Cost Analysis for Car-sharing in Stockholm

Author:**Xu Zhang**
Supervisor:**Mikael Möller**

June, 2008.

Abstract

In the modern world cars have become an indispensable part of daily life. With the number of cars increasing, more and more negative impacts have emerged, such as energy waste, environment pollution, traffic congestion and so on. As an alternative transportation paradigm, car-sharing has become increasingly popular, what this mean is that many people from a carpool in order to reduce the cost of car use. This thesis use the method of GLMM to deal with the data about car-sharing in Stockholm, which provided by Stockholm Bolpool, to form a model where the dependent variable depends on a number of factors. Several measures of effectiveness have been calculated to characterize the overall model performance.

key word

Car-sharing, GLMM, Newman-Keuls test

Contents

1 Introduction **4**

2 Data description **6**

3 Statistic method and analyze **9**

3.1 Analyze the Cost of Members 9

3.1.1 Theory and method 9

3.1.2 Model application 10

3.2 Analyze the Cost of Bilpool 13

3.2.1 Theory about Newman-Keuls Test 13

3.2.2 The Newman-Keuls Test 14

4 Conclusion **15**

1 Introduction

Since Henry Ford invented the automobiles in 1896, the world has undergone enormous changes and the cars take human being into the modern civilization community. However, 100 years after the invention of cars, the developed countries have deeply felt that the car is not only "happiness" but also "bane". Environment pollution, traffic congestion, excessive taxes and so on are all hindrance in automobile consumption. Auto industry can not be negative or develop blindly, this is a dilemma of choice. For consumers, the travel convenience and benefits are the eternal pursuits, thus, a new consumption pattern – car-sharing came into life.

The so-called car-sharing is, in essence a member of the car rental business, and each user pay a certain amount of membership fees to these companies and are then entitled to get a car when needed. No matter when and where, when you need. Just a phone call away, you are able to enjoy the conveniences through the Car-sharing service network.

Car-sharing can reduce the volume of cars and be conducive to environmental protection. It also ease the traffic congestion situation. Car-sharing is attractive because of its low cost. Users are charged according to the actual use of time and should return the cars in time, the cost is lower than ordinary car rental. Usually, if a car is rented three times a month, its cost is just about 40 dollars, and it includes gasoline, car maintenance costs, insurance premiums and so on.

What are the benefits of choosing "Carsharing "? First answer is oil resources: if all countries have as much cars per capita as in United States then the world's known oil resources will be exhausted in the coming 10 years. In United States, 6% of the oil consumption is used for transport, the growth rate of the number of cars and their fuel consumption is twice the population growth.

Second answer is environmental pollution. From the time the car is put into use until it is abandoned, it has always been a source of pollution on the environment. A report of United Nations Human Action Fund Organization warned that because of rapidly increasing number of cars, by 2025, the developing countries annually emission 15.06 billion tons of carbon dioxide to the atmosphere, if 1/4 of the world population all have their own cars, the ozone layer will be like a cheese full of hollow.

Last but not least, traffic congestion is another important aspect. According to a survey, because of city traffic blocking in Paris, people will waste 300,000 working days daily, according to new statistics from the related departments of EU (European Union), the European Union have a loss of 2% GDP because of traffic jam each year; If the EU countries put all the automobiles on the highway into a platoon, they even up to 40,000 km. When at rush hour, the peak speed will be not more than 10 km per hour in one third of EU's cities.

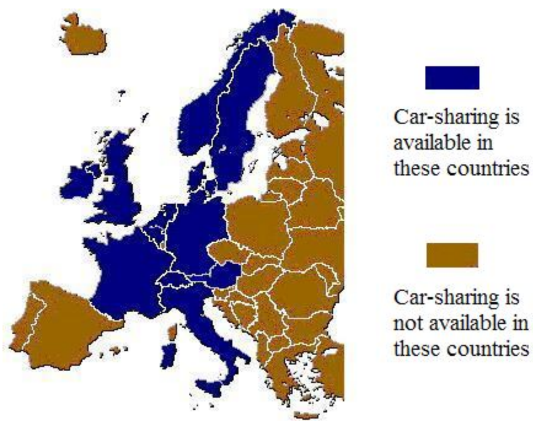

Figure 1: Car-sharing in Europe

Energy, pollution and traffic jam are the three major problems of car growth. The basic way to solve this problem is back to the old method, namely, restrict the growth of private cars. For this purpose, lots of major cities in developed countries take many approaches, such as raising parking fee, increasing road tolls and these methods is spreading through whole Europe, promoting consumers to change from "one person, one car" into "many persons, one car".

