OFFICE OF FAIR TRADING

# Valuation of Hackney carriage licence premiums

## Annexe E

## November 2003

**This study was commissioned for the market study. It is the responsibility of the authors and any views expressed in it are those of the authors and not necessarily of the OFT.**

## Further copies

Further copies of this paper can be downloaded from our website at www.oft.gov.uk/news/publications/leaflet+ordering.htm - look under 'Reports' and 'Competition policy' - or you can obtain hard copies by phoning OFT publications on the number shown below.

## OFT publications order line

OFT publications can be ordered free of charge on 0870 60 60 321.

## Information on the internet

You can also order OFT publications online and find advice and information on consumer rights and on the work of the OFT at www.oft.gov.uk

# Contents

# 1 INTRODUCTION

## 1.1 Objectives

1.1.1 This report forms part of a study into the impact of regulation of entry in taxi markets on consumers. The aim is to determine whether such regulation provides measurable benefit to consumers. Work reported elsewhere as part of this study has examined this issue directly. This report presents the results of an empirical exercise to assess whether there are measurable costs to consumers from entry control. The work is based on data collected from a series of demand and regulation studies on behalf of entry restricted local authorities.

1.1.2 The objectives of this study are set out as follows:

- to establish the apriori theoretical relationship between the resale value of hackney licence plates in entry restricted authorities and market characteristics,
- to explore these relationships with econometric models, and
- to draw conclusions about consumer detriment as a result of entry control.

1.1.3 A licence premium arises when the revenue from hackney carriage operation exceeds that associated with the market clearing price. Hackney markets are generally subject to heavy regulation with price and quality being controlled in addition to quantity. In an authority with a limit on the number of hackneys, consumer benefit will be maximised at the price associated with normal levels of profit to operators. In most markets competition between suppliers could be expected to bring about this optimum price. In the hackney market, however, the price mechanism is weak due to individual drivers' inability to reap the reward of cutting price. At ranks passengers are compelled to take the vehicle at the front of the queue, and if seeking a vehicle in the street, it is highly likely that the consumer will accept the first vehicle to arrive rather than seeking out the particular driver who offered a discounted fare on a previous occasion.

1.1.4 The implication is that in the presence of entry control, there is downward pressure on fares. In addition, it has traditionally been regarded as necessary to protect the public from the pricing policies of unscrupulous drivers. The combined effect of these two factors has been near universal fare regulation.

1.1.5 The consequence of entry control and price regulation together has been the emergence of a resale value attached to a hackney vehicle licence (the premium). Such values do not exist in authorities with no entry control suggesting that there is no tangible goodwill associated with hackney licences that could explain this value. Other things being equal it is to be expected that the main influences on licence values would include:

- price. The higher the level of fares, the higher the premium,
- service level. The lower the level of supply the higher the level of cab utilisation and the higher the revenue. There are three principal measures of service available:
    - passenger delay ,
    - a further more comprehensive measure of service performance, that takes account of other factors that affect service level, such as the presence of high demand late at night at weekends (ISUD),
    - per capita provision of hackneys, and
    - time. The value of the premium would be expected to increase over time with prices and/or incomes.

## 1.2 Overview of method

1.2.1 Two distinct types of analysis have been used to investigate the value of taxi licence premiums. The majority has been performed using econometric analysis of the results from over 14 years of hackney carriage demand studies to construct several models between lice premiums and a number of explanatory variables. This is supplemented by a series of interviews with taxi licensing officers, to explore the robustness of some of the key data.

## 1.3 Economic analysis

1.3.1 There are effectively three stages to the economic analysis, these are presented below.

(a) **Simple Data Search.** This part of the analysis involved searching through all 71 taxi demand studies for which data are available as well as other sources of information such as the Office for National Statistics. The aim was to compile a data set with the licence premiums as the dependent variable and as many potentially related independent variables as possible. The results of this work are detailed later in this chapter.

(b) **Mulitple Regression Analysis.** This stage of the work utilised an extensive least squares regression analysis to develop a number models explaining the resale value of licence premiums, this analysis had two distinct purposes; firstly, to determine which of the variables presented below, (from paragraph 1.5 onwards), have a statistical link with the value of licence premiums; and secondly to determine the best way to model these links.

(c) **Diagnostics.** This portion of the analysis tests the models constructed in the previous section for the common econometric problems of autocorrelation and heteroskedasticity, and applies corrective procedures where necessary. These problems and the standard corrections used to overcome them are explained in the appendix.

## 1.4 Interviews with licensing officers

1.4.1 The data on licence premiums has been derived from discussions with licensing officers. Trading in licence plates is generally discouraged by local authorities and is not therefore publicised by those involved. Licensing and enforcement officers do nevertheless obtain information on trading from several sources including,

- adverts for vehicles at inflated prices. These may be placed in local publications or even posted on the notice board in the licensing department,

- complaints from drivers who either feel they have been overcharged, or who wish to defend the value of the licence from policy changes such as proposals to issue more licences or tighten vehicle conditions,

- the general banter between officers and drivers that occurs on a day to day basis over, in some cases, many years and,

- direct involvement in the transaction as an arbitrator.

1.4.2 The interviews included several questions concerning licence valuations, how officers form an estimate of value and the confidence they had in the values they reported at the times of surveys. Of the authorities approached, all but one was willing to provide information on this subject. Officers depend almost exclusively on word of mouth for their information, though in one authority the officer had seen a classified advert that confirmed his information, and in another the officer actually witnessed a transaction taking place.

1.4.3 Officers generally rated their confidence in their information as 'high'. Of 10 officers consulted, eight gave this rating and two rated their confidence as 'medium'. Officers in derestricted authorities could quote historic values of licences from previous restricted eras with great confidence and apparent accuracy. Other officers could distinguish between the values for different classes of vehicle and different areas on the authority.

## 1.5 The data set

1.5.1 Halcrow has access to a part time series, part panel data set of 71 hackney carriage studies covering 59 licensing authorities between 1988 and 2002. Of this original data set 40 of the studies contain the sufficient level of information to be used in the construction of a causal model for the value of taxi licence premiums. The majority of studies that were excluded did not provide a monetary value for hackney carriage licences either as a result of being derestricted or the absence of a valuation by the licensing authority. In some of the older studies premiums were attached to licences however insufficient information was provided to explain how such values were derived.

1.5.2 **The Taxi Licence Premium.** In order to gain a broader appreciation of the factors determining the valuation of the taxi licence premium, Halcrow has used the study results to construct three different types of variables based on taxi licence premiums.

**(a)** **Taxi Licence Premium.** The resale value of a taxi licence premium, defined as the value the licensing authority estimated a taxi licence to be worth, at the time of study.

**(b)** **Monopoly Rent.** This variable is designed to represent a monopoly profit, or the rent that accrues directly from the existence of a limitation policy. Monopoly rent is calculated by multiplying by the number of hackney carriages within an authority by the licence premium.

(c) **Licence Premium per Trip.** Defined by the licence premiums for an authority divided by the annual number of taxi trips.

1.5.3 This is an empirical study and as such the premium has been analysed several different forms to gauge the strength of the relationship with market characteristics.

1.5.4 **Independent Variables.** Each of the 40 studies that have been included in the final data set contains detailed information about the dynamics of the taxi market within the authority where they have been conducted. This information has been used to produce a series of variables which could plausibly have an effect on the value attached to licence premiums, these variables are explained below.

1.5.5 **Total Revenue.** This is an estimate of annual revenue taken by all hackney carriages within a given licensing authority. It is calculated by multiplying the fare for a standard three mile taxi trip by the annual total number of trips derived from the Halcrow survey. The number of trips are calculated by annualising the results from Halcrow rank observations results and factoring up to include all methods of hire, (the factors for which are derived from the Halcrow public interview surveys).

1.5.6 **Real Revenue.** This is an index value of revenue designed to remove the effect of price inflation from the data set. It is calculated by dividing Total Revenue by the RPIX annual index value of inflation for the appropriate year.

1.5.7 **Revenue per Cab.** This is the figure for total revenue figure divided by the number of licensed hackney carriages within a given authority.

1.5.8 **Real Revenue per Cab.** This is analogous to the Real Revenue variable and is calculated by dividing the Real Revenue by the number of hackneys.

1.5.9 **Day/Night Ratio.** The proportional split between the trips made during the day and night, as reported by the Halcrow survey.

1.5.10 **Fare.** This is the standard fare paid for trips made within an authority, assumed to be equivalent to a single person three mile trip. This has been calculated using a weighted average to reflect the proportion of trips made during the day and night, and the price differential between standard daytime and night-time charges.

1.5.11 **Number of Passengers.** The annual number of hackney journeys made by passengers within a licensing authority, derived from the Halcrow survey.

1.5.12 **Cab/Taxi Journeys.** The annual number of fare earning journeys made by hackney carriages registered to a licensing authority, based on the Halcrow survey.

1.5.13 **Average Passenger Delay.** The average amount of time each passenger waits at a rank before gaining entry to a hackney carriage, as observed by Halcrow surveys.

1.5.14 **Per Capita Provision.** The ratio of the number of hackneys licensed by an authority to the population. Calculated by dividing the population by the number of hackneys.

1.5.15 **Average Cab Delay.** The average amount of time spent waiting a rank before a hackney secures a fare, as observed by Halcrow surveys.

1.5.16 **Excess Demand.** The proportion of hours observed during weekday daytimes where non negligible passenger queues were observed.

1.5.17 **ISUD.** The Index of Significant Unmet Demand (ISUD) was developed in the early 1990s and is based on the following formula. It has been further developed for 2003 by the addition of the seasonality.

**ISUD = APD x PF x GID x SSP**

Where:

APD = Average Passenger Delay calculated across the entire week.

PF = Peaking Factor. If passenger demand is highly peaked at night the factor takes the value of 0.5. If it is not peaked the value is 1. Following case law this provides dispensation for the effects of peaked demand on the ability of the Trade to meet that demand.

