PRINCIPLES OF ECONOMETRICS

5TH EDITION

ANSWERS TO ODD-NUMBERED EXERCISES IN CHAPTER 4

EXERCISE 4.1

(a) $R^2 = 0.843$

(b) $\hat{\sigma}^2 = 8.4249$

(c) $R^2 = 0.711$

EXERCISE 4.3

(a) $\hat{y}_0 = 4.4$

(b) $\text{se}(f) = 1.5875$

(c) $(-0.6520, 9.4520)$

(d) $(-4.8722, 13.6722)$

(e) $(-1.8189, 5.8189\ )$

EXERCISE 4.5

(a) $R^2 = 0.204$

(b) $r_{y\hat{y}} = r_{yx} = 0.4517$.

(c) $h_5 = 0.01131$, $h_{16} = 0.001351$, $h_{21} = 0.0063175$

(d) $\text{DFBETAS}_{2i} = -0.05237256$.

(e) $\text{DFFITS}_i = -0.06474186$

(f) $\hat{e}_i^{stu} = -0.71691235$

EXERCISE 4.7

(a) $R^2 = 0.\ 113$

(b) $\widehat{ENTERT} = 60.12$

(c) $[-52.95,\quad 173.19]$ or $[0,\quad 173.19]$

(d) $[-158.85,\quad 519.57]$ or $[0,\quad 519.57]$

EXERCISE 4.9

(a) The Jarque-Bera = 30.405483. The test statistic value is larger than the critical value and we reject the null hypothesis.

(b) In this case *JB* = 1.9153333. Thus we fail to reject the null.

(c) In this case *JB* = 0.88941667. Thus, we fail to reject the null hypothesis.

(d) The log-log model fits the data the best.

(e) (i) The magnitude of the correlation between y and x is the same as the correlation between y and a + bx.

(ii) The R^2 = 0.31315216.

(iii) *YHAT* and *RPRICE3* have an exact linear relationship.

(iv) 0.43046721

(v) 0.45616516

(f) The log-log model fits the data best.

EXERCISE 4.11

(a) [5.7059548, 42.134045].

(b) [0.61731625, 56.582684].

EXERCISE 4.13

(a) $\tilde{\alpha} = \dfrac{\sum \tilde{x}_i \tilde{y}_i}{\sum \tilde{x}_i^2}$. Substituting we have $\tilde{\alpha} = \dfrac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} = b_2$

(b) $\tilde{e}_i = \tilde{y}_i - \tilde{\alpha}\tilde{x}_i = (y_i - \bar{y}) - b_2(x_i - \bar{x}) = y_i - (\bar{y} - b_2\bar{x}) - b_2 x_i = \hat{e}_i$

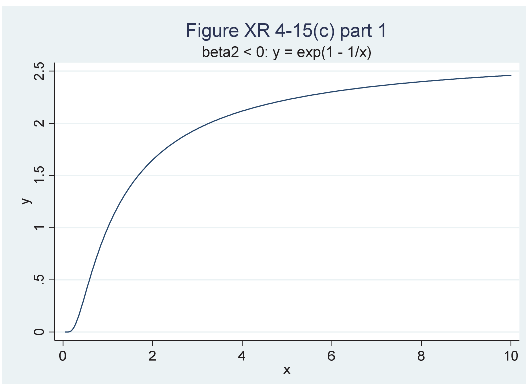
EXERCISE 4.15

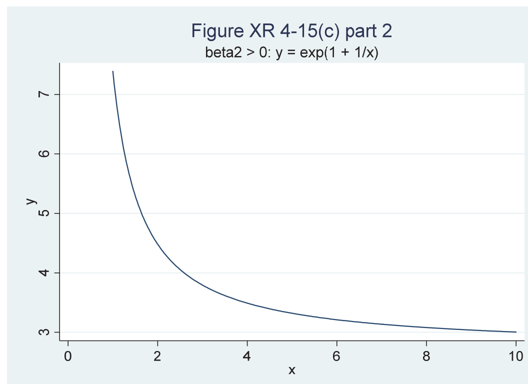
(a) For all values of x the dependent variable will be positive. An $x = 0$ will create an undefined value.

(b) $dy/dx = \exp\left[\beta_1 + \beta_2 (1/x)\right] \times \left(-\beta_2 x^{-2}\right)$

Assuming that $x > 0$ the slope will be positive if $\beta_2 < 0$.

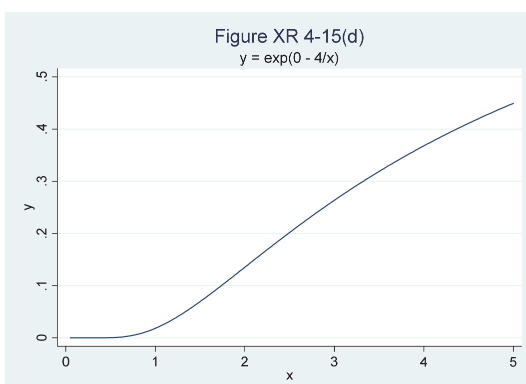
(c) When $\beta_2 < 0$, as x approaches zero from above, we see that $\beta_1 - \beta_2 (1/x) \to -\infty$ and $y = \exp\left[\beta_1 - \beta_2 (1/x)\right] \to 0$. If x approaches infinity $\beta_1 - \beta_2 (1/x) \to \beta_1$ and $y = \exp\left[\beta_1 - \beta_2 (1/x)\right] \to \exp(\beta_1)$. In Figure XR 4-15(c) part 1 we chose $\beta_1 = 1$, so $y \to e \cong 2.7182818$.





See Figure XR 4-15(c) part 2. If $\beta_2 > 0$, as x approaches zero from above, we see that $\beta_1 + \beta_2 (1/x) \to \infty$ and $y = \exp\left[\beta_1 + \beta_2 (1/x)\right] \to \infty$. If x approaches infinity $\beta_1 + \beta_2 (1/x) \to \beta_1$ and $y = \exp\left[\beta_1 + \beta_2 (1/x)\right] \to \exp(\beta_1)$. In Figure XR 4-15(c) part 2 we chose $\beta_1 = 1$, so $y \to e \cong 2.7182818$.

(d) See Figure XR 4-15(d).



The slopes at the x-values 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0 are

x	dy/dx
.5	.0053674
1	.07326256
1.5	.12352614
2	.13533528
2.5	.12921377
3	.11715428
3.5	.10413275
4	.09196986

We see that the slope increases but then begins to decrease. So, the answer is both.

