Analysing the Bayesian Mixed Media Model

Channel Strategy Investigation Report

Author: Ammar Mahmood

18. April 2025

Abstract

This report presents a Bayesian Mix Media Model (MMM) built using PyMC to evaluate the effectiveness of marketing efforts across seven advertising channels for a fictional online retail company. The aim is to quantify the impact of each channel on weekly sales revenue, while accounting for lagged marketing effects (adstock), seasonality, trends, and campaign spikes. This probabilistic model allows for uncertainty quantification, and helps guide future marketing investments by estimating each channel's return on investment (ROI).

Introduction

As businesses adapt to the rapid evolution of Generative AI, they are increasingly exploring innovative and sustainable strategies for growth and decision-making. The integration of Generative AI into areas such as supply chains, management systems, and marketing operations is accelerating at an unprecedented pace. At the core of these advancements lie probabilistic methods — including probability distributions and Markov Chain Monte Carlo (MCMC) — which form the mathematical foundation for next-word prediction, uncertainty modeling, and reasoning in today's large language models (LLMs).

Looking more closely at marketing operations, multinational companies face increasing complexity in managing multi-million dollar budgets. Allocating spend across various online and offline channels, evaluating their effectiveness, and estimating ROI under uncertainty is a significant challenge. This is precisely where Bayesian methods prove invaluable. Companies like HelloFresh, which invest over $300 million annually () in marketing across diverse channels, require data-driven strategies to optimise budget allocation. Techniques such as the Bayesian Marketing Mix Model (MMM) enable these organisations to quantify channel impact, forecast returns, and make informed investment decisions under uncertainty.

In this report, we focus on the Bayesian Marketing Mix Model (MMM) inspired by the approach proposed in Google Research (Jin et. al 2017). This probabilistic framework is used to analyse and model our dataset, enabling us to answer key questions related to channel-level ROI estimation, marketing performance, and effectiveness across media channels. Through this model, we explore insights such as the spread between prior and posterior distributions, the impact of adstock decay and lagged effects, and broader strategic patterns that inform effective marketing decision-making.

Methodology

A structured methodology was adopted by incorporating advanced feature engineering with a Bayesian probabilistic modeling framework. However before this, an exploratory data analysis (EDA) was performed to gain meaningful insights into the mean, mode and median as well as the seasonality and trends that the data followed. For feature engineering, we began with a weekly marketing dataset consisting of spend across seven media channels and corresponding revenue's. The following preprocessing steps were applied:

- **Adstock Transformation:** Each channel's spend was processed using a geometric adstock function to model lagged marketing effects over time with a maximum length (L_max) of 12 weeks.
- **Structural Zero Handling:** Channel 6 contained significant zero-heavy periods. Here binary indicators (channel_1_active, channel_6_active) were created to distinguish campaign-active vs inactive weeks.
- **Lagged Features:** For all channels, 1-week and 2-week lagged spend variables were generated to capture delayed responses not fully explained by adstock.
- **Normalization:** All spend features, including lagged values, were scaled by their respective mean values to stabilise the model estimation and ensure comparability across features.
- **Additional Controls:** A trend feature and a peak_flag binary indicator were added to account for temporal trends and revenue spikes.

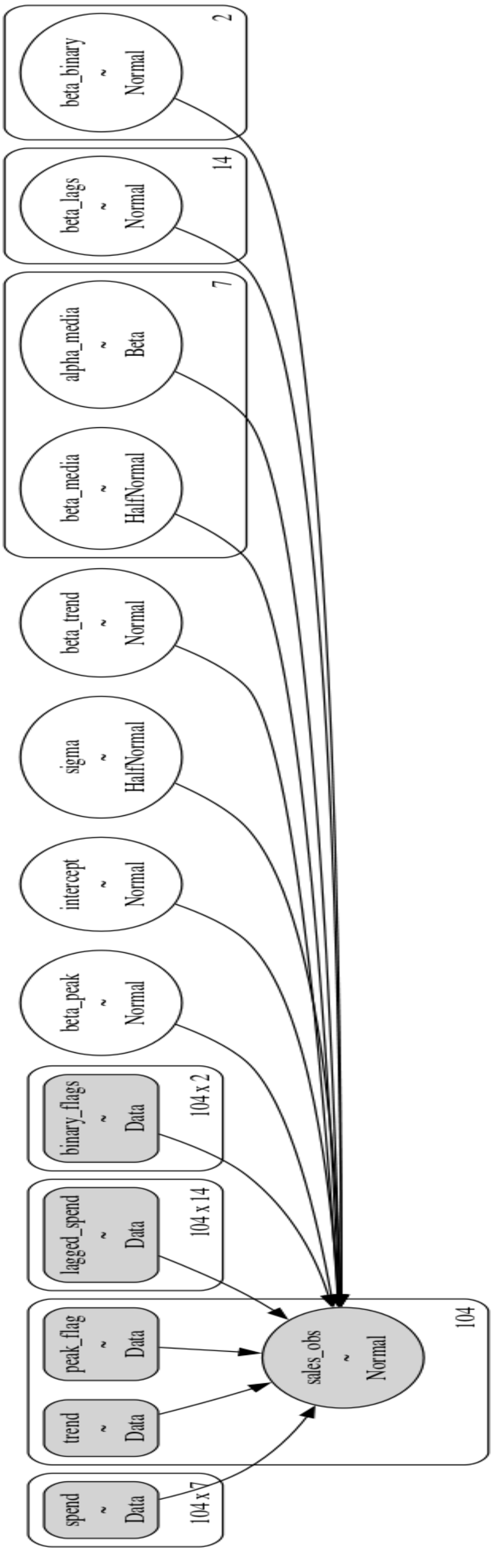


Figure 1: Graph Visualisation of the Bayesian Mixed Media Model

Following the pre-processing steps, the bayesian mix media model (with PyMC) was configured as shown in figure 1, according to the following key components:

- **Priors:** Weakly informative priors were applied using HalfNormal distributions for media effects (`beta_media`) and Normal priors for control effects. A Beta prior was used for adstock decay (`alpha_media`).
- **Likelihood:** The observed revenue was modeled as a normal distribution centered on the sum of baseline controls, media contributions (via adstock), lagged effects, and binary campaign indicators.
- **Sampling:** Inference was conducted using Hamiltonian Monte Carlo (HMC) via the No-U-Turn Sampler (NUTS), with 4 chains, 1000 draws each, and a high `target_accept` rate of 0.99 to ensure convergence and reduce divergences.

The following formulae's were used as part of the model for geometric adstock and sales revenue (y_t):

$$\text{Adstock}_{t,i} = \sum_{l=0}^{L-1} \alpha_i^l \cdot x_{t-l,i}$$

$$y_t = \mu + \sum_i \beta_i \cdot \text{Adstock}_{t,i} + \gamma_1 \cdot \text{trend}_t + \gamma_2 \cdot \text{peak}_t + \boldsymbol{\delta}^\top \cdot \text{lagged}_t + \boldsymbol{\theta}^\top \cdot \text{flags}_t + \epsilon_t$$

Results and Discussions

Seasonality and Trend Analysis

An exploratory data analysis was revealed major and minor cycles which incurred throughout the year (52 week period), these fluctuations significantly impact the results of the model hence they were plotted and a visual analysis was conducted to begin the analysis as shown in figures 2 and 3.

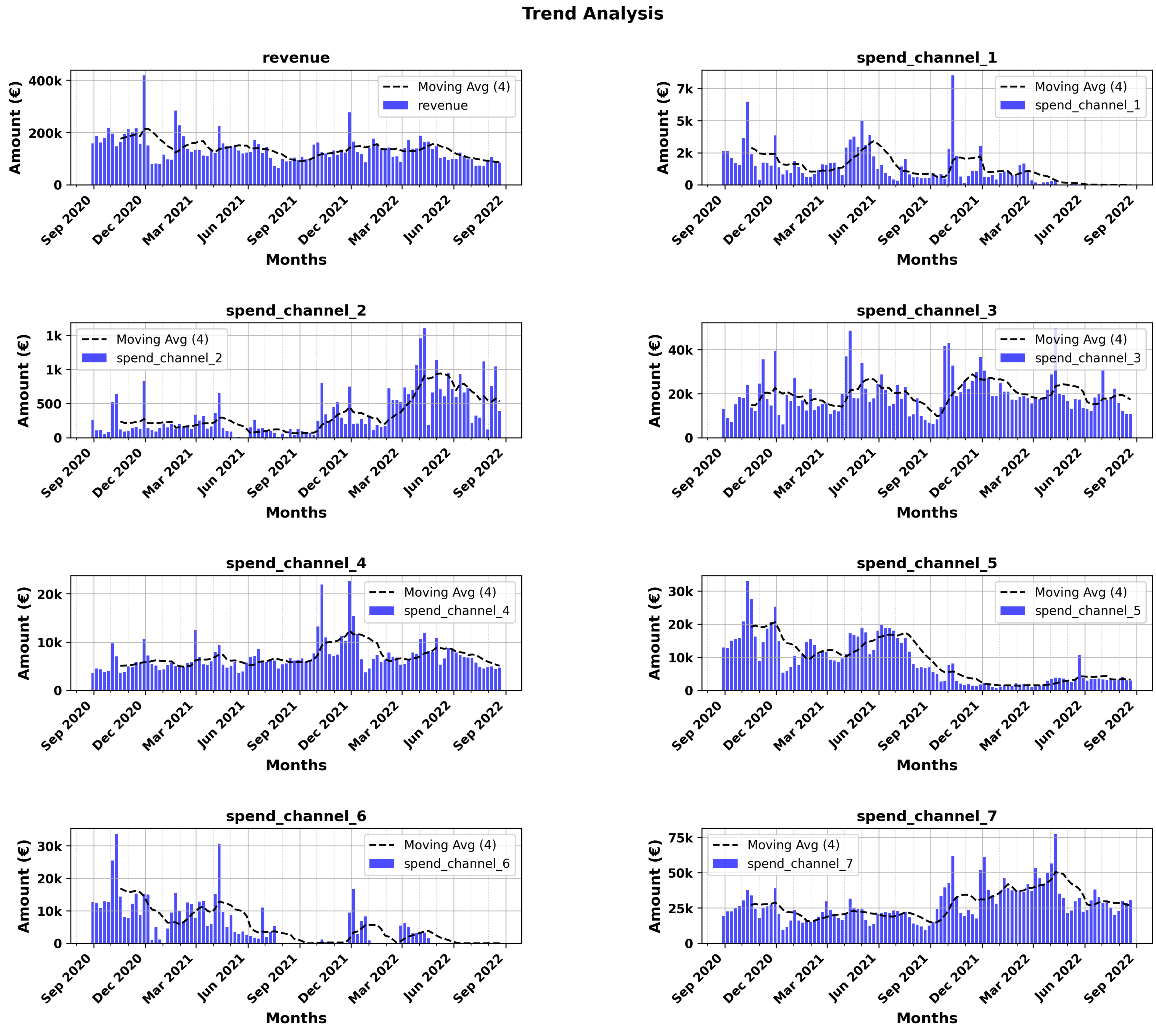


Figure 2: This is a time series trend plot (with a moving average of 8 weeks) depicting amount of money earned (revenue) or spent (spend channels 1-7) between 2020 and 2022 for company X.