GID = General Incidence of Delay. This is measured as the proportion of passengers who travel in hours where the delay exceeds one minute.

SSP = Steady State Performance. The corollary of providing dispensation during the peaks in demand is that it is necessary to focus on performance during 'normal' hours. This is measured by the proportion of hours during weekday daytimes when the market exhibits excess demand conditions (i.e. passenger queues form at ranks).

1.5.18 The product of these four measures provides an index value. The index is exponential and values above the 80 mark have been found to indicate significant unmet demand. This benchmark was defined by applying the factor to the 25 or so studies that had been conducted at the point it was developed. These earlier studies had used the same principles but in a less structured manner. The highest ISUD value for a study where a conclusion of no significant unmet demand had been found was 72. The threshold was therefore set at 80.

1.5.19 **Hackney/Private Hire Ratio.** The ratio of hackney carriages to private hire vehicles within a licensing authority.

1.5.20 **Population.** The number of people residing within the boundaries of a licensing authority, as listed by the most recent census at the time each study was conducted.

1.5.21 **Dummy Variables.** Incorporation of dummy variables into an economic model is a commonly used tool that is used to allow a model to expressly consider specific observations or time periods within the data which exhibit unique properties, or differ in some way from the rest of the data set. A set of dummy variables were included in the study of taxi licence premiums, these pertained to a number of different variables:

(a) **Zone.** Dummy variables were included to allow for the hypothesis that licensing authorities that are split into zones, such as Sunderland, may have different values of licence premiums to the rest of the country.

(b) **Calderdale.** A dummy was included to account for the hypothesis that Calderdale has an unexplainably high value for the licence premium in relation to other authorities.

(c) **Annual.** Annual dummy variables were included to allow for differences in licence premiums depending on the year of the study. For example the dummy, 'D96 onwards', refers to all years after 1995.

(d) **Quarter**. Dummy variables were incorporated to account for seasonal variations in the value of licence premiums.

1.5.22 **Wider Economic Data.** It is possible that national economic conditions beyond the scope of individual taxi markets will have a bearing on the valuation of licence premiums. Consequently a set of wider economic variables has been included, the source of these is the Office for National Statistics (ONS):

(a) **Unemployment.** The annual rate of unemployment.

(b) **GDP.** An annual index value for Gross Domestic Product.

(c) **Inflation.** The annual RPIX level of inflation.

(d) **Personal Earnings**. The annual index of personal earnings.

# 2 INITIAL DATA ANALYSIS

## 2.1 Overview

2.1.1 The first stage of the analysis of the data required the compilation of shortlist, from the list of variables presented in the previous section, of those variables that are correlated with the licence premium. This was done by building a large number of least squares regression models with premium as the dependent variable and permutations of variables, (from the list presented above), as the regressors. Variables were omitted from the remainder of the analysis if they were not found to be correlated with the licence premium, the lack of such a link is indicated by a P value above the 0.05 cut off level and often a tendency increase P values of otherwise statistically significant variables.

## 2.2 Omitted variables

2.2.1 Despite having strong apriori expectations of the major influences on licence premium values, it is useful to explore the impact of other data available as this can inform the subsequent analysis. This led to a number of variables being excluded on empirical as well as theoretical grounds. Table 1 illustrates the variables that were excluded from the remainder of the analysis as well as the best P values that could be found for each. Also included is an indication of whether the significance of other variables reduced when these variables were included.

2.2.2 Of the omitted variables presented in Table 1 the dummy variable for the second quarter of each year was the closest to having a significant effect on the licence premium, the P value is moderate at 0.065. When the quarterly dummy was included in the models however it regularly caused a fatal increase in the P values of variables that had been found to have a significant effect on the licence premium. Similarly to the quarterly dummy, the number of cab journeys was close to being significant, exhibiting a marginal P value, however as with the previous variable, it's inclusion was to the determent of other components in the model.

2.2.3 The decision to omit the remaining variables was straightforward as all P values were well above the critical values. To illustrate why this is the case Figures 1 and 2 below show scatter plots between the licence premium and the GDP, and the premium and excess demand, respectively. GDP is one of the better variables in the remaining group, this is indicated by a little positive correlation in the plot of premium and GDP. The majority of the data points however appear to be widely scattered to the extent that a meaningful relationship cannot be modelled. The proportion of time when the market is in a state of in excess demand, by contrast is the least significant of the variables with a high P value. Consequently when 'time in excess demand' is plotted against the licence premium there is no correlation, with data points apparently scattered randomly across the graph.

TABLE 1: REJECTED VARIABLES

| **Variable** | **Best P value** | **Renders other variables insignificant** |
|---|---|---|
| Day/Night Ratio | 0.214 | Yes |
| Passengers | 0.170 | Yes |
| Average Cab Delay | 0.439 | Yes |
| Cab Journeys | 0.076 | Yes |
| Excess Demand (% of Time) | 0.661 | Yes |
| Hackney/Private Hire Ratio | 0.075 | Yes |
| Quarter (Q2) | 0.065 | Yes |
| GDP | 0.130 | Yes |
| Inflation | 0.122 | Yes |
| Personal Earnings | 0.101 | Yes |

FIGURE 1: SCATTER PLOT OF GDP AGAINST LICENCE PREMIUM.

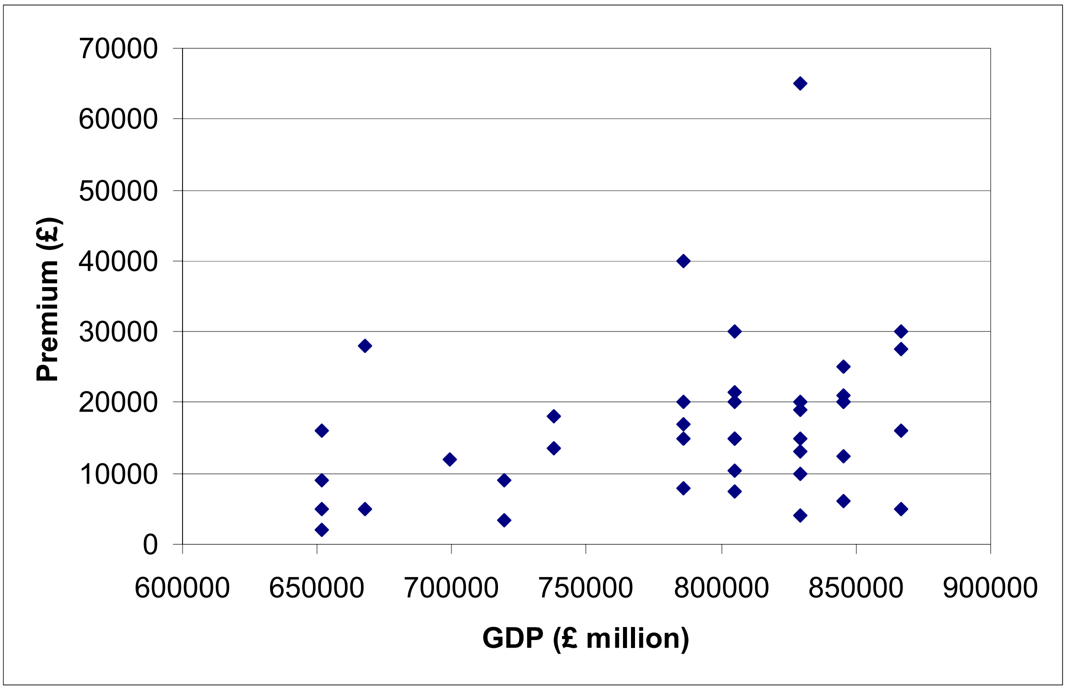
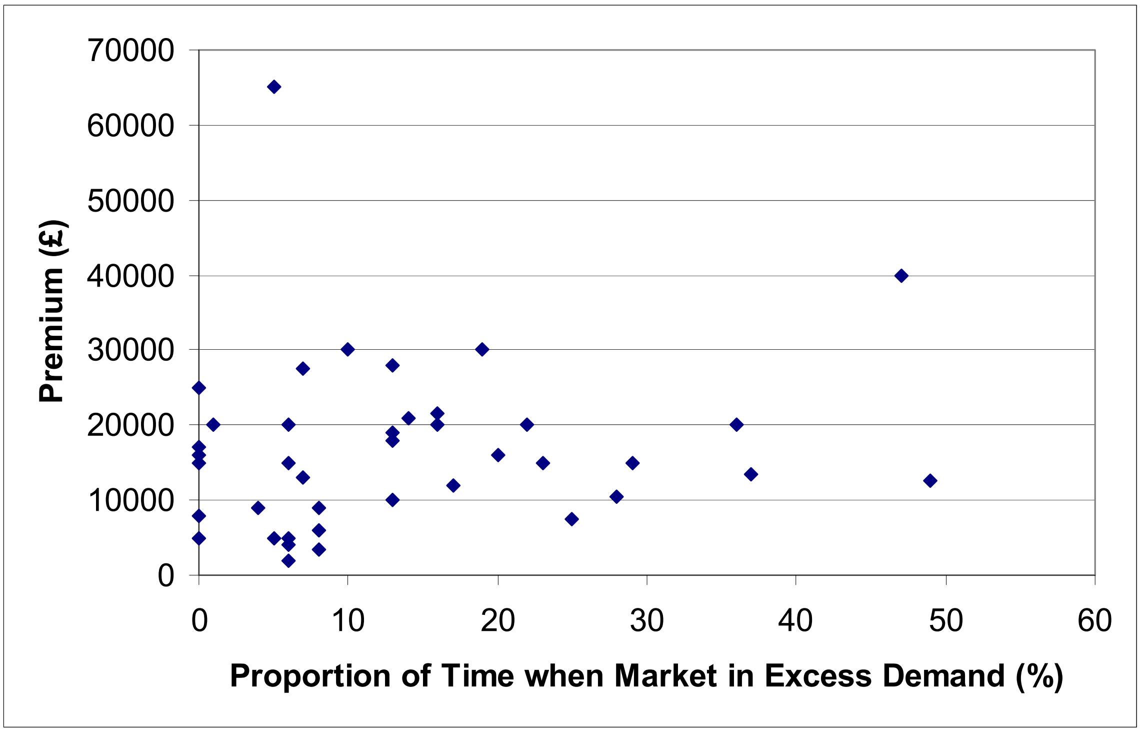



FIGURE 2: SCATTER PLOT OF TIME IN EXCESS DEMAND AGAINST LICENCE PREMIUM.



