

APPENDIX 2.5
(referred to in paragraph 2.66)

Regression analysis of United Kingdom ferry fares

1. We received views that cross-Solent ferry fares were 'the most expensive in the world' for the distance travelled. It was also, however, argued to us that one reason for the high level of fares was the relatively short distance of the crossings. This assumes the existence of a relationship between unit price and distance travelled. In order to test the extent to which ferry fares could be said to be related to the distance covered or the crossing time, we undertook an analysis of ferry services in United Kingdom waters.

2. Figures 2.10, 2.12, 2.14 and 2.16 plot actual values of pence per kilometre against distance and suggest the existence of a negative relationship.

3. For the purposes of the regression analysis, each unit price was given its logarithmic value in order to reduce the likelihood of error arising from wide variation in unit values (price per kilometre) and increase the degree of standardisation.

4. Details of the regression output undertaken to test that relationship are given in Table 1. The statistic which confirms the strength or weakness of a statistical relationship is R Squared. Where there is a perfect and positive relationship between two variables (perhaps, for example, the demand for cars and car tyres), R Squared will be equal to 1.0. Where there is a perfect but negative relationship between two variables, R Squared will equal -1.0. The closer R Squared is to 1.0, therefore, the more appropriate it is to use those data to make predictions about unit prices.

5. Table 1 shows the results achieved when attempting to test the strength of the relationship between price for a 4-metre car plus two adults and two children and distance for all United Kingdom ferry services for which data were received.

6. Where the strength of the relationship between distance and peak single fares was tested, an R Squared of 0.56 suggested the existence of a weak relationship. However, as shown in Table 2, when observations on subsidised Scottish services run by Caledonian MacBrayne were excluded from the analysis, R Squared fell to 0.15.

7. Where distance and time were tested against peak single fares equivalent results were obtained: 0.58 including observations on Caledonian MacBrayne, 0.15 when they were excluded.

8. Testing off-peak single fares against distance, however, gave the reverse result. When observations on Caledonian MacBrayne were included the value of R Squared was 0.33. Excluding those services, R Squared rose to 0.6.

9. The same result was found where distance and time were tested against off-peak single fares. When observations on Caledonian MacBrayne were included, R Squared was calculated at 0.38, when they were excluded R Squared rose to 0.62.

10. The analysis might therefore be used to suggest that for all unsubsidised ferry services a relationship between price and distance exists for off-peak single fares, and that a relationship exists between distance and peak single fares for all United Kingdom ferry services.

11. However, the caveats stated in paragraphs 2.66 and 2.67 and the small sample size, particularly when observations on Caledonian MacBrayne are excluded, undermine the likelihood that an analysis of this type will be capable of producing robust results.

Regression analysis of logarithmic log values of distance, time and price where price = 4m car + 2 adults + 2 children

TABLE 1 Regression analysis: all UK ferry services

Regression output:

Where X = Log distance; Y = Log peak single

| | |
|---------------------|----------|
| Constant | 2.622934 |
| Std err of Y est | 0.192633 |
| R Squared | 0.558879 |
| No. of observations | 54 |
| Degrees of freedom | 52 |
| X coefficient(s) | -0.31828 |
| Std err of coef | 0.039212 |

TABLE 2 Regression analysis excluding observations on Caledonian MacBrayne services

Regression output:

Where X = Log distance; Y = Log peak single

| | |
|---------------------|----------|
| Constant | 2.409967 |
| Std err of Y est | 0.218256 |
| R Squared | 0.149673 |
| No. of observations | 32 |
| Degrees of freedom | 30 |
| X coefficient(s) | -0.17904 |
| Std err of coef | 0.077912 |

Regression output:

Where X = Log distance; Y = Log off-peak single

| | |
|---------------------|----------|
| Constant | 2.510184 |
| Std err of Y est | 0.342168 |
| R Squared | 0.327591 |
| No. of observations | 54 |
| Degrees of freedom | 52 |
| X coefficient(s) | -0.35058 |
| Std err of coef | 0.069652 |

Regression output:

Where X = Log distance; Y = Log off-peak single

| | |
|---------------------|----------|
| Constant | 2.605841 |
| Std err of Y est | 0.158459 |
| R Squared | 0.603141 |
| No. of observations | 32 |
| Degrees of freedom | 30 |
| X coefficient(s) | -0.38195 |
| Std err of coef | 0.056566 |

Regression output:

Where X = Log distance and log time; Y = Log peak single

| | |
|---------------------|----------|
| Constant | 2.237623 |
| Std err of Y est | 0.189669 |
| R Squared | 0.580572 |
| No. of observations | 54 |
| Degrees of freedom | 51 |
| X coefficient(s) | -0.77898 |
| Std err of coef | 0.286287 |

Regression output:

Where X = Log distance and log time; Y = log peak single

| | |
|---------------------|----------|
| Constant | 2.292009 |
| Std err of Y est | 0.22154 |
| R Squared | 0.153093 |
| No. of observations | 32 |
| Degrees of freedom | 29 |
| X coefficient(s) | -0.37404 |
| Std err of coef | 0.575239 |

Regression output:

Where X = Log distance and log time; Y = Log off-peak single

| | |
|---------------------|----------|
| Constant | 3.393777 |
| Std err of Y est | 0.331045 |
| R Squared | 0.382703 |
| No. of observations | 54 |
| Degrees of freedom | 51 |
| X coefficient(s) | 0.705913 |
| Std err of coef | 0.499679 |

Regression output:

Where X = Log distance and log time; Y = Log off-peak single

| | |
|---------------------|----------|
| Constant | 2.313552 |
| Std err of Y est | 0.157347 |
| R Squared | 0.621738 |
| No. of observations | 32 |
| Degrees of freedom | 29 |
| X coefficient(s) | -0.86515 |
| Std err of coef | 0.408557 |

Source: MMC.