In figure 2, general information regarding expenditure fluctuations or stale periods can be be seen of different media channels for company x and also the peaks and drops can be observed for revenue as well. Revenue seems to be declining year on year and this can be attributed to any number of reasons such as effective channel strategy which will a the focus of our research in this report.

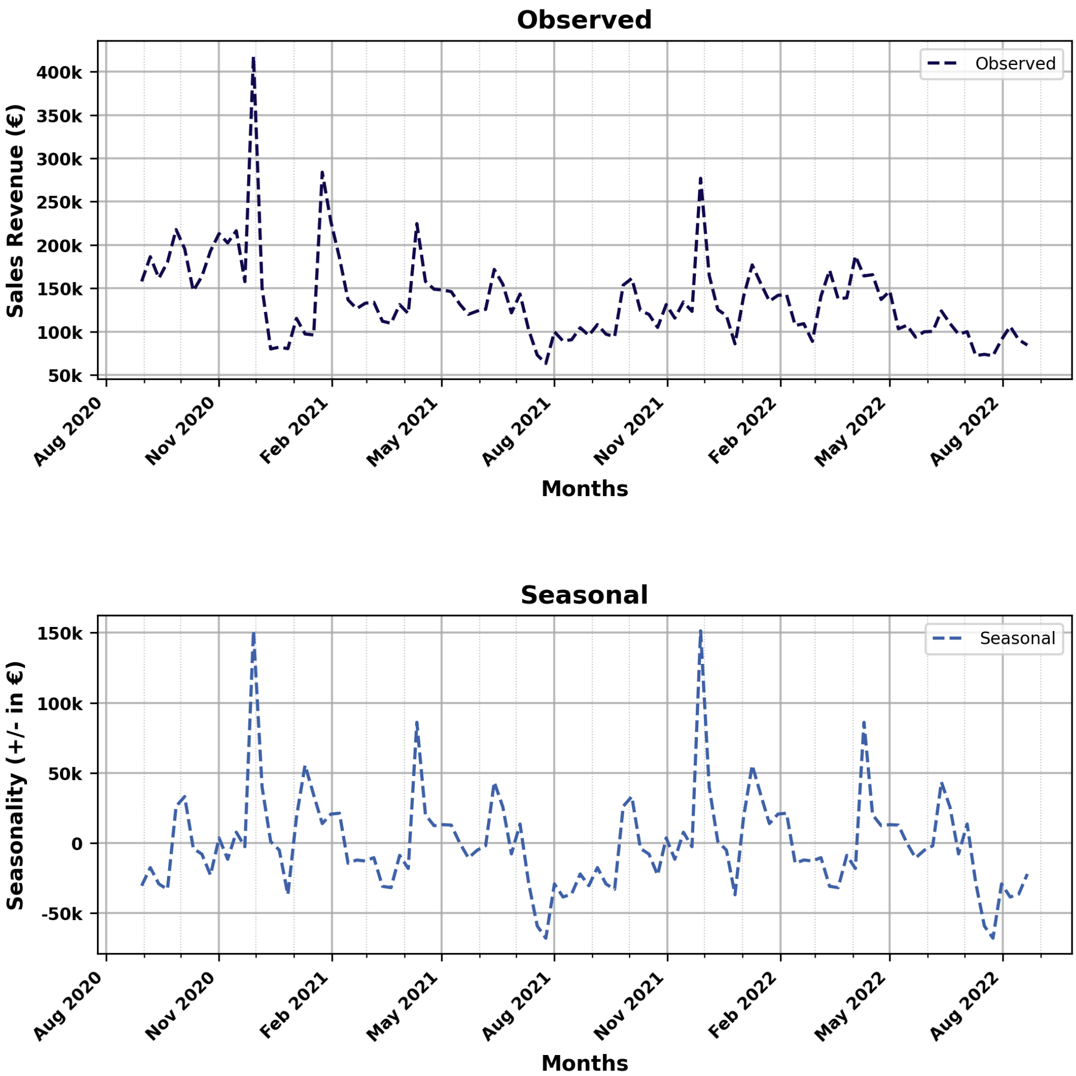


Figure 3: This a plot between observed revenue spikes and seasonality peaks and troughs between 2020 and 2022 for company X.

In figure 3, a seasonal plot is drawn which shows consistent peaks and dips throughout a year (52-week period). This clearly indicates business cycles and explains why seasonality is an important control variable in our model. In late November (29th), there is a clear indicator or sales spike perhaps due to the black friday event which occurs every year. Similarly, another peak can be observed in April which may be a result of easter event. This figure also shows dips in certain periods for instance during Christmas holidays when everybody is spending time with their families and not a lot of people are shopping from online retailers.