# 3 FORMAL DATA ANALYSIS

## 3.1 Model development and analysis

3.1.1 The data have been analysed in four tranches to explore the implications of entry control on premium expressed in different ways:

- explaining the variation in notional premium values (tranches 1 and 2)
- explaining the variation in monopoly rent (tranche 3), and
- explaining the variation in premium per trip (tranche 4).

## 3.2 Tranche 1 analysis. The relationship between premium per cab average fare and service performance

3.2.1 The first tranche of analysis seeks to model the relationship between the licence premium and those variables within a taxi market which would be expected to determine the premium levels. These include fare, two measures of demand and per capita taxi provision.

FARE AND AVERAGE PASSENGER DELAY

3.2.2 The first part of the analysis examines the relationship between the licence premium, fare and average passenger delay. Halcrow has developed six models to explain these relationships. The first is a simple version, with fare and a constant as independent variables and licence premiums as the dependent variable. The second and third build on the simple model by respectively incorporating passenger delay, and a dummy as well as passenger delay, as independent variables. The fourth model replaces fare with real fare, and the fifth model takes an alternative version of this approach by keeping nominal fare and incorporating an annual dummy. The sixth model is analogous to the third except that it removes the constant. The key results from all the models are illustrated in Table Two.

3.2.3 Each model has been expressed in four different functional forms,

- linear
- semi-logged (with the natural logs rather than the values of total revenue and revenue per cab)
- double logged (with natural logs of licence premiums as well as total revenue and revenue per cab), and
- exponential (with natural logs of licence premiums and values of the independent variables).

**Model 1**

3.2.4 This model has fare and a constant as independent variables and the licence premium as the dependent variable. In determining the model adequacy the convention is to look at some broad features of the results such as whether the signs of the estimated coefficients are in relation to their prior expectations, the model fit, given by the adjusted $R^2$ value and the significance of individual variables, given by p-values.

3.2.5 The estimated slope coefficient for fare indicates that it is positively correlated with the licence premium. This is intuitively correct as one would expect the resale value of a hackney licence to reflect some measure of drivers' income. The overall fit of the model however is poor which is indicated by an adjusted $R^2$ ranging between 0.05 and 0.1 depending on the functional form. In both the linear and semi logged model fare and the constant were found to be insignificant, this is indicated by p values above the critical level of 0.05. Fare and the constant are however significant in the double logged and exponential models.

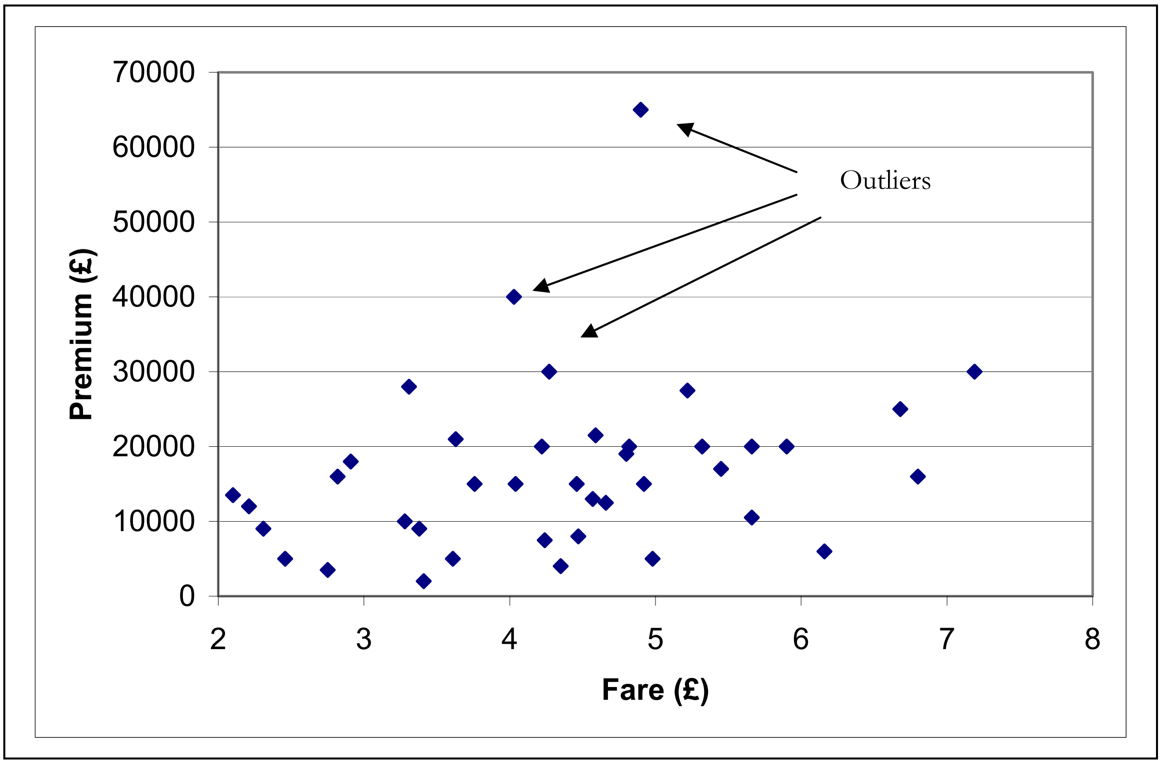
**Model 2**

3.2.6 Economic theory suggests that as well as fare the licence premium should be determined by some notion of supply of taxi trips, on this basis a second model was constructed which added passenger delay (as a proxy for supply) to the previous version.

3.2.7 Similarly to the previous model the signs of the estimated slope coefficients were both as economic theory would suggest; the model associates high fares and levels of demand with high premium values. Despite this the overall fit of the model was little better than the first one, with similar adjusted $R^2$ values, and all the independent variables were not consistently found to be significant under each functional form.

3.2.8 The failure to find an improved fit can largely be explained by the outliers within the data which differ from the underlying trend. The existence of such outliers can be seen when fare is plotted against licence premium in Figure 3. With the exception the three data points that are highlighted the data follows a shallow upward sloping pattern. The three points in question all have a higher licence premium than the fare would suggest, and correspond to the Calderdale, and two studies of Sunderland.

FIGURE 3: SCATTER PLOT OF FARE AND LICENCE PREMIUM



**Model 3**

3.2.9 Based on the previous analysis the expectation was that the poor fit of the first two models could largely be explained by licence premiums being higher in zoned authorities than in un–zoned authorities. To take this into account the third model added a dummy variable to represent the Calderdale authority, (in this model a dummy for all zoned authorities was found to be insignificant, although this was not the case in other models). The third model therefore has fare, passenger delay, a Calderdale dummy and a constant as the independent variables.

3.2.10 Inclusion of the dummy variable dramatically increased the fit of the model with adjusted $R^2$ values of 0.30 and 0.23 for the double log and exponential functional forms, and 0.53 and 0.50 for the linear and partial log versions. However while all the regressors are significant in the former two, they are not in the latter two versions, (which have the better model fit).

**Model 4**

3.2.11 The fourth model is analogous to the previous version except that the fare has been replaced with real fare. The expectation was that real fare would provide a more accurate explanation of the notional variation the licence premium.

3.2.12 The overall fit is of a similar magnitude to the previous model with adjusted $R^2$ values ranging between 0.52 and 0.22. However the p-values for real fare are above the critical value for each functional form and as a result it is not possible to conclude that there is a significant link between real fare and licence premium.

**Model 5**

3.2.13 The fifth model reinstated the nominal fare variable and incorporated year of study rather than expressly modelling real fare. This is in effect an alternative approach to the previous model.

3.2.14 The overall fit is very similar to the previous model with adjusted $R^2$ values ranging between 0.52 and 0.24. The year variable was not found to be significant with all the p-values falling above the critical limit.

**Model 6**

3.2.15 It is notable that for each of the previous models the slope coefficient is insignificant for both the linear and semi logged functional forms. As a result a sixth model has been constructed which is analogous to the third model, except that the constant has been removed and the origin used as the y-intercept.

3.2.16 Removing the constant has a positive effect on the linear form of the model. All the independent variables were found to be significant with p-values below the critical value of 0.05. This is an important finding as it means that the model concurs with economic theory by providing empirical backing to the link notional variation in the licence premium to the fare and the demand for travel by hackney carriages. This goes some way to fulfilling the aim of this tranche of the analysis, however the adjusted $R^2$ value of 0.53 is still relatively moderate and the model is sensitive to the presence of the extreme outlier, Calderdale. As a consequence it was necessary to expand the analysis to include other variables which could, determine the notional variation in licence premiums.

