$E(\varepsilon\varepsilon'|X) = \sigma^2 I$ e 1t 6t

1. (30 points) Consider a linear regression model $Y = X\beta + \epsilon$, where Y is a $n \times 1$ vector, X is a $n \times k$ matrix with the first column equals to one, β is $k \times 1$ vector, and ϵ is a $n \times 1$ vector with mean 0 and variance $\sigma^2 I$. The following sub-questions are based on this model setting.

(a) (5 mins) The OLS method provides the estimator of β by minimizing the sum of residual squares. Set up this minimization problem and derive the OLS estimator $\hat{\beta}_{ols}$. $= (X'X)^{-1}X'y$

$X'(y - X\hat{\beta}) = 0 = X'y - X'X\hat{\beta} = X'(y - X\hat{\beta})$

$E(e)$

(b) (5 mins) Using the least square normal equation to show for each regressor X_k of X, we have $X_k'e = 0$, where $e = Y - X\hat{\beta}_{ols}$ is the vector of residuals. What is the mean of e? Why $X_k'e = 0$ implies X_k and e are uncorrelated?

(c) (5 mins) Show $\hat{\beta}_{ols}$ is an unbiased estimator, i.e., $\mathrm{E}(\hat{\beta}_{ols}) = \beta$? List necessary assumptions and show how to use them in proving the unbiasedness of $\hat{\beta}_{ols}$.

(d) (5 mins) Show $\mathrm{var}(\hat{\beta}_{ols}) = \sigma^2 \mathrm{E}((X'X)^{-1})$.

(e) (5 mins) Define the *ridge regression* estimator

$$\hat{\beta}_{ridge} = (X'X + I_k\lambda)^{-1}X'y$$

where $\lambda > 0$ is a fixed coefficient. Find $\mathrm{E}(\hat{\beta}_{ridge}|X)$. Is $\hat{\beta}_{ridge}$ unbiased?

CLT

(f) (2 mins) What are the justifications of using the normal distribution as the sampling distribution of $\hat{\beta}_{ols}$ in the small sample case and in the large sample case?

normality assumption

2. (25 points) When studying cross-sectional data, such as individual, firm, city, or country level data in a regression model, the problem of heteroskedasiticity should be addressed.

(a) (2 mins) In a regression model, $Y = X\beta + \epsilon$, ϵ has mean 0 and variance Ω ($\sigma^2\Omega$), where Ω is a $n \times n$ square matrix. Provide a general specification of Ω for the heteroskedasticity problem.

white?

DW ?

(b) (5 mins) One may inspect the plot of residuals to detect the heteroskedasticity problem. Other than that, what formal tests can be used? (Note: you should name one test and provide details about this test)

(c) If the heteroskedasiticty problem is confirmed in the regression, answer the following true or false questions. No explanation is needed.

F (i) (1 min) The OLS estimator is biased. True or False?

F (ii) (1 min) The OLS estimator is efficient. True or False?

1\. a. $\min \sum e'e \iff \min_{\hat\beta} \sum (Y-X\hat\beta)'(Y-X\hat\beta) \iff \min_{\hat\beta} Y'Y - Y'X\hat\beta - \hat\beta'X'Y + \hat\beta'X'X\hat\beta$

FOC: $(X'X)\hat\beta - X'Y = 0$

$\Rightarrow \hat\beta = (X'X)^{-1}X'Y$

b. $X'(X\hat\beta - Y) = 0 \Rightarrow X'e = 0 \Rightarrow X_k'e = 0$

$E(e|X) = E(Y - X\hat\beta|X) = E(Y|X) - X E(\hat\beta|X)$

$= X\beta - X\beta = 0$

$\Rightarrow E(e) = 0$

$Cov(X_k, e) = E(X_k'e) - E(X_k)E(e)$

$= 0 - 0 = 0$

c. $E(\varepsilon|X) = 0$

?

d. $var(\hat\beta) = E((X'X)^{-1}X'\varepsilon\varepsilon'X(X'X)^{-1})$

$= E[(X'X)^{-1}X' \underbrace{E(\varepsilon\varepsilon'|X)}_{\sigma^2} X(X'X)^{-1}]$

$= \sigma^2 E((X'X)^{-1})$

7 e. $\hat\beta = (X'X + I_k\lambda)^{-1}X'y$

f. in small: Assume $\varepsilon \sim N(0, \sigma^2 I)$

in large: CLT

T F (iii) (1 min) When using HAC (Heteroskedasticity and Autocorrelation Consistent) estimation, the OLS estimates of regression slopes will not change. True or False? $\hat{\beta} = (X'X)^{-1}X'y$

T (iv) (1 min) Asymptotically, the OLS estimator is still consistent and has a limiting normal distribution. True or False?

T (v) (1 min) The Generalized Least Square (GLS) estimator is more efficient than OLS estimator. True or False?

