Assignment 4: Marketing campaign optimization and Dynamic pricing

Advanced Econometrics

2022

Econometrics and Operations Research

Valentina Guldberg 2779376 *v.guldberg@student.vu.nl*

Mario Pirez-Pinacho 2780066 *m.pirez@student.vu.nl*

Benjamin Weggelaar 2682475 *b.d.weggelaar@student.vu.nl*

Jelmer de Jong 2782922 *j.a.de.jong3@student.vu.nl*

Group 23

18-10-22

Question 1

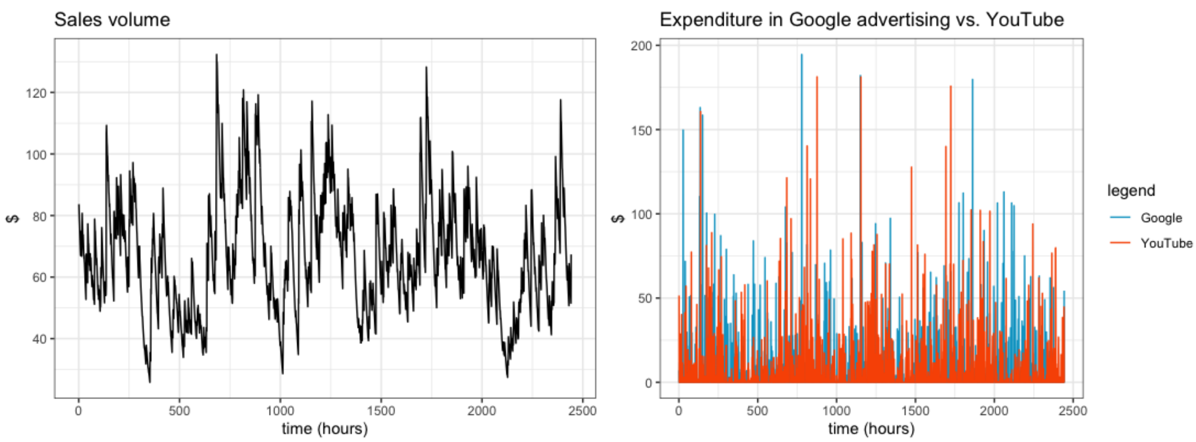

Figure 1: Time series of sales and expenditure in Google and YouTube Advertisement

Question 2

We estimate the following nonlinear dynamic time-varying parameter model:

$$s_t = \mu + \phi_1 gads_t^{\delta_1} + \phi_2 yads_t^{\delta_2} + \varepsilon_t \tag{1}$$

$$gads_t = \beta_1 gads_{t-1} + \alpha_1 g_t \tag{2}$$

$$yads_t = \beta_2 yads_{t-1} + \alpha_2 y_t \tag{3}$$

Table 1: Results for Least-square estimation of the model

$\hat{\mu}$	$\hat{\phi}_1$	$\hat{\phi}_2$	$\hat{\delta}_1$	$\hat{\delta}_2$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\beta}_1$	$\hat{\beta}_2$
1,070	1.761	1,926	0.221	0.560	5.075	5.097	0.908	0.945

The estimation results showed small differences in decimals depending on the computer used. Nevertheless, the results converge, and we decide to use these estimates for the rest of the questions in this part of the assignment.

