Determining Optimal Lyft Offers

MSAN 631 - Design of Experiments

Arda A., Mikaela H.S., Vincent R., Zefeng Z., Tim Z.

Abstract

With the recent media controversy surrounding Uber, a sizable portion of commuters and travelers will likely be seeking out alternative ride-sharing services. Lyft, a long-time competitor of Uber in the Bay Area, wants to take advantage of this in order to grow their user base. To help persuade people to transition, Lyft is trying out new promotional offers. This report attempts to find the optimal offer that results in the highest booking rate.

Experimental Design and Data Collection

Since we had no prior information, and thus no intuition as to the significance of each factor, we first tested Lyft's offer structure using a full 2^4 factorial design, shown in Figure 1.

Factor	Low	High
Discount Amount ($X1$)	10%	50%
Offer Duration ($X2$)	1 Day	5 Days
Delivery Method ($X3$)	In-App	Email
Ride Type ($X4$)	Standard	Shared

Figure 1: Offer Tier Structure

Using a $(-1, 1)$ encoding system, a 16×16 design matrix containing every possible combination of experimental factors was constructed. Using this design, 1,000 replications were performed, and the proportion of replications in which a booking occurred was calculated, for each condition. The analysis of these results can be found in the subsequent sections.

Based on these results, we chose to conduct a second experiment of the form 2^3 where our three factors corresponded to discount amount, offer duration, and ride type. In our choice to exclude the third factor, our experimental conditions were defined differently, so we wanted to check that these results were similar to the ones obtained in the first experiment. We conducted this experiment in the same manner as before, but this time with an 8×8 design matrix. The analysis of these results can also be found in the subsequent sections.

Methodology and Assumptions

Various linear regression models were fit using booking rate as response and factors/their interactions as predictors. This methodology allows for quantification of the significance of main and interaction effects by noting the significance of the regression coefficients.

To avoid choosing over-determined models that disallow standard error estimates, heuristic approaches were used to assess the significance of various factors and interactions. One approach used is the quantile-quantile plot. Assuming the null hypothesis $H_0 : \beta_j = 0$ for all j is true, the residuals of the estimates should align with a normal distribution with zero mean. Hence, estimates that correspond to deviations from the normal line should be seen as significantly different from zero.

Main and interaction effect plots were also used to heuristically determine significant effects. Steeper slopes in main effect plots correspond to more significant effects; factors which had noticeable slopes were more likely to be kept in the model. Similarly, in interaction effect plots, parallel effect lines indicate a lack of interaction, which helped determine the significance of the associated interaction effect.

Exploratory Results of First Experiment

The four main effect plots and six 2-factor interaction effect plots can be found in Figures 4 and 5 in the Appendix, respectively (we do not show higher-order interaction effect plots since there is no easy way to visualize these).

In Figure 4, all effects except for delivery method appear to be significant. In Figure 5, only the discount amount-offer duration and discount amount-ride type interactions appear to be significant.

In order to formally test the apparent insignificance of delivery method, a regression model was fit to all factors and interactions. However, this full model would be over-saturated, meaning no conclusions about statistical significance could be made. To rectify this, a quantile-quantile plot for the full model was generated, which can be found in Figure 6 in the Appendix.

From the plot, we see that there are some points which lie on or close to the diagonal line – these points most likely correspond to factors or interactions that are insignificant. In order to identify these points, we created a list ordered by distance from the diagonal line:

1	2	3	4	5	6	7	8	9
$X4$	$X1X4$	$X1$	$X2$	$X1X2$	$X1X2X4$	$X2X4$	$X1X2X3$	$X1X3X4$

10	11	12	13	14	15
$X3X4$	$X2X3$	$X1X2X3X4$	$X2X3X4$	$X1X3$	$X3$

The top seven only involve $X1$, $X2$, and $X4$, while the bottom eight all involve $X3$, which corresponds to delivery method. These results are consistent with Figures 4 and 5; discount amount, offer duration, and ride type appear to be significant, while delivery method is insignificant.