TABLE 2: LICENCE PREMIUM AGAINST FARE AND AVERAGE PASSENGER DELAY

| | | Linear | | Semi Log | | Double Log | | Exponential | |
|---|---|---|---|---|---|---|---|---|---|
| Model | Variables | Coefficient | P -Value | Coefficient | P-Value | Coefficient | P-Value | Coefficient | P- Value |
| 1 | Intercept | 5697.23 | 0.38 | 1456.37 | 0.86 | 8.38 | 0.00 | 8.66 | 0.00 |
| | Fare | 2533.60 | 0.80 | 10707.38 | 0.07 | 0.79 | 0.03 | 0.19 | 0.03 |
| | Fit, Adjusted $R^2$ | 0.05 | | 0.06 | | 0.10 | | 0.10 | |
| 2 | Intercept | 4505.29 | 0.50 | 1366.04 | 0.88 | 8.55 | 0.00 | 8.48 | 0.00 |
| | Fare | 2455.03 | 0.09 | 10759.72 | 0.08 | 0.69 | 0.06 | 0.18 | 0.33 |
| | Passenger delay | 1283.11 | 0.45 | -71.60 | 0.97 | 0.13 | 0.26 | 0.20 | 0.48 |
| | Fit, Adjusted $R^2$ | 0.04 | | 0.03 | | 0.10 | | 0.17 | |
| 3 | Intercept | 4157.60 | 0.38 | 5303.58 | 0.42 | 8.68 | 0.00 | 8.47 | 0.00 |
| | Fare | 1956.36 | 0.58 | 7318.21 | 0.10 | 0.57 | 0.09 | 0.16 | 0.04 |
| | Passenger delay | 2333.18 | 0.56 | 1607.14 | 0.26 | 0.19 | 0.09 | 0.23 | 0.01 |
| | Calderdale dummy | 50603.78 | 0.00 | 50013.78 | 0.00 | 1.73 | 0.01 | 1.75 | 0.01 |
| | Fit, Adjusted $R^2$ | 0.53 | | 0.50 | | 0.23 | | 0.30 | |
| 4 | Intercept | 2770.065 | 0.62 | 46765.15 | 0.00 | 11.67 | 0.00 | 20.17 | 0.00 |
| | Real Fare | 369041.70 | 0.07 | 8443.33 | 0.12 | 0.59 | 0.15 | 1.91 | 0.7 |
| | Passenger delay | 2425.81 | 0.05 | 1804.04 | 0.20 | 0.21 | 0.06 | 2.69 | 0.01 |
| | Calderdale dummy | 50933.21 | 0.00 | 50493.28 | 0.02 | 1.78 | 0.01 | 2.90 | 0.01 |
| | Fit, Adjusted $R^2$ | 0.52 | | 0.49 | | 0.22 | | 0.28 | |

| Model | Variables | Linear Coefficient | Linear P - Value | Semi-Log Coefficient | Semi-Log P – Value | Double-Log Coefficient | Double-Log P - Value | Exponential Coefficient | Exponential P - Value |
|---|---|---|---|---|---|---|---|---|---|
| 5 | Intercept | -10629.25 | 0.86 | -16243.46 | 0.80 | 3.17 | 0.51 | 0.89 | 0.38 |
| | Fare | 1650.03 | 0.30 | 5556.49 | 0.43 | 0.11 | 0.84 | 0.07 | 0.56 |
| | Passenger delay | 2254.64 | 0.08 | 1441.46 | 0.35 | 0.15 | 0.20 | 0.21 | 0.03 |
| | Calderdale dummy | 50391.14 | 0.00 | 49677.71 | 0.00 | 1.64 | 0.02 | 1.68 | 0.01 |
| | Year | 165.23 | 0.80 | 245.72 | 0.74 | 0.06 | 0.25 | 0.52 | 0.29 |
| | Fit, Adjusted $R^2$ | 0.52 | | 0.48 | | 0.24 | | 0.31 | |
| 6 | Fare | 2771.95 | 0.00 | 10908.69 | 0.00 | 6.35 | 0.00 | 1.83 | 0.00 |
| | Passenger delay | 2580.13 | 0.03 | 1273.20 | 0.35 | -0.36 | 0.26 | 0.74 | 0.05 |
| | Calderdale Dummy | 50691.03 | 0.00 | 49285.85 | 0.00 | 0.54 | 0.79 | 1.92 | 0.45 |
| | Fit, Adjusted $R^2$ Taken from model 3 (with constant.) | 0.53 | | 0.50 | | 0.23 | | 0.30 | |

**FARE AND INDEX OF SIGNIFICANT UNMET DEMAND, (ISUD)**

3.2.17 Similarly to the previous analysis this section examines the relationship between the licence premium, fare and the Index of Significant Unmet Demand. The aim is to improve on the fit of the previous model on the basis that ISUD is a more accurate reflection of the demand for taxi travel, (it should be noted that ISUD and passenger delay could not be included in the same model). The approach taken is identical to the previous section, constructing three versions of the original model by progressively adding ISUD, the Calderdale dummy and removing the constant. To avoid repetition the commentary is less detailed than for the previous analysis, concentrating on the preferred model. Each model has been expressed in linear and exponential functional form, it was not possible to estimate semi log or double log versions since some values of ISUD are zero. The key results from all the models are illustrated in Table Three. Similarly to the previous section Halcrow also examined models which included real fare and an annual dummy, as before neither were found to be significant consequently these models have not been presented.

TABLE 3: LICENCE PREMIUM AGAINST FARE AND 'ISUD' FACTOR

| | | **Linear** | | **Exponential** | |
|---|---|---|---|---|---|
| **Model** | **Variable** | **Coefficient** | **P -Value** | **Coefficient** | **P -Value** |
| 7 | Intercept | 4934.31 | 0.45 | 8.58 | 0.00 |
| | Fare | 2505.45 | 0.08 | 0.19 | 0.03 |
| | ISUD | 0.94 | 0.28 | 0.01 | 0.04 |
| | Fit, Adjusted $R^2$ | 0.06 | | 0.17 | |
| 8 | Intercept | 5324.12 | 0.25 | 8.59 | 0.00 |
| | Fare | 2072.05 | 0.04 | 0.18 | 0.03 |
| | ISUD | 1.213 | 0.05 | 0.01 | 0.02 |
| | Calderdale dummy | 49522.83 | 0.00 | 1.63 | 0.01 |
| | Fit, Adjusted $R^2$ | 0.53 | | 0.29 | |
| 9 | Fare | 3185.51 | 0.00 | 1.97 | 0.00 |
| | ISUD | 1.29 | 0.04 | 0.00 | 0.23 |
| | Calderdale dummy | 49390.99 | 0.00 | 1.42 | 0.59 |

| | | |
|---|---|---|
| Fit, Adjusted $R^2$ | 0.53 | 0.29 |

**Model 9**

3.2.18 The coefficients for all the variables are all positive, this is intuitively correct as it implies that unmet demand and the level of fare are both positively correlated with the licence premium, it also shows Calderdale to have a strong positive effect on the licence premium, this is known to be correct from the previous analysis. In the linear version the p-values for ISUD and fare are above the critical value, which indicates that both generate a significant effect on licence premium values.

3.2.19 The linear model exhibits a moderate level of accuracy with an adjusted $R^2$ figure of 0.53, this is exactly the same as the corresponding model in the previous section of analysis.

**FARE AND PER CAPITA PROVISION**

3.2.20 This section of analysis examined the relationship between the licence premium, fare and the population per Hackney, which is a proxy for the Per Capita Provision of taxis (PCP). This approach can be thought of an alternative way of incorporating demand into the model, the aim was therefore to improve fit of the previous two versions. The approach taken is identical to the previous two analyses, constructing three versions of the original model by progressively adding PCP, the Calderdale dummy and removing the constant. Each model has been expressed in four different functional forms, these are linear, semi log, double log and exponential. The key results from all the models are illustrated in Table 4.

3.2.21 For the most part the signs of the estimated coefficient are as economic theory would suggest, with licence premiums high licence premiums associated with large numbers of people per taxi. The other results however are less positive as there is no improvement of the overall fit of the models here compared to those in previous sections and often the p-value for PCP is above the critical value.

TABLE 4: LICENCE PREMIUM AGAINST FARE AND PER CAPITA PROVISION (PCP)

| | | Linear | | Semi Log | | Double Log | | Exponential | |
|---|---|---|---|---|---|---|---|---|---|
| Model | Variable | Coefficient | P -Value | Coefficient | P -Value | Coefficient | P -Value | Coefficient | P -Value |
| 10 | Intercept | 5041.61 | 0.56 | -975.20 | 0.97 | 8.61 | 0.00 | 8.62 | 0.00 |
| | Fare | 2592.18 | 0.10 | 10915.32 | 0.09 | 0.77 | 0.05 | 0.20 | 0.04 |
| | PCP | 0.25 | 0.91 | 295.97 | 0.93 | -0.03 | 0.89 | 0.01 | 0.90 |
| | Fit, Adjusted $R^2$ | 0.03 | | 0.03 | | 0.07 | | 0.07 | |
| 11 | Intercept | 10800.47 | 0.10 | 26164.29 | 0.22 | 9.47 | 0.00 | 8.80 | 0.00 |
| | Fare | 1702.55 | 0.13 | 6592.93 | 0.16 | 0.63 | 0.09 | 0.17 | 0.06 |
| | PCP | -1.72 | 0.30 | -2785.28 | 0.26 | -0.13 | 0.52 | -0.01 | 0.72 |
| | Calderdale dummy | 50168.62 | 0.00 | 50159.16 | 0.00 | 1.59 | 0.02 | 1.57 | 0.03 |
| | Fit, Adjusted $R^2$ | 0.49 | | 0.50 | | 0.18 | | 0.17 | |
| 12 | Fare | 3407.45 | 0.00 | 10238.05 | 0.01 | 1.95 | 0.00 | 1.56 | 0.00 |
| | PCP | 0.15 | 0.90 | 115.21 | 0.87 | 0.92 | 0.00 | 0.01 | 0.00 |
| | Calderdale dummy | 47928.81 | 0.00 | 47827.54 | 0.00 | 0.75 | 0.41 | -0.26 | 0.90 |
| | Fit, Adjusted $R^2$ | 0.49 | | 0.50 | | 0.18 | | 0.17 | |

**MODEL CHOICE, DIAGNOSTIC STATISTICS AND INTERPRETATION**

3.2.22 Two models have the highest adjusted $R^2$ values, i.e. the best explanatory power. These are the linear version of model six presented in Table 2, which has independent variables of fare, passenger delay and the Calderdale dummy, and the linear version of Model 9 in Table 3 which includes fare, ISUD and the Calderdale dummy. In both models all the variables are significant with p-values below the critical values and each show the least evidence of any mis-specification.