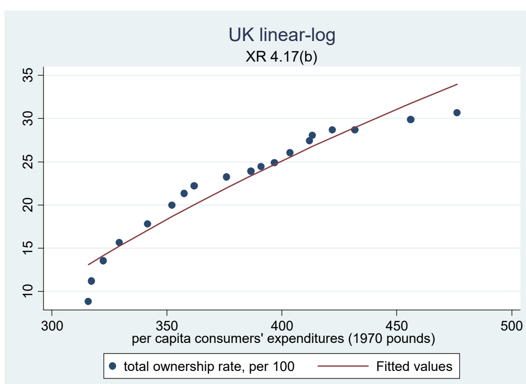
(e) The first derivative is $dy/dx = \exp\left[\beta_1 + \beta_2(1/x)\right] \times \left(-\beta_2 x^{-2}\right)$. The second derivative, uses Derivative Rule 6. To solve for the point where the second derivative equals zero we set the second term to zero and solve for x:

$$x = \frac{-\beta_2}{2}$$

EXERCISE 4.17

(a) A possible match is the linear-log model, shown in Figure 4.5(f).

(b)
$$\widehat{RATE\_UK} = -278.834 + 50.729\ln(SPEND\_UK) \quad R^2 = 0.9276$$
$$\text{(se)} \qquad (20.4216) \qquad (3.4378)$$

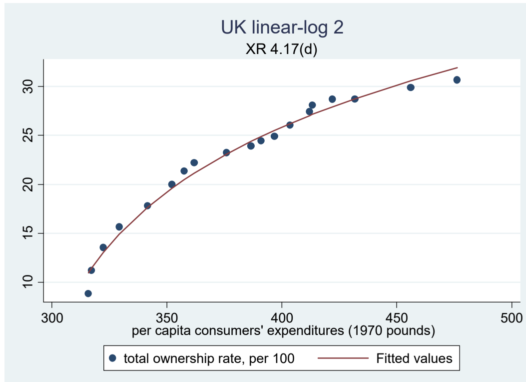


In this case the log-linear model fits the data well, except at the very low and very high ends.

(c) If $SPEND\_UK = 1$, then $\ln(SPEND\_UK) = 0$, then $E\left(RATE\_UK \mid SPEND\_UK = 1\right) = \beta_1$

(d)

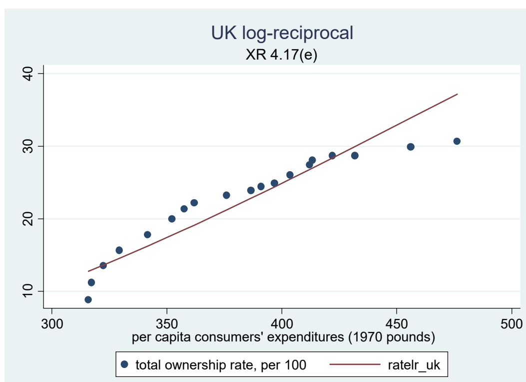
$$\widehat{RATE\_UK} = -33.0221 + 12.296\ \ln(SPEND\_UK - 280) \quad R^2 = 0.9840$$
$$\text{(se)} \qquad (1.7274) \quad (0.3805)$$



The adjustment improves the fit in the lower and upper ends.

(e)

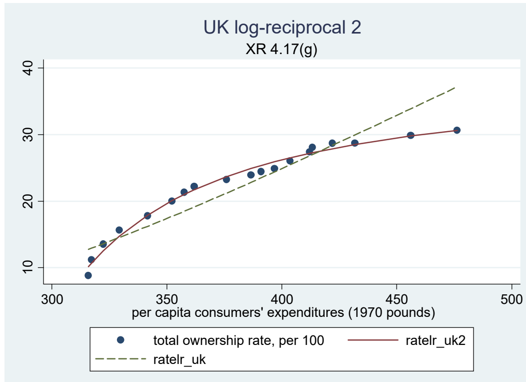
$$\widehat{\ln(RATE\_UK)} = 5.72032 - 1002.266(1/SPEND\_UK) \quad R^2 = 0.8642$$
$$\text{(se)} \qquad (0\ .2574) \quad (\ 96.3413)$$



For these data the log-reciprocal model is almost a straight line, missing the curvature of the data.

(f) The problem is that in the reciprocal the value of the explanatory variable x, here $SPEND\_UK$, becomes large, which makes $1/SPEND\_UK$ very small.

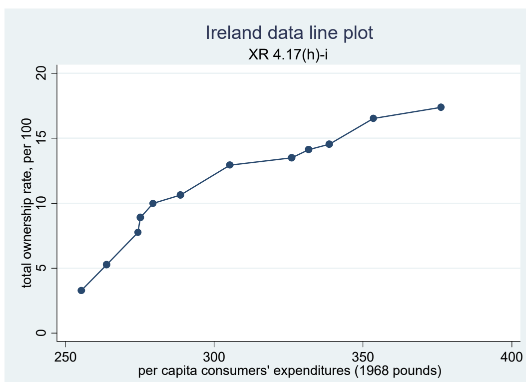
(g)
$$\widehat{\ln(RATE\_UK)} = 3.6684 - 48.3067\left(1/\left[SPEND\_UK - 280\right]\right) \quad R^2 = 0.9818$$
$$(\text{se}) \qquad (0.0229) \quad (1.5957)$$



The solid fitted line if from the modified model. For comparison we plot the dashed fitted values from part (e).

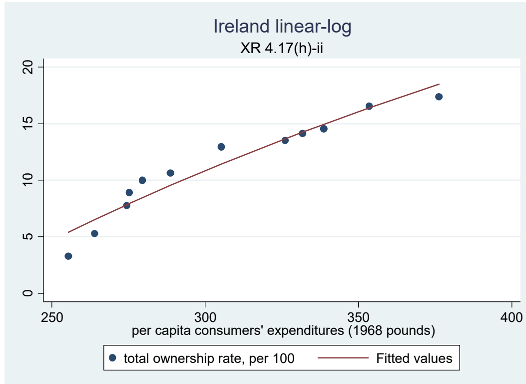
(h)

(i) The data for Ireland is shown below.



