VILNIUS UNIVERSITY
FACULTY OF MATHEMATICS AND INFORMATICS
MASTER'S STUDY PROGRAMME
(DATA SCIENCE)

Marketing Mix Modelling using Bayesian statistics

Master's thesis

Author: Rokas Rimša
Supervisor: Marijus Radavičius

Vilnius
2024

CONTENTS

Abstract

Marketing Mix Modelling is used by advertisers to understand the effect of marketing channels spending on sales and optimize media budget. In most cases, marketing modelling accounts for lagged effects of advertisement and diminishing returns. This thesis proposed a marketing mix model with Bayesian approach, using prior knowledge and assumptions. In addition, Bayesian regression fits adstock and diminishing return parameters of functional forms in the regression which let decision makers learn about marketing channels' features. The thesis goes through sensitivity check, posterior predictive check to assess model's performance. The study finds that TV marketing channel is the biggest contributor to sales when comparing to other channels. However, TV has relatively low adstock rate indicating that impressions in previous weeks poorly directs to sales of current week. Moreover, according to learned diminishing returns parameters, TV is slow saturation channel which is effective as others only when high number of impressions is achieved. When calculating contribution to sales from the model, for advertisers it is applicable to calculate Return on Advertisement Spend which helps optimizing advertisement budget.

Key words: Marketing Mix Modelling; MMM; Bayesian statistics; Adstock; Carryover effect; Shape effects; Diminishing returns.

Introduction

Marketing mix modelling (MMM) is implemented to understand the effect of marketing channels spending on sales in order to optimize future media investments with respect to revenue. These models are usually based on weekly or monthly aggregated national or geo level data. The data may include sales, price, company's special events, media spend in different channels, and external factors such as lockdowns due to COVID-19, macroeconomic figures, weather, seasonality, and market competition. These models are typically regression models that infer causation from correlation. MMM also need adstock and saturation transformations applied on marketing spending data. Adstock approach accounts for lagged effect of advertisement on consumer's behaviour. For instance, a marketer expects that spendings on advertisements will have effect not only first week but some weeks later, as well. Saturation function is implemented in MMM to account for diminishing returns of increased amount of advertising.

This thesis was also conducted with business partner AB Telia Lietuva. The company specializes mainly in telecommunication services. The partner provided company's data which is relevant for the research. Thesis was also supported by the specialists from Telia by providing knowledge about company's products, marketing field and MMM.

The purpose of the thesis is primarily to build marketing mix model applying Bayesian statistics. Even though literature with research in MMM has been trending in the last decades, literature with specifically Bayesian approach in MMM is still maturing; the field is not fully covered, yet, and needs more researches to be conducted. The model output will expose the effect of various marketing channels, such as TV, radio, social, display, etc. on sales. Parallel thesis objective is to find optimal carryover effect rate and diminishing returns transformation parameters for every media channel to understand better effect on consumers' behaviour. Additional goal for the thesis was to master Stan library which is used for Bayesian modelling and inference. Stan is widely used for Bayesian tasks, as well as, PyMC library, however Stan was preferred since it has its version in R programming. RStan is C++ syntax library and is quite complex.

Originally, MMM may be applied with simple OLS regression, however, Bayesian statistics has its own benefits. Firstly, Bayesian approach has ability to set constraints on model parameters using priors, for instance, making sure that media channels coefficients are non-negative. Also, one can incorporate existing knowledge about the effects of advertising, meaning, priors can be set based on the expert opinion or experience. Additionally, simple linear regression in Guadagni and Little (1983) linear relationship between sales and marketing predictors is assumed. Thus, such methodology is not able to account for saturation and adstock effects. Bayesian approach benefits here since it has ability to produce saturation and decay parameters as part of the model fitting process. Finally, with posterior distributions, the uncertainty can be assessed.

Further, Bayesian inference is described which is relevant to understand before going further. Later, Literature review follows which mainly focusses on “Bayesian Methods for Media Mix Modelling with Carryover and Shape Effects” by Jin and Wang (2017) article from Google. Also, explanatory analysis of modelling data is conducted. After data analysis, MMM methodology is described; results of the modelling are discussed. Finally, conclusion remarks are presented including summarizing results, limitations and further research suggestions.

The study provided some evidence that TV marketing channel is biggest contributor to sales. However, TV has relatively low adstock rate indicating that impressions in previous weeks poorly directs to sales of current week. Moreover, according to learned diminishing returns parameters, TV is slow saturation channel which is effective as others only when high number of impressions is achieved. Display channel saturates at highest pace, however, its contribution on sales is 4^{th} largest. Interestingly, Consumer Confidence Index (CCI) controls for large part of the sales during full of uncertainty COVID-19 pandemic period.

Bayesian Inference

Current section provides an overview of Bayesian inference. Firstly, it covers different types of information used when performing statistical inference – the data and the model. Furthermore, it is explained that in a Bayesian framework, even more information may be incorporated into the model by adding prior distributions to the parameters which are aimed to be estimated. Thereafter, the foundation of Bayesian inference – Bayes theorem – is explained. It is then made clear that Bayesian models in some cases, depending on the choice of likelihood and prior distributions, are complex to estimate. This is due to the challenge of finding a closed form of the posterior distribution. The review of mentioned summarized material was conducted by Karlsson (2022) and presented below.

Statistical inference relies on information to procure estimates of various unknown quantities. One source of information is observed data. However, the data by itself does not suffice to make accurate estimates – the data needs context. By casting the data into a model, the data is enriched with additional information. The model formalizes how the observed data is related to the unknown quantities which are aimed to be estimated. A simple linear regression model defines a structure between the observed covariates and the target variable. The target variable y_t is assumed to equal the weighted sum of the observed covariates $x_{i,t}$ alongside with Gaussian noise ε_t . Each covariate is weighted by an unknown quantity β_i, which is aimed to be estimated. Notably, the model lacks additional information on the nature of these weights. Bayesian inference offers a way of incorporating prior knowledge of the unknown quantities, further enriching the model with information. This is done by adding a prior probability of the unknown quantities to the model. These can be viewed as the beliefs of the unknown quantities, before any data is observed. For example, if it is assumed that the weights β_i are strictly positive, these can be assumed to be drawn from a distribution which is defined over the positive real line.

In the setting of Bayesian inference, the goal is to derive estimates of unknown quantities, in light of both the data and any prior assumptions. The foundation of the inference lies in Bayes theorem as described in Eq. (1), which is used to update the probability of a parameter, as more data is available (Rossi, Allenby, and McCulloch, 2012). Let θ denote the unknown quantity which is aimed to be estimated, and let x denote the data. The posterior probability $p(x;\ \theta)$ of θ given x is the result of two entities - the likelihood function of the data given the parameter $p(\theta;\ x)$ and the prior probability of the parameter $\pi(\theta)$. The posterior distribution, $p(x;\ \theta)$ is the focal point of Bayesian inference, as it discloses information on how probable the model parameters θ are, based on the data x.

$$\mathrm{p}(x;\theta) = \frac{p(\theta;x)}{p(x)}\pi(\theta) \propto p(\theta;x)\pi(\theta). \tag{1}$$

By sampling from $p(x;\ \theta)$, it is possible to estimate the distribution of θ given the data and prior probability. However, in order to do so $p(x;\ \theta)$ must be derived. In cases where $\pi(\theta)$ is conjugate with $p(x;\ \theta)$, meaning that $\pi(\theta)$ and $\mathrm{p}(x;\ \theta)$ have the same density form (Raiffa and Schlaifer, 1961), $p(x;\ \theta)$ can be obtained analytically. Or, in cases where θ is low-dimensional, it is possible to use numerical integration to estimate the integrals. However, in most situations the nature of the model results in a more challenging problem where it is no longer possible derive a closed form of $p(x)$, making it impossible to sample directly from $p(\theta;\ x)$ (Congdon, 2019).

Literature Review

Bayesian application in MMM area is still new approach of modelling and it is indicated by relatively low supply of literature. However, among this field the most popular research is "Bayesian Methods for Media Mix Modelling with Carryover and Shape Effects" by Jin and Wang (2017) from Google. The model is applied to shampoo advertiser data. Like in most MMM models, advertising variables are transformed by specific functions to account for lag effect and diminishing returns. These transformations will be discussed later. Paper applies Bayesian approach to use prior knowledge collected from previous studies, models and business knowledge.

The paper proposed a media mix model with flexible functional forms for media carryover and shape effects and show how the model parameters can be estimated in a Bayesian framework. As mentioned, in MMM advertisement spend on each channel is transformed to account for carryover and saturation effects. Regarding carryover effect, the paper suggested Weibull adstock function.

$$adstock(x_{t-L+1,m}, \dots, x_{t,m}; w_m; L) = \frac{\sum_{l=0}^{L-1} w_m(l) x_{t-1,m}}{\sum_{l=0}^{L-1} w_m(l)} . \quad (2)$$

Here w_m is a nonnegative weight function. The cumulative media effect is a weighted average of media spend in the current week and previous *L–1* weeks. *L* is the maximum duration of carryover effect assumed for a medium. Two Weight functions were applied: *geometric decay* and *delayed adstock*. Weibull adstock function is complex and difficult to interpret for experts less familiar with the MMM. Also, the later function is more computationally intensive. Our research implemented simpler and more intuitive adstock transformation which will be described later. Advantages of each of the adstock functions is presented in thesis's methodology section.

Shape effect is accounted by Hill function which was applied both in the reviewed paper and the thesis. Hill function covers saturation effect in majority of MMM literature. Its form and parameters are discussed in thesis's methodology part. For modelling, Bayesian linear regression is run. Firstly, marketing expenditure data has adstock transformation and on top of that Hill function is applied.

Michael Johns, Zhenyu Wang, Bruno Dupont, and Luca Fiaschi (2020) research captured saturation effects with logistic function. Carryover effect is captured with geometric weight function.

Also, in the Google paper it was assumed no synergy effect between media, thus, in the regression, media effects are additive. Regarding priors, for media's coefficients it was chosen not to implement strict priors and, traditionally, positively truncated (half normal (0,1)) priors were selected. Overall, the Jin et al. (2017) did research on shampoo data in U.S., thus, the results about media channels are not comparable with market in Lithuania and with completely different product. Although, the paper introduces techniques how to evaluate priors and model's other parameters.

Data analysis

For the purpose of the research in this paper, data was provided by the business entity. Data frequency is weekly. For MMM weekly data frequency is most often used in practice. The data for this research contains: sales (sold post-paid MHS contracts), company's marketing impressions on specific media channel, one of the economic environment indicator - Consumer Confidence Index (CCI; index is responsible for economic uncertainty, unemployment and inflation rates), event time information about Black Fridays, time information of Christmas (also includes New Year's eve weeks; binary variables), also binary variables containing information about imposed 4 different COVID-19 lockdowns (period 1: 2020-03-16 – 2020-05-17, period 2: 2020-05-18 – 2020-06-17, period 3: 2020-11-17 – 2020-12-16, period 4: 2020-12-17 – 2021-06-30), information about different brand new phone launches and pre-order dates, number of entity's active stores over time, competitors' received marketing impressions on their media channels, log transformed return of S&P 500 index. Additionally, there were generated manually binary variable that indicate last week of the month to account for payday effect. Also, seasonality variable with respect to response variable was generated from package Prophet powered by Meta.

