

Assessing Direct Marketing Effectiveness at a Retailer's: a Mixed-Effects Model of Heterogeneous Responsiveness to Personalized Promotions

Roko Pedisic¹

Abstract

This study proposed a mixed linear modeling framework for revealing the average treatment effects of coupon promotions as well as for figuring out what kind of customers are more responsive to coupon promotions. The analysis used detailed data on 2,469 households who are frequent shoppers at a retailer over one year. A series of models starting from fixed parameters OLS regression to more flexible random coefficient models were built. The results showed that receiving more coupon promotions does not lead to higher spending among households. None of the available demographic characteristics was significantly linearly related to the responsiveness to coupon promotions according to the initial model with cross-level interaction terms. However, exploratory analysis of random slopes with the help of a regression tree allowed to correct the specification, which allowed confirming that people aged 25+ with annual household income exceeding \$25,000 are more responsive, yet their response is insignificantly different from zero.

JEL classification numbers: C02, C10, D10, M30, M31.

Keywords: Mixed-Effects Model, Multilevel Analysis, Heterogeneous Responsiveness, Promotions.

¹ University of Bologna (Alma Mater Studiorum - Università di Bologna).

1. Introduction

With the growth of loyalty programs, it has become increasingly easy for retailers to track information on marketing exposures, such as coupon promotions. However, translating this seemingly rich data into measures of marketing response has proven difficult, largely because of possible endogeneity of promotional efforts as well as because of the lack of statistical framework that practitioners could adopt to study the effects of limited-time offers. Hereinafter, for brevity, I will use the word "coupon", implying various offers having a limited validity period.

Traditionally scanner data has been widely used for brand/SKU choice modeling using data on actual purchases following the logit modeling approach described in Guadagni and Little (1) and its extensions. Jain, Vilcassim, and Chintagunta (2) were the first to use a random parameters logit model to study the demand for yogurt brands. Fader and Hardie (3) proposed switching from brand-level to SKU-level choice modeling. Thunström (4) assessed the strength of habit persistence in the cereal market, while Hoffman and Bronnmann (5) applied a similar random coefficients logit model to model preference and response heterogeneity in the German carbonated soft drink market.

While SKU/brand choice models are of special interest to manufacturers, retailers are less interested in how people make their choice in a specific product category and more - in the effects of their direct marketing campaigns on customer value, i.e. in figuring out whether households who received coupon promotions became more active buyers overall. This paper proposes a framework for the identification of causal effects of coupon promotions on customer value outcomes.

I show how the problem of coupon effects estimation can be addressed and illustrate it using a dataset of household purchases and coupons sent to them. I test the following hypothesis: response to coupon promotions is heterogeneous across households and can be explained with household characteristics. I develop a hierarchical mixed-effects model for coupon promotion response that allows us to control for random household as well as time-fixed effects to mitigate the possible endogeneity problem caused by more intense promotional activity in periods of high demand (holidays, weekends, etc.) and by a possible confounding effect of more intense in-store promotions in such periods. In addition, being a random-slope model, the model allows the responsiveness to coupon promotions to vary across households depending on their demographic characteristics. Empirical Bayes estimates of individual reaction to coupon promotions are obtained and can be used for targeting customers who are less responsive to coupon promotions differently compared to those who are responsive to them.

2. Data

For my analysis, I used detailed data on 2,469 households who are frequent shoppers at a retailer over one year, which originates from the 84.51° "Complete Journey 2.0" R source files and also includes useful metadata on products, coupons, campaigns,

and promotions². The dataset contains all of each household's purchases, not just those from a limited number of categories. For certain 801 households, demographic information as well as direct marketing contact history are included. These households are used in my analysis.

Over the year 2017, there were 27 coupon campaigns of 3 types (A, B, or C). Each customer participating in a Type A campaign received 16 coupons out of the pool. The 16 coupons were selected based on the customer's prior purchase behavior. Identifying the specific 16 coupons that each customer received is outside the scope of this database. For campaign Type B and Type C, all customers participating in a campaign receives all coupons pertaining to that campaign.

The store's profit from all purchases made by household i on day t would be a perfect dependent variable. However, while profit margins are typically known to retailers, they are rarely available for academic researchers using secondary data. In my study I use sales revenue received from household i in period t as the dependent variable.