We then fit a reduced model that excluded the 4-factor interaction, allowing a degree of freedom for estimating the statistical significance of the remaining effects. These results indicated that the only significant factors and interactions were $X1$, $X2$, $X4$, $X1X2$, and $X1X4$. These match the top five in our list above perfectly, and conveniently do not include $X3$.

Additional regression models were fit to the data to strengthen our confidence in these results. To help with model selection, the main effects were first considered individually. That is, regressions of the form

$$Y = \beta_i X_i + \varepsilon, i = 1, \ldots, 4$$

were fit, and the significance and magnitude of the coefficient estimates were compared. It was found that the delivery method ($X3$), had an effect that was essentially 0 (though with little statistical significance).

To further obviate the effect of delivery method, ANOVA tests were run on various models containing $X3$, either as a main effect and/or an interaction effect. In all cases, the effect of any terms involving $X3$, interaction or main, were not considered significant.

With all this evidence, the choice was made to exclude $X3$ and any associated interaction terms from further experimentation and analysis.

Exploratory Results of Second Experiment

A second experiment was then conducted using only $X1, X2$, and $X4$.

The three main effect plots and three 2-factor interaction effect plots can be found in Figures 7 and 8 in the Appendix, respectively. In Figure 7, all effects appear to be significant. In Figure 5, discount amount-offer duration and discount amount-ride type interactions appear to be the only significant interactions. These results are consistent with findings from the first experiment.

Again, the full regression model was over-saturated, so a quantile-quantile plot was used in order to figure out which terms were most and least significant; this can be found in Figure 9 in the Appendix. The points, ordered by distance from the diagonal line, correspond to the terms as follows:

1	2	3	4	5	6	7
$X4$	$X2$	$X1X4$	$X1$	$X1X2$	$X1X2X4$	$X2X4$

While this is a slightly different ordering from the results from the first experiment, the top seven terms are still the same. Given these consistent results, we are confident that our second experiment was essentially equivalent to the first, and in turn concluded that delivery method was indeed negligible. Our final model, discussed in the next section, includes $X1$, $X2$, $X4$, and their associated interactions.

Final Model and Interpretation

Running the model `lm(Y = X1*X2*X4)`, the regression coefficients are shown in Figure 2 (the inclusion of a factor in the coefficient refers to that factor taking on a level of "High", or 1).

Coefficient	Estimate	SE	p-value
(Intercept)	7.00000	0.01976	$<$ 2e-16
$X1$	0.22500	0.01976	3.20e-06
$X2$	0.18750	0.01976	1.26e-05
$X4$	-0.25000	0.01976	1.43e-06
$X1 : X2$	0.13750	0.01976	0.000118
$X1 : X4$	0.22500	0.01976	3.20e-06
$X2 : X4$	-0.03750	0.01976	0.094350
$X1 : X2 : X4$	-0.03750	0.01976	0.094350

Figure 2: Regression Coefficients

From this table, we can see that high discount rate and duration have a positive main effect, while applying the offer to shared services has a negative main effect. All but one of the interaction terms corroborates this; the only exception is the interaction term $X1 : X4$, which is positive. However, the overall net effect is negative due to the effects of $X1$ and $X4$.

Thus it is believed that setting $X1 = 1, X2 = 1, X4 = -1$, corresponding to a 50% discount over 10 days on standard rides, will produce the maximal booking rate.

As stated earlier, main effect and interaction terms involving $X3$ produced near-zero effect, so the choice is irrelevant; if one had to choose, the regression showed that the effect was on the negative side of zero when the delivery method was by email. Thus, we will choose to send the offer through the app itself.

Running the model, this results in a fitted value of 7.65% booking rate. The 95% prediction interval is $(7.43, 7.87)$, which agrees with the value obtained from the actual experiment.