3.2.23 Diagnostic tests on the chosen models uncovered the econometric problems of first order autocorrelation and heteroskedasticity. Such occurrences however are common in this type of work and there are several types of corrective procedures available[1]. A side effect of applying these corrections is that it is necessary to adjust the coefficients as well as the standard errors, the revised figures are presented in Table 3. Neither the coefficients or the p-values have altered significantly, in particular the p-values for passenger delay and ISUD have remained below the critical value.

3.2.24 After the corrections have been applied passenger delay has a slightly lower p-value than ISUD and therefore provides a marginally better explanation of the licence premium. Consequently model 6 is preferred to model 9.

**TABLE 5: SUMMARY OF THE REVISED PREFERRED MODELS FROM THE TRANCHE 1 ANALYSIS**

| Model | Variable | Revised Coefficient | Revised P-Value |
|---|---|---|---|
| Table 2, Model 6 (Linear) | Fare | 2824.12 | 0.00 |
| | Passenger Delay | 2260.88 | 0.03 |
| | Calderdale Dummy | 57330.11 | 0.00 |
| | Fit, Adjusted $R^2$ | 0.53 | |
| Table 3, Model 9 | Fare | 3185.40 | 0.00 |
| (Linear) | ISUD | 1.29 | 0.04 |
| | Calderdale Dummy | 49391.56 | 0.00 |
| | Fit, Adjusted $R^2$ | 0.53 | |

1 The Beach-Mackinnon method was used to correct for autocorrelation and the 'White' correction was used to estimate consistent standard errors in the presence of heteroskedasticity.

## 3.3 Conclusions

3.3.1 The following conclusions can be drawn from the initial analyses:

- the positive correlation between fare and the premium as well as demand and the premium provide evidence of consumer detriment, to the extent that the unwanted scenario of high fares and high levels of (excess) demand is synonymous with high premiums,
- the models are weakened when an allowance is made for time trends, and
- the relatively low level of explanatory power suggests that further exploration of the data set could provide some more insight.

## 3.4 Tranche 2 analysis: relationship between premium per cab and revenue

3.4.1 The second tranche of analysis seeks to explore the relationship between the premium value and measures of supply, price and revenue. Halcrow has developed four models to explain the values attributed to taxi licence premiums. The first is a simple model with the total revenue, revenue per cab and a constant as independent variables. The second and third models add dummy variables to the first one in order to explain individual departures from the underlying trend, and as a consequence improve the accuracy of the model. The fourth model is analogous to the third except that it replaces real revenue per cab with unemployment. (The three models are detailed in Table 6). In each model revenue has been deflated by the RPI. This proved preferable to including a time trend separately.

### Model 13

3.4.2 This model has total revenue, revenue per cab and a constant as the independent variables (both have been calculated on the basis that all cab journeys are made by a single person over a distance of three miles). The sign of the coefficients for revenue per cab and total revenue indicate that both are positively correlated with the value of licence premiums. This is intuitively correct since revenue per cab is closely related to taxi drivers' income, and total revenue is linked to the size of the market. The p-values for total revenue and revenue per cab are below the critical value of 0.05, which means that both have a significant effect on licence values.

3.4.3 The accuracy of the model itself however is poor. This is indicated by a low adjusted $R^2$ statistic ranging between 0.261 and 0.315.

**Model 14**

3.4.4 The inaccuracy of Model 13 can be largely explained by a number of outliers or observations, which differ from the underlying trend. The existence of these outliers has been illustrated previously in Figure 3. The outliers in question correspond to Calderdale, and two studies of Sunderland, these are all zoned licensing authorities. A dummy variable has therefore been added to the previous model to allow for zoning to have an effect on licence premiums.

3.4.5 It is clear from Table 6 that incorporating the zone dummy variable has increased the accuracy of the model, with $R^2$ statistics ranging from 0.371 to 0.647 depending on the functional form. The zone dummy itself is significant with a low p-value, and a large positive coefficient indicates that zoned authorities will place a much greater value on licence premiums (around £26,000).

3.4.6 The revenue per cab and total revenue variables have similar coefficients and t-statistics to those of the first model, this indicates that both are robust and confirms that the previous inaccuracy was largely caused by the existence of the zoned authorities.

**Model 15**

3.4.7 Despite the relative success of the second model, the £65,000 licence premium in Calderdale is extremely high and cannot be justified by the model even with the existence of the zone dummy variable. As a consequence the model will not be as accurate as it could be. Model 15 has therefore added a dummy variable to allow for the revenue generated per cab in Calderdale to have a different effect on the values of the licence premium than that of the rest of the sample.

3.4.8 Inclusion of the Calderdale revenue per cab dummy has produced a small improvement in accuracy with adjusted $R^2$ statistics as high as 0.710 depending on the functional form. The p-values show that the Calderdale dummy is significant.

3.4.9 The total revenue variable has very similar coefficients and p-values to those of the previous models this indicates that it is a very robust variable. Revenue per cab is significant in three of the four functional forms and although the coefficients have dropped, the sum of this and the coefficients for Calderdale revenue per cab is approximately equal to those for revenue per cab in the first model, this is intuitively correct. The zone dummies have remained significant, however the coefficients have reduced.

**Model 16**

3.4.10 Model 16 builds on the previous version replacing real revenue per cab, (which has a marginal p-value), with unemployment. The coefficients for real revenue, zone dummy and Calderdale dummy are similar to the previous model and the p-values are all below the critical level. The coefficient for unemployment is negative, this means higher unemployment is associated with lower licence premiums. This can be justified in terms of an increase in unemployment reducing individuals pre-disposition to leisure travel.

3.4.11 The model is slightly more accurate than previous version with $R^2$ statistics ranging between 0.389 and 0.725 depending on the functional form.

TABLE 6: SUMMARY OF THE INITIAL MODELS.

| Model | Variables | Linear | | Semi Logged | | Double Logged | | Exponential | |
|---|---|---|---|---|---|---|---|---|---|
| | | Coefficient | p-value | Coefficient | p-value | Coefficient | p-value | Coefficient | p-value |
| 13 | Constant | Removed - Insignificant | - | -63289.011 | 0.007 | 3.812 | 0.005 | 8.629 | 0.000 |
| | Real Total Rev | 0.124 | 0.200 | 3918.524 | 0.024 | 0.321 | 0.002 | 0.009 | 0.008 |
| | Real Rev per Cab | 87.2 | 0.000 | 8579.281 | 0.013 | 0.527 | 0.008 | 0.004 | 0.002 |
| | R-Bar Sq | 0.289 | | 0.261 | | 0.315 | | 0.287 | |
| 14 | Constant | Removed - Insignificant | | -51399.877 | 0.001 | 4.188 | 0.002 | 8.675 | 0.000 |
| | Real Total Rev | 0.146 | 0.00 | 3711.802 | 0.001 | 0.314 | 0.003 | 0.00001 | 0.004 |
| | Real Rev per Cab | 69.041 | 0.00 | 6142.898 | 0.016 | 0.450 | 0.011 | 0.003 | 0.010 |
| | Zone | 26021.491 | 0.00 | 27434.213 | 0.013 | 0.866 | 0.000 | 0.856 | 0.020 |
| | R-Bar Sq | 0.647 | | 0.626 | | 0.408 | | 0.371 | |

| | | Linear | | Semi Logged | | Double Logged | | Exponential | |
|---|---|---|---|---|---|---|---|---|---|
| Model | Variables | Coefficient | p-value | Coefficient | p-value | Coefficient | p-value | Coefficient | p-value |
| 15 | Constant | 6464.690 | 0.027 | -40961.094 | 0.006 | 4.369 | 0.001 | 8.677 | 0.000 |
| | Real Total Rev | 0.126 | 0.002 | 3884.954 | 0.001 | 0.317 | 0.001 | 0.00001 | 0.004 |
| | Real Rev per Cab | 32.513 | 0.055 | 3659.767 | 0.099 | 0.407 | 0.040 | 0.003 | 0.029 |
| | Zone | 20682.577 | 0.000 | 19082.906 | 0.000 | 0.722 | 0.080 | 0.853 | 0.047 |
| | C-dale rev/cab dummy | 52.251 | 0.021 | 4522.513 | 0.002 | 0.008 | 0.513 | 0.00004 | 0.987 |
| | R-Bar Sq | 0.689 | | 0.710 | | 0.398 | | 0.353 | |
| 16 | Constant | 18244.008 | 0.000 | Removed - Insignificant | | 7.691 | 0.000 | 9.826 | 0.000 |
| | Real Total Revenue | 0.09 | 0.027 | 2162.149 | 0.000 | 0.232 | 0.022 | 0.000006 | 0.057 |
| | Zone | 22260.588 | 0.000 | 21449.162 | 0.000 | 0.894 | 0.033 | 0.998 | 0.020 |
| | C-dale rev/cab dummy | 65.687 | 0.001 | 65.139 | 0.001 | 0.001 | 0.363 | 0.001 | 0.417 |
| | Unemployment | -1221.924 | 0.014 | -1271.899 | 0.002 | -0.103 | 0.021 | -0.116 | 0.009 |
| | R-Bar Sq | 0.709 | | 0.725 | | 0.418 | | 0.389 | |

**Model choice, diagnostic statistics and interpretation**

3.4.12 The linear version of Model 16 presented in Table 6 provides the best explanation of how licence premiums are valued. This model has the second highest level of overall accuracy with an adjusted $R^2$ of 0.709, and is preferable to the semi-logged version, which despite having a higher adjusted $R^2$, cannot include the constant as the p-value is above the critical level.