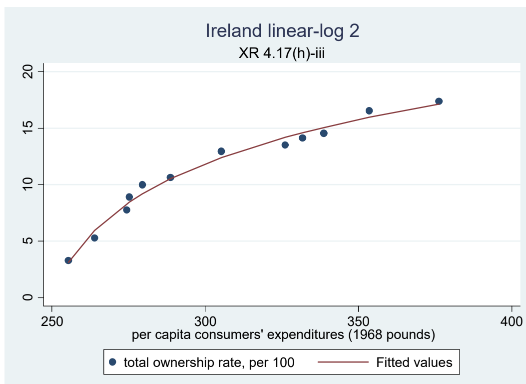
(ii) The linear log model for Ireland is

$$\widehat{RATE\_IR} = -181.762 + 33.768\ln(SPEND\_IR) \quad R^2 = 0.9326$$
$$(\text{se}) \qquad (16.412) \quad (2.871)$$



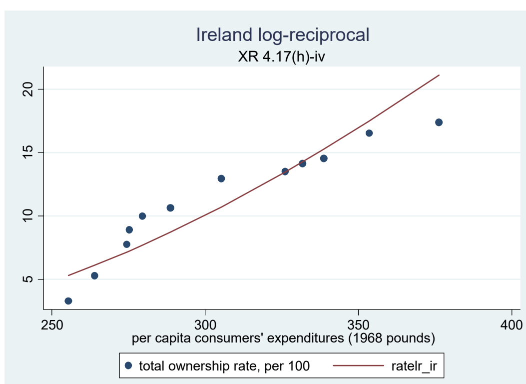
(iii) The modified linear-log model is

$$\widehat{RATE\_IR} = -14.4554 + 6.4279\ \ln(SPEND\_IR - 240) \quad R^2 = 0.9845$$
$$\text{(se)} \quad (1.0317)\ (0.2548)$$



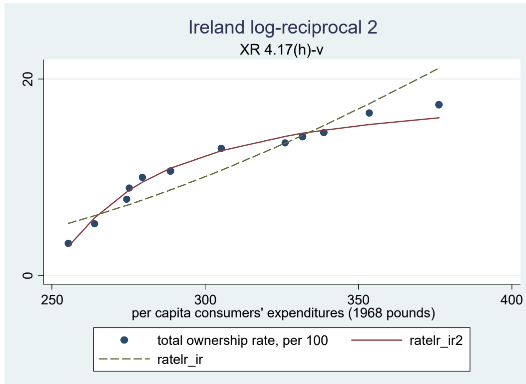
(iv) The log-reciprocal model for Ireland is

$$\widehat{\ln\left(RATE\_IR\right)} = 5.9641 - 1096.784\left(1/SPEND\_IR\right) \quad R^2 = 0.8190$$
$$\text{(se)} \quad (0.5448)\ \ (163.0473)$$



(v) The modified log reciprocal model for Ireland is

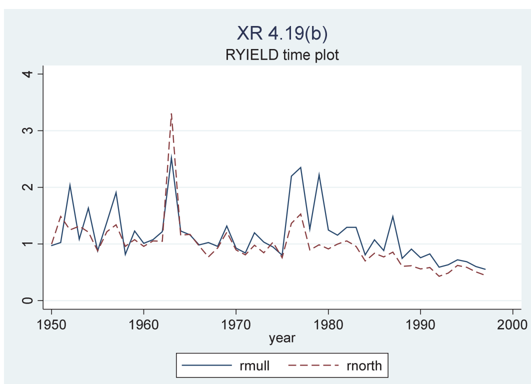
$$\widehat{\ln(RATE\_IR)} = 2.9895 - 29.0909\left(1/[SPEND\_IR - 240]\right) \qquad R^2 = 0.9824$$
$$\text{(se)} \qquad (0.0345) \quad (1.2325)$$



EXERCISE 4.19

(a) *RYIELD* can be interpreted as the number of hectares needed to produce one tonne of wheat.

(b) In the figure *RMULL* is the *RYIELD* series for Mullewa and *RNORTH* is the RYIELD series for Northampton. In both shires amount of land required to produce show a spike in 1963 and for Mullewa again in 1976-1977 and 1979.



There is an outlier in 1963 in the two shires, implying that a greater number of hectares was needed to produce one tonne of wheat than in any other year. There were similar but less extreme outliers in Mullewa in 1976, 1977 and 1979. Wheat production in Western Australia is highly dependent on rainfall, and so one would suspect that rainfall was low in the above years. A check of rainfall data at http://www.bom.gov.au/climate/data/ reveals that rainfall was lower than usual in 1976 and 1977, but higher than normal in 1963. Thus, it is difficult to assess why 1963 was a bad year; excess rainfall may have caused rust or other disease problems during the growing season, or rain at harvest time may have led to a deterioration in wheat quality.

(c) The estimated equations are given in Table XR4-19c

Table XR4-19c *RYIELD* estimations

	(1) *MULLEWA*	(2) *NORTHMAPTON*
C	1.4552	1.3934
	(11.1903)	(12.8220)
TIME	-0.0121	-0.0169
	(-2.6287)	(-4.3871)
N	48	48

t statistics in parentheses

(d) In each case the null hypothesis is rejected indicating that the required number of hectares is decreasing over time.

(e) The threshold for leverage is $2(2/N) = 0.08333333$. For studentized residuals the threshold is 2. For DFBETAS it is $2/\sqrt{N} = 0.28867513$ and for DFFITS it is $2\sqrt{K/N} = 0.40824829$. For Mullewa the values exceeding thresholds are given in Table XR 4-19(e). In the table RESID is the OLS residual, and STU_RESID is the studentized residual.

Table XR 4-19 (e) MULLEWA MEASURES

YEAR	RYIELD	LEVER	RESID	STU_RESID	DFBETAS	DFFITS
1952	2.038321	.0710124	.6195211	1.467464	-.3410547	.4057231
1963	2.522068	.0328014	1.236875	3.08792	-.3434951	.5686624
1976	2.198769	.0215118	1.071473	2.589605	.0681897	.3839671
1977	2.351835	.0221631	1.236684	3.066973	.1131014	.4617343
1979	2.222222	.0241171	1.131364	2.762376	.1602395	.4342566

For Northampton only 1963 produced unusual results.