$(X'X)^{-1}X'y$

(d) (5 mins) When Ω is known, we can implement GLS estimation by the weighted least square (WLS) method. Given Ω that you specify in part (a), discuss how to implement the WLS method. $W_i = \frac{1}{\sqrt{w_i}}$ $W = \Omega^{-1}$

(e) (5 mins) Consider the OLS and GLS estimators $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ and $\tilde{\beta} = (\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Omega}^{-1}\mathbf{y}$. Compute the (conditional) covariance for $\hat{\beta} - \tilde{\beta}$:

$$E\left(\left(\hat{\beta} - \tilde{\beta}\right)\left(\hat{\beta} - \tilde{\beta}\right)' \Big| \mathbf{X}\right).$$

3. (20 points) Suppose the random variable y is generated from the density of **log-normal** distribution

$$f(y_i|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}y_i} \exp\left[-\frac{(\ln(y_i) - \mu)^2}{2\sigma^2}\right], \quad y_i > 0$$

(a) (5 mins) Given a total of n observations, i.e., $y_1, \cdots, y_n$, write down the log-likelihood function for the unknown parameters μ and σ^2.

(b) (5 mins) Derive the ML estimator for μ and σ^2. outer product estimator

(c) (5 mins) Derive the BHHH (Berndt–Hall–Hall–Hausman) variance estimator for the ML estimator $\hat{\mu}_{ml}$.

Var

(d) (2 mins) Is the ML estimator an efficient estimator for this model? Explain.

4. (25 points) The following questions are based on the Simultaneous Equations Model

Supply $$hours = \beta_{10} + \alpha_1 \log(wage) + \beta_{11}edu + \beta_{12}age + \beta_{13}kidslt6 + \beta_{14}otherincome + u_1$$

Demand $$\log(wage) = \beta_{20} + \alpha_2 hours + \beta_{21}edu + \beta_{22}exper + \beta_{23}exper^2 + u_2$$

The first and second equations aim to capture labor supply and labor demand of married female workers, respectively. The variable *hour* stands for working hours; $\log(wage)$ stands for the log hourly wage rate; *kidslt6* stands for a dummy of having kids under age 6.

3. a.

$$\ln: \frac{1}{\sqrt{2\pi\sigma^2}y_i}\exp\left[-\frac{(\ln(y_i)-\mu)^2}{2\sigma^2}\right], \quad y_i > 0$$

$$= -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\sigma^2 - \ln(y_i) - \frac{[\ln(y_i)-\mu]^2}{2\sigma^2}$$

$$l(\mu,\sigma^2) = -\frac{1}{2}n\ln 2\pi - \frac{1}{2}n\ln\sigma^2 - \sum\ln(y_i)$$

$$-\frac{1}{2\sigma^2}\cdot \cancel{[}\underline{\sum(\ln(y_i)-\mu)^2}\cancel{]^2} \quad (\ln y-\mu)'(\ln y-\mu)$$

b. FOC to μ: $-\frac{1}{2\cancel{\sigma^2}}\cdot 2[\sum\ln(y_i)-n\mu]\cdot(\cancel{-1}) = 0 \Rightarrow \hat{\mu} = \frac{\sum\ln y_i}{n}$

可消掉 $\quad -1$

to σ^2: $-\cancel{\frac{1}{2}}n\cdot\frac{1}{\sigma^2} + \cancel{[\sum\ln(y_i)-n\mu]^2}\cdot\cancel{\frac{1}{2}}\cdot\frac{1}{\sigma^4} = 0$

$\sigma^2 \quad (\ln y-\mu)'(\ln y-\mu)$

$$\Rightarrow \cancel{\hat{\sigma}^2 = \frac{\sum\ln(y_i)-n\mu^2}{n}}$$

$$\hat{\sigma}^2 = \frac{(\ln y-\mu)'(\ln y-\mu)}{n}$$

$$l^c(\mu) = -\frac{n}{2}\ln 2\pi - \frac{n}{2}\ln\left(\frac{1}{n}(\ln y-\mu)'(\ln y-\mu)\right) - \frac{1}{2}$$

$$\text{FOC to } \mu = -\frac{n}{2}\cdot\frac{}{\frac{1}{n}(\ln y-\mu)'(\ln y-\mu)}$$

2 a.

$$\Omega = \begin{bmatrix} \sigma^2 w_1 & & & 0 \\ & \sigma^2 w_2 & & \\ & & \ddots & \\ 0 & & & \sigma^2 w_n \end{bmatrix}$$

b. White test: e^2 對 x_i, x_i^2, $x_i x_j$ 跑回歸

c. F

F

T

T

T

d. Let $P' = C\Lambda^{-\frac{1}{2}}$, then $\Omega^{-1} = P'P$

$$Py = PX\beta + P\varepsilon \Rightarrow y^* = X^*\beta + \varepsilon^*$$

$$\Rightarrow \hat{\beta} = (X^{*\prime}X^*)^{-1}X^{*\prime}y^* = (X'\Omega^{-1}X)^{-1}(X'\Omega^{-1}y)$$

$$\Omega^{-1} = \begin{bmatrix} \frac{1}{w_1} & & & \\ & \frac{1}{w_2} & & \\ & & \ddots & \\ & & & \frac{1}{w_n} \end{bmatrix}$$