Before analysing posteriors, figure 4 indicates the priors that are considered for the model configuration. A half normal distribution and beta distribution prior is shown here.

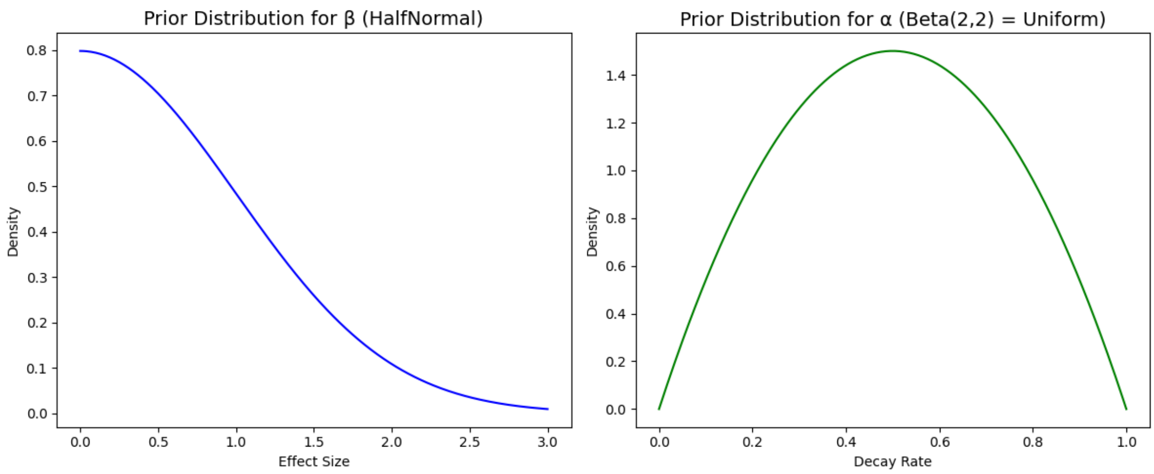


Figure 4: These are priors for the model. On the left, is a half normal distribution for beta_media (β_i) and on the right, a beta distribution for alpha_media (α_i)

Following model convergence, figure 5 indicates the posterior plots for both beta_media and alpha_media coefficients. From these plots, it can be observed that beta media coefficients for channels 1 and 2 are relatively lower effect size, which is how much a unit spend in these channels increases sales revenue. The beta coefficients of channels 3 and 7 are the impactful and have strong signals overall due to their higher effect size of 0.123 and 0.132. Channels 4 to 6 have a moderate impact contributing to the overall advertising spent. Our prior distribution for the alpha coefficients follows a bell shaped behaviour which is more suitable to choose as we do not have much information about our spending channels. By observing the