Table XR 4-19 (e) NORTHHAMPTON MEASURES

YEAR	RYIELD	LEVER	RESID	STU_RESID	DFBETAS	DFFITS
1963	3.306878	.0328014	2.150595	11.8371	-1.31674	2.179886

(f) The year 1963 had three measures above the usual thresholds. Other years did not have as many, so we choose 1963 as the most unusual year.

(g) Omitting 1963

Mullewa

$$\widehat{RYIELD} = 1.3929 - 0.0107TIME$$
$$\text{(t)} \qquad (11.499) \; (-2.5033)$$

Northampton

$$\widehat{RYIELD} = 1.2850 - 0.0144TIME$$
$$\text{(t)} \qquad (23.389) \; (-7.4383)$$

EXERCISE 4.21

(a) Malwai is located in southeast Africa and is bordered by Zambia to the northwest, Tanzania to the northeast and Mozambique to the east, south and west. See, for example, https://en.wikipedia.org/wiki/Malawi. As of April 2017, the exchange rate was 1 Malawian Kwacha to 0.0014 US$. The 2015 population is just over 17 million according to the world bank, http://data.worldbank.org/country/malawi. The main industry is agriculture.

(b) The fitted model is

$$\widehat{PFOOD} = 0.5179241 - 0.0726372\ln(TOTEXP) \qquad R^2 = 0.1180$$
$$\text{(se)} \qquad (0.0112454)(0.0057363)$$

We estimate $\beta_2 < 0$, so that as total expenditure rises the share devoted to food declines. A 95% interval estimate of β_2 is

$$b_2 \pm t_{(0.975,N-2)}\text{se}(b_2) = -0.0726372 \pm 1.96(0.0057363) = [-0.0838916, \quad -0.0613828]$$

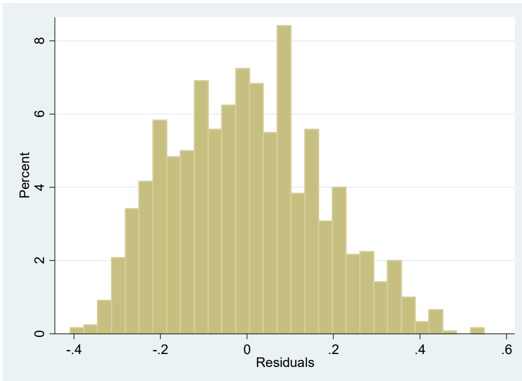
The interval estimate is relatively narrow due to the small standard error of the coefficient estimate. We have estimated β_2 relatively precisely.

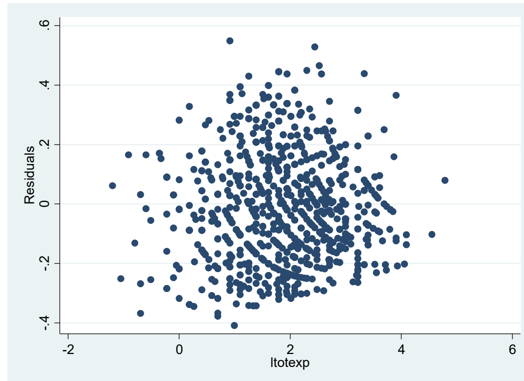
(c) The 5th percentile of *TOTEXP* is 1.5 and the 75th percentile is 10.

$\hat{\varepsilon}_5 = 0.8512972, \quad \text{se}(\hat{\varepsilon}_5) = 0.0095097$

$\hat{\varepsilon}_{75} = 0.7928621, \quad \text{se}(\hat{\varepsilon}_{75}) = 0.0184522$

(d)





The histogram shows a not quite bell-shaped distribution. The residual plot shows no strong "spray" or other pattern.

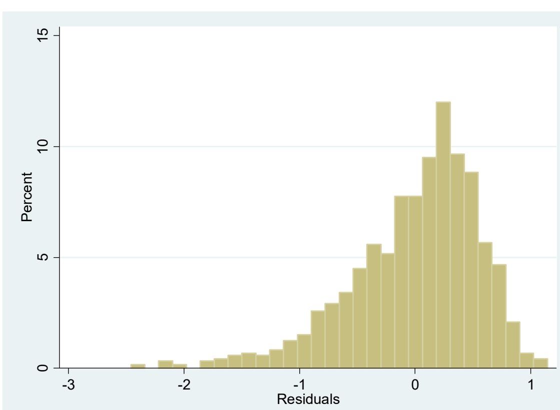
The values for the skewness and kurtosis are approximately 0.261 and 2.519, and the calculated JB test statistic value is 25.192. Thus, we reject the null hypothesis that the regression errors are normally distributed at the 1% level.

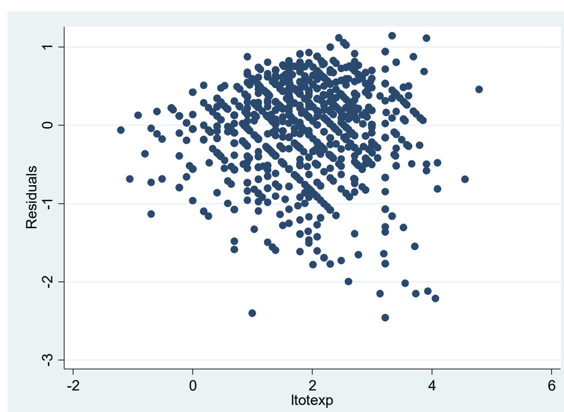
(e) The estimated log-log model is

$$\widehat{\ln(FOOD)} = -0.647313 + 0.7498553 \ln(TOTEXP) \quad R^2 = 0.5785$$
$$(\text{se}) \qquad (0.036251) \quad (.0184918)$$

The log-log model is a constant elasticity function.

(f)





The value of the Jarque-Bera statistic is 298.16 and far exceeds $\chi^2_{(0.99,2)} = 9.210$. We reject the normality of the regression errors.

