Problem Set 3 | **AREC/ECON 535** | *Fall 2023*
Key | **Applied Econometrics** | *S.R. Koontz*

Part A: Serial Correlation

1. OLS results for the beef demand equation.

$$y_t = 6.4612 - 0.2712\ x_{1t} + 0.9957\ x_{2t} - 0.4063\ x_{3t} - 0.4237\ x_{4t} + e_t$$

se	(1.4199)	(0.1261)	(0.1621)	(0.1446)	(0.1124)
p	(0.0001)	(0.0357)	(0.0001)	(0.0067)	(0.0004)

$R^2 = 0.6859$ $F_{4,58} = 31.66$ P-Value = 0.0001 $\sigma = 0.0834$

2. DW-d = 0.683   where N=63 and k'=4

$d_l = 1.444$ and $d_u = 1.727$ (N=60, k'=4, and α=5%)
$d_l = 1.471$ and $d_u = 1.731$ (N=65, k'=4, and α=5%)

Reject the null hypothesis of no first order serial correlation. There is evidence of significant positive serial correlation.

3. $BG = (63 - 2) \times 0.5214 = 31.8054 > 5.991 = \chi^2$ (df=2 α=5%)

and $F_{2,54} = 26.18$ (p-value = 0.0001).

Significant first order but not second order autocorrelation.

4. EGLS Results.

$$y_t = 4.2564 - 0.4027\ x_{1t} + 0.0946\ x_{2t} - 0.0318\ x_{3t} + 0.2010\ x_{4t} + e_t$$

(1.2074)	(0.0740)	(0.0565)	(0.0557)	(0.1153)
(0.0008)	(0.0001)	(0.0994)	(0.5707)	(0.0865)

$$e_t = 0.9856\ e_{t-1} + u_t$$

(0.0224)
(0.0001)

$F_{4,57} = 9.01$ P-Value = 0.0001 Total $R^2 = 0.9693$ $\sigma = 0.0263$

Regression $R^2 = 0.3873$

5. Coefficients are supposed to not change much and standard errors are supposed to be larger (relative to the beta) – most p-values less significant. F-statistic smaller, R^2 larger, and root error variance smaller. Economic conclusions more reasonable?

6. $y_t = \beta_0 + \Sigma \beta_i x_{it} + e_t$ and $e_t = \rho e_{t-1} + u_t$

$y_t = \beta_0 + \Sigma \beta_i x_{it} + \rho e_{t-1} + u_t$

$E(y_t) = \beta_0 + \Sigma \beta_i x_{it} + \rho e_{t-1}$

$E(y_{2010}) = \beta_0 + \Sigma \beta_i x_{i\ 2010} + \rho e_{2009}$

$E(y_{2010}) = 4.2564 - 0.4027 \ln(432.7) + 0.0946 \ln(307.3) - 0.0318 \ln(124.7)$

$+ 0.2010 \ln(36180) + 0.9856 (-0.026102) = 4.31038$ ➔ PCQ = 61.4588 #/cap

7. OLS Structural Change. $F_{5,53} = 24.73$ p-value = 0.0001 Reject.

EGLS Structural Change. $F_{5,52} = 3.58$ p-value = 0.0074 Reject.

(If using P-W the df2 is 53.) Correcting the serial correlation makes the F-test for structural change smaller. But structural change is still present.

8. OLS

$R^2 = 0.9209$ $F_{7,55} = 91.47$ P-Value = 0.0001 DW-d = 0.629

H_0: $\beta_5 = \beta_6 = \beta_7 = 0$. $F_{3,55} = 54.47$ P-Value = 0.0001 Reject.

EGLS

$y_t = 2.4685 - 0.4307 x_{1t} + 0.0770 x_{2t} - 0.0565 x_{3t} + 0.4433 x_{4t}$

(1.2288) (0.0603) (0.0458) (0.0472) (0.1328)
(0.0496) (0.0001) (0.0986) (0.2364) (0.0015)

$+ 0.0249 t - 0.001527 t^2 + 0.0000158 t^3 + e_t$

(0.0085) (0.000275) (0.000003)
(0.0050) (0.0001) (0.0001)

$e_t = 0.9321 e_{t-1} + u_t$

(0.0493)
(0.0001)

Total $R^2 = 0.9822$ $F_{7,54} = 15.16$ P-Value = 0.0001 $\sigma = 0.0206$

Regression $R^2 = 0.6628$

H_0: $\beta_5 = \beta_6 = \beta_7 = 0$. $F_{3,54} = 19.48$ P-Value = 0.0001 Reject.

Trend variables are useful but vague.

9. Elasticity Table

	OLS	EGLS	OLS w/ t	EGLS w/ t
Beef	-0.2712 (0.1261)	-0.4027 (0.0740)	-0.1706 (0.0821)	-0.4307 (0.0603)
Pork	+0.9957 (0.1621)	+0.0946 (0.0565)	+0.2458 (0.1123)	+0.0770 (0.0458)
Chicken	-0.4063 (0.1446)	-0.0318 (0.0557)	-0.1099 (0.1125)	-0.0565 (0.0472)
Income	-0.4237 (0.1124)	+0.2010 (0.1153)	+0.9562 (0.1878)	+0.4433 (0.1328)
Rho		0.9856 (0.0224)		0.9321 (0.0511)

Which is the better model?

Trend and serial correlation are measuring something similar. Without the trend variables, the model appears to be misspecified.

Elasticities more reasonable in trend model. But what is the trend variable measuring?

(When trends and dummies are used together, and serial correlation is corrected for, then it is the trend that wins. The own-price elasticity become more inelastic.)