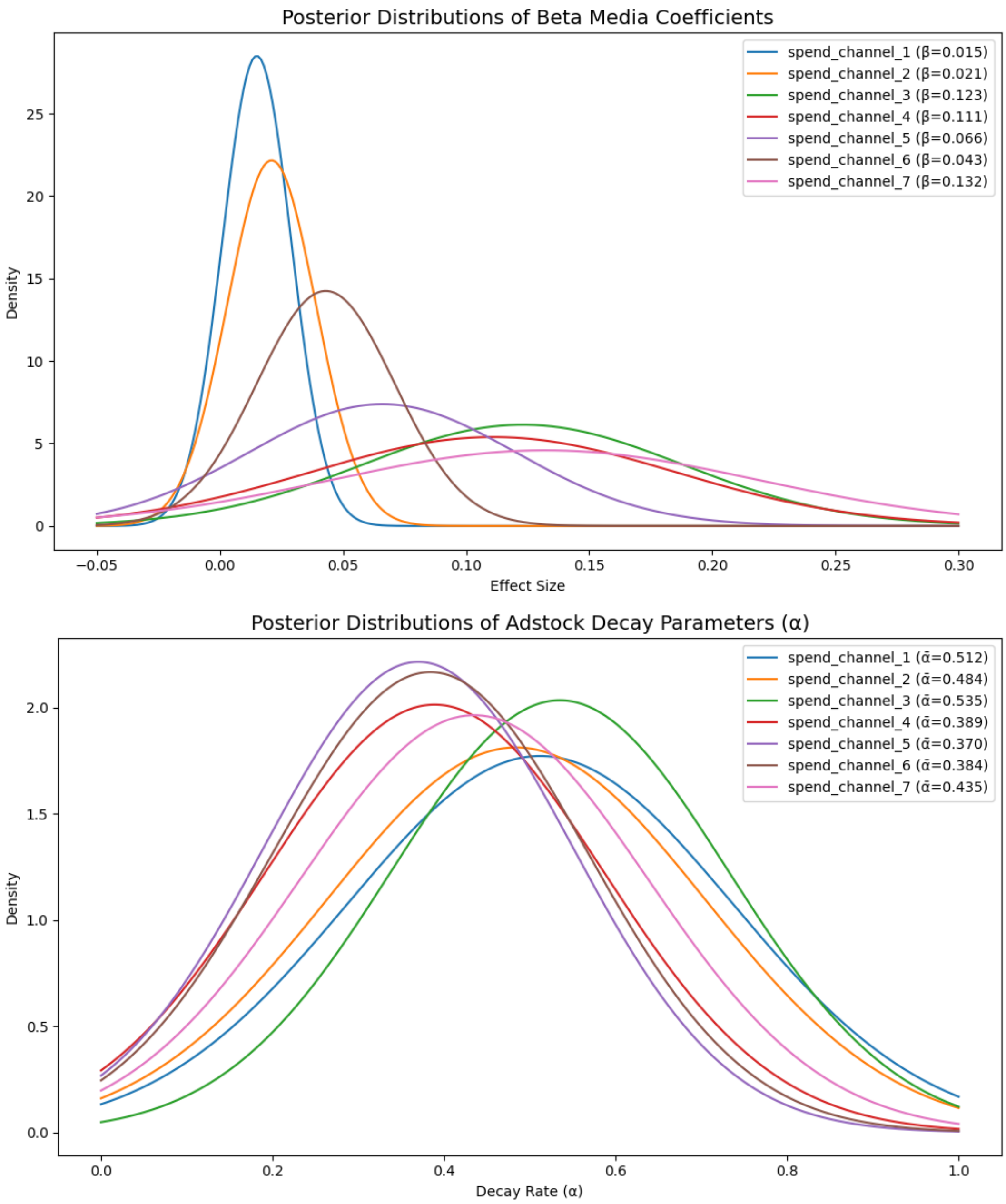


Figure 5: These plots depict posterior probability distributions known as Kernel Density Estimations (KDEs), for each media channels beta and alpha coefficients (and their mean values).

posterior distribution for alpha coefficients, some channels are a wider bell indicating long lasting effects of advertisements on these channels than others. For instance, channel 3 has the longest carry over of 0.535 compared to channel 5 and 6 with the smallest decay rate between 0.37-0.38. Similarly, channel 1 has a rate of 0.512 and channel 7 is moderate with a rate of 0.435. Here larger alpha values indicate longer lasting spend which are important for campaign planning and budgeting.

However in order to truly understand the results and the implications of the model, the Return On Investment (ROI) was calculated based on the following formulae.

$$\text{ROI}_i = \frac{\beta_i}{\text{MeanSpend}_i} \cdot \text{RevenueScaler}$$

Here the mean spend is the mean unscaled revenue for channel i and the revenue scaler is the average raw revenue used to rescale the predictions.

Assuming a normal distribution (whereby 95% of values lie within +/- 1.96 standard deviations) for ROI estimates, the following formulae was implemented in order to establish credible intervals for lower and upper bounds.

$$\text{ROI}_{3\%} = \mu_{\text{ROI}} - 1.96 \cdot \sigma_{\text{ROI}}$$

$$\text{ROI}_{97\%} = \mu_{\text{ROI}} + 1.96 \cdot \sigma_{\text{ROI}}$$

This gives a quick and interpretable approximation to the true 94 % Highest Density Interval. Here mean and standard deviation are calculated using the the ROI formulae given before by replacing beta_mean and beta_sd (Refer to Appendix A given below).

Mixed Media Model ROI and Uncertainty Data

channel	beta (posterior mean)	beta_sd	mean_spend	ROI_mean	ROI_sd	ROI_3%	ROI_97%
spend_channel_1	0,015	0,014	1245,60	1,64	1,53	(1,36)	4,65
spend_channel_2	0,021	0,018	343,64	8,34	7,15	(5,67)	22,35
spend_channel_3	0,123	0,065	19507,18	0,86	0,45	(0,03)	1,75
spend_channel_4	0,111	0,074	6915,14	2,19	1,46	(0,67)	5,05
spend_channel_5	0,066	0,054	8575,61	1,05	0,86	(0,63)	2,74
spend_channel_6	0,043	0,028	5063,70	1,16	0,75	(0,32)	2,64
spend_channel_7	0,132	0,087	27701,37	0,65	0,43	(0,19)	1,49

Table 1: This table highlights the ROI and its uncertainty based on the beta posteriors from the model for each channel.