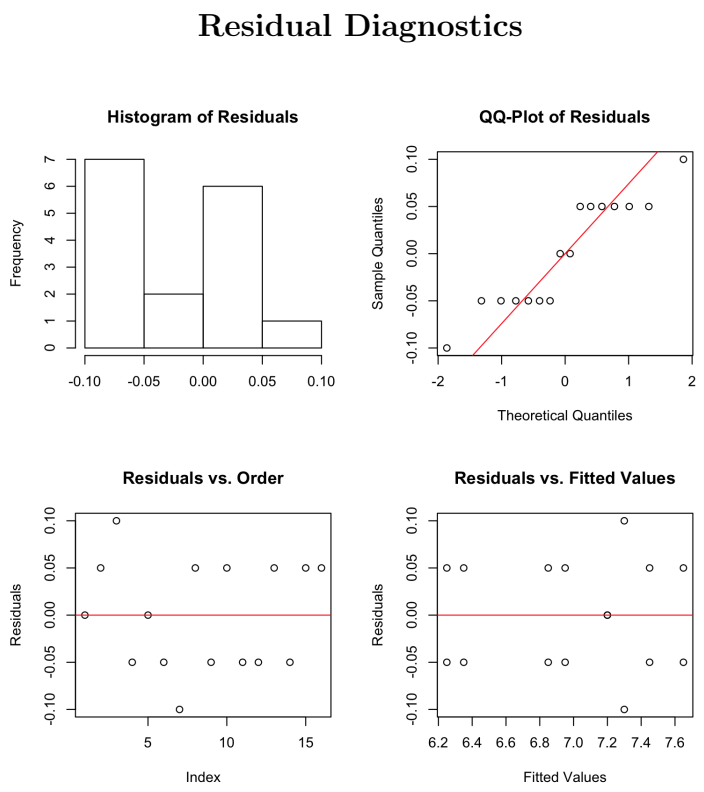

Figure 3: Residual diagnostics for the final model. Given the limited number of points, it is difficult to interpret.

There are four assumptions that linear regression requires, and we can heuristically test these assumptions by considering plots of model residuals. Given that our response is a proportion based off a much larger number of data points (1,000 to be exact), this leaves us with a very limited number of residuals to consider. While this makes it difficult to interpret the plots, we include this analysis here to have some intuition as to the validity of our model.

From the histogram and quantile-quantile plot in Figure 3, it is difficult to assess whether the residuals are normally distributed. Given the limited number of points and the fact that residuals are relatively robust to non-normality, we do not consider this an issue in this case. From the remaining two plots in Figure 3, we see that the residuals appear to be uncorrelated since there are no obvious patterns, have mean zero since they are randomly scattered about zero, and have constant variance since their spread is at the same level throughout.

Based on the limited information we can get from these residual diagnostics, we consider these four assumptions satisfied and deem our model as valid.