3.4.13 Diagnostic tests on the chosen model uncovered the econometric problems of first order autocorrelation and heteroskedasticity. The revised figures are presented in Table 7. Neither the coefficients nor the p-values have altered significantly, in particular the p-value for revenue has remained below the critical value.

TABLE 7: SUMMARY OF THE REVISED PREFERRED REVENUE MODEL.

| | **Revised Model 16 - Linear** | **Revised p-values** |
|---|---|---|
| Constant | 16997.786 | 0.000 |
| Real Total Rev | 0.092 | 0.031 |
| Unemployment | -1118.767 | 0.013 |
| Zone Dummy | 22258.588 | 0.000 |
| C-dale rev/cab dummy | 61.819 | 0.002 |
| R-Bar Sq | 0.709 | |

## 3.5 Tranche 3 analysis: the relationship between monopoly rent measures of premium and service characteristics

3.5.1 The premium value can be considered to represent a measure of monopoly profit, or monopoly rent in a local taxi market. This rent accrues directly from the existence of a limitation policy. The purpose of this tranche of analysis is to demonstrate that the level of monopoly rent can be successfully explained via measures of market revenue. The analysis is at market level and therefore reintroduces supply considerations.

3.5.2 Figure 4 shows a strong positive relationship between the total value of premium in a market (premium * cabs) and the total annual revenue (annual demand * average fare), and Table 8 records the regression of this relationship. The model produces an insignificant intercept term so the model was re-run through the origin. The resulting coefficient value and p-value is included in Table 6 in parentheses. The model suggests that as total revenue increases by £1 per annum, the value of the monopoly rent increases by almost the same amount. Analysis at a market level appears to allow more robust estimation than at the premium per hackney level.

FIGURE 4: PREMIUM*CABS AGAINST TOTAL ANNUAL REVENUE.

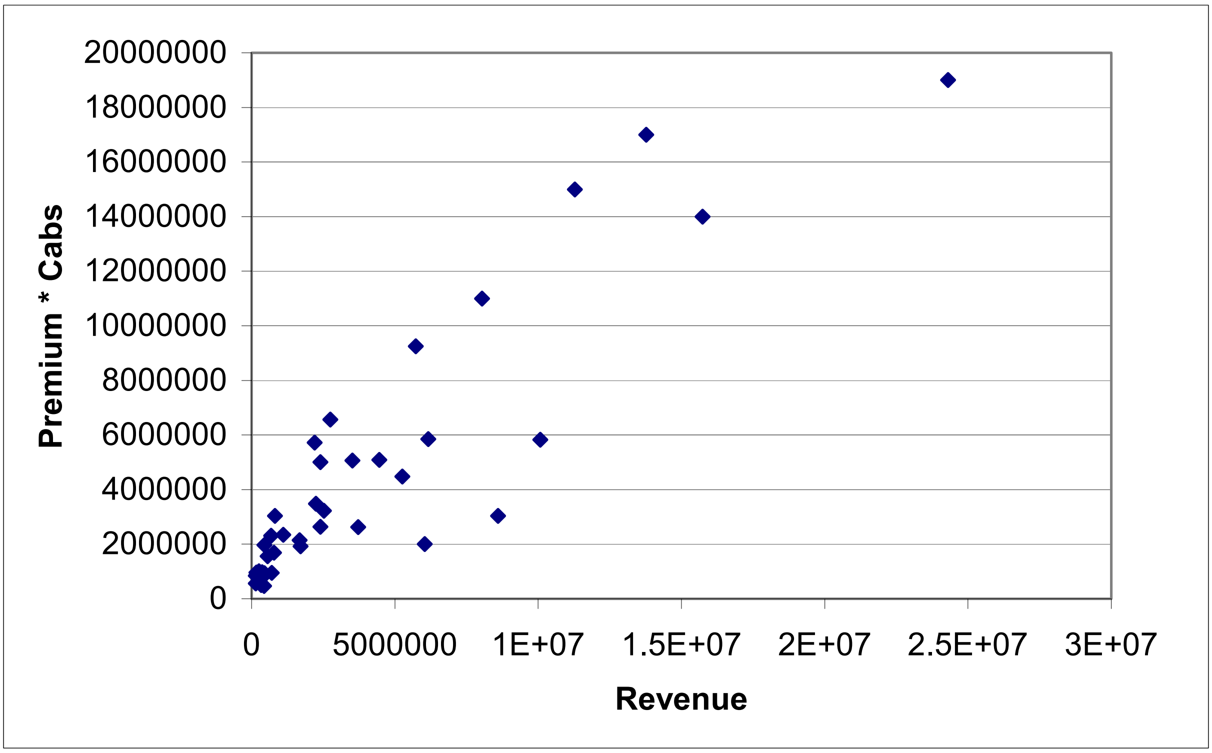


TABLE 8: MODEL 17 - SIMPLE MONOPOLY RENT MODEL[2]

| Variable | Coefficient | p-value |
|---|---|---|
| Constant | -425,823 | 0.299 |
| Total Revenue | 0.980 | 0.00 |

Dependent variable (Mrent) = premium*cabs

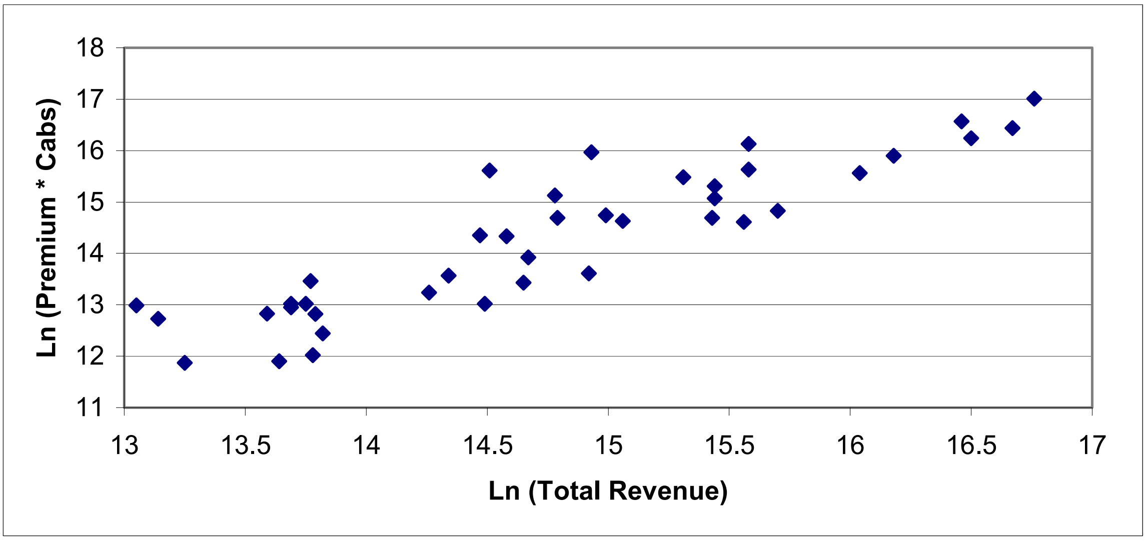
N=40

R-Bar Sq = 0.82

Values in parentheses are for regression through the origin

3.5.3 A strong relationship has been also been derived between the natural log of Mrent and market size (population) and total annual revenue (demand multiplied by average fare), as illustrated in Figure 5 below. Table 9 gives the preferred specification.

FIGURE 5: LOG OF PREMIUM*CABS AGAINST LOG OF TOTAL REVENUE




2 Corrected for Heteroskedasticity and Autocorrelation.

## FIGURE 6: LOG OF PREMIUM*CABS AGAINST LOG OF POPULATION

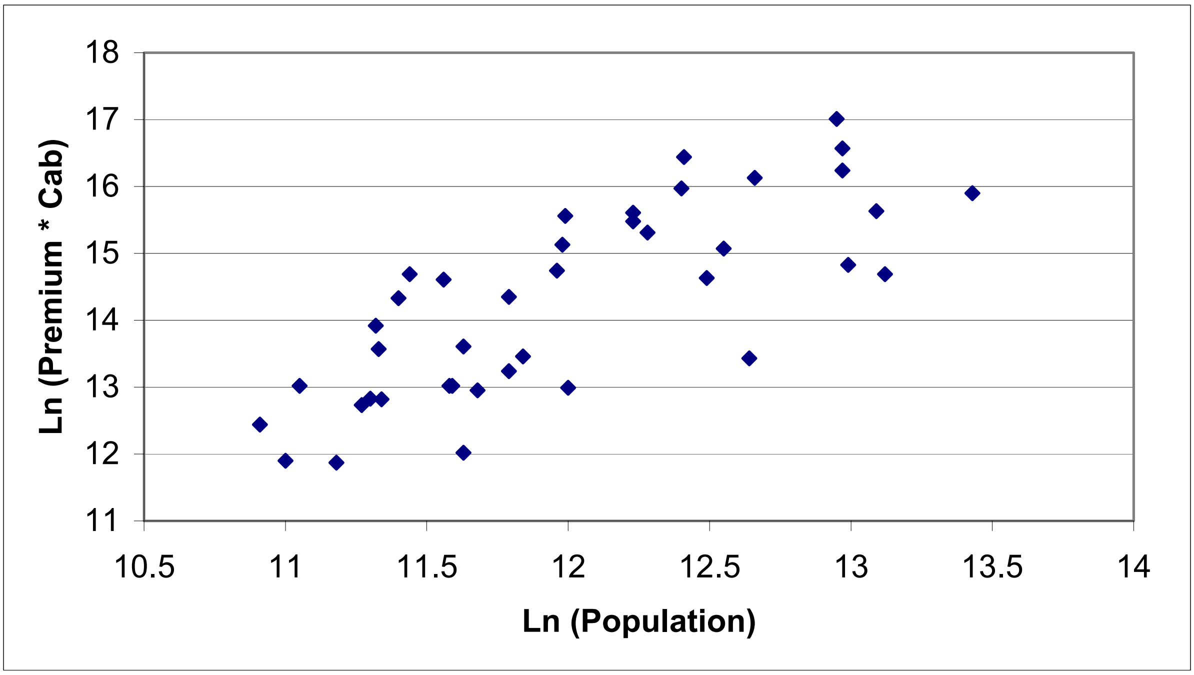



TABLE 9: MODEL 18- MONOPOLY RENT MODEL A [3]

| Variable | Coefficient | p-value |
|---|---|---|
| Constant | -6.335 | 0.000 |
| Zone | 0.957 | 0.006 |
| LN(Population) | 0.455 | 0029 |
| Ln(Annual revenue) | 1.021 | 0.000 |

Dependent variable Ln(Mrent)

N=40

R-Bar Sq = 0.85

3 Corrected for Heteroskedasticity and Autocorrelation.

3.5.4 With monetary values appearing on both sides of the equation the influence of a time trend is eliminated and it is not possible to successfully include any of the temporal variables. The model is also much less sensitive to the inclusion of the zoned authorities than the initial specification and achieves a much higher degree of fit to the data.