I will work with repeated measurements (level 1) nested within the individual (level 2). Step-by-step instructions on converting raw data to the balanced panel format are provided below.

1. Convert transactional data into a balanced panel data format, where there is information about revenue brought by each household each day or each week (depending on the available computing resources). I aggregated data to the level of weeks. The resulting table will be referred to as *spending*. In this case it contains 128388 rows (2469 households*52 weeks)
2. Knowing each campaign's start and end date, create a dataset containing information on how many coupons of each type were available to each household. If the coupon has been withdrawn by household, the day of withdrawal is the last day when it was available to the household. In my example, there were 3 types of campaigns. I will refer to this table as *coupons* (Figure 2). This data frame contains the same number of rows as *spending* - 128388 rows (2469 households*52 weeks)
3. Join *spending* and *coupons* data, as well as append demographic characteristics of households for which demographic details are available (Figure 3). Households without any demographics data available were excluded from my analysis. This resulted in 801 households observed over 52 weeks (a total of 41652 observations).

A description of all variables contained in the resulting dataset is provided in Table 1. Age, income, household size and kids count were converted to numeric variables (same variable names with “_num” suffix by assigning the *midpoint* of each category).

² <https://CRAN.R-project.org/package=completejourney>

Table 1: Description of variables used in the analysis

Variable name	Variable definition/levels
Household_id	Household ID (the clustering variable)
Week	Week number (1-52)
Campaigns_per_day	Average number of coupon campaigns that the household was exposed to on week t
Age	19-24, 25-34, 35-44, 45-54, 55-64, 65+
Income	Under 15K, 15-24K, 25-34K, 35-49K, 50-74K, 75-99K, 100-124K, 125-149K, 150-174K, 175-199K, 200-249K, 250K+
Household_comp	1 Adult Kids, 1 Adult No Kids, 2 Adults Kids, 2 Adults No Kids, Unknown
Kids_count	0, 1, 2, 3+

3. Methods

A series of models starting from fixed parameters OLS regression to more flexible random coefficient models were built. The baseline model is a pooled OLS regression model with time-fixed effects (Model 1).

$$\log(S_{it} + 1) = \mathbf{X}_{it}\boldsymbol{\beta} + e_{it}, \quad (1)$$

where S_{it} is the average daily amount spent by household i in week t , e_{it} is the idiosyncratic error term, \mathbf{X}_{it} is a row vector of variables describing coupon availability for household i on day t (the first element of \mathbf{X}_{it} is always 1 to account for the presence of the intercept in the model), while $\boldsymbol{\beta}$ is a column vector of the corresponding coefficients (including the intercept). In this study \mathbf{X}_{it} will be represented by a single variable – the number of coupons available for household i on day t and time effects are captured by variable *week* and the equation can be written as follows: $\log(S_{it} + 1) = \beta_0 + \beta_1 \text{campaigns_per_day}_{it} + \beta_2 \text{week}_t + \varepsilon_{it}$

Models 2-5 were estimated using the restricted maximum likelihood method (REML). Model 2 is a model with time-fixed effects (δ_t) and random effects of households (v_i). These effects account for differences in baseline spending, but the responsiveness of households to coupon promotions remains constant:

$$\log(S_{it} + 1) = \mathbf{X}_{it}\boldsymbol{\beta} + \delta_t + v_i + e_{it} \quad (2)$$

In this study the specification can be written as follows: $\log(S_{it} + 1) = \beta_0 + \beta_1 \text{campaigns_per_day}_{it} + \beta_2 \text{week}_t + v_i + \varepsilon_{it}$

Model 3 allows the slope to vary across households as well:

$$\log(S_{it} + 1) = \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{X}_{it}\boldsymbol{\lambda}_i + \delta_t + v_i + u_{it} \quad (3)$$

In this study this general specification is reduced to the following equation:

$$\log(S_{it} + 1) = \beta_0 + \beta_1 \text{campaigns_per_day}_{it} + \beta_2 \text{week}_t + \lambda_i X_{it} + v_i + \varepsilon_{it}$$