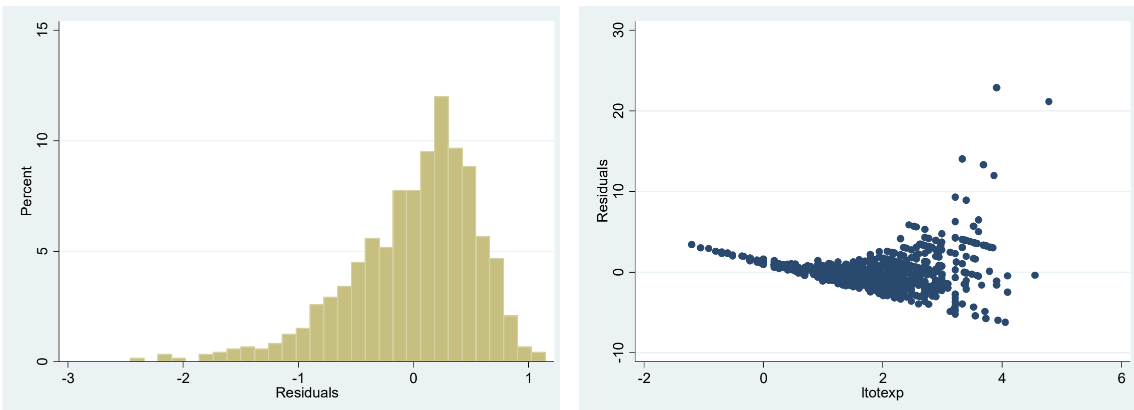
(g) The estimated model is

$$\widehat{FOOD} = -0.7754008 + 2.017512\ \ln(TOTEXP) \quad R^2 = 0.4426$$
$$(\text{se}) \qquad (0.1282382) \quad (0.0654148)$$

$$\hat{\varepsilon}_{50} = 0.7373919, \quad \text{se}(\hat{\varepsilon}_{50}) = 0.0286551$$

$$\hat{\varepsilon}_{75} = 0.5213085\ , \quad \text{se}(\hat{\varepsilon}_{75}) = 0.0143217$$

(h)



The estimated skewness is 3.422 and the estimated kurtosis is 34.168. The value of the Jarque-Bera statistic is 50913.77 which is greater than the critical value 9.210. We reject the normality of the model random errors.

(i) The correlations with *FOOD* for the three models are 0.7153, 0.7162 and 0.6653, respectively. The first of the models seems the best choice.

EXERCISE 4.23

(a) The estimates are given in Table XR 4.23

Table XR 4.23

	i	ii	iii	iv	v	vi
	Telephone		Clothes		Fuel	
C	-0.0160	-3.2731	0.0490	-1.6186	0.0666	-2.8179
	(0.0031)	(0.2443)	(0.0071)	(0.0975)	(0.0033)	(0.0523)
LTOTEXP	0.0186	1.1042	0.0111	0.6918	-0.0063	0.8046
	(0.0016)	(0.0930)	(0.0036)	(0.0448)	(0.0017)	(0.0265)
N	1200	245	1200	556	1200	1159
R^2	0.105	0.367	0.008	0.301	0.012	0.443
ε at p25	3.5634	1.1042	1.1762	0.6918	0.8928	0.8046
ε at p75	1.6945	1.1042	1.1487	0.6918	0.8792	0.8046

Standard errors in parentheses

In model (i), we estimate that a 1% increase in total expenditure leads to about a 0.000186 increase in the proportion of expenditures devoted to telephone services. In model (iii) we estimate that a 1% increase in total expenditure leads to about a 0.000111 increase in the proportion of expenditures devoted to clothing. And in model (v) we estimate that a 1% increase in total expenditure leads to about a -0.000063 decrease in the proportion of expenditures devoted to fuel. In models (ii), (iv) and (vi) the coefficients are elasticities which are discussed in the next part of the solution.

(b) In the log-log models, (ii), (iv) and (vi), the estimated coefficients are constant elasticities; they do not vary. For the telephone expenditures the elasticity is 1.1042, suggesting that it is a luxury item. For the clothing expenditures the elasticity is 0.6918, suggesting that it is a necessity item. For the fuel expenditures the elasticity is 0.8046, suggesting that it is a necessity item.

(c) The variable *TOTEXP* is "Total household expenditure last month, in thousands of Malawian Kwacha". The median expenditure is 5.7 and the mean 8.5. The 25th percentile is 3.5 and the 75th percentile is 10. As we can see the elasticities vary quite a bit across total expenditure levels. For telephone services the elasticity at the 25th percentile is 3.56, and at the 75th percentile it is 1.69. In both cases we would classify telephone services as a luxury, but the responsiveness of household's budget share is quite different. For clothing and fuel there are slight differences across the percentiles, but not very much. And the elasticities are quite similar to the constant elasticity estimated by the log-log model, for fuel.

EXERCISE 4.25

(a)

Table XR4.25 Solutions

	Model (1)		Model (2)		(3)	
	ln(*PRICE*)		ln(*PRICE*)		*PRICE*	
C	4.3939	(0.0433)	2.0497	(0.1580)	-115.4236	(13.0882)
SQFT	0.0360	(0.0015)			13.4029	(0.4492)
ln(*SQFT*)			1.0248	(0.0484)		
N	500		500		500	
R^2	0.542		0.474		0.641	
R_g^2	0.662		0.645		0.641	
JB	26.6790		14.2787		221.1115	
JB *p*-value	0.0000		0.0008		0.0000	
$\hat{\sigma}$	0.34001		0.36432		102.8	

Standard errors in parentheses

(b) We estimate a 1% increase in *SQFT* to increase expected priced by 1.02%. At the mean values, we estimate that an added 100 square feet of interior space increases expected price by $9,400.

(c) The linear model results are in Table XR4.25, Model (3). The $R^2 = 0.641$, or 64.1% of the variation in *PRICE* about its mean is explained by the model. The generalized R^2 for the log-linear model is largest, indicating that it fits the data better than the other two models, at least based on this measure.

(d) In each case we reject normality of the regression errors at any usual level of significance.

(e) It is fair to say the logarithmic models have less clear evidence about the violation of the homoskedasticity assumption.