3.5.5 The model yields elasticities directly. A one per cent increase in district population is associated with a 0.46 per cent increase in monopoly rent. This confirms the widely held view that hackneys prosper in the more densely populated urban areas. A one per cent increase in total revenue is consistent with a one per cent increase in the monopoly rent (illustrated in Model 18 above).

3.5.6 Since all districts in the dataset are restricted, it is not possible to derive a meaningful estimate of the reduction in fare that is consistent with zero monopoly rent. There is high collinearity between LPOP and LTOTREV (correlation = 0.77). Despite this both variables are significant.

TABLE 10: MODEL 19- MONOPOLY RENT MODEL B [4]

| **Variable** | **Coefficient** | **p-value** |
|---|---|---|
| Constant | -4.058 | 0.004 |
| Zone | 0.830 | 0.032 |
| Passenger delay | 0.201 | 0.036 |
| LN(Annual revenue) | 0.948 | 0.000 |

Dependent variable Ln(Mrent)

N=40

R-Bar Sq = 0.85

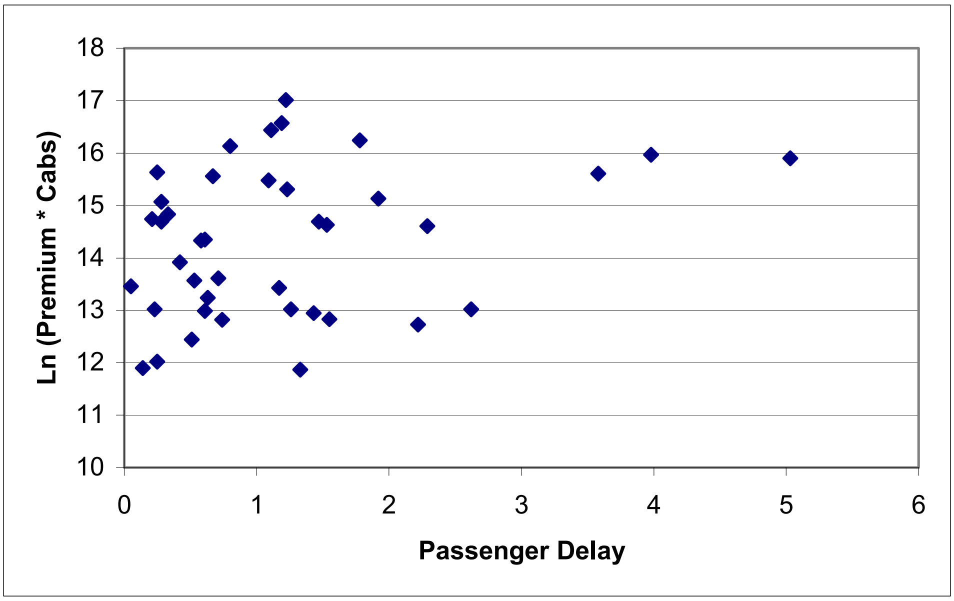
4 Corrected for Heteroskedasticity and Autocorrelation.

3.5.7 This model is similar to the previous version as it is not possible to successfully include any of the temporal variables, and the model is less sensitive to the inclusion of the zoned authorities than earlier specifications as well as achieving a much higher degree of fit.

3.5.8 Similarly to the last model this version yields the elasticity of monopoly rent to total revenue directly, as before a one per cent increase in total revenue is consistent with a one per cent increase in the monopoly rent. The model also implies that the elasticity of passenger delay to monopoly rent is 0.24, in other words a one per cent increase in passenger delay is associated with a quarter percent increase in monopoly rent.

3.5.9 The major drawback with this model is that, despite the good overall level of fit, passenger delay is not as good an explanatory variable as the population variable in the previous model. This is apparent in the relative p-values as well as the lack of correlation in the plot of passenger delay against the log of premium * cabs, illustrated below in Figure 7

FIGURE 7:    LOG OF PREMIUM*CABS AGAINST PASSENGER DELAY




## 3.6 Tranche 4 analysis: premium per trip

3.6.1 In a further attempt to explore the effect of supply changes on licence premia in a more successful way than in the initial analysis, a specification using the premium value per annual trip has been tried. The per trip premium would be expected to be negatively related to supply. This is indeed the case (see Figure 8). Two specifications of the model produce similarly performing models and both are presented in Table below (see Figure 9 and 10 and Table 11).