The data initially is available from 2019-07-01 to 2021-12-27. The period contains major fluctuations of sales caused by COVID-19 pandemic. Table 1 below shows whole available period for analysis. From the graph it is visible that just before the first COVID-19 wave, there was a spike in sales going up which was followed by significant decrease caused by COVID-19 shock. Nevertheless, second COVID-19 wave fluctuations in the end of year 2020 is also visible, although, relatively less significant. For those reasons, to avoid wide fluctuations sales variables were floored at 500 value and applied ceiling was 1500 sold contracts. In the given period there are 131 weekly data points.

Sold contracts over time

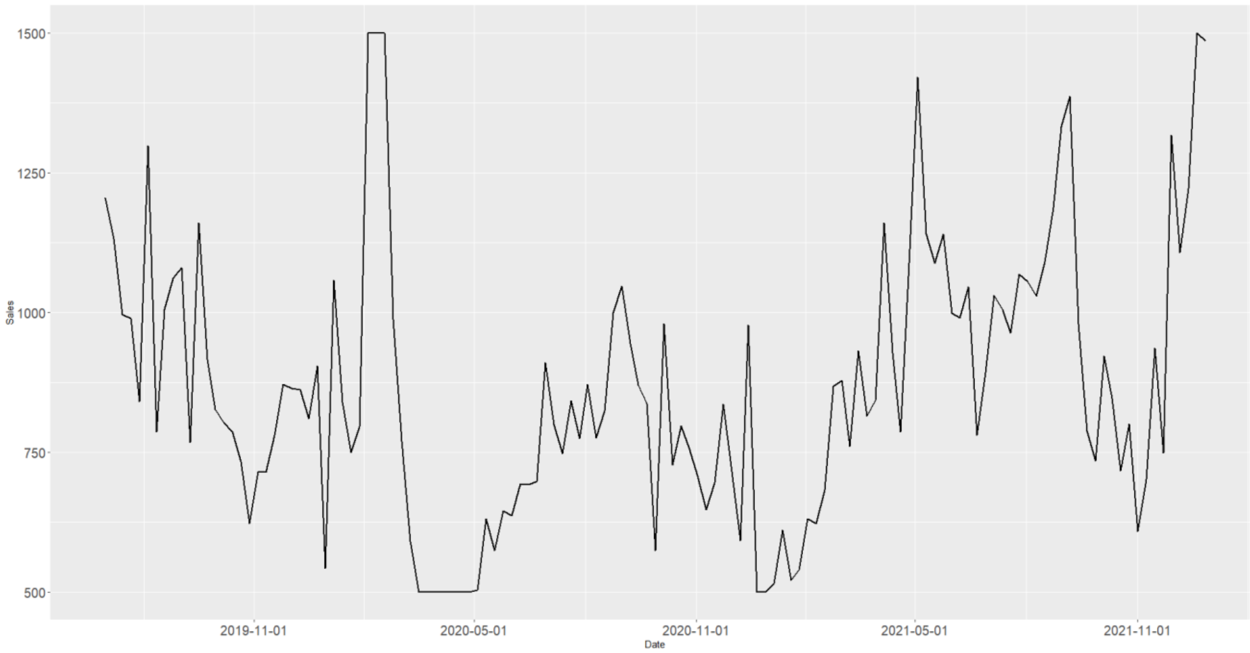


Table 1

The main MMM focus is on explanatory media channels variables. For modelling purpose 8 marketing channels were given: social media, display advertising, Google, TV, printed media, radio, video-on-demand (VOD), out-of-home (OOH) advertising. There are three main marketing channels' activity indicators: spend, impressions and gross rating point (GRP). However, GRP and impression attributes are not viable for all 8 media channels, only spending variable is found for all 8 marketing channels. GRP and impressions for some channels are impossible to count. The shortage of the spending variable as measure of marketing channels impact on sales is that itself it includes inflation factor. For instance, increasing inflation could lead to increased spending, however, it does not mean that it would impact the marketing exposure. GRP measure is viable only for radio and TV. Impressions can be calculated for all channels except OOH. Thus, in this research marketing channels will be depicted by collected impressions.

Table 2 below shows the descriptive statistics for a set of response variable and marketing channels' impression variables. The table contains the following information for each variable: the number of observations (N), the mean, standard deviation (Std. Dev.), minimum value, 25th percentile (Pctl. 25), 75th percentile (Pctl. 75), and maximum value.

Descriptive statistics of model variables

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Sales	131	866	245	500	704	1002	1500
Social	131	4574865	2451112	173711	2765550	6292244	12409976
Display	131	13104913	7814504	1580540	7587390	16210804	33486143
Google	131	99403	73905	8268	19265	124962	309438
TV	131	34575944	15310636	5690186	23044421	43270042	85814973
Print	131	144676	246762	0	0	217257	1098737
Radio	131	9234911	8843287	0	2229302	15378169	44926451
VOD	131	1102080	1066678	0	227425	1624082	4288517

Table 2

The sales variable has a mean value of 866, a standard deviation of 246. Display advertising, radio and TV have much larger amount of generated impression than any other variable, indicating they have the highest reach among all marketing channels. Number of impressions in marketing is defined as the number of times your content is displayed. Least impressions on average were generated from Google channel. Importantly, Print channel in some weeks have 0 impressions since company decides not to invest in particular week as they expect that earlier investments in the channel will have delayed effect on sales (adstock effect). The variables of phone launches (indicate if at the specific week there was at least one phone launch or pre-order), paydays, Christmas, black Friday, COVID-19, and seasonality are binary ones. The variable CCI has a mean of -0.5, indicating that on given period consumers were pessimistic about the economic conditions in Lithuania. Finally, for competitors generated impressions regarding mobility sector, one can clearly see that competitor 1 and its subsidiary generated significantly more than competitor 2. However, impressions collected by main entity and its competitors are incomparable due to different data collection approaches.

The correlation matrix below provides insights on the relationships between explanatory variables and the target sales variable. Media channels have a moderate to strong positive correlation with sales. Radio has the strongest correlation with sales, followed by display and TV. The correlation between print channel and sales is weak, as well as with Google channel. Furthermore, there are several strong positive correlations between the marketing explanatory variables themselves, including a high correlation between social media and VOD, display, and Google. The control variables have little correlation with the media variables. It is important to consider the correlation between variables as it can affect the performance and accuracy of the model. Variables with high correlation can lead to multicollinearity, which can inflate standard errors and make it difficult to interpret the coefficients. Noticeably, CCI has quite strong positive correlation with Sales and none or weakly negative correlation with marketing channels. Also, Number of open company's stores is not positively correlated with sales which was not expected since increasing quantity of open stores should increase number of sales since availability of physical stores is wider. Number of open stores positively correlates with first COVID-19 lockdown as during first day of pandemics company closed some of its kiosks. First lockdown as well has strong negative relationship with CCI and sales which was expected due to high uncertainty at that time. Additionally, impressions and spends within the same channel have very strong correlation. It could indicate that inflation factor does not impact significantly spends variables' representativeness on impressions.

Correlation table of model variables

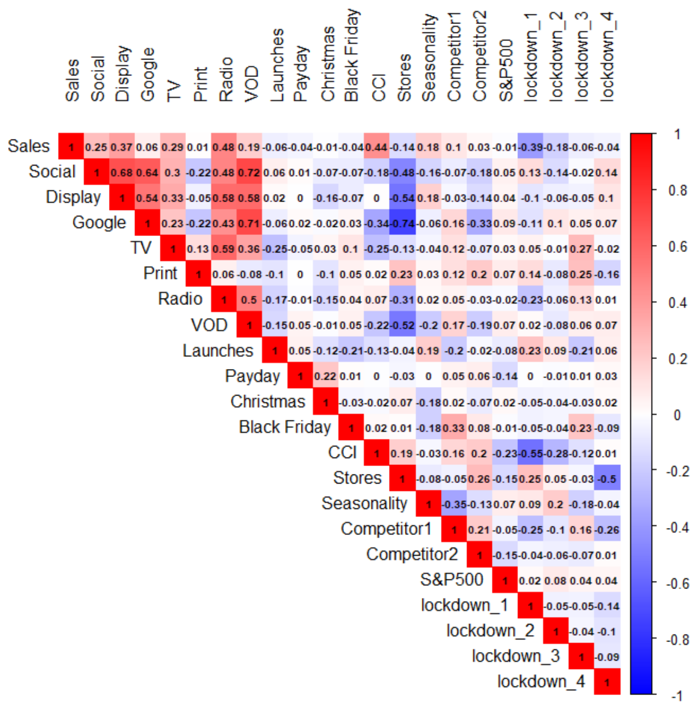


Table 3

Below, the graphs present marketing channels impressions over analysis period. For majority of channels number of impressions increased in the second part of 2022. It was correlating with marketing channels spend at that time period. Also, radio, VOD and print channels stand out as at some week point they did not generate any impressions, as mentioned, companies at those weeks expect that previous investments would have effect later. In addition, from the graphs it is visible high variance for majority of the channel impressions. Obviously, print impressions were generated more before COVID-19 pandemic while digital channels' impressions increased during the pandemic.