(f) The predictions are

Linear model: $246,456

Log-linear model “Natural predictor”: $214,234

Log-linear model “Corrected predictor”: $226,982

Log-log model “Natural predictor”: $227,539

Log-log model “Corrected predictor”: $243,151

(g) The 95% prediction intervals are

Linear model: $44,277 to $448,634

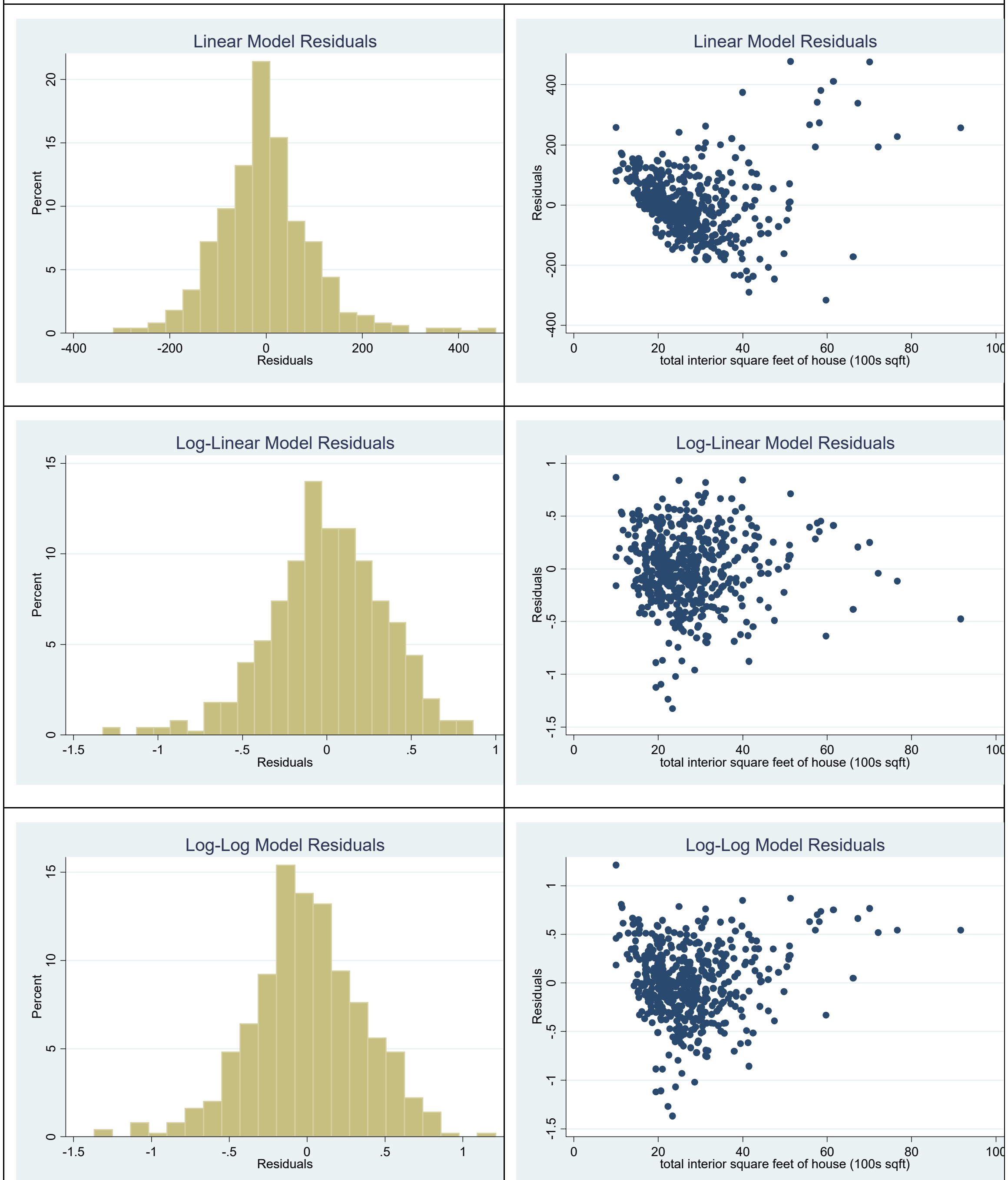
Log-linear model: $109,769 to $418,117

Log-log model: $111,141 to $465,841

In each case the intervals are quite wide and not too informative. House prices depend on much more than their size.

(h) There is little difference between the log-log and log-linear model results. The log-linear model has a slightly higher generalized R^2 but the log-log model has a slightly smaller Jarque –Bera Statistic. The slopes and elasticities are similar. At this point either model is preferable to the linear relationship.

Figure XR 4.25



EXERCISE 4.27

(a) The summary statistics for each partition are shown below. We observe that the mean and median wage of white males is highest, followed by white female, black male and black female. Female wages have a higher standard deviation than males: the standard deviation is lowest for black male, followed by black female, white male and white female.

White male Wage							
	N	mean	Std. Dev.	CV	min	p50	max
	4740	24.9481	15.2378	61.07801	2.5	21	221.1

White female Wage							
	N	mean	Std. Dev.	CV	min	p50	max
	3674	22.06724	17.6279	79.88264	2.56	18	466

Black male Wage							
	N	mean	Std. Dev.	CV	min	p50	max
	404	19.81181	11.22209	56.64343	5.49	16.83	72.13

Black female Wage							
	N	mean	Std. Dev.	CV	min	p50	max
	473	19.67019	13.51805	68.72353	2.75	16.15	173.13

(b) The coefficient of variation is given in the part (a) tables, CV. We see that the variation in female wages, scaled by the mean, is higher than for males. The variation for white females is higher than the variation for black females, and the variation for white males is higher than the variation for black males.

(c) The estimated models are in Table XR 4.27.

Table XR 4.27

	(1) White male	(2) White female	(3) Black male	(4) Black female
C	1.7274	1.3944	1.5160	1.3998
	(0.0359)	(0.0452)	(0.1292)	(0.1270)
EDUC	0.0951	0.1058	0.0983	0.1015
	(0.0025)	(0.0031)	(0.0093)	(0.0089)
N	4740	3674	404	473
R^2	0.231	0.244	0.216	0.216

Standard errors in parentheses

(d) Interval estimates are given in Table XR 4.27(d).

Table XR 4.27(d)

	(1) White male	(2) White female	(3) Black male	(4) Black female
C	1.727 [1.6570,1.7977]	1.394 [1.3057,1.4830]	1.516 [1.2620,1.7700]	1.400 [1.1503,1.6493]
EDUC	0.0951 [0.0902,0.1001]	0.106 [0.0998,0.1118]	0.0983 [0.0799,0.1166]	0.101 [0.0840,0.1190]
N	4740	3674	404	473
R^2	0.231	0.244	0.216	0.216

95% confidence intervals in brackets

The interval estimates for the slope coefficients all overlap. The returns to education are similar in the sense that the range of values are not distinct. If we were to test the hypothesis that the slope parameter $\beta_2 = c$, where c is a value in another interval, we would find values that we would not reject.

(e) For the intercept parameter, the white male and white female intervals do not overlap. The white male and black female intervals do not overlap.