Appendix

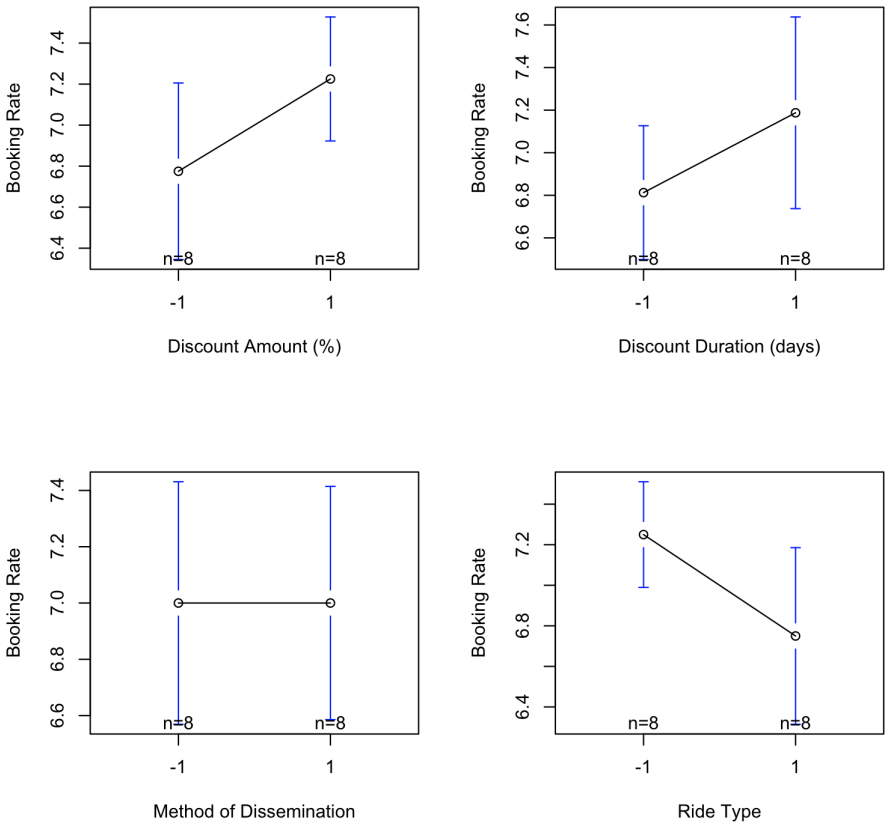

Figure 4: Main effect plots for each of the four factors in the first experiment. All factors appear to be significant except for delivery method since the slope is constant.

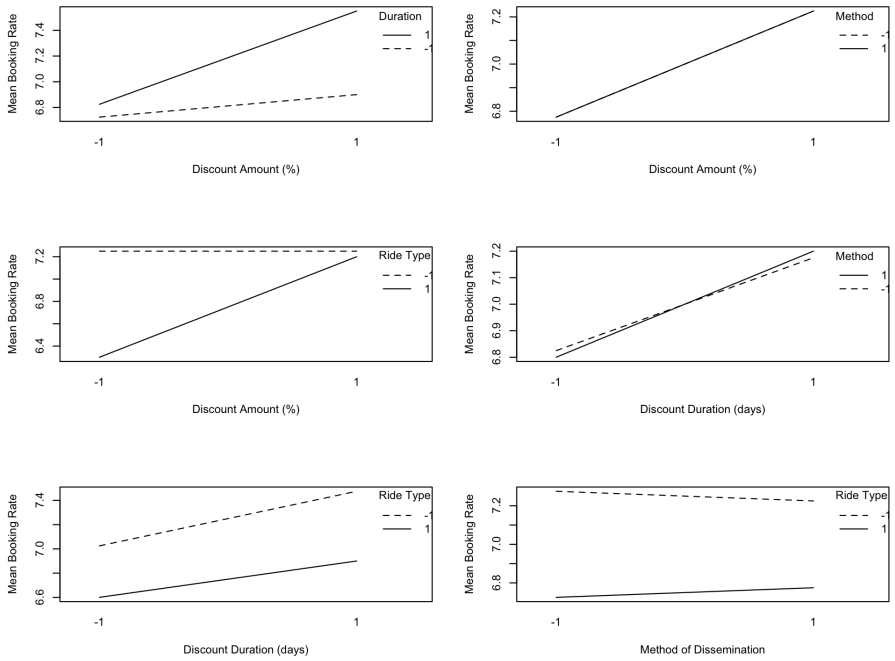

Figure 5: Interaction plots for each of the six 2-factor interactions in the first experiment. The discount amount-offer duration and discount amount-ride type interactions appear to be the most significant since the remaining interactions correspond to parallel or nearly parallel lines.