Marketing impressions over time

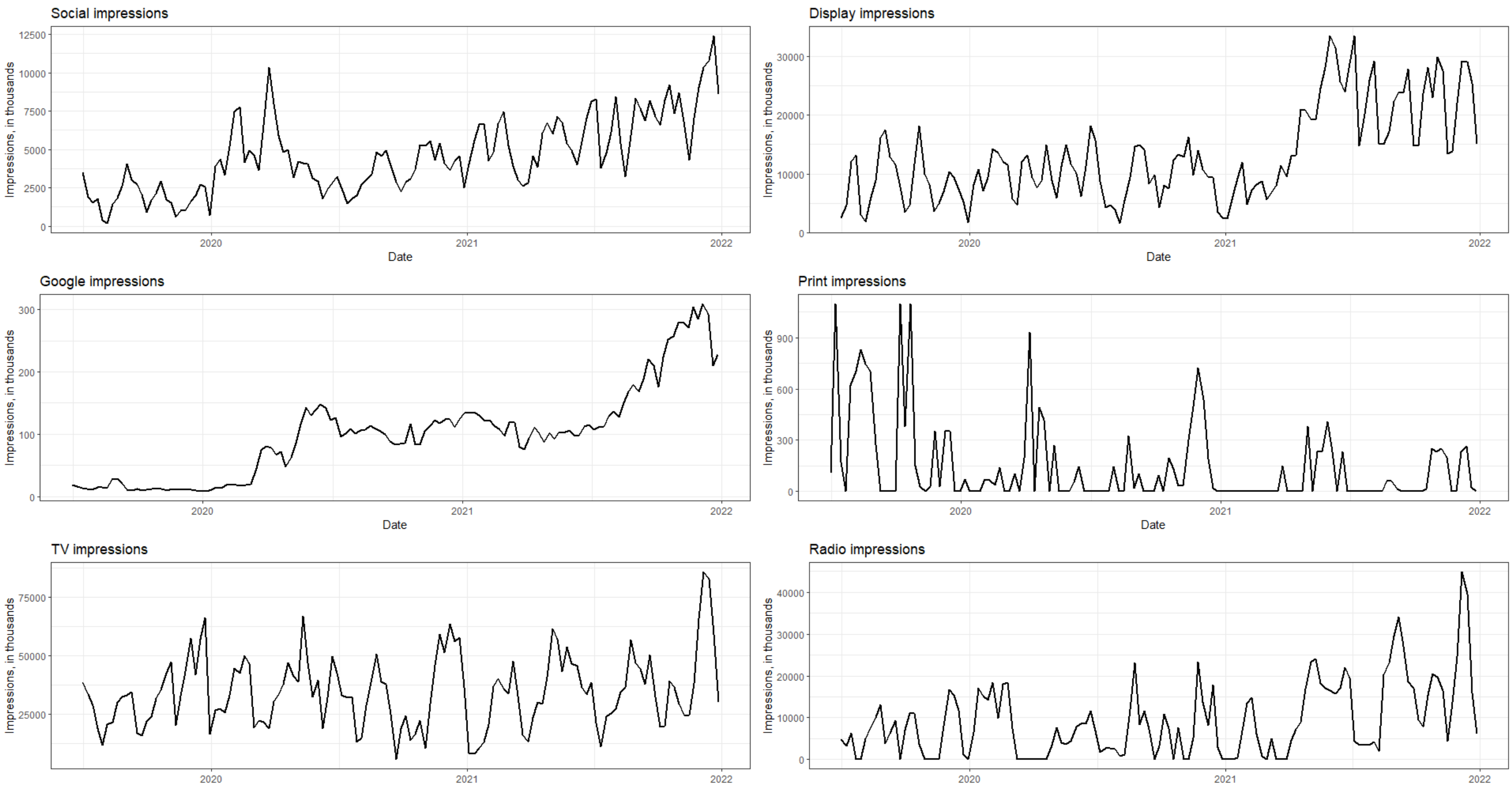


Table 4

Methodology

For the modelling part, Bayesian regression is run with Stan application using No-U-Turn Sampler (NUTS; Hoffman & Gelman, 2014) which extends a type of MCMC algorithm known as Hamiltonian Monte Carlo (HMC; Duane et. al., 1987; Neal, 2011) Thus, we can incorporate a-priori business knowledge into the modelling. The prior distribution is combined with the data (likelihood) to get a posterior distribution, which is the updated belief about the parameters after seeing the data. The posterior distribution can be used to make predictions, estimate uncertainty, and test hypotheses. Just by incorporating prior knowledge, a Bayesian style regression model can produce results that are in-line with the real world using less data which is the case for the thesis. Specifically for MMM, company's marketing spending each week in different specific channels are used to predict weakly number of sold post-paid contracts (company's specific product). As mentioned, in MMM variables of media channels are transformed.

Adstock effect

Adstock is applied to account for the awareness of an advertisement that gradually decays after its ending instead of instantly disappearing. For instance, an advertisement picture seen by customer on Monday might contribute to sales in current and later weeks as well. The adstock function can take various forms such as the geometric adstock or the delayed adstock (Jin et al., 2017). Equation 3 shows an example of the shapes of the adstock functions. The geometric model is a prevalent one and mathematically it is defined as a geometric decay function.

$$adstock\left(x_{t,m};\, \lambda\right) = \lambda x_{t-1,m} + x_{t,m}\ . \tag{3}$$

Parameter λ is called retain rate. It stands for lagged advertisement effect. Retain parameter also has prior distribution set and goes through Bayesian optimization for every channel. This transformation is less complex than proposed by Jin et al. (2017). Geometric adstock function does not have weight functions which tend to put more weight on recent weeks and less weight on much later weeks. Also, geometric adstock does not include maximum duration of carryover effect and delay of the peak effect parameters.

Saturation effect

Saturation transformation helps to deal with the diminishing returns of marketing activities. Saturation effect on advertisement spending in marketing was discussed by Farris et al. 2015. For instance, if we run a TV ad campaign for our brand, we might see an initial boost in sales, but we might also see a plateau or decline after a certain point as people get tired of seeing the same ad over and over again. To model the saturation effect of advertising, the spending needs to be transformed through a curvature function also called *response curve*. It will illustrate the relationship between company's sales and marketing effort on different channel. As mentioned above, Hill function is selected, following practices from the literature (Jin et al., 2017). It provides a flexible functional form:

$$Hill(x_{t,m}; K_m; S_m) = \frac{1}{1+(\frac{x_{t,m}}{K_m})^{-S_m}} . \qquad (4)$$

Here $S_m > 0$ is the shape parameter which is also referred to as slope, and $K_m > 0$ is the half saturation point, it controls the inflection point. The Hill saturation function assumes that the input variable falls within a range of 0 to 1, which means that the input variable must be normalized. Both mentioned parameters in Bayesian regression get optimized.

Bayesian regression

The response variable is log transformed number of sold contracts. We have weekly data, weeks are noted as $t = 1, . . . ,T$. There are M media channels in the media mix, and $x_{t,m}$ is the media spend of channel m at week t. Media variables included into the model is: TV, search, Google, radio, print, social. Media variables were scaled from 0 to 1 with function:

$$x_{scaled} = \frac{x-x_{min}}{x_{max}-x_{min}} . \qquad (5)$$

We apply the same assumption as by Jin et al. (2017) that there is no synergy effect between media channels, thus, its effects are additive in the regression. Media variables firstly had adstock transformation applied and later saturation transformation. Finally, MMM Bayesian multivariable regression is as follows:

$$y_t = \tau + \sum_{m=1}^{M} \beta_m Hill(adstock(x_{t,m};\ \lambda); K_m, S_m) + \sum_{v=1}^{V} \beta_v z_{t,v} + \sum_{c=1}^{C} \beta_c \rho_{t,c} + \epsilon_t . \qquad (6)$$

Here τ is the baseline sales, β_c is the coefficient for positively truncated control variables, β_v is the effect of the rest control variable z_c which may be negative or positive and ϵ_t is white noise that is assumed to be uncorrelated with the other variables in the model and to have constant variance. Coefficient β_m is the effect of channel m on sales. It is assumed that there is a linear relationship between the control variables and the response. For the likelihood, normal (Gaussian) distribution is selected (Jin et al., 2017).

For thesis relevant Bayesian regression output is posterior distributions. For further calculations single number will be required, thus, mean of the distributions are taken (alternatives – median, mode). Also, a credible interval can be derived which could help by estimating uncertainty. In case of data shortage posterior distribution can depend a lot on the priors set. In instances where the available data lacks substantial information for the estimation of a parameter, the prior exerts a considerable influence on the posterior distribution.

Choice of priors

Priors are chosen based on expert knowledge, gained experience from the past or strong beliefs. Priors can also be truncated letting coefficient be, for our case, nonnegative. Selection of prior distributions is presented below:

Priors and its parameters

Parameter	Prior
λ	beta(3,3)
K	beta(2,2)
S	gamma(3,1)
β	half normal (0,1)
τ	half normal(0, 5)
γ	normal(0, 1)
π	half normal(0, 1)
Var(ϵ)	normal(0,1)

Table 5

The retention rate, denoted as λ, is subject to a constraint within the interval [0, 1), necessitating a prior distribution defined on the same interval, such as a beta or uniform distribution. If there exists substantial a priori knowledge regarding the retention rate, the support may be further restricted to a subset within [0, 1). However, at the moment researcher is not provided with confident prior knowledge, thus, the interval is not restricted and the prior distribution is selected the same as by Jin et al. (2017). In the context of the parameter S_m, a prevalent choice for a prior distribution is a gamma distribution having a positive mode. Furthermore, as already mentioned above, parameter K_m values must be between media spend values range. Thus, since media variables are scaled from 0 to 1, parameter K_m has to be between 0 and 1 with beta distribution applied for prior. In the case of the regression coefficients denoted as β_m, nonnegative priors, such as the half-normal distribution is employed, reflecting the a priori belief that the media effect is nonnegative. Since, Bayesian MMM for telecommunications product in Lithuania has not been in practice before, it was decided not to implement strongly informative priors, resulting that for every media coefficient prior is normal (0,1).

Regarding control variables, it was selected which control variables must have positive impact to sales, thus, positively truncated prior applied. These variables are events of new phones launches, week of the payday, Christmas and black Friday events, consumer confidence index, number of open company's kiosks and seasonality. For these control variables weakly informative prior was chosen since no further valuable information is available. Control variables that have no truncation are COVID-19 lockdowns and two competitors' spend on marketing activities.

Intercept coefficient prior is positively truncated since baseline sales are expected to be positive. The prior is weakly informative since no concrete information was known. White noise is also normally distributed by default. Jin et al. (2017) suggested using inverse gamma distribution; however, we selected normal distribution with large variance to reflect the uncertainty.

Discussion of results

Bayesian regression was run with 4 chains, each with 3000 iterations. Each chain has different initial value. For warmup there was 1500 iterations selected. It took 20 minutes to run the regression on 8GB RAM machine. Summary of the results are presented below in Table 6. As already mentioned, mean as measure was selected to represent regression outputs in one number. Regression produced standard deviation of the posterior distribution, as well as quantiles of the distribution.

Summary of the model results

Variable	Mean	SD	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
Social	0.26	0.17	0.03	0.14	0.22	0.33	0.69	2115	1
Display	0.11	0.08	0.01	0.05	0.10	0.15	0.32	3158	1
Google	0.05	0.05	0.00	0.02	0.04	0.07	0.18	3584	1
TV	0.45	0.16	0.20	0.34	0.43	0.53	0.77	2340	1
Print	0.06	0.05	0.00	0.02	0.05	0.09	0.20	4703	1
Radio	0.15	0.11	0.01	0.07	0.13	0.21	0.42	3569	1
Launches	0.07	0.03	0.01	0.05	0.07	0.09	0.12	3975	1
Payday	0.02	0.01	0.00	0.00	0.01	0.02	0.05	6565	1
Christmas	0.06	0.05	0.00	0.02	0.05	0.09	0.18	6854	1
Black Friday	0.08	0.06	0.00	0.04	0.07	0.12	0.21	5046	1
CCI	0.58	0.11	0.36	0.51	0.58	0.65	0.78	3423	1
Kiosks	0.07	0.05	0.00	0.03	0.07	0.11	0.20	4095	1
Seasonality	0.16	0.03	0.09	0.13	0.16	0.18	0.22	4055	1
Competitor1	0.03	0.07	-0.12	-0.02	0.03	0.08	0.17	4257	1
Competitor2	0.07	0.08	-0.10	0.01	0.07	0.13	0.24	5111	1
S&P 500	0.22	0.11	0.01	0.14	0.22	0.30	0.44	5082	1
Lockdown 1	-0.38	0.08	-0.53	-0.43	-0.38	-0.33	-0.22	3533	1
Lockdown 2	-0.20	0.08	-0.35	-0.26	-0.20	-0.15	-0.05	4800	1
Lockdown 3	-0.16	0.09	-0.34	-0.22	-0.16	-0.11	0.00	4840	1
Lockdown 4	-0.08	0.04	-0.16	-0.11	-0.08	-0.05	0.00	3819	1
Intercept	5.75	0.18	5.33	5.65	5.76	5.86	6.05	2069	1
Noise	0.15	0.01	0.14	0.15	0.15	0.16	0.17	4983	1

Table 6

First 6 variables displayed in Table 6 representing media channels impressions affect to sales are central part of MMM. Social media channel has the second biggest effect on sales among marketing channels. Social media plays vital role in consumers' routine worldwide, thus, such impact could have been expected. Display has lower affect to sales, than social media. Another internet related channel – Google, has the smallest impact than any other channel with mean of the posterior distribution – 0.05. Unsurprisingly, advertisement on TV is the biggest factor on sales. Print, similarly to Google, has low effect while Radio has the third highest posterior mean. Regarding control variables, CCI variable that is responsible in the model for economic environment has by far the biggest effect on sales. CCI strongly reflects the mood of the consumers, especially during COVID-19 times where uncertainty was extremely high. Unsurprisingly, variables that marked 4 different COVID-19 lockdown periods had negative posterior mean, which claim that implementing lockdowns had negative effect on company's sold contracts.