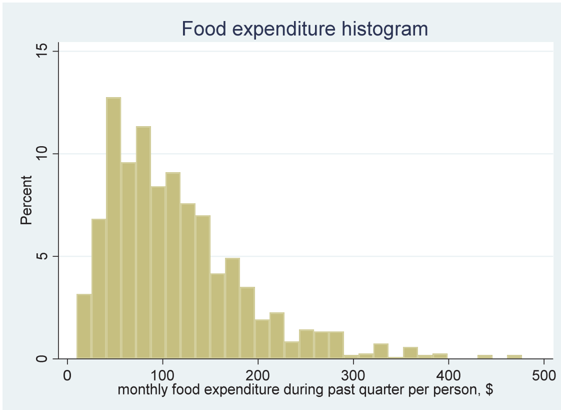
(f) The models fit similarly well for all groups, but the fit is higher for whites than blacks, and is higher for white females than white males.

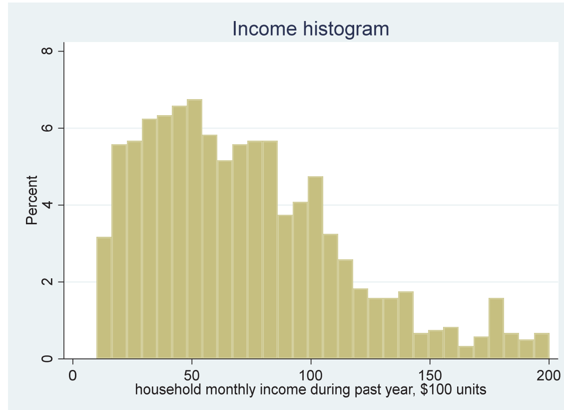
EXERCISE 4.29

(a) The summary statistics are

variable	N	mean	p50	min	max	Std. Dev.
FOOD	1200	114.4431	99.8	9.63	476.67	72.6575
INCOME	1200	72.14264	65.29	10	200	41.65228

The histograms are





Both distributions are positively skewed with mean's greater than medians. They are not bell shaped or symmetrical. For *INCOME* the Jarque-Bera statistic is 148.21 and for *FOOD* expenditure it is 648.65. The critical value for a test at the 5% level is 5.99. We reject the null hypothesis of normality for each variable.

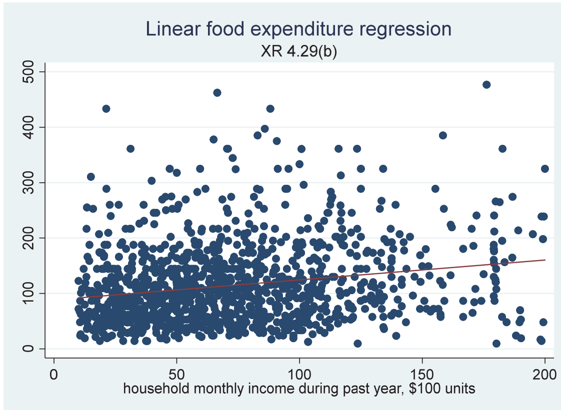
(b) The estimation results are in Table XR 4.29. The linear relation estimates are in column (1).

Table XR 4.29

	(1) Linear relation	(2) Log-Log relation	(3) Linear-Log relation
C	88.5665	3.7789	23.5685
	(4.1082)	(0.1203)	(13.3696)
INCOME	0.3587		
	(0.0493)		
ln(*INCOME*)		0.1863	22.1874
		(0.0290)	(3.2253)
N	1200	1200	1200
R^2	0.042	0.033	0.038

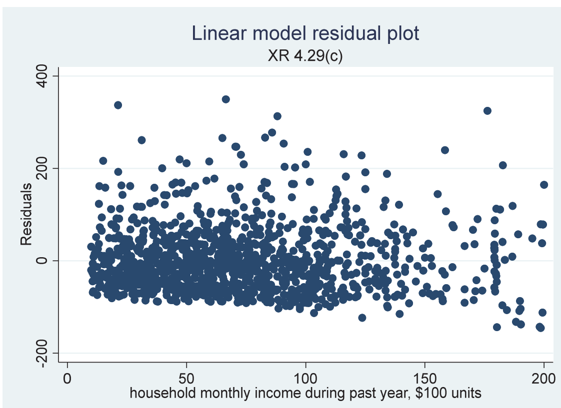
Standard errors in parentheses

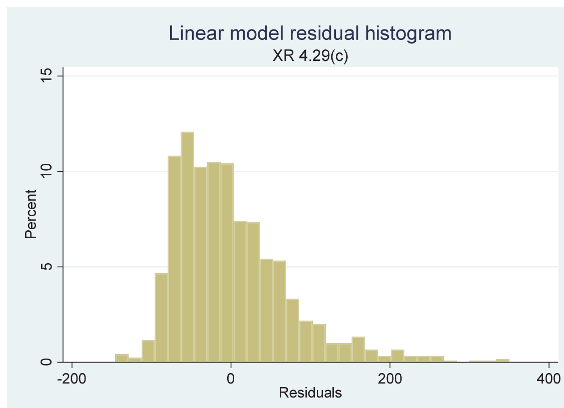
The scatter plot and fitted line are in Figure XR 4.29(b).



The 95% interval estimate of the slope is [0.2619, 0.4555].

(c) The least squares residuals are plotted in Figure XR 4.29(c). The positive skew at each income is clear. There is not a clear "spray" pattern except at high incomes. The residual histogram shows the skewness. The Jarque-Bera statistic is 624.186, which is far greater than the 5% critical value 5.99.