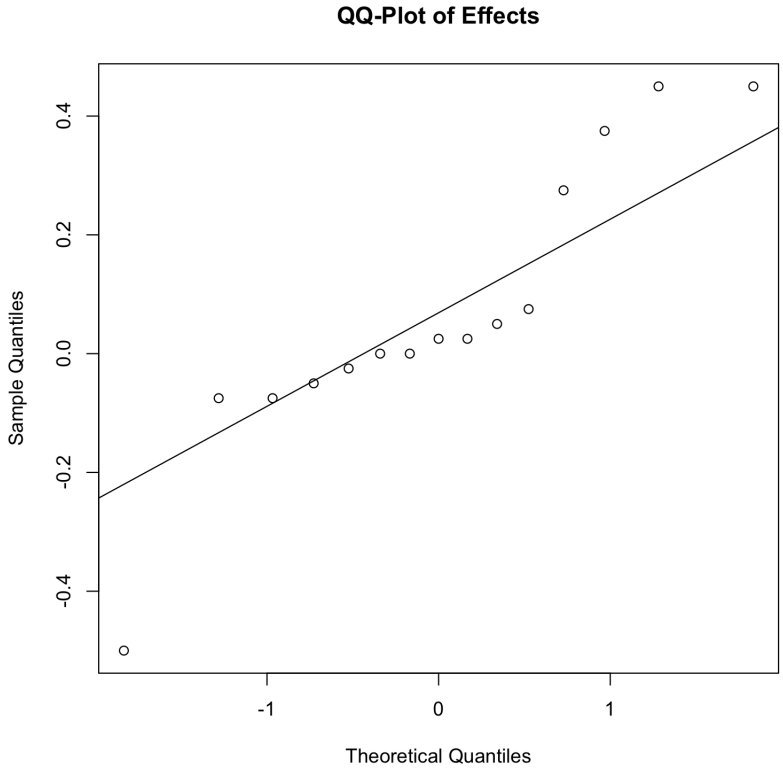

Figure 6: QQ-plot for the regression model that included all factors and interactions. The points on or close to the diagonal line correspond to factors or interactions that are insignificant.

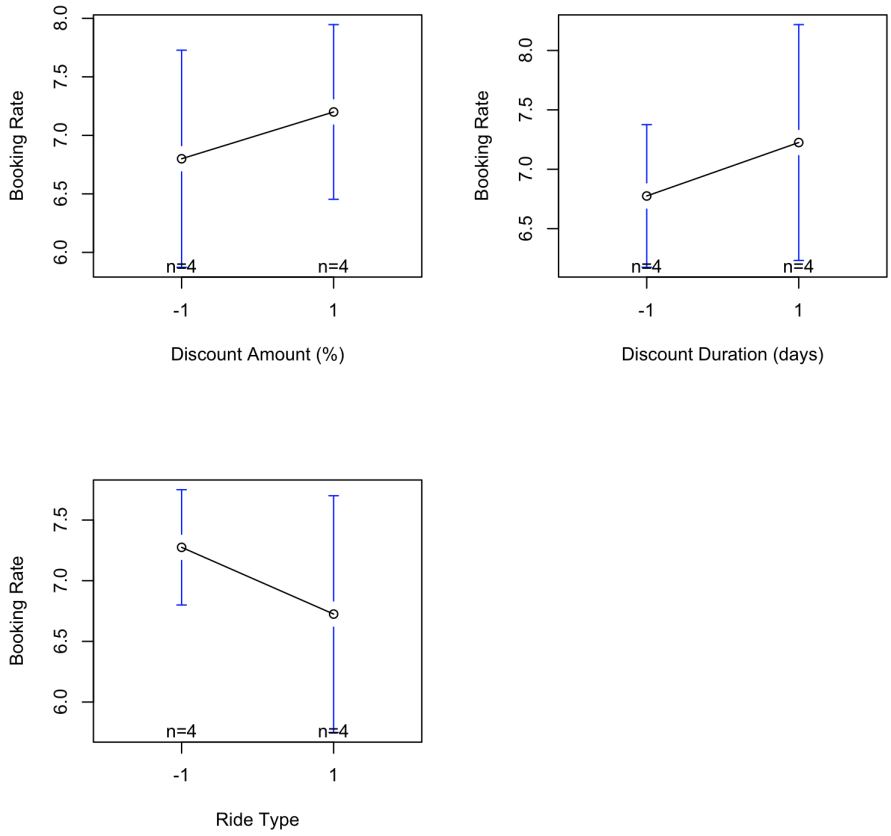

Figure 7: Main effect plots for each of the three factors in the second experiment. All factors appear to be significant since they have non-constant slopes.