The fact that TV has the biggest effect on sales and radio has third largest effect, it corresponds well with funnel effect. It is well described by Johns et al. (2020) claiming that offline marketing can have a significant downstream effect on certain online channels. For example, after seeing a TV ad, a prospective customer would need to go online, where they might encounter search engine marketing (Google). Even though both channels have generated impression which is visible at our dataset, the bigger role of sold contract should be assigned to TV channel (offline marketing). On top of the funnel is considered to be offline channels such as TV or radio, mid-level is social media and, lastly, search engines such as Google.

Below in Table 7 there are plotted posterior distributions of marketing channels. One of the MMM with Bayesian statistics benefit is presenting uncertainty of the marketing channels' effect. For example, wide tails of the posterior distribution indicates there is similar probability for whole interval of values occurring, while narrow distribution means that one can be certain that coefficient value will likely occur between very short interval of values. The later problem is important for company's decision makers to know how certain one is about some model outputs. Table 3 shows Social and TV media channels have the widest tales meaning uncertainty is the highest, for instance, there is high probability that effect of the channel is lower than posterior's median or mean value shows. Media channels with lowest means have posteriors with the narrowest distributions.

Posterior distributions of marketing variables

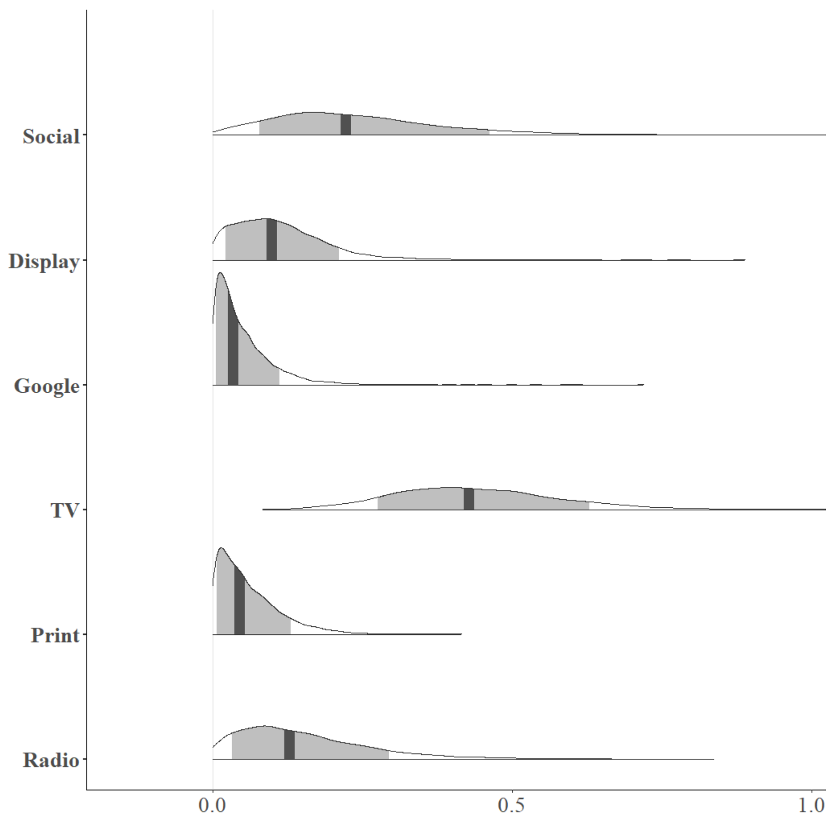


Table 7

There was also computed correlation between posteriors of the marketing channels coefficients. To save space, control variables were not included into the table.

Correlation table of posterior distributions

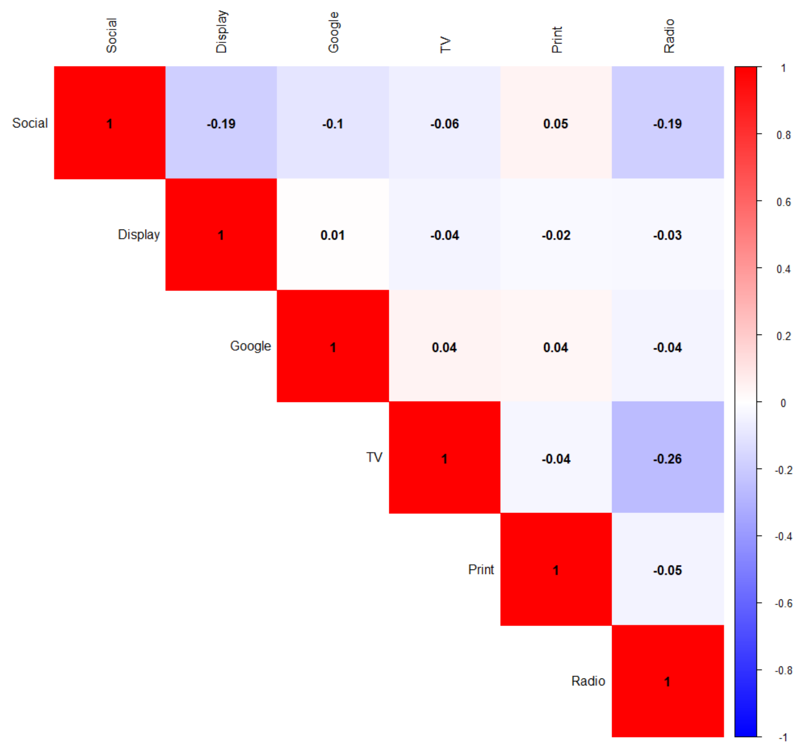


Table 8

Majority of correlations are negative and generally weak, indicating either weak associations or little correlation between those specific pairs of parameters. Needless to say, correlation between posterior parameters and correlation between raw impressions of marketing channels (see Data analysis section) differ since posterior beta coefficients indicate marketing channels impressions' effect on sales.

Regarding model performance, there were produced several indicators. Rhat is potential scale reduction statistic, also known as the Gelman-Rubin statistic, it compares the between- and within-chain estimates for model parameters. If chains have not mixed well, R-hat is larger than 1. Stan reports R-hat which is the maximum of rank normalized split-R-hat and rank normalized folded-split-R-hat, which works for thick tailed distributions and is sensitive also to differences in scale. According to the model results, Rhat for all the parameters is 1 which is acceptable. Moreover, effective sample size (ESS) is a metric used to assess the quality of the samples generated by MCMC algorithms. Due to autocorrelation between successive samples, not all of these samples are considered independent and contribute equally to the estimation of posterior properties. The effective sample size is an estimate of the number of independent samples that would provide the same amount of information as the actual correlated samples. In practice, for all parameters number of effective sample size is high enough to be confident with the model. Additionally, traceplots are plotted in Appendix 5 table. Initially traceplots show values varying from initial value, however, it quickly stabilizes, and it indicates that the Markov chain has converged to the posterior distribution. Initial values of majority of chains are way above means of posteriors. Also, no auto-correlation observed meaning that samples are independent. Overall, any chain does not stand out which is a good sign of convergence, the results are reliable.

Common practice in Bayesian modelling is to apply posterior predictive check. Detailed application and derivations are presented by Stan. If model is fitted well, it should be able to regenerate data that is very similar to the realized one. The data is generated from posterior predictive distribution. This is the distribution of the outcome variable implied by a model after using the observed data y (a vector of N outcome values) to update our beliefs about unknown model parameters θ. The posterior predictive distribution for observation $\tilde{y}$ can be written as

$$p(\tilde{y}|y) = \int p(\tilde{y}|\theta)p(\theta|y)d\theta. \tag{7}$$

Calculations are repeated on 500 simulated data sets. Graphical posterior predictive check regarding replications of the outcome *y* and observed *y* are displayed in the plot below.

Correlation table of posterior distributions

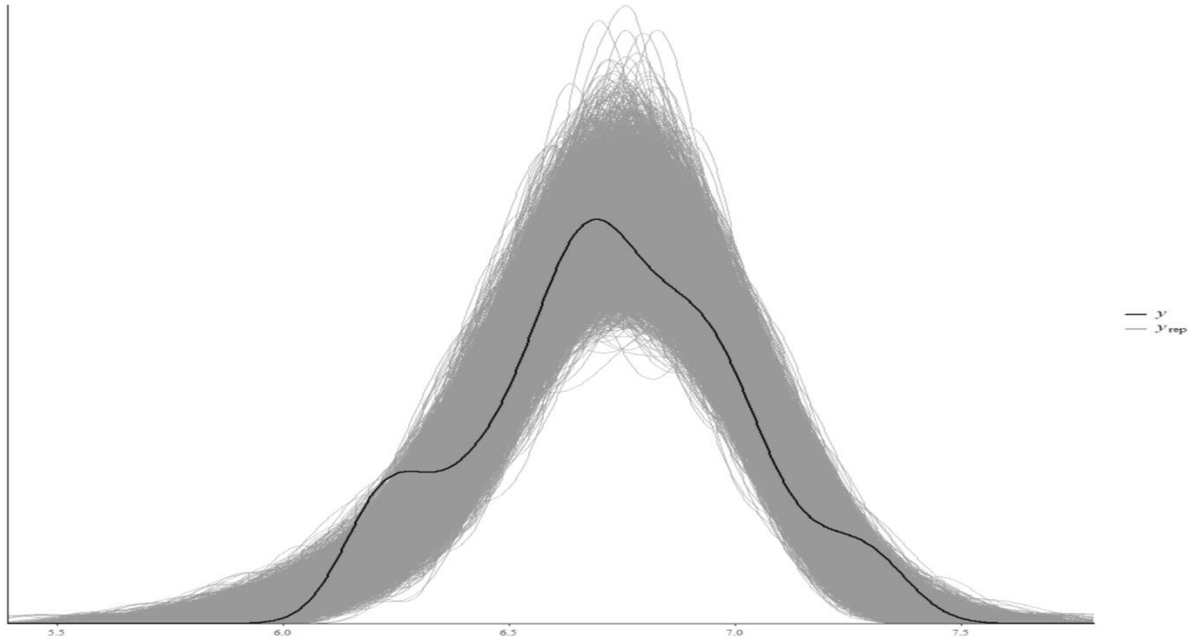


Table 9

Here the black solid line is the distribution of the observed outcomes y and each of the 50 grey lines is the kernel density estimate of one of the replications of *y* from the posterior predictive distribution. This plot makes it easy to see that this model accounts for the large proportion of realized *y* values. Furthermore, mean of replications of *y* posterior predictive distribution and 95% credible intervals were calculated for every time *t*. Together with realized *y* it was plotted in Table 10 for full train sample period. The graph indicates that the model does not catch some fluctuations over time. Realized data during pandemic times has high variation.