So, $\hat{\beta} = \left[\sum \frac{1}{w_i} x_i x_i'\right]^{-1} \left[\sum \frac{1}{w_i} x_i y_i\right]$.

e. $E\left(\left[(X'X)^{-1}X'\varepsilon - (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\varepsilon\right]\left[(X'X)^{-1}X'\varepsilon - (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}\varepsilon\right]'\right)$

$=$?

Yes, 就是硬展開

回頭看出內生?

(a) (2 mins) Explain why log($wage$) is an endogenous variable in the labor supply equation.

Yes, because we have 2 exogenous variables excluded in the first equation

(b) (2 mins) Discuss if the first equation satisfies the order condition or not.

(c) (5 mins) Discuss the required coefficient constraints for the first equation to satisfy the rank condition for identification.

(d) (2 mins) Estimation result of the first equation by OLS is reported in Figure 1 Panel (a). Comment on the estimated coefficient for log($wage$).

over-justifi

(e) (2 mins) Estimation result of the first equation by 2SLS using variables experience and experience square as instruments is reported in Figure 1 Panel (b). Compare the estimated coefficient for log($wage$) with the OLS result.

(f) (2 mins) Some information of the first stage in the 2SLS estimation is reported in Figure 1 Panel (c). Comment on whether we have a weak instrument problem or not.

Yes, $F = 9.33 < 10 \Rightarrow$ weak instrument problem.

```
> regress hours lnwage  educ age kidslt6 otherincome;
```

Source	SS	df	MS
Model	9290528.56	5	1858105.71
Residual	248020491	422	587726.283
Total	257311020	427	602601.92

Number of obs	=	428
F(5, 422)	=	3.16
Prob > F	=	0.0082
R-squared	=	0.0361
Adj R-squared	=	0.0247
Root MSE	=	766.63

wage ↑ 1%
hour ↓ 2 unit

hours	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
lnwage	-2.046796	54.88015	-0.04	0.970	-109.9193	105.8257
educ	-6.62187	18.11627	-0.37	0.715	-42.23123	28.98749
age	.5622541	5.140012	0.11	0.913	-9.540961	10.66547
kidslt6	-328.8584	101.4573	-3.24	0.001	-528.2831	-129.4338
otherincome	-5.918459	3.683341	-1.61	0.109	-13.15844	1.321522
_cons	1523.775	305.5755	4.99	0.000	923.1352	2124.414

(a)

```
. ivregress 2sls hours educ age kidslt6 otherincome (lnwage=exper expersq);
```

Instrumental variables (2SLS) regression

Number of obs	=	428
Wald chi2(5)	=	17.45
Prob > chi2	=	0.0037
R-squared	=	.
Root MSE	=	1344.7

$hours = (\widehat{lnwage}) + \beta Z_1 \ldots + edu \ldots$

hours	Coef.	Std. Err.	z	P>|z|	[95% Conf. Interval]	
lnwage	1639.556	467.2656	3.51	0.000	723.7318	2555.379
educ	-183.7513	58.68409	-3.13	0.002	-298.77	-68.73257
age	-7.806092	9.312048	-0.84	0.402	-26.05737	10.44519
kidslt6	-198.1543	181.6424	-1.09	0.275	-554.1669	157.8583
otherincome	-10.16959	6.568215	-1.55	0.122	-23.04306	2.703873
_cons	2225.662	570.5226	3.90	0.000	1107.458	3343.866

Instrumented: lnwage
Instruments: educ age kidslt6 otherincome exper expersq

(b)

```
estat firststage
```

First-stage regression summary statistics

(lnwage)
$Y = \pi_1 Z_1 + \pi_2 Z_2 + \ldots$

若 $\pi_1, \pi_2 \neq 0$

H_0: all 0.

H_1: not all 0

Variable	R-sq.	Adjusted R-sq.	Partial R-sq.	F(2,421)	Prob > F
lnwage	0.1633	0.1514	0.0424	9.32933	0.0001

Minimum eigenvalue statistic = 9.32933

Critical Values # of endogenous regressors: 1
Ho: Instruments are weak # of excluded instruments: 2

	5%	10%	20%	30%
2SLS relative bias	(not available)			

	10%	15%	20%	25%
2SLS Size of nominal 5% Wald test	19.93	11.59	8.75	7.25
LIML Size of nominal 5% Wald test	8.68	5.33	4.42	3.92

(c)

Figure 1