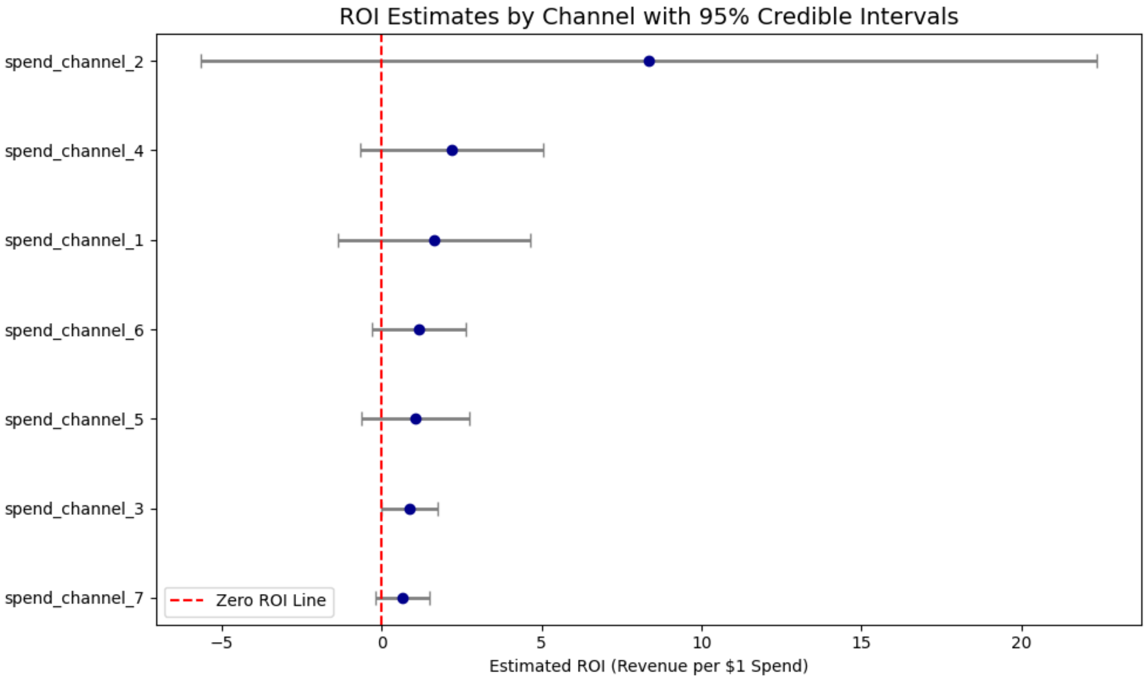


Figure 6: This is a forest plot for the ROI (Revenue per dollar spend) for each channel.

The ROI table and forest plot is shown in table 1 and figure 6, along with its uncertainty. From the data shown in figure 6 and table 1, it can be incurred that the model higher ROI mean for channel 2 of 8.34, however with very large credible interval size from -5,67 % to 22,35 %. This variance between the ROI estimates for this channel make difficult to invest into it, as a higher profit or a huge loss can happen due to the uncertainty of the channel. On the contrary, channels 3-7 have a higher potential to perform better with some negative bounds however mostly positive returns. The safest investment is channel 3 which barely touches the break even or the zero return on investment line. Channel 1 also performs better compared to channel 2 however their is a risk associated with it.

There are some limitations that need to be addressed as well regarding the model. The model assumes a linear spend–response relationship without saturation (diminishing-returns) curves, which may mis-estimate ROI at high spend levels. It treats each channel independently with no hierarchical pooling, reducing robustness for sparse data, and omits interaction terms that could capture synergies. Media-effect and carryover parameters are static, ignoring shifts in channel performance over time.

Conclusion

In summary, the Bayesian MMM converged cleanly and clearly identified **Channel 3** as the top investment (highest and most reliable ROI), followed by **Channel 7** and then Channels 4–6 (all showing positive expected returns). Channels 1 and 2 delivered low or highly uncertain ROI—Channel 2 in particular carries too much risk for heavy spend. These insights point to reallocating budget toward Channels 3 and 7 for maximum and predictable return, while treating Channels 1–2 as lower priorities or subject to further testing.

References

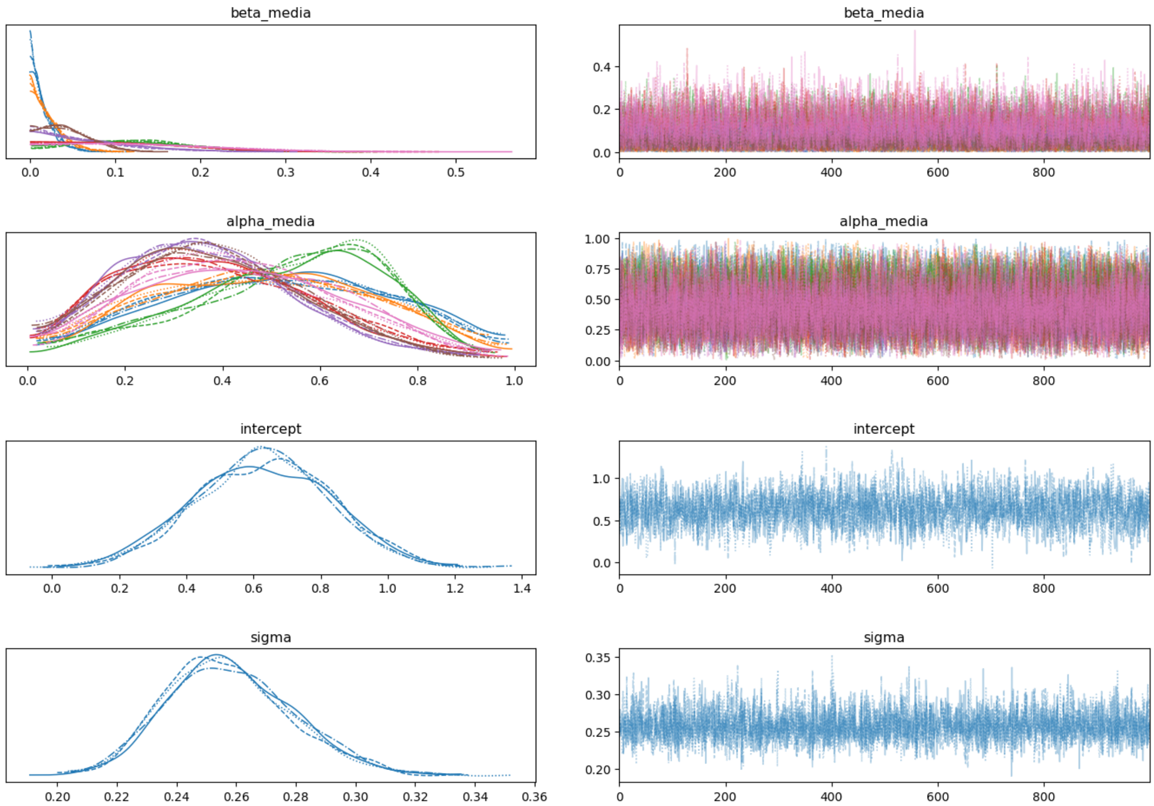
1. Jin, Y., Wang, Y., Sun, Y., Chan, D., & Koehler, J. (2017). Bayesian methods for media mix modeling with carryover and shape effects. *Google Research*, 1-34.
2. *Bayesian Marketing Mix Models: State of the art and their future* (2022) *PyMC Labs*. Available at: https://www.pymc-labs.com/blog-posts/2022-11-11-HelloFresh/ (Accessed: 18 April 2025).