The rest of this paper is organized as follows. In section 2, we briefly discuss the explanatory variables which will be used in this paper. In section 3, we discuss a generalized linear model for car-sharing and show the relationship between depend variables and response variable. Through the estimates of the coefficients which were estimated, study the impact of each factor for its effect on customers cost and then analyze company's cost with Newman-Keuls test. Section 4 summarizes the paper and discusses potential extensions.

2 Data description

This paper uses data that were provided by the association Stockholm Bilpool for the year 2007[1]. Part of these original data is displayed in Appendix, which includes the membership number, the using duration, time duration, driving mile, type of car used, location of members, cars' plate *NO*. etc. It is recorded from Jan to Dec in 2007. Since the original data contains some information which is out of concern for this paper, the first step is to filter the irrelevant data and sort out the relevant ones. Moreover, some of the variables I used in the model may be calculated from the original data directly.

There are some problems with the raw data which hence needs to be adjusted. There are some variables in raw data, which are not related to the cost of members, such as reservation-start and reservation-stop, thus they will be erased. Furthermore, there are some member' addresses which were not found on the map and some cars locations were not contained in the Charges fee table (Appendix B), these records will be erased as well. Moreover some members' using durations were equal to 0 (maybe caused by record error), this does not match with the actual situation, so they were taken away too. Some components of member's cost were not given in the original data, such as oil prices.

This paper mainly focus on how to reduce the costs for users and companies from two different views of point. From the users' perspectives, we know which factor that will have a significant effect to cost with a GLMM model. In this paper we select the cost of using cars as response variable. As explanatory variables we use distance (the distance between the member's address and car's location), booking time (also seen as a dummy variables), car's size (divide into Large, Medium, Small), and length of trip (include long-trip and short-trip) is another important independent variable as well, I will tell the reason why I choose these variables in the following paragraph. While in company's perspective, according to frequency of cars use for each month, corresponding adjusting the number of cars owned by the company's is a feasible way to reduce its costs. Here is a figure about using frequency, which clearly revealed the changes trends of using frequency in 2007.

[1]Information about Stockholm Bilpool may be found from the website of http://www.stockholmsbilpool.nu/start.shtml

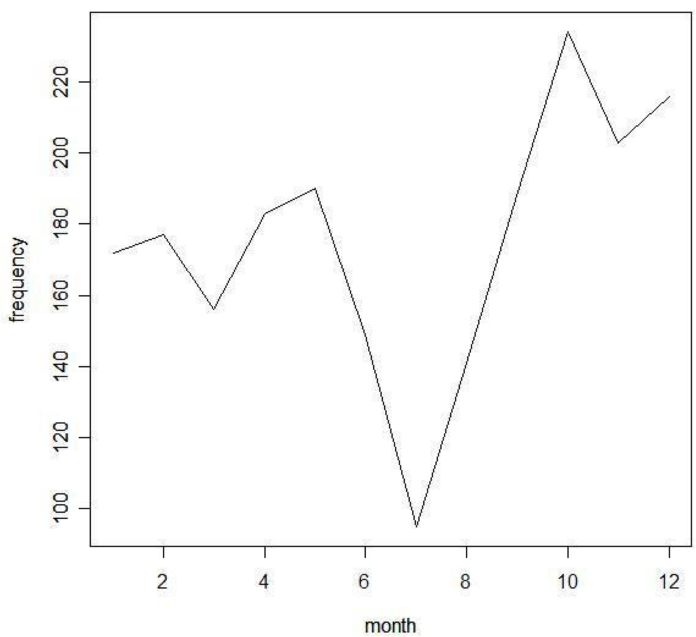

Figure 2: Using frequency in 2007

The variable distance measures the distance between the place where a user live and the nearest location to one of Bilpools cars. As a variable[2] (Appendix A) it directly reflects the convenience of car-sharing services. the shorter the distance is, the more convenient the service is. Maybe users will have more booking frequency than other people who have a distance. Moreover, this distance is not straight distance, because there are too many islands in Stockholm, the straight distance can not accurately show this variable. In order to get a necessary car, people will spend time or money to reach the location of the car on foot, by bike or even by bus if the distance is too long (i.e. the distance of NO.234 is 17.2 km), so it maybe take a member's some time or money to get the car.