Question 3

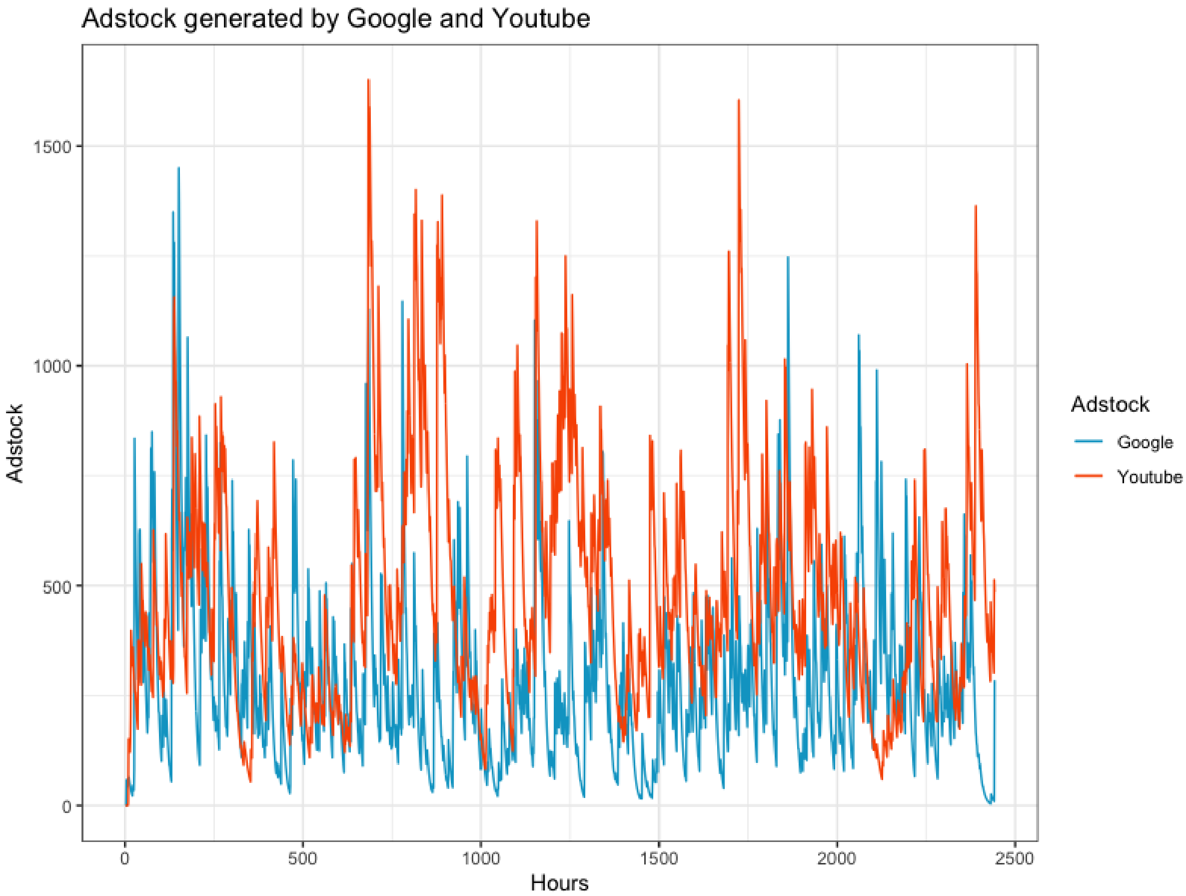

Figure 2: Filtered adstock generated by marketing on Google search and YouTube

Notes: The provided graph represents the filtered Adstock created by an advertisement campaign in Google search (blue) and Youtube(red). The observation time is 2.500 hours which is shown in x-axis, while the y-axis represents the amount of Adstock generated.