Appendix A

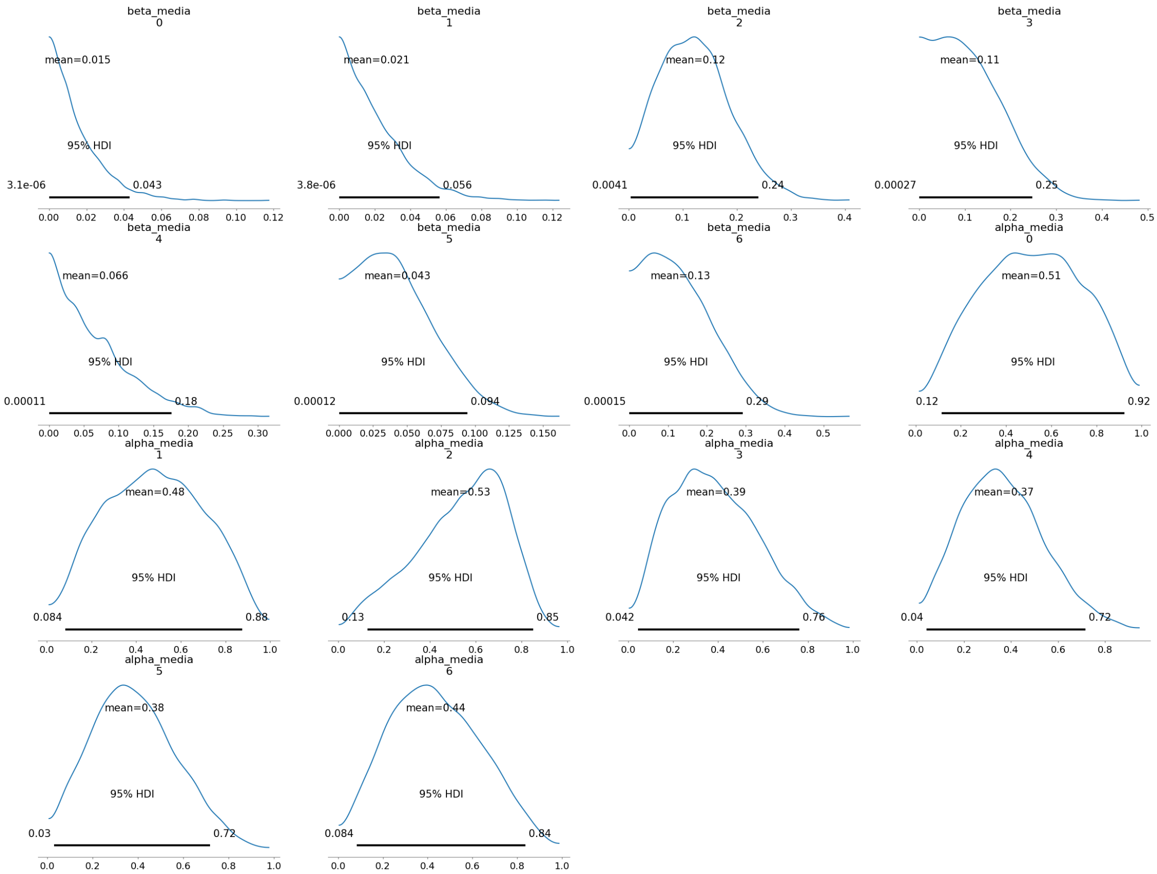
Model Summary

	mean	sd	hdi_3%	hdi_97%	mcse_mean	mcse_sd	ess_bulk	ess_tail	r_hat
beta_media[0]	0,015	0,014	0,0	0,041	0,0	0,0	3172,0	2013,0	1,0
beta_media[1]	0,021	0,018	0,0	0,053	0,0	0,0	3541,0	2313,0	1,0
beta_media[2]	0,123	0,065	0,006	0,235	0,001	0,001	2744,0	1184,0	1,0
beta_media[3]	0,111	0,074	0,0	0,24	0,001	0,001	2425,0	1795,0	1,0
beta_media[4]	0,066	0,054	0,0	0,168	0,001	0,001	2687,0	2257,0	1,0
beta_media[5]	0,043	0,028	0,0	0,092	0,001	0,0	2270,0	1740,0	1,0
beta_media[6]	0,132	0,087	0,0	0,283	0,002	0,001	2292,0	1789,0	1,0
alpha_media[0]	0,512	0,225	0,113	0,902	0,003	0,003	6175,0	3236,0	1,0
alpha_media[1]	0,484	0,22	0,094	0,866	0,003	0,003	4934,0	2680,0	1,0
alpha_media[2]	0,535	0,196	0,136	0,841	0,003	0,002	3440,0	2736,0	1,0
alpha_media[3]	0,389	0,198	0,061	0,758	0,003	0,003	4613,0	2684,0	1,0
alpha_media[4]	0,37	0,18	0,029	0,681	0,003	0,003	4052,0	2516,0	1,0
alpha_media[5]	0,384	0,184	0,055	0,721	0,003	0,003	3987,0	2599,0	1,0
alpha_media[6]	0,435	0,203	0,064	0,793	0,003	0,003	4332,0	2459,0	1,0
intercept	0,628	0,2	0,233	0,984	0,004	0,003	2677,0	2784,0	1,0
sigma	0,258	0,021	0,22	0,298	0,0	0,0	4438,0	2993,0	1,0