The length of trip is also very important, we have two kinds of trip, long-trip and short-trip, I take it as a dummy variables named X_k

$$X_k = \begin{cases} 1 & if\ short\ trip \\ 0 & if\ long\ trip \end{cases} \tag{1}$$

[2]we used the website http://www.hitta.se to find this distance

Here a short trip means driving within the city, because the given data were not enough, we need to find a suitable length of trip to define it short or long. In this paper, we mainly talk about car-sharing in Stockholm, so I defined if the driving distance is shorter than city scope (we take the whole city as a circle and use its diameter $45km = \sqrt{6519km^2 \div \pi}$ to express it), we can think they were driving inside the city and denote it as short-trip, otherwise is long-trip. we are interesting wether kinds of trip have a significant effect to user's costs.

Cars' size also have an effect on the cost, costs are differences because different cars' sizes, focuses on how the different car's sizes affect the cost. All the cars provided in Stockholm Bilpool can be divided into three size: small, medium and large. Generally speaking, with the increase of cars' size, there is an increase in the cost, more bigger the car's size is, more higher the cost is.

Booking time, we use the time of trip-start in the raw data table to express it. In order to make it more relevance to practice, we may separate all the times into two groups: one is leisure time, and the other is working time. Leisure time contains weekends, holidays, nights (8 : 00 AM to 8 : 00 PM) and code it to one, the other times are taken to be zero.

When people chose car-sharing instead of private cars, driving cost is a very important factor. Compared with the private automobile, car-sharing costs must be much lower, mainly including booking fee, annual fees and driving fees. Because annual fees was charged by month: $Annualfee = 20SEK \times 12$, we mainly focus on driving fees which have two different parts one is $(Price(SEK/min) \times DrivingTime$, and the other is $Price(SEK/km) \times DrivingMileage$, namely by time and distance-based respectively. Here in my paper, we plus them together as total cost. In order to make this variable more convincing, in the article I use the unit cost SEK/km, the follow equation can show how to calculate it:

$$Cost = [(Price(SEK/min) \times Time + Price(SEK/km) \times Mileage) + \\ Bookingfee + Annualfee \div Frequency] \div Mileage$$

3 Statistic method and analyze

3.1 Analyze the Cost of Members

3.1.1 Theory and method

In this section, I will introduce some basic knowledge about Generalized Linear Mixed Model (GLMM) which is relative to this paper[3]. Generalized linear mixed models provide a flexible way to model production traits which do not satisfy the assumptions of a linear mixed model. This flexibility allows us to select an more appropriate model as opposed to finding manipulations to make the data fit a restricted class of models. Figure 3 provides a symbolic representation of a generalized linear mixed model. The same as a linear mixed model, a generalized linear mixed model also includes fixed effects β; random effects u $\sim$ N(0 , G); design matrices X and Z; and a vector of observations y for which the conditional distribution given the random effects has mean μ, and covariance matrix R. In addition, a generalized linear mixed model includes a linear predictor η, and a link or inverse link function. In addition, the conditional mean μ, depends on the linear predictor through an inverse link function, h($\cdot$) and the covariance matrix R, depends on μ through a variance function.