Posterior predictive check, response variable

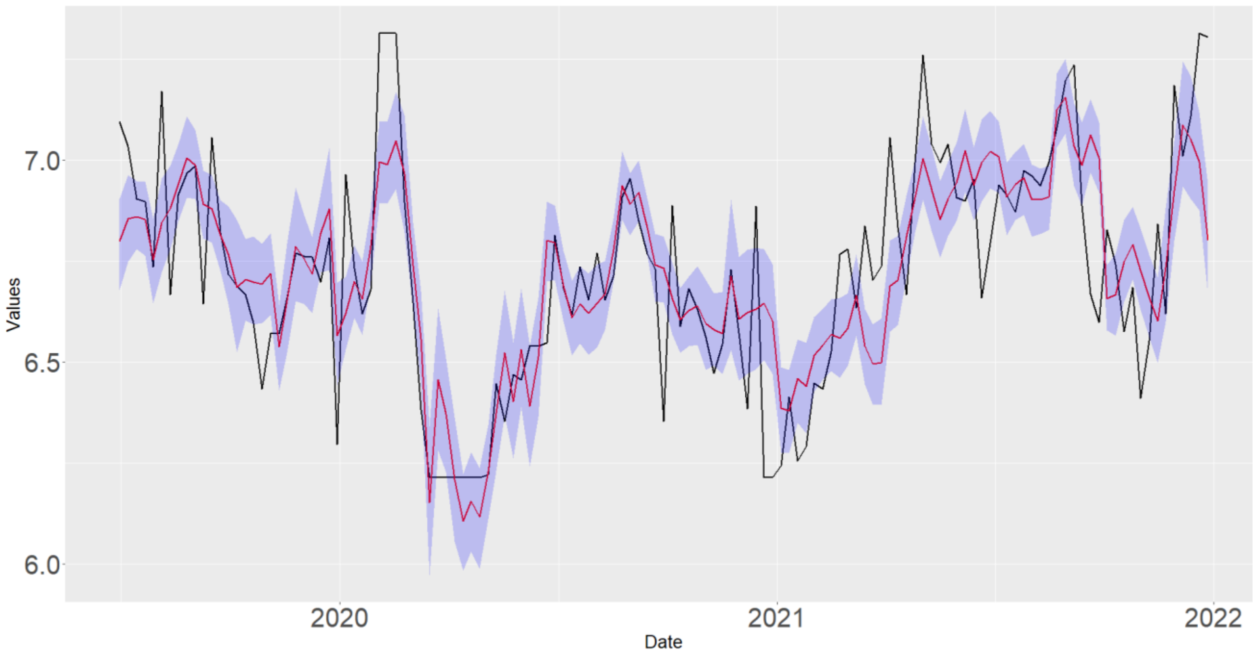


Table 10

Sensitivity check was conducted, mainly directed at marketing channels. As discussed in previous sections, marketing channels' impressions beta priors are positively truncated, thus, there was model run without truncated priors but with the same mean and standard deviation. The output of the regression is available in the Appendix 2 table. So, with prior normal(0,1) for every marketing channel, we observe increased standard deviation of the posterior distribution which is expected. Importantly, 4 marketing channels (Social, Google, Print, Radio) have negative posterior means which contradicts assumption of MMM. These results claim that spending on the marketing channels has negative effect on sales for some of the media channels. Average of TV and Display channels' posterior estimate has even increased comparing to the benchmark model. MAPE of the model on training data is 1.9%, slightly higher than for the main model. Moreover, number effective sample size has decreased for several marketing channels indicating that the MCMC samples are not providing as much information as the total number of iterations might suggest. There is a chance that the samples are not exploring parameters' space efficiently, leading to redundant information. Also, it may have high autocorrelation within drawn samples.

Additionally, Bayesian MMM regression was run with wide, however, positively truncated, marketing channels' priors. Prior for all the marketing channels is positively truncated normal with mean 0 and standard deviation 10. Positively truncated and not truncated control variables have prior normal (0,10), intercept has prior normal (0,10) and white noise – normal (0,10).

Appendix 4 table shows regression results. In fact, means of the posterior distributions of the marketing variables, as well as standard deviations, are very similar to the benchmark model that is described at the Methodology section. Noticeably, only TV parameter's posterior standard deviation is larger than in the main model. As a result, the observed data is providing strong information, and the prior's influence diminishes as the data becomes more informative. Both models have their posteriors influenced by the likelihood rather than priors (except truncation). Estimation of explanatory variables, especially marketing ones, is not sensitive to prior standard deviation in the given model. Thus, priors in this case help only to implement assumption that effect of marketing impression on sales is positive.

Adstock and saturation parameters

As already mentioned, MMM with Bayesian application additionally benefits from having its vital functions' parameters optimized through Bayesian regression. In the same manner as beta coefficients for marketing channels or any other predictor, retain rate from adstock function, slope and half saturation parameters are learned for each channel as part of fitting process. Table 11 shows model's posterior results of adstock and saturation parameters.

Posterior results of transformation parameters, main model

Variable	Mean	SD	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
Social retain rate	0.39	0.15	0.12	0.28	0.38	0.49	0.71	4219	1
Display retain rate	0.39	0.17	0.10	0.27	0.39	0.51	0.75	4072	1
Google retain rate	0.50	0.19	0.14	0.35	0.50	0.64	0.85	6185	1
TV retain rate	0.27	0.09	0.10	0.21	0.27	0.33	0.46	4651	1
Print retain rate	0.40	0.19	0.10	0.26	0.38	0.53	0.84	4201	1
Radio retain rate	0.39	0.16	0.11	0.27	0.38	0.50	0.70	5474	1
Social slope	2.37	1.39	0.47	1.39	2.09	3.07	5.87	4241	1
Display slope	3.03	1.81	0.49	1.68	2.74	4.03	7.34	5472	1
Google slope	2.70	1.74	0.39	1.41	2.36	3.62	7.06	6711	1
TV slope	3.89	1.62	0.98	2.80	3.74	4.84	7.53	3738	1
Print slope	2.67	1.70	0.47	1.39	2.32	3.59	6.92	6816	1
Radio slope	3.07	1.68	0.61	1.86	2.79	3.95	7.03	5318	1
Social 1/2 saturation point	0.62	0.20	0.20	0.47	0.63	0.78	0.94	4347	1
Display 1/2 saturation point	0.43	0.21	0.10	0.26	0.39	0.57	0.88	3694	1
Google 1/2 saturation point	0.53	0.23	0.09	0.36	0.54	0.71	0.92	5144	1
TV 1/2 saturation point	0.78	0.12	0.50	0.71	0.80	0.87	0.97	3206	1
Print 1/2 saturation point	0.52	0.24	0.05	0.35	0.54	0.71	0.92	4144	1
Radio 1/2 saturation point	0.62	0.20	0.18	0.50	0.64	0.77	0.93	4590	1

Table 11

Mean of posterior distribution for retain rate is lowest for TV channel. In practice, this indicates that carryover effect for latter channel is the weakest among other channels. For company it means that marketing effort at TV channel at week *t*-1 converts poorly to the next periods' sales. On the other hand, Google (last level of funnel effect) has largest carryover effect. Intuitively, if potential customer receives advertisement on Google during product searching, it means customer is already interested in the product and it may lead to potential sales in the next weeks.

Regarding Hill function, TV channel have the highest saturation slope, indicating that its effectiveness saturates quickly with increased spending. Also, TV reaches half saturation at a relatively high spending level. It is important for decision makers since marketers may need to carefully assess the cost-effectiveness of TV advertising, considering its higher saturation threshold. For instance, if advertisers seek for marketing channels with quick saturation, channels with low half saturation values should be selected. In fact, display marketing channel with lowest inflection point has quick saturation. From the model with given posterior statistics (Table 11) about saturation Hill function parameters (half saturation point and slope), saturation curve is plotted below.

Saturation curves

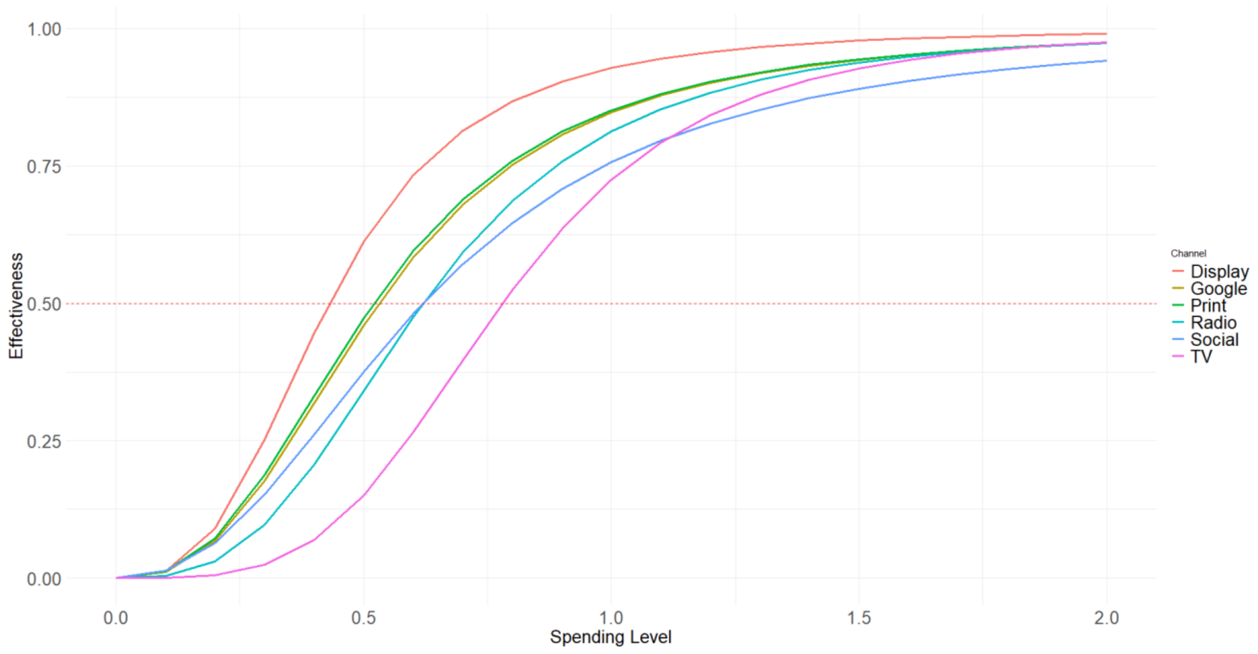


Table 12

Plot clearly corresponds what has been mentioned about quick effectiveness with low spending where display channel is clear favorite. Even though advertisement at TV is not efficient at the beginning, high level of spending could lead to satisfied results. Obviously, credible intervals could show uncertainty of the model. The distributions are plotted in the Appendix tables 6-8. Indeed, majority of posterior distributions (except TV channel) have relatively flat tales. Distributions are wide enough to warn decision makers about wide range of possible values of the parameter.

Contribution analysis

One of the main products of the MMM is contribution graph. It shows how much each of the channels (and control variables) contributed to sales at each week. Contribution is calculated as product of posterior mean of beta coefficient of the predictor and channel impressions. Exponential function is applied on the product since previously modelling was done with response variable log transformed. Decomposed sales on training data sample are presented As mentioned from the beginning CCI accounts for huge part of sales comparing among other predictors. Regarding marketing channels, one can clearly see that TV and social media channels are most dominant in terms of contribution. Also, it is visible that from the second part of 2020 (since COVID-19 pandemic started) social media channels contributed more to sales when comparing with the end of 2019. On the other hand, print marketing channel contributed to sales more before the pandemic started.