Appendix B

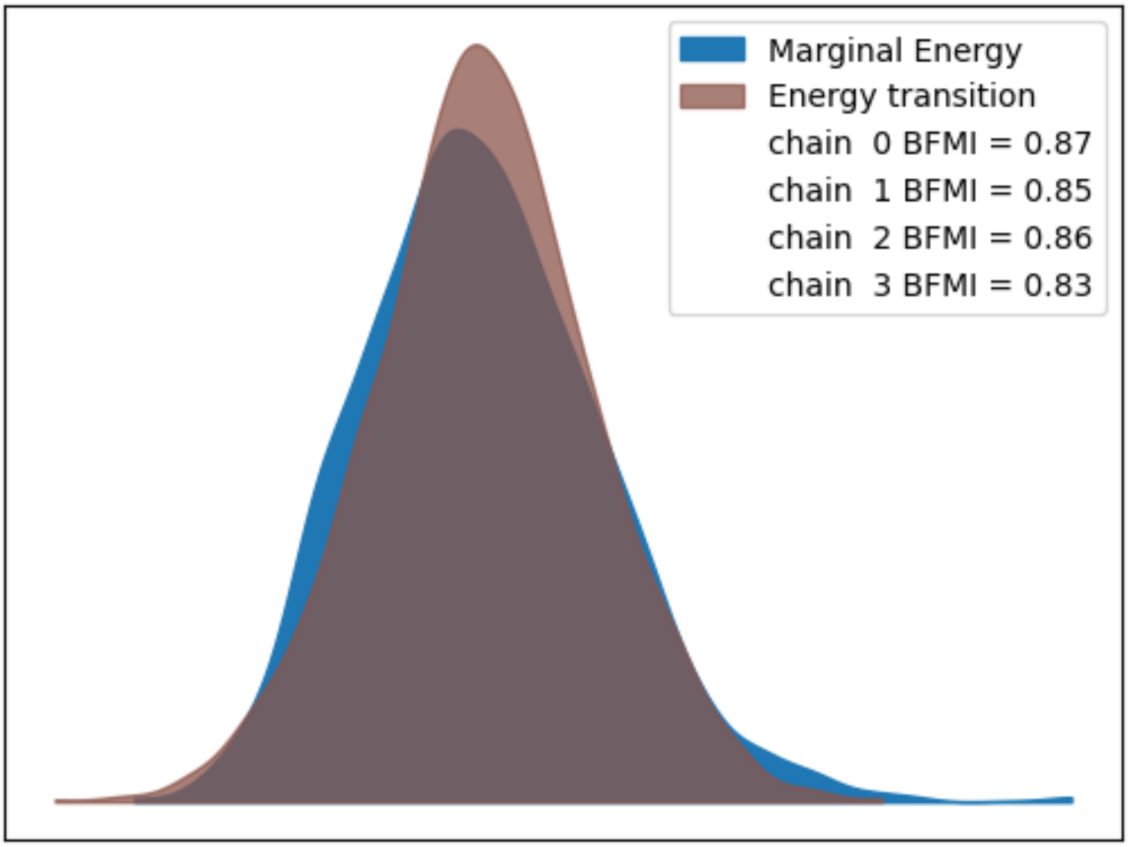

beta_media
beta_media
alpha_media
alpha_media
intercept
intercept
sigma
sigma

Appendix C



beta_media
0
mean=0.015
95% HDI
3.1e-06
0.043
beta_media
1
mean=0.021
95% HDI
3.8e-06
0.056
beta_media
2
mean=0.12
95% HDI
0.0041
0.24
beta_media
3
mean=0.11
95% HDI
0.00027
0.25
beta_media
4
mean=0.066
95% HDI
0.00011
0.18
beta_media
5
mean=0.043
95% HDI
0.00012
0.094
beta_media
6
mean=0.13
95% HDI
0.00015
0.29
alpha_media
0
mean=0.51
95% HDI
0.12
0.92
alpha_media
1
mean=0.48
95% HDI
0.084
0.88
alpha_media
2
mean=0.53
95% HDI
0.13
0.85
alpha_media
3
mean=0.39
95% HDI
0.042
0.76
alpha_media
4
mean=0.37
95% HDI
0.04
0.72
alpha_media
5
mean=0.38
95% HDI
0.03
0.72
alpha_media
6
mean=0.44
95% HDI
0.084
0.84

Appendix D



Marginal Energy
Energy transition
chain 0 BFMI = 0.87
chain 1 BFMI = 0.85
chain 2 BFMI = 0.86
chain 3 BFMI = 0.83