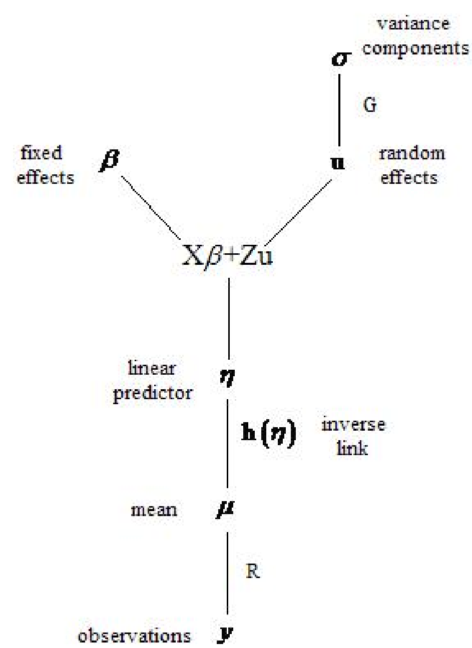

Figure 3: Symbolic Representation of GLMM

Summarize up, a generalized linear model is consist of three parts. First, a linear predictor, $\eta = X * \beta + Z * u$, is used to model the relationship between the fixed effects and random effects, the residual variability contained in the residual e. Second, an inverse link function $\mu_i = h(\eta_i)$, is used to model the relationship between the linear predictor and the conditional mean of the observed values. Under normal circum-

[3]You can learn more information in <AN INTRODUCTION TO GENERALIZED LINEAR MIXED MODELS> was written by Stephen D. Kachman

stances, the link function should be selected both simply and reasonably. Third, the variance function $v(\mu_i, \phi)$, is used to model the residual variability. Selection of the variance function is typically dictated by the error distribution that was chosen. In addition, observed residual variability is often greater than expected due to sampling and needs to be accounted for with an overdispersion parameter.

3.1.2 Model application

We take cost as response variables, because the cost did not follow the Normal distribution. We can transform the response variables to make it follow the normality distribution. To modeling the data, first thing is to find what kind of data it is and what method is suitable for modeling it. Due to the question of this paper, the response variable is an continuous variables and after the "log" transformation, it becomes Normal distributed. The following histogram and Normal Q-Q Plot (figure 4) confirms this. Thus, generalized linear model with gaussian family is suitable for the data analysis. Moreover, because the variable *usernumbers* can be regard as groups, some people live together (maybe they are family members) but have different numbers, and the records of cost are the repeated measures for one person, we can take *usernumber* as random effect part. As mentioned in the "Theory and method"section, mixed model is a powerful tool for grouped data analyze. Therefore, generalized linear mixed model can be quoted in the model construction.

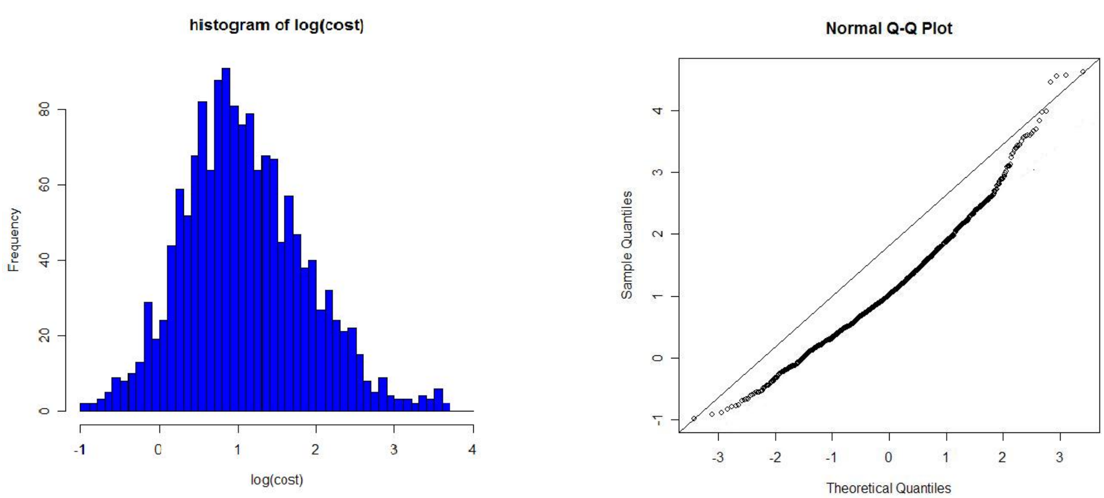

Figure 4:

In this paper, we analyze the relationship between response variable $log(Cost)$ and the independent variables with generalized linear mixed model and find out which factor has a significant effect to response variable, and all the data was shown in Appendix C. We will mainly use the R package “lme4”to model this data, all the factors except Usernumber are taken as fixed effects, while the factor of Usernumber is random effects. Then compare with models in gaussian family with different links, then select identity link model with the smallest value of AIC, finally construct a mixed model as follows:

$$log(Cost) = \alpha + \beta_i \times Carsize + \beta_j \times Distance + \beta_k \times Bookingtime + \beta_l \times Trip + \beta_m \times Month + \beta_n \times Usernumber + \epsilon$$

Where $i = Small, Medium, Large$, three sizes of cars; $k = 1$ when booking a car in leisure time, otherwise $k = 0; l$ has two levels 1 means $Short - trip$ and 0 means $Long - trip$; $o = Jan, Feb, Mar \ldots Dec$, it expresses the levels of month. Finally, we get a result in Table 1.

Table 1: Estimated result of GLMM model

Fixed effects:			
	Estimate	Std.Error	t-value
(Intercept)	1.2384	0.0918	13.493
SizeMedium	-0.1209	0.0581	-2.082
SizeSmall	-0.3284	0.0606	-5.421
Distance	-0.0006	0.0062	-0.101
Trip	0.6196	0.0381	16.273
Bookingtime	-0.0788	0.0351	-2.247
Month(1:12)	Included in appendix D		

Random effects:			
Groups	Name	Variance	Std.Dev.
Usernumber	(Intercept)	0.16980	0.41207
Residual		0.37326	0.61095
number of obs:	1528	groups:	Usernumber 113

AIC	BIC	logLik	MLdeviance
3109	3205	-1537	3002

Here we examining the model with the help of function $qqnorm(residuals())$ we check the residuals of this model wether follows normal distribution The following picture is the QQ-plot of the residuals. It tells that the residuals of this model follow the normal distribution.

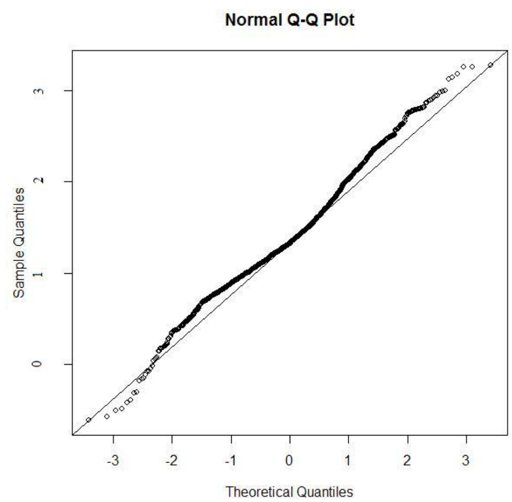

Figure 5: QQ-plot of the residuals

From the above table we see that the AIC of this model is 3109. From Table 1, also shows that the estimators of factor *Distance* is not significant (for the t-value: -0.101). This factor have not a significant effect to the *Cost*. By contrast, the estimator of *Bookingtime*and *carsize* is significant, specify both of them are important ingredients of the using cost. Because the estimators are all negative and the estimator of *SizeSmall* is lower than *SizeMedium* $(-0.3284 < -0.1209)$, that means the cost of using a medium car may be significantly higher than using a small car(correspond with the actual situation). So in order to reduce the rental cost, maybe the cost of using large or medium-sized cars can be lowered charged.

With analyzing the estimated coefficient of these variables, Distance has a positive effect to cost, the use of cost be rose when distance increasing. Reduce distance is also an feasible way to reduce the cost, there are two methods which I suggest to decrease distance: First, the bilpool can increase the number of locations to reduce distance between members and automobiles; second, provide a variety of sizes in a location to ensure that each user will be able to get the car as they demand. The cost for booking a car at working time will be obviously less than leisure time. The variable of *Trip*, which has a significant effect to cost either. the cost of long-term travel is significantly higher than short-term. From this point, Bilpool should offer a certain price discount to users who take a long-term travel in order to reduce its cost.