Question 4

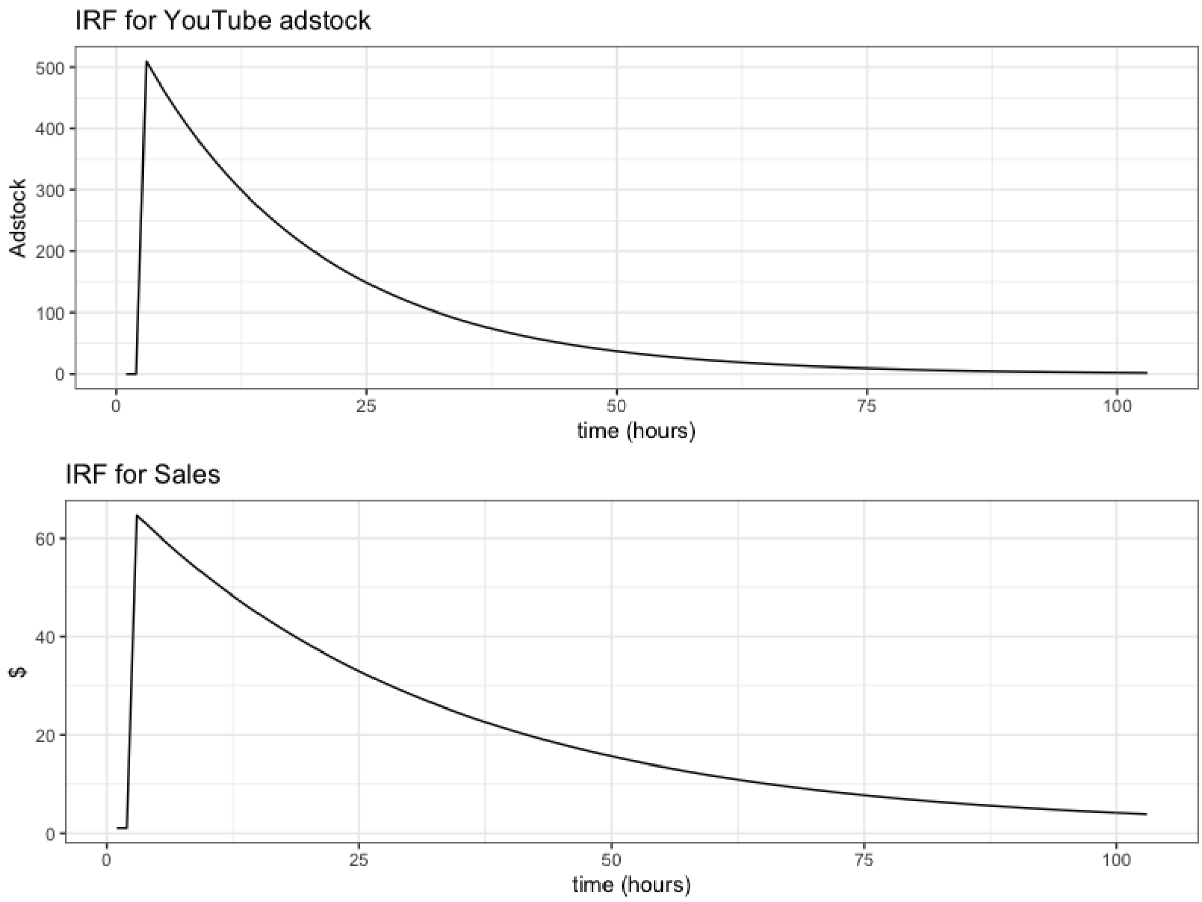

Figure 3: Impulse Response Function for Sales and Youtube Adstock

Notes: The Figure 3 represents the Impulse Response Funciton generated after spending 100 euros in marketing in Youtube. The first graph shows the IRF in Youtube adstock while the second one captures the changes on the sales. The observation time is 100 hours, which is shown in x-axis, while the y-axis indicates the money generated by this shock.

Question 5

The Impulse Response function shows positive results for the company. Making an expenditure of 100 euros in Youtube advertisement provokes an increasing in the adstock (from 0 to 510), and in consequence, a 98% growth in sales in the first hour after the shock. The adstock value lasts less over time, since its fall is greater than the sales obtained per hour. For example, 25 hours after the shock, the adstock felt by 75% , while sales have only decreased by 50% . Both metrics continue to show positive values by the end of the analysis, since at hour 100 after the shock has taken place, their value continues to be higher than the one they have prior to the investment.

Question 6

We calculated the sum of excess sales generated from the increase of 100 dollars in YouTube Advertisement, from the initial shock until sales reaches its pre-shock level. The sum is equal to 2076.24. This means that, conditional on our model, we can expect that in increase in 100 dollars in YouTube advertisement will lead to an increase of 2076.24 dollars in sales.

Question 7

We apply the same procedure as in the previous question, but this time an increase of 300 dollars in YouTube advertisement. The expected increase in sales is 3835.15 dollars. This shows that increasing the expenditure in YouTube advertisement by a factor of 3 does not triple the sales. This is explained by our model specification (equation 1), and the estimation results. Since $\hat{\delta}_2$ is less than 1, we have decreasing returns to Adstock (holding all else constant). Doubling (or tripling) the expenditure in YouTube leads to a less than doubling (or tripling) in sales.

Question 8

Since we suspect simultaneity between marketing expenditure and sales, creating a model using observational data would be a good predictive model, but it will be unable to asses the effect marketing expenditure has on sales, the causal relationship. If we ignore this complex relationship between sales and marketing expenditure, and run a simple regression, we will obtain an approximate model for the conditional expectation: $\mathbb{E}(sales_t|exp_t)$. If we use this model for structural purposes, the estimates would be biased due to endogeneity: $\mathbb{E}(sales_t|u_t) \neq 0$. Therefore, running an AB testing procedure and generating experimental data is crucial for analyzing the structural effect of marketing expenditure on sales.

