1 < 0 \end{cases}$$

$$\text{t-stat}_{avgprc} = \frac{\hat{\beta}}{S.E} = \frac{0.425}{0.176} = 2.41$$

t-stat = 2.41 > 1.96 at 5% significance level. so reject H_0. It's significant.

(b) It is commonly thought that prices are jointly determined with quantity in equilibrium where demand equals supply. What are the consequences of this simultaneity for the properties of the OLS estimator? In your answer you may want to provide the system of equations that determine quantity and prices and demonstrate these consequences.

simultaneity → A3 violation → OLS estimator is biased and inconsistent.
The system of equations

Structural form : $\begin{cases} q_i = \alpha_1 p_i + \beta_1 z_i + u_{1i} & \text{: supply} \\ q_i = \alpha_2 p_i + u_{2i} & \text{: demand} \end{cases}$

Reduced form : $\alpha_1 p_i + \beta_1 z_i + u_{1i} = \alpha_2 p_i + u_{2i}$

Assume $\mathrm{Cov}(z_i, u_{1i}) = \mathrm{Cov}(z_i, u_{2i}) = 0$

$$\begin{aligned} p_i &= \frac{\beta_1}{\alpha_2 - \alpha_1} z_i + \frac{1}{\alpha_2 - \alpha_1}(u_{1i} - u_{2i}) \\ &= \phi_p z_i + v_{pi} \\ &\quad \text{with } \phi_p = \frac{\beta_1}{\alpha_2 - \alpha_1} \text{and} v_{pi} = \frac{1}{\alpha_2 - \alpha_1}(u_{1i} - u_{2i}) \end{aligned}$$

Denote $\mathrm{Cov}\,(u_{1i}, u_{2i}) = \sigma_{12}, \mathrm{Var}\,(u_{1i}) = \sigma_1^2$ and $\mathrm{Var}\,(u_{2i})\,\sigma_2^2$
Using the reduced form for p_i
- Supply: $\mathrm{Cov}\,(p_i, u_{1i}) \neq 0$

$$= \mathrm{Cov}\left(\pi_p z_i + \frac{1}{\alpha_2 - \alpha_1}\left(u_{1i} - u_{2i}\right), u_{1i}\right) = \frac{1}{\alpha_2 - \alpha_1}\left(\sigma_1^2 - \sigma_{12}\right) \neq 0$$

- Demand: $\mathrm{Cov}\,(p_i, u_{2i}) \neq 0$

$$= \mathrm{Cov}\left(\pi_p z_i + \frac{1}{\alpha_2 - \alpha_1}\left(u_{1i} - u_{2i}\right), u_{2i}\right) = \frac{1}{\alpha_2 - \alpha_1}\left(\sigma_{12}^2 - \sigma_2\right) \neq 0$$

(c) The variables $wave2_t$ and $wave3_t$ are measures of ocean wave heights over the past several days. In view of your answer in (b), what two assumptions do we need to make in order to use $wave2_t$ and $wave3_t$ as instruments for $\log(avgprc_t)$ in estimating the demand equation? Discuss whether these assumptions are reasonable.

If $wave2_t$ and $wave3_t$ are proper instrumental variables, they must satisfy validity(orthogonal with error) and relevance(correlated with $\log(avgprc_t)$) conditions.

(d) Below we report two sets of regression results, where the dependent variable is $\log(avgprc_t)$. Are $wave2_t$ and $wave3_t$ jointly significant? State the test statistic and rejection rule. How is your finding related to your answer in (c)?

	Regressor						
Dependent Variable $\log(avgprc_t)$	constant	wave2	wave3	day-of-the-week dummies	R^2	RSS	n
Regression (2.2)	-1.022 (.144)	-1.022 (.144)	-1.022 (.144)	yes	.3041	10.934	97
Regression (2.3)	$-.276$ (.092)	-	-	yes	.0088	15.576	97

By F-test,

The hypothesis $\begin{cases} H_0 : \text{The coefficient of } wave2_t \text{ and } wave3_t = 0 \\ H_1 : \text{not } H_0 \end{cases}$

$$F = \frac{(RRSS - URSS)/2}{URSS/(n-7)} = \frac{(15.576 - 10.934)/2}{10.934/(97-7)} = 18.37 \quad \sim F_{(2,90)}$$

Thus reject H_0 at 5% significance level. It means $wave2_t$ and $wave3_t$ are jointly significant.
And $wave2_t$ and $wave3_t$ have relevance, but they don't have high relevance, since the $R^2 = 0.3041$

(e) the following IV results were obtained in regression:

$$\widehat{log(totqty_t)} = \underset{(.182)}{8.164} - \underset{(.327)}{.815}\, log(avgprc_t) - \underset{(.229)}{.307}\, mon_t - \underset{(.226)}{.685}\, tues_t - \underset{(.223)}{.521}\, wed_t + \underset{(.225)}{.095}\, thurs_t$$

Discuss how these results can be obtained using Two Stage Least Squares (2SLS).
1st stage : $\widehat{log(avgprc_t)} = \hat{\alpha}_0 + \hat{\alpha}_1 wave2_t + \hat{\alpha}_2 wave3_t + \hat{\alpha}_3 mon_t + \hat{\alpha}_4 tues_t + \hat{\alpha}_5 wed_t + \hat{\alpha}_6 thurs_t$
2nd stage : $\widehat{log(totqty_t)} = \hat{\beta}_0 + \hat{\beta}_1 \widehat{log(avgprc_t)} + \hat{\beta}_2 mon_t + \hat{\beta}_3 tues_t + \hat{\beta}_4 wed_t + \hat{\beta}_5 thurs_t$
The 2SLS estimator will be biased, but consistent.