3.2 Analyze the Cost of Bilpool

3.2.1 Theory about Newman-Keuls Test

According to the above section we have know a few useful information about members' cost, while in this part we will focus on how to control the company's operating costs. First, the company's operating costs is mainly determined by the number of cars owned by Bilpool, booking frequency have a significant effect on number of cars as shown in Appendix E, we take this problem as a time serial regression, separate the whole year into twelve months and discuss whether the booking frequency have a significant change with month changed, then decide how many cars exactly should be owned by Stockholm Bilpool for each month. First of all, we should give a hypothesis test whether the booking frequency have a significant change with month changed

- ▶ The null hypothesis $H_0 : p_1 = p_2 = p_3 = p_4 = \ldots = p_{12} = \frac{1}{12}$
- ▶ The alternative hypothesis H_1: at least one month is different from the others. where p_i means the probability of booking frequency for ith month.

Test Result: Chi-squared test for given probabilities
X-squared = 68.43, df = 11, p-value = 2.874e-12

Thus we reject null hypothesis, there must have at least one month is different from the others significantly. Then we find out which month is higher than the average level and which one is lower with the Newman-Keuls test. Here is a simple introduction to the Newman-Keuls test in the next.

the Newman-Keuls test is very common "multiple comparisons "method, this test have been developed for the purpose of analyzing the reasons that made ANOVA reject the null hypothesis[4]. The Newman-Keuls test makes pairwise comparisons of group means after ANOVA has rejected the null hypothesis.
More specifically :

♦ If the test has declared means μ_i and μ_j ($\mu_i < \mu_j$) not significantly different.

♦ Any pair of means μ_l and μ_n within $\mu_i \le \mu_l \le \mu_n \le \mu_j$ will also be found "significantly different"by the Newman-Keuls test.

The result of the test is as series of pairs of groups, the means in each pair being

[4]Here use non-parameter (Pearson's Chi-square test) to instead of ANOVA test. For the frequency data in paper doesn't follows normal distribution so not satisfy ANOVA test's requirement, moreover these two tests always have the same result

considered “significantly different” by the test at a chosen a significance level. For any pair of groups, the Newman-Keuls test produces a value $q_{observed}$ of the test statistic. This value is compared to a theoretical critical value found in Newman-Keuls tables. For each pair of groups, this critical value depends on:

◇ The sizes of the groups,

◇ The difference of the ranks of the means of the compared groups. For example, if the means are found to be in the order $\mu_1 < \mu_2 < \mu_4 < \mu_3$, then comparing μ_1 to μ_3 will use the value $K = 4$ in the Newman-Keuls table.

◇ The significance level α of the global test.

So a new critical value has to be used for each and every new comparison. If, for a comparison, $q_{observed}$ is larger than the $q_{critical}$ as read in the table, then the hypothesis that the means of the corresponding two populations are equal is rejected. If the test fails to reject the hypothesis of equality of a certain pair of means, then it is garanteed that it will also fail to reject this hypothesis for any pair of means in-between the two original means. The test is then usually not conducted for these pairs.

3.2.2 The Newman-Keuls Test

We can get a frequency table (see Appendix E) from raw data, and take this table for further processing, we are able to have a new table which contains all enough data for the Newman-Keuls test as below:

Month	*Jan*	*Feb*	*Mar*	*Apr*	*May*	*Jun*	*Jul*	*Aug*	*Sep*	*Oct*	*Nov*	*Dec*
Sum	125	137	117	131	144	116	69	107	142	188	156	163
Mean	1.09	1.19	1.02	1.14	1.25	1.01	0.60	0.93	1.23	1.63	1.36	1.42
Variance	2.90	4.35	2.46	3.10	3.01	2.04	0.87	2.59	2.97	4.36	2.78	4.32
Cars	11	12	11	13	14	11	11	13	14	13	15	14

Here are steps of the Newman-Keuls test (significant level is $\alpha = 0.05$):

⧫ At first, ordered all the group means in increasing order.