Question 9

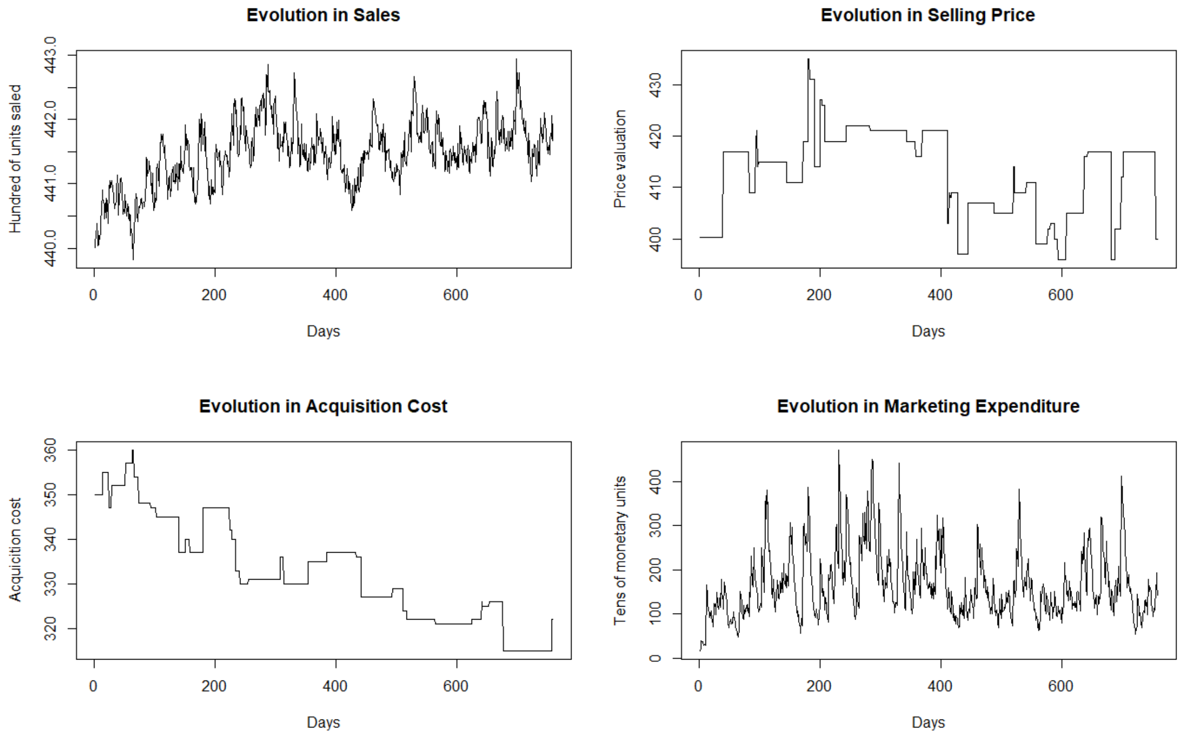

Figure 4: Daily data on sales, selling price, acquisition cost and marketing expenditures

Notes: The plots show the daily data for hundred of sales (top left panel), selling price(top right panel), acquisition cost (bottom left panel) and marketing expenditure (bottom right panel). The x-axis shows the total quantity (scaled for sales and marketing), whereas the y-axis displays a daily time axis.

Question 10

Regression model: $s_t = \alpha + \beta p_t + \epsilon_t$

Table 2: Regression results of sales on prices

Variables	Sales
Constant	43551.246
Selling price	1.443

Question 11

Usefulness of this regression as a predictive model for sales: by definition a good regression model has independent variables that are exogenous from from the dependent variable, i.e., exogeneity implies that $Cov(\epsilon_i, x_i) = 0$, where ϵ_i is the error and x_i is the independent variable. In our case, the regression model of sales on prices ($s_t = \alpha + \beta p_t + \epsilon_t$) would be useful if prices are not correlated with the error or innovation term. However, $Cov(\epsilon_t, p_t) \neq 0$ and that is telling us that price is an endogenous variable. The economic interpretation is that the price can be changed by the seller depending to the demand, so, here we would have a reverse causality problem since number of sales can determine the price. As a result of that we see that the estimate coefficient of prices (β) is positive with value 1.443, that is to say that an increase in prices by one monetary unit would increase the number of sales by 1.443; whereas economic theory holds that demand has a negative slope function with respect to prices, meaning to say that an increase in prices has negative effect on sales.

Supposing my estimator is consistent, if we are using this model for predictive purposes in the limit it will converge in probability to the true parameter, i.e., $\hat{\beta}_T \xrightarrow{p} \beta_0$ as $T \to \infty$ whose properties are that my estimator delivers the best predictive accuracy for the model.

Question 12

As a structural model, we could assume that the independent variable prices is exogenous from sales, in other words, $Cov(\epsilon_t, p_t) = 0$. Under this assumption, we should have a negative causal parameter ($\beta < 0$) that describes how an increase in prices affects demand supporting by economic theory that higher prices should have a negative effect in number of sales.