Part B: Heteroskedasticity

0. Summary statistics and correlation coefficients.

Name	N	Mean	Std Dev	Min	Max
exp	72	262.53	318.05	9.58	1898.03
age	72	31.437	7.1527	20.00	55.00
rnt	72	0.3750	0.4875	0.0	1.0
inc	72	3.4371	1.6995	1.5	55.0
d	72	0.5139	0.5033	0.0	1.0

1. OLS Results.

$$exp_i = -237.1 - 3.0818\ age_i + 27.94\ rnt_i + 234.3\ inc_i - 14.9968\ inc^2_i + e_i$$

(199.3)	(5.5147)	(82.92)	(80.37)	(7.4693)
(0.2384)	(0.5781)	(0.7372)	(0.0048)	(0.0487)

$F_{4,67} = 5.39$ P-Value = 0.0008 $R^2 = 0.2436$ $\sigma = 284.7508$

2. Error variance does not appear to be constant. The error variance is higher with higher income and maybe with age.

3. Park Test

$$\ln(e_t^2) = 2.9 + 0.0617\ age_i - 0.2455\ rnt_i + 2.2417\ inc_i - 0.2272\ inc^2_i + e_i$$

(1.3612)	(0.0377)	(0.5662)	(0.5187)	(0.0510)
(0.0381)	(0.1061)	(0.6660)	(0.0001)	(0.0001)

$F_{4,67} = 6.77$ P-Value = 0.0001 $R^2 = 0.2877$ $\sigma = 1.9443$

Park: F-test = 6.77 > $F_{4,67}(5\%)$. P-value = 0.0001. Reject.

Park for specific forms:
Income and squared: F-test = 12.07 (0.0001) Reject.
Income only: t-test = 0.21 (0.8354) Fail to reject.

Goldfeld-Quandt: (RSS and no dropped center obs so df1=31 df2=31)
F-test = 4894130 / 326247 = 15.00 > 1.90 = $F_{31,31}(5\%)$.
(dummy variable) t-test = 2.33 p-value = 0.0226.

BP: (98.8539)/2 = 49.4270 > 9.48773 = $\chi^2_4(5\%)$.

White: (72) × 0.1540 = 11.088 < 16.9190 = $\chi^2_9(5\%)$.
White: (72) × 0.1990 = 14.328 < 21.0261 = $\chi^2_{12}(5\%)$.

So heteroskedasticity is there but... and it’s related to...

4. White Results.

$$exp_i = -237.1 - 3.0818\ age_i + 27.94\ rnt_i + 234.3\ inc_i - 14.9968\ inc^2_i + e_i$$

(213.0)	(3.3017)	(92.19)	(88.87)	(6.9446)
(0.2655)	(0.3506)	(0.7618)	(0.0084)	(0.0308)

5. Weighted Least Squares - divide each variable by 1/inc.

$$exp_i/inc_i = -114.1\ (1/inc_i) - 2.6942\ age_i/inc_i + 60.45\ rnt_i/inc_i$$

(139.5)	(3.8073)	(58.55)
(0.4169)	(0.4816)	(0.3056)

$$+ 158.43 - 7.2493\ inc_i + e_i$$

(76.39)	(9.7243)
(0.0419)	(0.4586)

$F_{4,67} = 1.06$ P-Value = 0.3846 $R^2 = 0.0594$ $\sigma = 70.0984$

or report as

$$exp_i = -114.1 - 2.6942\ age_i + 60.45\ rnt_i + 158.43\ inc_i - 7.2493\ inc^2_i + e_i$$

(139.5)	(3.8073)	(58.55)	(76.39)	(9.7243)
(0.4169)	(0.4816)	(0.3056)	(0.0419)	(0.4586)

$F_{4,67} = 1.06$ P-Value = 0.3846 $R^2 = 0.0594$ $\sigma = 70.0984$

Coefficients change. Standard errors larger and p-values less significant. F-statistic smaller, R^2 smaller, and root error variance different.

6. Maximum Likelihood

$$exp_i = 50.2234 - 3.5262\ age_i + 94.52\ rnt_i + 56.77\ inc_i + 4.8148\ inc^2_i + e_i$$

(283.6) (2.7728) (36.64) (165.1) (4.8148)
(0.8594) (0.2035) (0.0099) (0.7310) (0.8596)

$$\sigma^2_i = \exp(\ 64.4122 + 0.3688\ inc_i + 1.8857\ D_i\ )$$

(20.05) (0.2527) (0.6045)
(0.0013) (0.1445) (0.0018)

$\chi^2_2 = 36.1953$ P-Value = 0.0001 $R^2 = 0.1269$

(Not with income but a jump...)

Two-Step Process

$$exp_i = -196.0 - 3.9134\ age_i + 56.23\ rnt_i + 218.0\ inc_i - 13.6682\ inc^2_i + e_i$$

(170.7) (4.7162) (74.66) (69.52) (6.0831)
(0.2549) (0.4096) (0.4540) (0.0025) (0.0279)

$F_{4,67} = 7.40$ P-Value = 0.0001 $R^2 = 0.3063$

$$\ln(\sigma^2_i) = 9.3234 - 0.3725\ inc_i + 2.0383\ D_i$$

(0.5822) (0.2011) (0.6791)
(0.0001) (0.0683) (0.0037)

$F_{2,69} = 4.53$ P-Value = 0.0142 $R^2 = 0.116$ $\sigma = 2.1343$

Error variance jump versus gradual change?

(A jump and maybe declining with income...)

7. Changes in coefficients and changes in significance between OLS and models which correct for heteroskedasticity. Changes again between weighted least squares and maximum likelihood models. Fits are very similar.