$\mu_7 < \mu_8 < \mu_6 < \mu_3 < \mu_1 < \mu_4 < \mu_2 < \mu_9 < \mu_5 < \mu_{11} < \mu_{12} < \mu_{10}$

⧫ Second compares the extreme group means, *i.e.* μ_7 and μ_{10} ($K = 12$) $q_{observed} > q_{critical}$. Then μ_7 and μ_{10} are declared significantly different (at chosen significance level). And next, compared μ_7 with μ_{12} etc. In the end, we found that μ_7 didn't have

a significantly different to only μ_8. Then compare μ_8 with $\mu_{10} \ldots \mu_6$, and continued until found μ_4 and μ_2 have no significantly different.
⧫ Finally, according to the table we separate twelve months into 3 groups: Jan, Feb, Mar, Apr, May, Jun as medium group; Jul, Aug as less group; Sep, Oct, Nov, Dec as much group. You can see more clear in the following figure,

Thus, we may know if the Stockholm bilpool want to reduce its operating costs, can adjust the number of cars owned by bilpool flexible with selling or buying cars at the appropriate time. At a word, selling few cars in Jun, buying cars in a large number in Aug, selling a bit of cars in Dec to prepare for next year.

4 Conclusion

To summarize what I have done in my paper, I study the factors which have a effect on the cost of members with the method of Generalize Linear Mixed Model (GLMM) and analyze the cost of Stockholm bilpool itself by the Newman-Keuls test. Through analyzing the original data, we know $Carsize$ have a significant negative effect to users' costs. With sizes increasing the cost should also have a significant rising, exactly means that the cost of using a medium car may be nearly three times more than using a small car. The unit cost of Long-trip travel was significantly lower than short-trip, so in order to reduce this cost, we can give a few discounts to these users who take a short-trip travel.

By studying the booking frequency for each month, we can clearly separate the whole year twelve months into three groups with the Newman-Keuls test, then have a decision how to deal with cars owned by Stockholm bilpool, selling or buying cars to keep appropriate number in order to minimal cost of Stockholm bilpool itself.

References

[1] **Ulf Olsson** ,*Generalized Linear Models - An Applied Approach*, 2002: Uppsala University

[2] **Peter Dalgaard** ,*Introductory Statistics With R*, 2002: Department of Biostatistic University of Copenhagen

[3] **Stephen D. Kachman**, *An Introduction To Generalized Linear Mixed Models*, 1998: Department of Biometry, University of NebraskaCLincoln

[4] **Hao Li**, *Does A Booking Fee Hold Back Members Short Trip In Falu Bilpool: A GLMM Approach*, 2007: Dalarna University.

[5] **Shanshan Yang**, *Carsharing C An Integral Part Of An Intermodal Transport System*, 2007: Dalarna University.

[6] *The Impact of Driving Fee on the Car Usage*, January 2001: STOA - Scientific and Technological Options Assessment .