FIGURE 8: PREMIUM PER TRIP AGAINST CAB SUPPLY

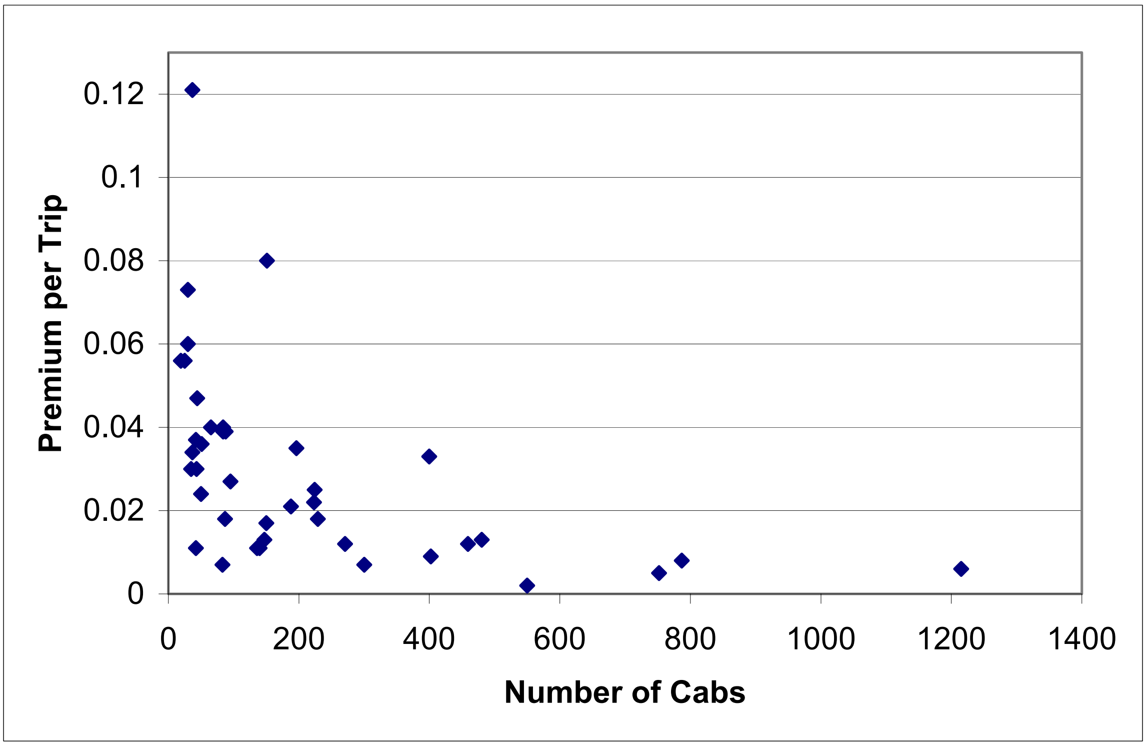


FIGURE 9: LOG OF PREMIUM PER ANNUAL TRIP AGAINST LOG OF CABS

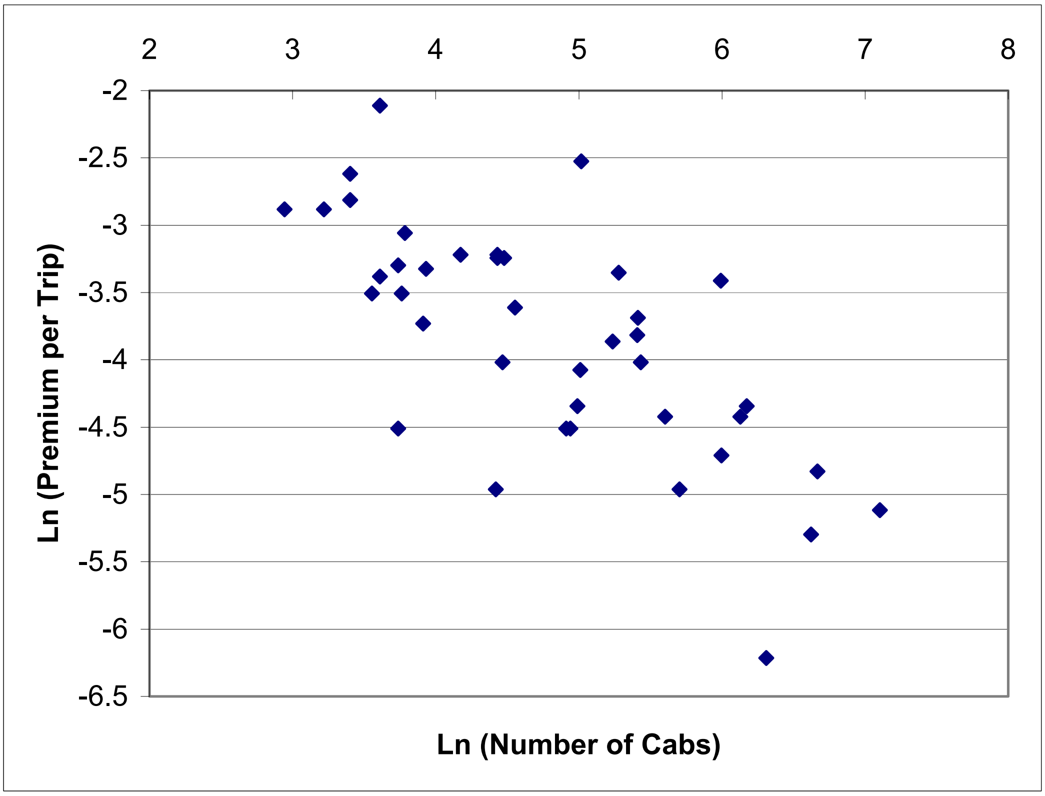


FIGURE 10: PREMIUM PER ANNUAL TRIP AGAINST LOG OF CABS

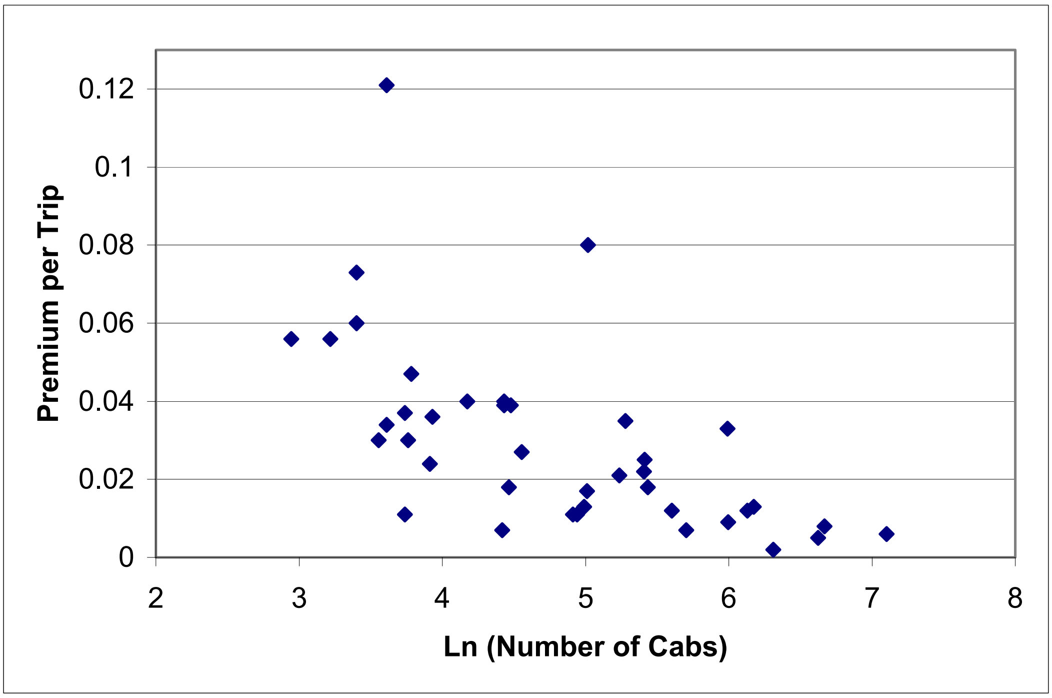



TABLE 11: MODEL OF PREMIUM PER ANNUAL TRIP[5]

| Model 20 Semi Log | | | Model 21– Double Log | |
|---|---|---|---|---|
| **Variable** | **Coefficient** | **P -value** | **Coefficient** | **p-value** |
| Constant | 0.083 | 0.000 | -1.513 | 0.000 |
| Zone | 0.047 | 0.000 | 1.061 | 0.000 |
| Ln(Cabs) | -0.015 | 0.000 | -0.616 | 0.000 |
| D96onwards | 0.018 | 0.000 | 0.697 | 0.000 |
| Dependent variable Premtp | | | Dependent variable = | |
| N=40 | | | LnPremtp = | |
| R-Bar Sq = 0.71 | | | Ln(Premium/annual trips) | |
| | | | N=40 | |
| | | | R-Bar sq=0.69 | |

5 Corrected for Heteroskedasticity and Autocorrelation.

3.6.2 The models offer a semi-log and constant elasticity formulation respectively, both explaining approximately 70 per cent of the variation in the data. The inclusion of monetary values on one side of the equation only means that the time trend is required. The model was found to perform marginally better with a dummy variable than a continuous variable. D96onwards takes a value of one for years 1996 and after and zero otherwise. Model 21 yields elasticities directly. It implies that a one per cent increase in supply is associated with a 0.62 per cent reduction in the premium per trip in restricted markets. Despite the market expanding, the scarcity value of the licence remains intact for as long a limit is retained. Halcrow's experience is that even where an authority establishes a policy of issuing 40 or 50 licences per year over, say, five years, the premium value is resilient (e.g. Leeds and Manchester).

3.6.3 Model 20 implies an elasticity of premium per annual trips to supply of –0.58 at current average levels of premium per annual trip (2.4 pence). The elasticity increases as the premium falls, rising to -1.16 at half the average value of premium per annual trip.

# 4 SUMMARY AND CONCLUSION

## 4.1 Overview

4.1.1 The results of 40 demand and regulation studies conducted on behalf of local authorities with quantity controlled hackney carriage markets have provided the database for this investigation. Each study provided an estimate of the resale value (premium) of a hackney vehicle licence plus measures of demand and passenger delay from direct observation of market activity. The studies were carried out between 1992 and 2002.

## 4.2 Stages of analysis

4.2.1 The analysis of the hackney carriage licence premium has been conducted at 3 levels to fully explore the data,

- the first and second tranches of analysis examined the relationship between the licence premium per cab and direct measures of travel cost, revenue and exogenous variables,
- the third tranche of analysis looked at the relationship between the total value of premia in the market (premium * cabs), aggregate measures of revenue and exogenous variables, and
- the fourth tranche of analysis examined the relationship between premium per trip and levels of supply.

## 4.3 Findings

4.3.1 Evidence was found linking the licence premium per cab with direct measures of travel cost, namely,

- authorities with higher fares have higher licence premiums. If fares are £1 higher the premium is around £3,000 higher,
- authorities with higher passenger delay have higher licence premiums. If the average delay to passengers is one minute higher the premium is around £2,000 higher.

4.3.2 Evidence was found linking the licence premium per cab with aggregate measures of revenue and exogenous economic variables, in particular,

- higher unemployment is associated with lower premium values. If unemployment is one per cent higher the premium is around £1,000 lower,

- higher total annual revenue is associated with higher premiums. If total annual revenue is £1 higher the licence premium is nine pence higher.

- licence premiums can be shown to be positively related to the presence of zoning with premium values in the principal zones being some £22,000 higher than in non-zoned authorities.

4.3.3 The econometric models used to derive these relationships however are not particularly robust. As a result two further sets of analysis were conducted using different dependent variables. The most robust of these used premium *cabs, i.e. the value of premiums for the entire market, and explained 85 per cent of the variation in the data. The preferred explanatory variables are total revenue, population and the presence of zoning. The keys results of this model are:

- in restricted hackney markets there is unitary elasticity between annual revenue and the value of the premium,

- premia are higher in authorities with larger populations and the zoning of licences.

## 4.4 Conclusion

4.4.1 The data analysis has demonstrated the evidence of a clear relationship between premium values and measures of travel cost. This is particularly strong for measures of price (revenue). Evidence of a link between premium values and passenger delay has also been found, but these models are generally inferior on statistical grounds to those that include exogenous variables such as population and unemployment instead of passenger delay.

# 5 APPENDIX

## 5.1 Autocorrelation

5.1.1 Autocorrelation refers to situations where observations on the dependent variable are not independently drawn, such that the standard errors from one time period to the next are correlated. This problem occurs predominantly in time series data and causes the model estimators to be unreliable. There are in existence a number of diagnostic tests, as well as several correction procedures.

## 5.2 Durbin-Watson test for autocorrelation

5.2.1 The standard test for autocorrelation is the Durbin-Watson (DW) test, and was the version used during this study. There is a statistical table related directly to this test which reports upper and lower critical values, based on the degrees of freedom within the model. Positive autocorrelation exists if the test statistic falls between these values.

## 5.3 Heteroskedasticity

5.3.1 In regression analysis, heteroskedasticity refers to situations in which the conditional variance of the independent variable given the dependent variables is not constant across the values of the dependent variables. This is a common problem with cross-sectional data and invalidates the standard errors and the results of the model tests. There are a number of tests for heteroskedasticity as well as several available correction procedures.

## 5.4 White test for heteroskedasticity

5.4.1 The White test was used to check for Heteroskedasticity in this study. There are a number of stages in order to apply the test:

- Run a regression and save the error terms (residuals).
- Square the residuals.
- Run separate auxiliary regressions between error terms and each of the independent variables.

- Check the T-statistics that are reported. Heteroskedasticity exists within the model if any of these exceed the critical value of 2.

## 5.5 Weighted least squares corrections

5.5.1 Both Heteoskedasticity and Autocorrelation can be corrected using a technique called weighted least squares. There are several different ways to apply this method, the version used to correct the standard errors in the presence of the former problem is known as the White correction, and the type used to correct the latter problem is called the Beach-Mackinnon Correction.