Question 13

Regression of selling prices on acquisition cost: $p_t = \delta + \gamma c_t + u_t$. This is the first stage of regression

Table 3: Regression results of prices on cost

Variables	Price
Constant	365.381
Cost	0.142

Regression of sales on predicted price: $s_t = \alpha + \beta \hat{p}_t + \epsilon_t$

Table 4: Regression results of sales on predicted price

Variables	Sales
Constant	52000.355
Predicted price	-19.029

Under instrumental variable approach where acquisition cost can explain prices and therefore sales through prices at the time that acquisition cost is exogenous from sales $Cov(\epsilon_t, s_t) = 0$, we obtained a negative causal parameter ($\beta < 0$) that describes how an increase in prices affects negatively the demand. This is supported by the economic theory that higher prices decrease sales ($\uparrow p_t \rightarrow \downarrow s_t$). To rephrase it, the effect of prices in demand is negative, then the coefficient estimate on prices should be negative $\beta < 0$ which, indeed, it is with value -19.029.

Question 14

Now let us investigate whether the regression model in Question 10 actually contains suffers from endogeneity. To do this we us the Durbin Wu Hausman test. This test compares the estimators from two models, one from the simple regression model in Question 10 and the estimators for the 2SLS regression model in Question 13. Under the null hypothesis endogeneity is no problem in the simple model and both the estimates from the simple and the 2SLS model are consistent, however the estimate from the simple model is more efficient. Under the alternative hypothesis the estimates from the simple model are inconsistent but the estimates from the 2SLS model are consistent.

The test statistic fro this test is defined as $H^T H = T(\tilde{\theta} - \hat{\theta})^T (\tilde{\Sigma} - \hat{\Sigma})^{-1} (\tilde{\theta} - \hat{\theta})$, with

$\hat{\theta}$ and $\tilde{\theta}$ being the estimates for the simple model and 2SLS model respectively. The expression in the middle represents the difference between the variance matrices from the 2SLS model and the simple model. This test statistic has a Chi-squared distribution $\chi^2(k)$ with degrees of freedom k equal to the rank of the variance difference matrix in the test statistic.

We find a value for the test statistic equal to 21393.29. The values for the Chi-squared distribution with degrees of freedom 2 and probability levels 1%, 5% and 10% are $9.21, 5.99$ and 4.61. Therefore, based on our value for the test statistic we can reject the null hypothesis for all probability levels. We can thus conclude that on a 1 percent significance level there is evidence that our model in Question 10 suffers from endogeneity.

Question 15

Regression of sales on predicted price and marketing expenditure: $s_t = \alpha + \beta \hat{p}_t + \omega m_t + \epsilon_t$

Table 5: Regression results of sales on predicted price and marketing expenditure

Variables	Sales
Constant	51265.642
Predicted price	-17.429
Marketing Expenditure	0.045

Question 16

Use the estimates from the model before to calculate the casual impact on expected profit of a unit increase in price, where:

$$E(\pi_t|c_t) = E(s_t) \cdot (p_t - c_t)$$
$$E(s_t) = \alpha + \beta p_t + \omega m_t$$
$$\text{Therefore, } E(\pi_t|c_t) = (\alpha + \beta p_t + \omega m_t) \cdot (p_t - c_t)$$

Substituting parameters from the previous model you get the expected profits that varies through times as:

$$E(\pi_t|c_t) = (51265.642 - 17.429 p_t + 0.045 m_t) \cdot (p_t - c_t)$$

The expected effect of prices on profits at given t comes from the partial derivative of expected profits over the price:

$$\partial E(\pi_t|c_t)/\partial p_t = [(51265.642 - 17.429p_t + 0.045m_t) \cdot (p_t - c_t)]/\partial p_t =$$
$$(51265.642 - 17.429p_t + 0.045m_t) - 17.429 \cdot (p_t - c_t)$$

Using the last observed values, the effect of price on profits is:

$$\partial E(\pi_{759}|c_{759})/\partial p_{759} = (51265.642 - 17.429 \cdot 400 + 0.045 \cdot 1425) - 17.429 \cdot (400 - 322) =$$
$$42998.45$$

Going back to the question, since $\partial E(\pi_{759}|c_{759})/\partial p_{759} = 42998.45 > 0$, the effect of the last observed prices on profits is positive which implies that increase prices would increase profits under ceteris paribus assumption, i.e., we recommend to increase the price.