[7] http://www.aiaccess.net/English/Glossaries/GlosMod/e_gm_newman_keuls.htm

Appendix A

Part of the distance for members in Stockholm bilpool

User	Street	location	km
293	Voxnegrand 12, 2 tr	Medborgarplatsen, Blekingeg - DELVIS NYCKEL	8.1
180	Kristinelundsvagen 34	Bagartorp, Bagartorpsringen	7.2
408	Magnus Ladulasgatan 21	Rosenlunds sjukhus, Tideliusg - DELVIS NYCKEL	0.98
146	Dalagatan 40	Fridhemsplan, Arbetargatan - Nyckel	1.7
320	Dalagatan 40	Fridhemsplan, Arbetargatan - Nyckel	1.7
409	Katarina Bangata 66	Medborgarplatsen, Blekingeg - DELVIS NYCKEL	2.7
435	Brahevagen 7	Bagartorp, Bagartorpsringen	4
430	Brahevagen 7	Bagartorp, Bagartorpsringen	4
199	Krukmakargatan 34A	Medborgarplatsen, Blekingeg - DELVIS NYCKEL	1.6
278	Rindagatan 48	Gardet, Vartavagen - NYCKEL	1
218	Bordsvagen 62	Gubbangen, Knektvagen	0.46
55	Skanegatan 53A	Medborgarplatsen, Blekingeg - DELVIS NYCKEL	2.3
9	Skanegatan 61	Medborgarplatsen, Blekingeg - DELVIS NYCKEL	2.3
420	Ringvagen 11 E	Rosenlunds sjukhus, Tideliusg - DELVIS NYCKEL	0.43
187	Magnus Ladulasgatan 21	Rosenlunds sjukhus, Tideliusg - DELVIS NYCKEL	0.98
258	Ankdammsgatan 24 B 6 tr	Solna C, Hanebergsgatan	0.75
207	Tranebergsvagen 78	Fridhemsplan, Arbetargatan - Nyckel	3.4
268	Styrmansgatan 45	Gardet, Vartavagen - NYCKEL	1.6
125	Grevturegatan 75	Gardet, Vartavagen - NYCKEL	2.4
120	Swedenborgsgatan 26	Medborgarplatsen, Blekingeg - DELVIS NYCKEL	2.2
232	Odd Fellowvagen 32	Medborgarplatsen, Blekingeg - DELVIS NYCKEL	10.9
233	Bagartorpsringen 74 2TR	Bagartorp, Bagartorpsringen	0.16
250	Nytorget 8	Medborgarplatsen, Blekingeg - DELVIS NYCKEL	2.3
17	Helgalunden 5	Medborgarplatsen, Blekingeg - DELVIS NYCKEL	3.3
418	Ronnebyvagen 40	Medborgarplatsen, Blekingeg - DELVIS NYCKEL	6

Appendix B

Charges for various types of cars

Bilmodell	Gub, Sol, Bag	Med, Ros, Gar	Fri	
Avensis	**0.283**	**-**	**-**	kr/min
	1.95	**-**	**-**	kr/km
Corolla, Focus, Astra	**0.242**	**0.333**	**0.367**	kr/min
	1.75	**1.75**	**1.75**	kr/km
Aygo, Corsa, Fiesta	**-**	**0.292**	**0.325**	kr/min
	-	**1.60**	**1.60**	kr/km

The missing values in this table means there were no such type cars in this location

Appendix C

Part of the raw data in Stockholm bilpool

Month	**user**	**size**	**dis**	**trip**	**booking**	**cost**
Dec	9	Large	0.28	0	0	29.746
Dec	9	Large	0.03	0	0	28.172
Dec	9	Large	0.1	0	0	128.737
Dec	9	Large	0.41	1	1	112.980
Dec	9	Small	0.19	1	0	86.240
Feb	9	Large	8.3	1	1	88.570
Feb	9	Large	7.7	1	1	87.390
Feb	9	Large	6.3	1	0	72.300
Mar	9	Large	5.8	1	1	84.660
May	9	Large	3.4	0	1	50.094
May	9	Large	3	0	0	101.900
Oct	9	Large	1.1	0	1	53.799
Oct	9	Large	0.89	1	0	109.000
Sep	9	Small	1.7	0	1	33.479
Sep	9	Large	1.7	0	0	27.475
Sep	9	Large	1.4	1	0	76.250
Sep	9	Large	1.7	1	0	33.320

Appendix D

	Fixed	**effects:**	
	Estimate	Std.Error	t-value
(Intercept)	5.575	0.1314	42.42
MonthAug	0.1771	0.1082	1.64
MonthDec	-0.04901	0.1082	-0.45
MonthFeb	-0.3193	0.1093	-2.92
MonthJan	-0.4257	0.1773	-2.4
MonthJul	0.5509	0.1216	4.53
MonthJun	-0.03856	0.1038	-0.37
MonthMar	0.04426	0.1026	0.43
MonthMay	-0.0534	0.09564	-0.56
MonthNov	-0.06783	0.1069	-0.63
MonthOct	-0.1598	0.1019	-1.57
MonthSep	-0.1376	0.1031	-1.33

Appendix E

The relationship between Cars and Frequency.

call	**lm (formula = Cars $\sim$ Frequency)**			
coefficients:				
	Estimate	**Std.Error**	**t-value**	**Pr($>|t|$)**
(Intercept)	8.862708	1.31625	6.733	5.15E-05
Frequency	0.029075	0.009765	2.977	0.0139
F-statistic:	8.864 on 1 and 10 DF,		p-value:	0.01387
Residual standard error:	1.096 on 10 degrees of freedom			