Contribution to sales

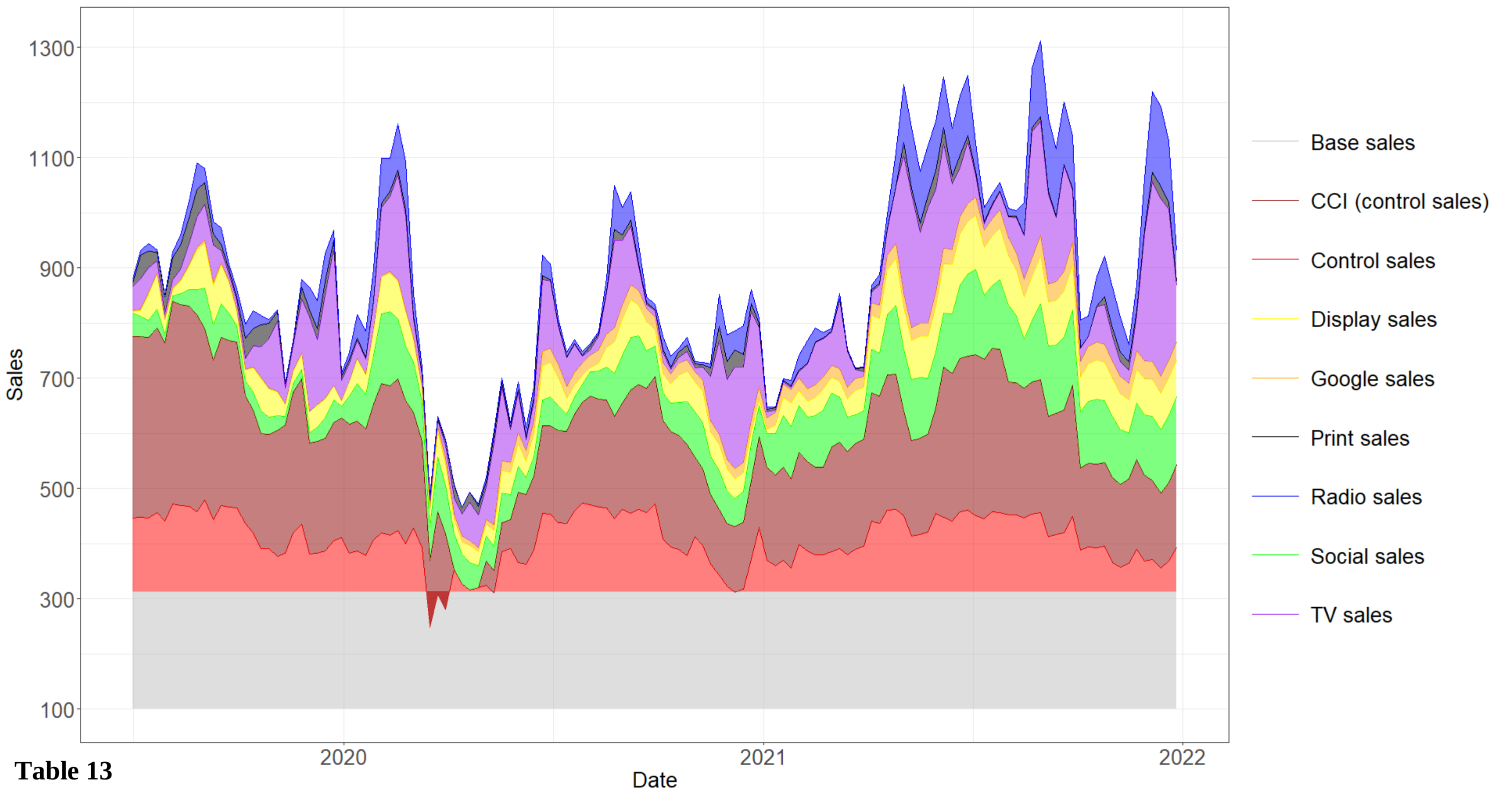


Table 13

With calculated contributions, it is possible to calculate Return on Advertisement Spend (ROAS). ROAS in general answers how much one makes for each dollar one spends on advertising campaign. For this study, it was calculated for each marketing channel m as:

$$ROAS_{m,t} = \frac{\beta_m * x_{m,t}}{Spending_{m,t}} \,. \tag{8}$$

Here x_m is weekly impression transformed by RStan with adstock and saturation functions. Spending variable represent weekly spends on channel m. Calculated ROAS as presented in the table 14 below:

ROAS for each channel

Social	Display	Google	TV	Print	Radio
1.73	1.4	2.23	1.55	19.19	5.97

Table 14

Print and Radio channels have outstanding figures comparing to others channels. It is mainly due to low spend on them. Among the rest, Google has the highest ROAS. Even though TV channel had the biggest effect on sales, however, ROAS is not the highest among others indicating high advertising costs of the channel.

Conclusion

Marketing Mix Modelling (MMM) using Bayesian methods is still young and undiscovered area. It is visible by low supply of researches in the literature. Medium website environment also contains limited number of mini articles with Bayesian applications in MMM. Mainly used open source programs for Bayesian inference (working with MMM) are Stan and PyMC. For R programing language, RStan application is mostly used. In this thesis, RStan application is used to run Bayesian Marketing Mix Modeling. The study applied Bayesian workflow to build the model, adjust priors, conduct sensitivity analysis, posterior predictive check, check model diagnostics results and interpret credible intervals.

Like every approach, MMM with Bayesian statistics has downsides. Firstly, Bayesian regression takes more time to estimate parameters than simple linear regression. It is relevant when regression is run with high number of predictors. Also, setting strict priors requires high confidence in industry. Wrongly selected priors may lead to false predictions. Thus, for this thesis weakly informative priors were selected.

On the positive side, Bayesian statistics offers more flexible and robust understanding in MMM interpretation. It allows incorporate knowledge from previous studies and expert knowledge into the model, thus, it accounts both for data and prior knowledge leading to single source of truth. Also, Bayesian methods help us deal with shorter, limited and inconsistent data which is important for smaller businesses. Model by providing posterior distributions of regression parameters informs decision makers about uncertainty of the estimate. Together with point estimates (mean) it is vital to consider credible intervals. The larger dataset, the more certainty model provides about estimates. Contemplating a range of potential outcomes, marketers can adjust decision making. Also, Bayesian regression can run complexed models, for example, in this thesis case it helped fitting parameters of the transformations (adstock and saturation).

The research successfully built marketing mix model with Bayesian regression. The model provided relevant estimates for MMM. To begin with, beta coefficient for every marketing channel was calculating which represented channels' contribution to sales. Since analysis period was mainly pandemic times, huge role in the model plays Consumer Confidence Index (CCI) which reflects high uncertainty at that time. The biggest contributor to sales is TV. Surprisingly, TV has relatively low retain rate calculated. While for other channels retain rate is higher (mean of the posterior), the posterior distribution has flat tales indicating likely possibility of wide range of values. Additionally, diminishing returns parameters from Hill function were learned from the model. All in all, quick saturation channel is regarded display channel, while slow saturation channel is TV. Only with high number of impressions TV channels become as effective as other channels.

Finalizing the study, a few research downsides and limitations are emphasized. First, model could include more informative predictors. For instance, company's promotions and price changes are expected to have an impact on sales. However, such data was lacking for this research. Secondly, posterior predictive check showed that model was not fully able to fit well such noisy sales data during COVID-19 pandemic period. Thus, student's t-distribution for the likelihood function should have been considered which has heavier tales, indicating higher probability of outliers. In addition, even though, priors in this study helped truncating beta coefficients, it was decided not to incorporate strict priors. As a result, the full potential of Bayesian statistics power in MMM has not been applied.

Summarizing model limitations, for further researches, it is suggested explore opportunities to implement more variables such as price and promotions. For noisy data, it is recommended to test different likelihood functions. Moreover, since one more MMM research in Lithuanian environment has been conducted, one can try applying stricter priors in the modelling phase.

References

Congdon, P.. *Bayesian Hierarchical Models: With Applications Using R, Second Edition.* 2017. CRC Press.

Facebook. (2023). *Prophet: Automatic Forecasting Procedure*.

Jin, Y., Wang, Y., Sun, Y., Chan, D., and Koehler, J.. *Bayesian Methods for Media Mix Modeling with Carryover and Shape Effects*. 2017. Google Inc.

Johns, M., Wang, Z., Dupont, B., and Fiaschi, L.. *Bayesian Media Mix Modeling using PyMC3, for Fun and Profit*. 2020. Medium.

Guadagni, P. M. & Little, J. D.. *A logit model of brand choice calibrated on scanner data*. 1983. Marketing science, 2 (3), 203–238.

Karlsson, J. Bayesian Structural Time Series in Marketing Mix Modelling. 2022. KTH.

Raiffa, Howard and Robert Schlaifer. *Applied statistical decision theory*. 1961. English. Division of Research, Graduate School of Business Adminitration, Harvard University Boston, 356 p.

Rossi A., Peter A., Greg A., and McCulloch R.. *Bayesian Statistics and Marketing*. Wiley Series in Probability and Statistics. 2012.

Stan Development Team. *Stan modeling language user's guide and reference manual*. 2015. version 2.10.0. Retrieved from http://mc-stan.org/.

Appendix

STAN Sampler Codes

Below is provided STAN code that runs MMM with Bayesian inference.

```
functions {

  real Hill(real t, real ec, real slope) {

      return 1 / (1 + (t / ec)^(-slope)); }}

data {

  int<lower=1> N;

  real<lower=0> Y[N];

  int<lower=1> K;

  int<lower=1> K2;

  int<lower=1> K3;

  row_vector[K2] X2[N];

  row_vector[K3] X3[N];

    matrix[N, K] X;}

parameters {

  real<lower=0> noise_var;

  real<lower=0> tau;

  vector<lower=0>[K] beta_medias;

  vector<lower=0>[K2] beta_pos;

  vector[K3] beta_neg;

  vector<lower=0,upper=1>[K] retain_rate;

  row_vector<lower=0,upper=1>[K] ec;

  vector<lower=0>[K] slope;}

transformed parameters {

  real mu[N];

    matrix[N,K] X_adstock;

 matrix[N,K] X_satur;

     X_adstock=X;

        for (media in 1 : K) {   for (nn in 2:N) {
```

```
    X_adstock[nn, media] = X_adstock[nn - 1,media] * retain_rate[media] +
X_adstock[nn, media];}}

    for (media in 1 : K) {

    for (nn in 1:N) {

     X_satur [nn, media] = Hill( X_adstock[nn, media], ec[media],
slope[media]);    } }

       for (nn in 1:N) {                  mu[nn] = tau +

      dot_product(X_satur[nn,], beta_medias) +

      dot_product(X2[nn], beta_pos)    +

      dot_product(X3[nn], beta_neg); }}

model {

  retain_rate ~ beta(3,3);

  slope ~ gamma(3, 1);

  ec ~ beta(2,2);

  tau ~ normal(0,3);

    for (media_index in 1 : K) {

    beta_medias[media_index] ~ normal(0,1);    }

  for (ctrl_index1 in 1 : K2) {

    beta_pos[ctrl_index1] ~ normal(0,0.3);  }

    for (ctrl_index2 in 1 : K3) {

    beta_neg[ctrl_index2] ~ normal(0,0.3);  }

 noise_var ~ normal(0,1);

    Y ~ normal(mu, (noise_var));}

generated quantities {

  vector[N] y_rep;

  for (nn in 1:N) {

    y_rep[nn] = normal_rng(tau +

      dot_product(X_satur[nn,], beta_medias) +

      dot_product(X2[nn], beta_pos)    +

      dot_product(X3[nn], beta_neg), (noise_var));  }}
```