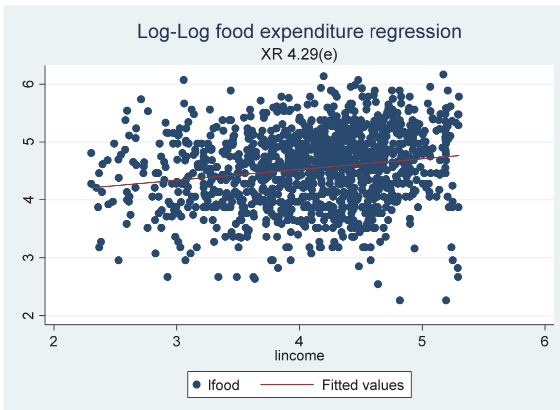




(d) The estimated elasticity and intervals are

INCOME	$b_1 + b_2 INCOME$	$\hat{\varepsilon}$	$se(\hat{\varepsilon})$	LB	UB
19	95.3815	0.0715	0.009825	0.0522	0.09073
65	111.8811	0.2084	0.02865	0.1522	0.2646
160	145.9564	0.3932	0.0541	0.2871	0.4993

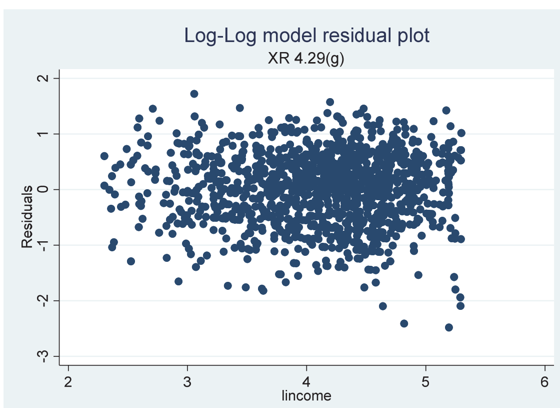
(e) The log-log model estimation results are in column (2) of Table XR 4.29. The data and fitted relationship are in Figure XR 2.29(e)



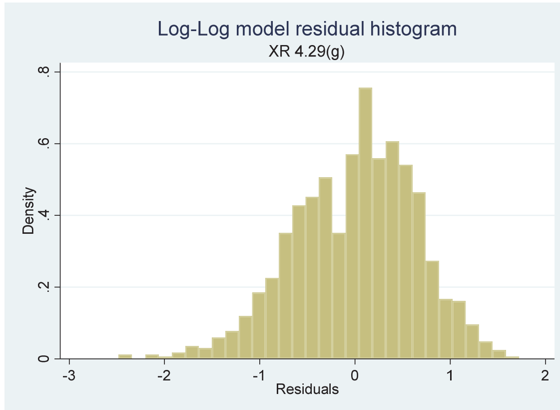
The generalized R^2 is 0.03965 which is slightly smaller than the R^2 from the linear model.

(f) The 95% interval estimate of the elasticity is [0.1293, 0.2433].

(g) The residual scatter from the log-log model is shown in Figure XR 4.29(g).

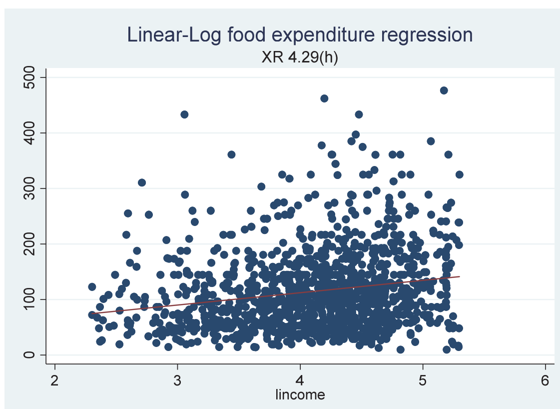


The residual histogram is given in Figure XR 4.29(g).



There is a slight negative skew (Skewness = -0.3577) and Kurtosis is 3.0719. The Jarque-Bera statistic is 25.85 which is greater than the 5% critical value 5.99. So, we reject the null hypothesis that the log-log regression errors are normal.

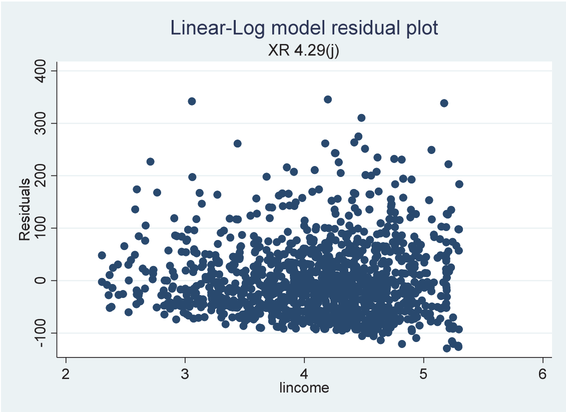
(h) The estimated linear-log model is in column (3) of Table XR 4.29. The data and fitted line are in Figure XR 4.29(h). The figure is much like that for the linear model, and not as well defined as that for the log-log model. The $R^2 = 0.038$, which is smaller than that of the linear model, and smaller than the generalized R^2 from the log-log model.

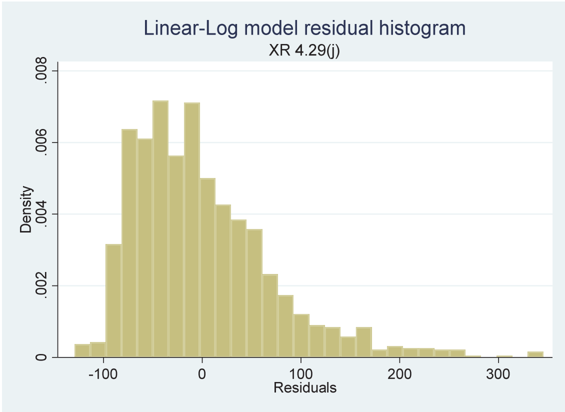


(i) At the given income values the estimates and intervals are:

INCOME	$\hat{\alpha}_1 + \hat{\alpha}_2 \ln(INCOME)$	$\hat{\varepsilon}$	$\text{se}(\hat{\varepsilon})$	LB	UB
19	88.8979	0.2496	0.0363	0.1784	0.3208
65	116.1872	0.1910	0.0278	0.1365	0.2454
160	136.1733	0.1629	0.0237	0.1165	0.2094

(j) The residual diagrams for the linear-log model are in Figures XR 4.29(j). The residual scatter shows positive skewness at each income level and overall. The Jarque-Bera statistic is 628.07 which is far greater than the 5.99 critical value. We reject the normality of the model errors. The data scatter suggests a slight "spray" pattern.





(k) The linear model is counter-intuitive with increasing income elasticity. The linear-log model certainly satisfies economic reasoning, but the residual pattern is not an ideal random scatter. The log-log model implies that the income elasticity is constant for all income levels, which is not impossible to imagine, and the residual scatter is the most random, and the residuals are the least non-normal, based on skewness and kurtosis. On these grounds the log-log model seems like a good choice.