5.3 Consider the following model that relates the percentage of a household's budget spent on alcohol *WALC* to total expenditure *TOTEXP*, age of the household head *AGE*, and the number of children in the household *NK*.

$$WALC = \beta_1 + \beta_2 \ln(TOTEXP) + \beta_3 NK + \beta_4 AGE + e$$

This model was estimated using 1200 observations from London. An incomplete version of this output is provided in Table 5.6.

TABLE 5.6 **Output for Exercise 5.3**

Dependent Variable: *WALC*				
Included observations: 1200				
Variable	**Coefficient**	**Std. Error**	***t*-Statistic**	**Prob.**
C	1.4515	2.2019		0.5099
ln(*TOTEXP*)	2.7648		5.7103	0.0000
NK		0.3695	−3.9376	0.0001
AGE	−0.1503	0.0235	−6.4019	0.0000
R-squared		Mean dependent var		6.19434
S.E. of regression		S.D. dependent var		6.39547
Sum squared resid	46221.62			

a. Fill in the following blank spaces that appear in this table.
  i. The *t*-statistic for b_1.
  ii. The standard error for b_2.
  iii. The estimate b_3.
  iv. R^2.
  v. $\hat{\sigma}$.

b. Interpret each of the estimates b_2, b_3, and b_4.

c. Compute a 95% interval estimate for β_4. What does this interval tell you?

d. Are each of the coefficient estimates significant at a 5% level? Why?

e. Test the hypothesis that the addition of an extra child decreases the mean budget share of alcohol by 2 percentage points against the alternative that the decrease is not equal to 2 percentage points. Use a 5% significance level.

a. i. $t = \dfrac{b_1}{SE(b_1)} = \dfrac{1.4515}{2.2019} = 0.6592$

ii. $SE(b_2) = \dfrac{b_2}{t} = \dfrac{2.7648}{5.7103} = 0.4842$

iii. $b_3 = t \times SE(b_3) = (-3.9376) \times 0.3695 = -1.4549$

iv. $R^2 = 1 - \dfrac{SSR}{SST} = 1 - \dfrac{46{,}221.62}{49{,}041.54} = 0.0575$

$SST = (N-1) S_y^2 = (1200-1) \times (6.39547)^2 = 49{,}041.5418$

v. $\hat{\sigma}^2 = \dfrac{SSR}{N-K} = \dfrac{46{,}221.62}{1200-4} = 38.6468$

$\hat{\sigma} = \sqrt{38.6468} = 6.2167$

b. b2 =2.7648 (coefficient for ln(TOTEXP)):
Interpretation: For a 1% increase in total expenditure (TOTEXP), since the variable is in log form (ln(TOTEXP)), the percentage of the budget spent on alcohol increases by approximately 2.7648÷100=0.027648 percentage points, holding NK and AGE constant. For example, if TOTEXP increases by 1%, WALC increases by about 0.0276 percentage points.

b3 =−1.4549 (coefficient for NK):
Interpretation: For each additional child in the household (NK), the percentage of the budget spent on alcohol decreases by 1.4549 percentage points, holding ln(TOTEXP) and AGE constant. For example, adding one child reduces the alcohol budget share by about 1.4549 percentage points.

b4 =−0.1503 (coefficient for AGE):
Interpretation: For each additional year in the age of the household head (AGE), the percentage of the budget spent on alcohol decreases by 0.1503 percentage points, holding ln(TOTEXP) and NK constant. For example, if the household head is one year older, the alcohol budget share decreases by about 0.1503 percentage points.

c. A 95% confidence interval for β4 is b4 ± t_critical*SE(b4)
For a 95% confidence interval, the critical t-value (t_critical) with 1196 degrees of freedom is approximately 1.96.
Lower bound: -0.1503 - 1.96*0.0235 = −0.19636
Upper bound: -0.1503 + 1.96*0.0235 = −0.10424

```
> qt(1-0.05/2,1196)
[1] 1.961949
```

Thus, the 95% confidence interval for β4 is approximately (-0.1964, -0.1042).
This means we are 95% confident that the true effect of a one-year increase in AGE on WALC lies between a decrease of 0.1964 percentage points and a decrease of 0.1042 percentage points, holding other variables constant.

d. A 5% significance level (α=0.05) means we reject H0 if the p-value is less than 0.05.
A coefficient is significant if the p-value is less than 0.05, meaning we reject the null hypothesis that the coefficient is zero.
For b1 , the p-value (0.5099) is greater than 0.05, so it's not significant different from zero at a 5% level.
For b2, b3, and b4, the p-value is less than 0.05, indicating that the coefficient estimates are significantly different from zero at a 5% level.

e. Null Hypothesis (H0): β3 =−2 (an extra child decreases WALC by 2 percentage points).
Alternative Hypothesis (H1): β3 ≠ −2 (the decrease is not equal to 2 percentage points).
Significance level: α=0.05.
The critical t-value for a two-tailed test at α=0.05 with 1196 degrees of freedom is approximately 1.96.
The calculated t-value is t = [−1.4549-(-2)]÷0.3695 = 1.4752

```
> qt(1-0.05/2,1196)
[1] 1.961949
```

We fail to reject H0 because | 1.4752 | < 1.96.
Therefore, the data does not provide sufficient evidence to conclude that having an extra child leads to a decline in the alcohol budget share that is significantly different from 2 percentage points.