Descriptive statistics of all variables

Variable	N	Mean	Std. Dev.	Min	Pctl. 25	Pctl. 75	Max
Launches	131	0.59	0.49	0	0	1	1
Payday	131	0.23	0.42	0	0	0	1
Christmas	131	0.031	0.17	0	0	0	1
Black Friday	131	0.031	0.17	0	0	0	1
CCI	131	-0.5	4.5	-16	-3	3	7
Stores	131	46	2.2	44	44	48	51
Seasonality	131	1.1	0.05	1	1	1.1	1.1
Competitor1	131	13647	11891	38	4238	19456	54067
Competitor2	131	2330	5542	0	0	1015	33627
S&P500	131	0.0037	0.03	-0.16	-0.0074	0.017	0.11
lockdown_1	131	0.069	0.25	0	0	0	1
lockdown_2	131	0.038	0.19	0	0	0	1
lockdown_3	131	0.031	0.17	0	0	0	1
lockdown_4	131	0.21	0.41	0	0	0	1

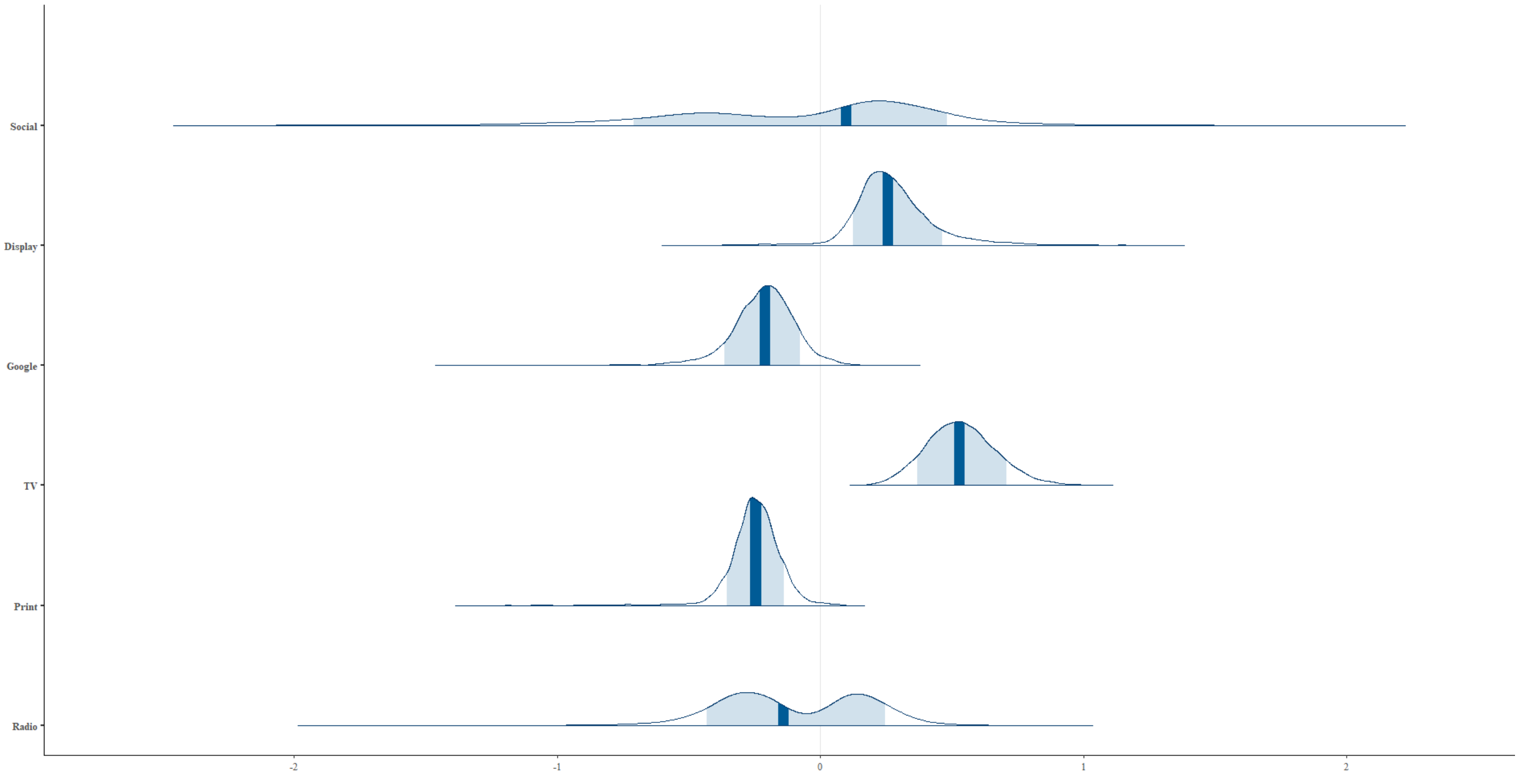
Appendix 1

Summary of the model results, marketing priors with no truncation

Variable	Mean	SD	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
Social	0.00	0.50	-1.10	-0.35	0.13	0.33	0.82	309	1.01
Display	0.28	0.18	0.00	0.18	0.26	0.35	0.71	970	1.00
Google	-0.22	0.13	-0.53	-0.28	-0.21	-0.15	0.00	1565	1.00
TV	0.54	0.14	0.29	0.45	0.54	0.63	0.83	1746	1.00
Print	-0.24	0.10	-0.42	-0.29	-0.24	-0.19	-0.08	821	1.01
Radio	-0.13	0.28	-0.62	-0.33	-0.18	0.11	0.36	317	1.01
Launches	0.07	0.03	0.01	0.05	0.07	0.09	0.13	2654	1.00
Payday	0.01	0.01	0.00	0.00	0.01	0.02	0.05	7865	1.00
Christmas	0.07	0.05	0.00	0.03	0.06	0.10	0.18	5137	1.00
Black Friday	0.08	0.05	0.00	0.03	0.07	0.11	0.20	5042	1.00
CCI	0.39	0.13	0.14	0.30	0.39	0.48	0.65	1403	1.00
Kiosks	0.04	0.03	0.00	0.01	0.03	0.05	0.13	5929	1.00
Seasonality	0.16	0.04	0.08	0.13	0.16	0.19	0.24	1851	1.00
Competitor1	0.10	0.08	-0.06	0.04	0.10	0.15	0.26	3553	1.00
Competitor2	0.01	0.09	-0.16	-0.04	0.02	0.07	0.18	3735	1.00
S&P 500	0.28	0.12	0.05	0.20	0.28	0.36	0.50	5328	1.00
Lockdown 1	-0.37	0.10	-0.56	-0.43	-0.36	-0.30	-0.19	755	1.00
Lockdown 2	-0.20	0.08	-0.35	-0.25	-0.20	-0.14	-0.04	3577	1.00
Lockdown 3	-0.02	0.10	-0.21	-0.08	-0.02	0.05	0.17	2726	1.00
Lockdown 4	-0.04	0.05	-0.14	-0.07	-0.04	-0.01	0.06	1111	1.00
Intercept	6.36	0.37	5.75	6.13	6.32	6.55	7.29	582	1.01
Noise	0.14	0.01	0.13	0.14	0.14	0.15	0.16	3589	1.00

Appendix 2

Posterior distributions, marketing priors with no truncation



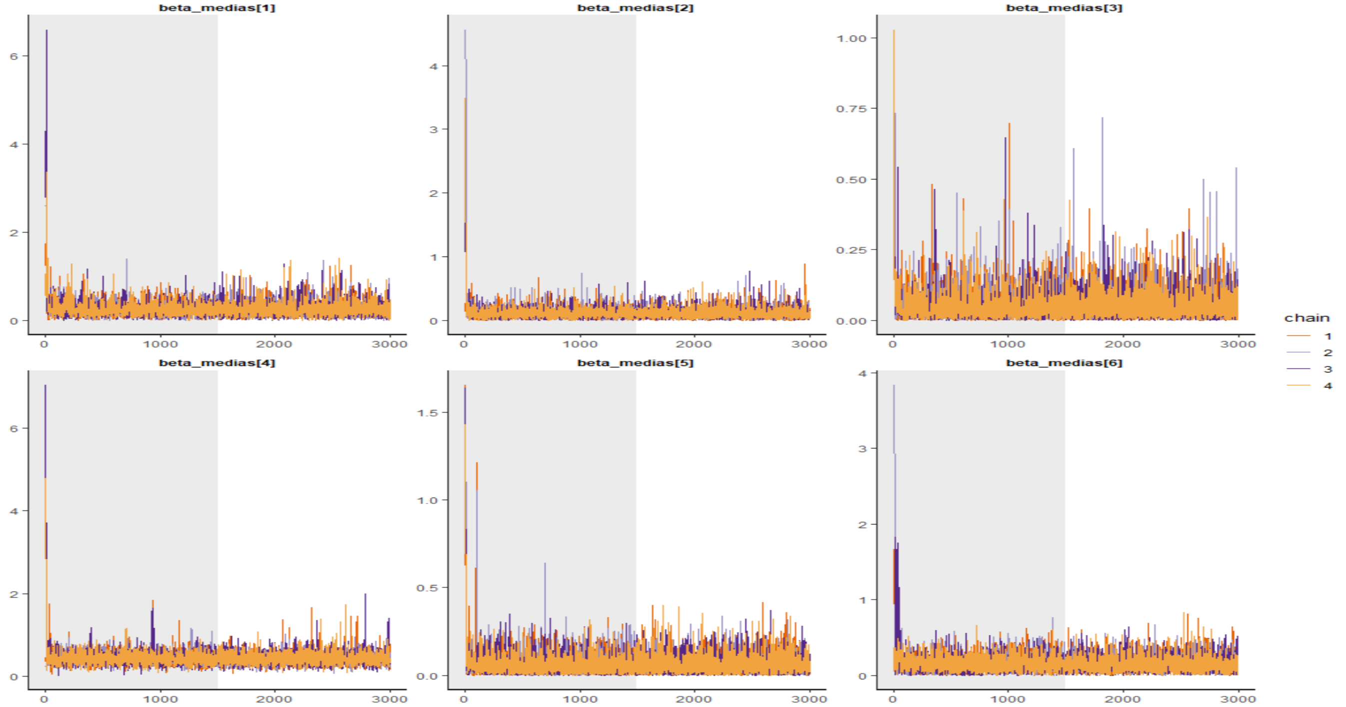
Appendix 3

Summary of the model results, flat priors

Variable	Mean	SD	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
Noise	0.27	0.17	0.03	0.16	0.24	0.35	0.69	1367	1.00
Intercept	0.11	0.08	0.01	0.05	0.09	0.15	0.30	1809	1.00
Social	0.06	0.08	0.00	0.02	0.04	0.08	0.23	608	1.00
Display	0.50	0.26	0.24	0.38	0.46	0.57	0.93	268	1.00
Google	0.06	0.06	0.00	0.02	0.05	0.09	0.21	2187	1.00
TV	0.13	0.11	0.01	0.05	0.11	0.18	0.41	1523	1.00
Print	0.07	0.03	0.01	0.05	0.07	0.09	0.13	1895	1.00
Radio	0.02	0.01	0.00	0.00	0.01	0.02	0.05	4100	1.00
Launches	0.06	0.05	0.00	0.02	0.05	0.09	0.18	3009	1.00
Payday	0.09	0.06	0.01	0.04	0.08	0.13	0.23	2481	1.00
Christmas	0.62	0.11	0.40	0.55	0.63	0.70	0.84	1469	1.00
Black Friday	0.08	0.06	0.00	0.03	0.07	0.11	0.21	1408	1.00
CCI	0.16	0.03	0.09	0.13	0.16	0.18	0.22	1968	1.00
Kiosks	0.02	0.08	-0.14	-0.04	0.02	0.07	0.17	2139	1.00
Seasonality	0.08	0.09	-0.09	0.02	0.08	0.14	0.25	3347	1.00
Competitor1	0.27	0.12	0.03	0.19	0.27	0.36	0.51	2416	1.00
Competitor2	-0.38	0.08	-0.54	-0.44	-0.38	-0.33	-0.22	1645	1.00
S&P 500	-0.21	0.08	-0.37	-0.26	-0.21	-0.16	-0.06	2597	1.00
Lockdown 1	-0.18	0.09	-0.36	-0.24	-0.18	-0.12	0.00	2506	1.00
Lockdown 2	-0.09	0.04	-0.17	-0.11	-0.09	-0.06	0.00	1725	1.00
Lockdown 3	5.65	0.22	5.15	5.55	5.68	5.79	6.00	360	1.01
Lockdown 4	0.15	0.01	0.14	0.15	0.15	0.16	0.17	2578	1.00