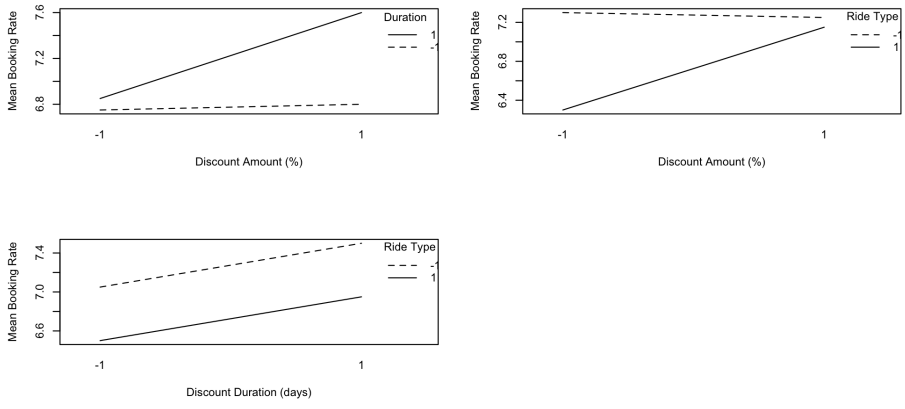

Figure 8: Interaction plots for each of the three 2-factor interactions in the second experiment. The discount amount-offer duration and discount amount-ride type interactions appear to be significant, while the offer duration-ride type interaction does not since the lines are parallel (consistent with Figure 4).

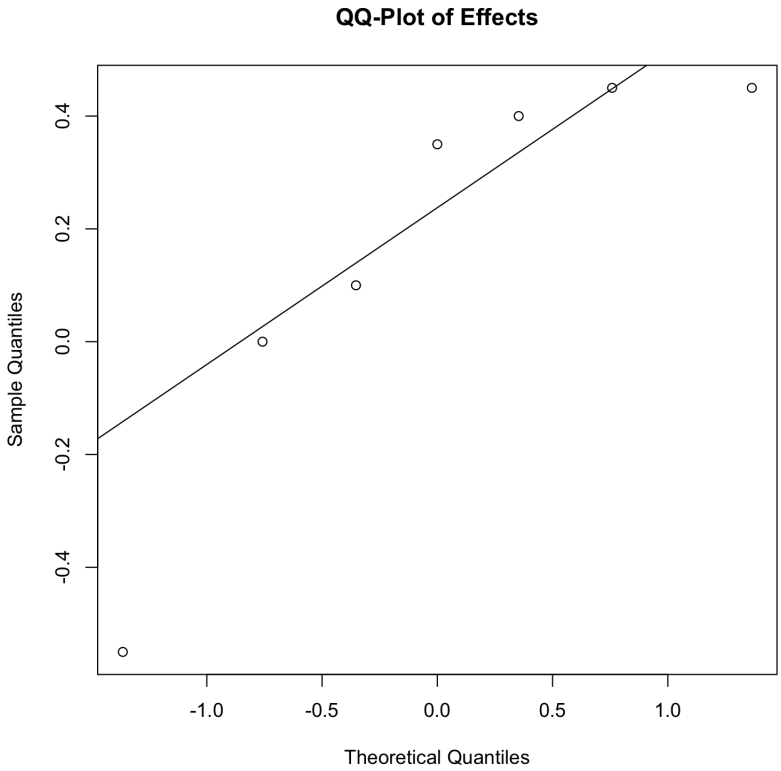

Figure 9: QQ-plot for the regression model that included $X1$, $X2$, $X4$, and associated interactions. The points on or close to the diagonal line correspond to factors or interactions that are insignificant.