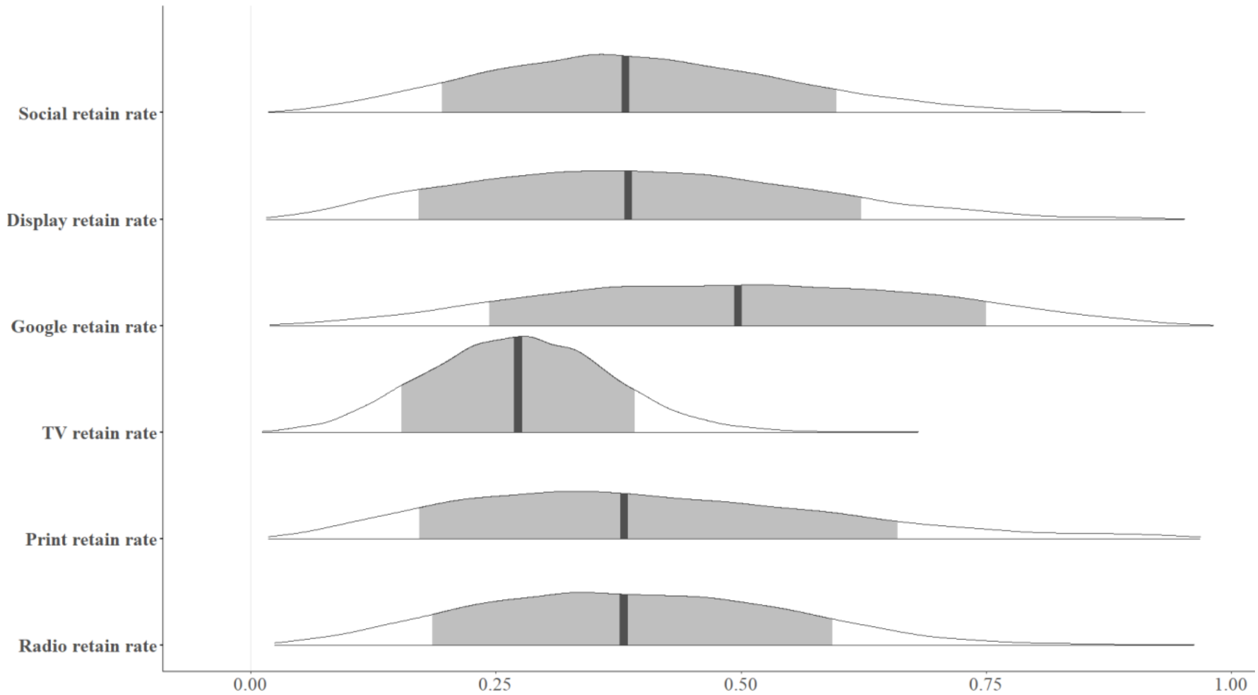
Appendix 4

Traceplot of media coefficients, main model



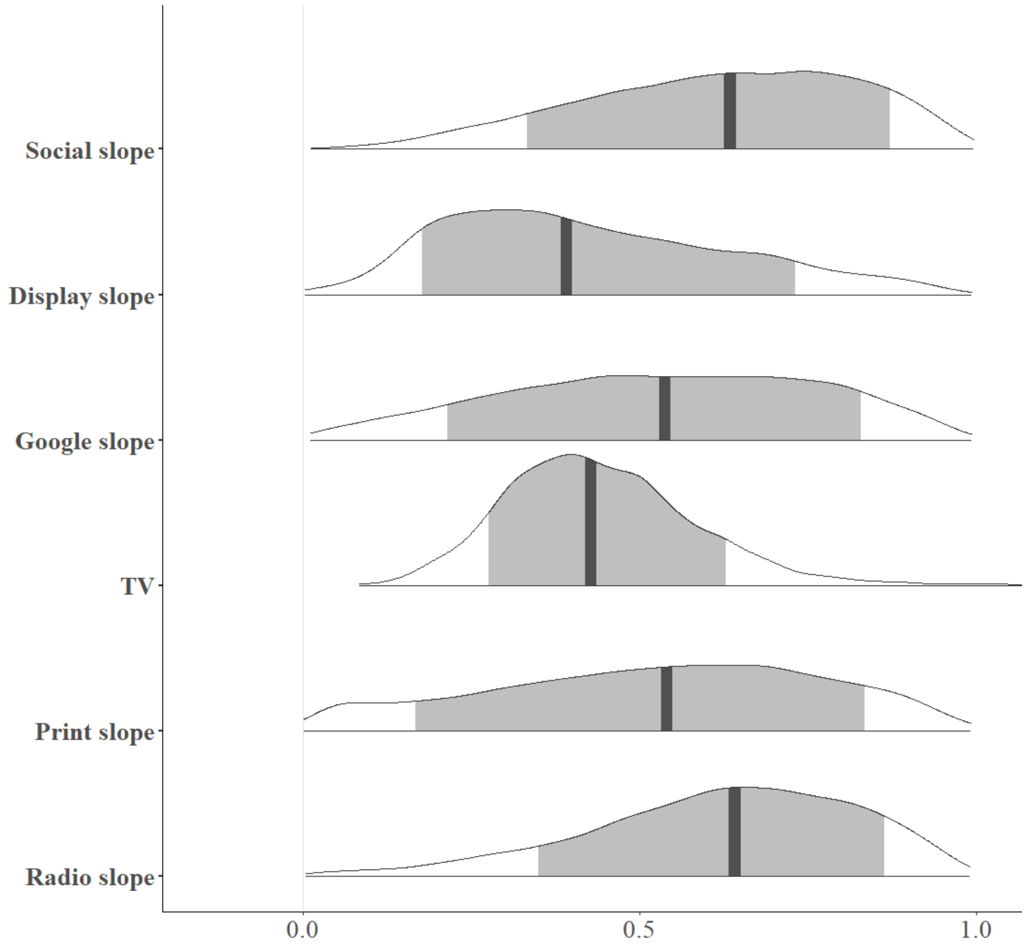
Appendix 5

Posterior distributions of retain rate, main model



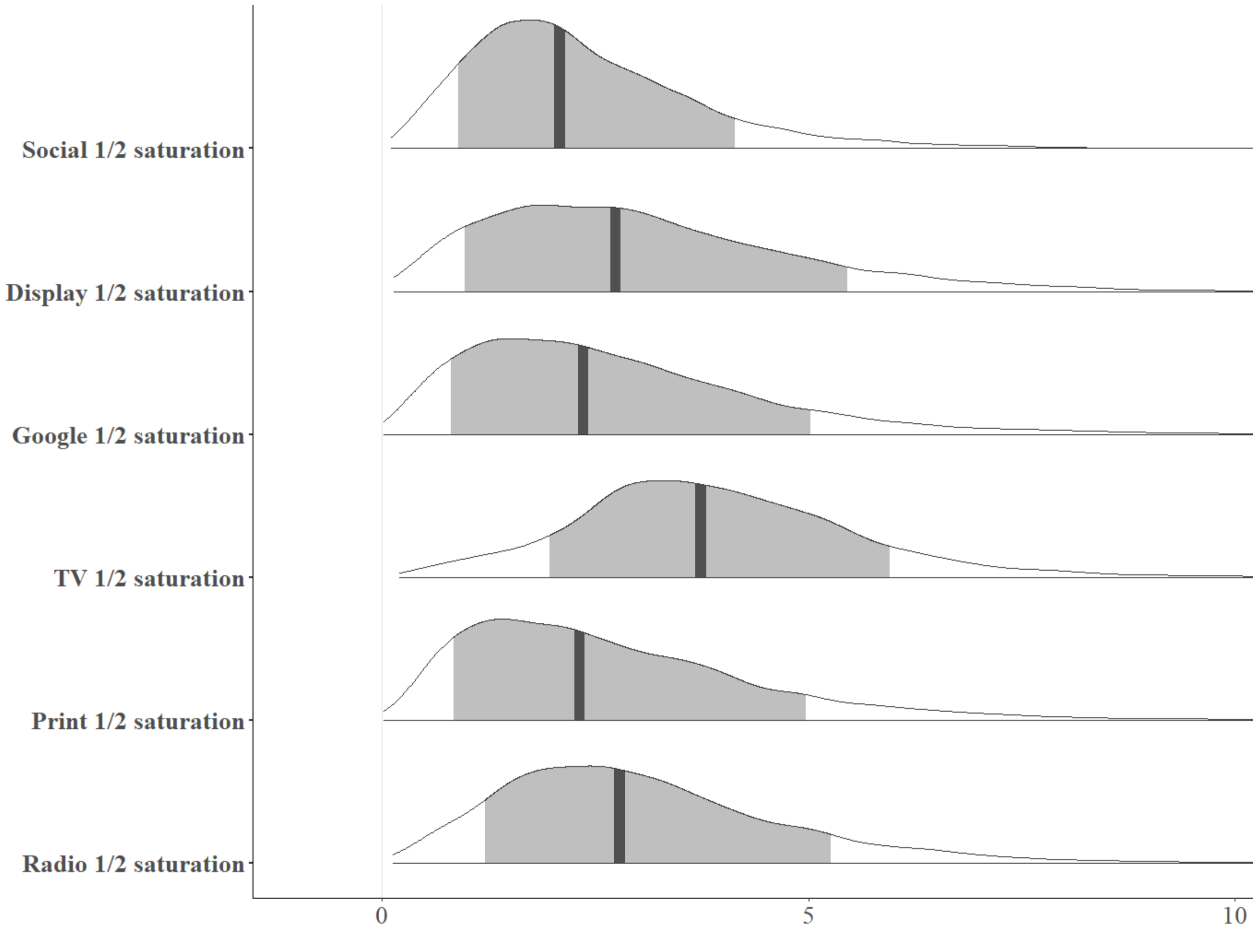
Appendix 6

Posterior distributions of the slope, main model



Appendix 7

Posterior distributions of the half saturation point, main model



Appendix 8