Finally, I allow β_i to vary depending on time-invariant demographic characteristics (vector \mathbf{D}_i) of households according to the level-2 model: $\boldsymbol{\beta}_i = \mathbf{D}_i\boldsymbol{\gamma} + \lambda_i$. This results in the following model (Model 4), which includes main effects of promotion (\mathbf{X}_{it}) and demographic variables (\mathbf{D}_i), as well as cross-level interaction effects (\mathbf{I}_{it}).

$$\log(S_{it}) = \mathbf{X}_{it}\boldsymbol{\beta} + \mathbf{D}_i\boldsymbol{\alpha} + \mathbf{I}_{it}\boldsymbol{\gamma} + \mathbf{X}_{it}\boldsymbol{\lambda}_i + \delta_t + v_i + u_{it} \quad (4)$$

Model 4 allows making explicit conclusions on how responsiveness to coupon promotions is influenced by demographic characteristics. In the case, model 4 is the same as model 3, but with additional fixed effects: the main effects of demographic characteristics and the interactions between demographic characteristics and *campaigns_per_day* added to the equation.

Model 4 can be built only for households with non-missing demographic characteristics. At the same time, a retailer could benefit from having information on coupon promotion responsiveness of all its clients. This responsiveness can be inferred by extracting random effects of the slope (λ_i) from Model 3. Random effects λ_i can be predicted for each household using Empirical Bayes estimation and can then be used in various segmentation and targeting tasks as a measure of coupon promotion responsiveness with zero corresponding to the average level of responsiveness.

All models were estimated using lme4 package in R (6). While the use of p-values in linear mixed models is a debatable issue (7), I report p-values obtained using the Satterthwaite's method. Given the large sample size and the balanced structure of the data, they can be expected to give a reasonable approximation of true p-values, while other methods are prohibitively computer-intensive for my large dataset.

4. Results

While no pronounced relationship between average daily spending and week number was found, the dynamics of the average daily campaigns available to households is cyclical and had a tendency to increase throughout the year (Figure 1). Accounting for time effects will help us account for potential confounding factors such as more intensive in-store/TV/radio/etc. promotions on certain weeks.

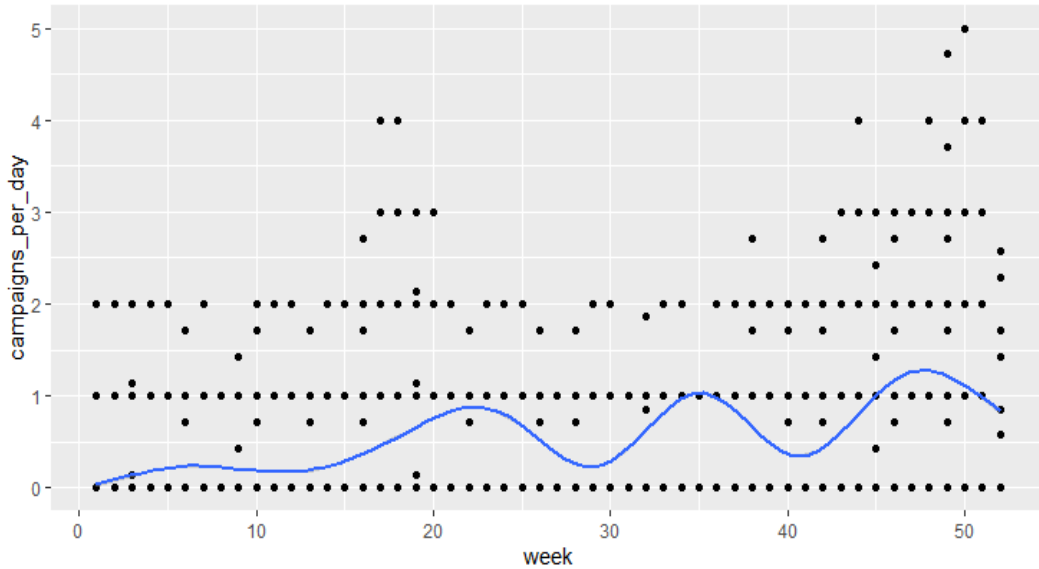


Figure 1: Relationship between average number of available campaigns and the week number (smoothing based on a general additive model)

The validity of the linearity assumption is supported by the nonparametric general additive smoothing of the relationship between the log-transformed average number of daily campaigns and the daily amount spent: the estimated relationship is almost perfectly linear (Figure 2).

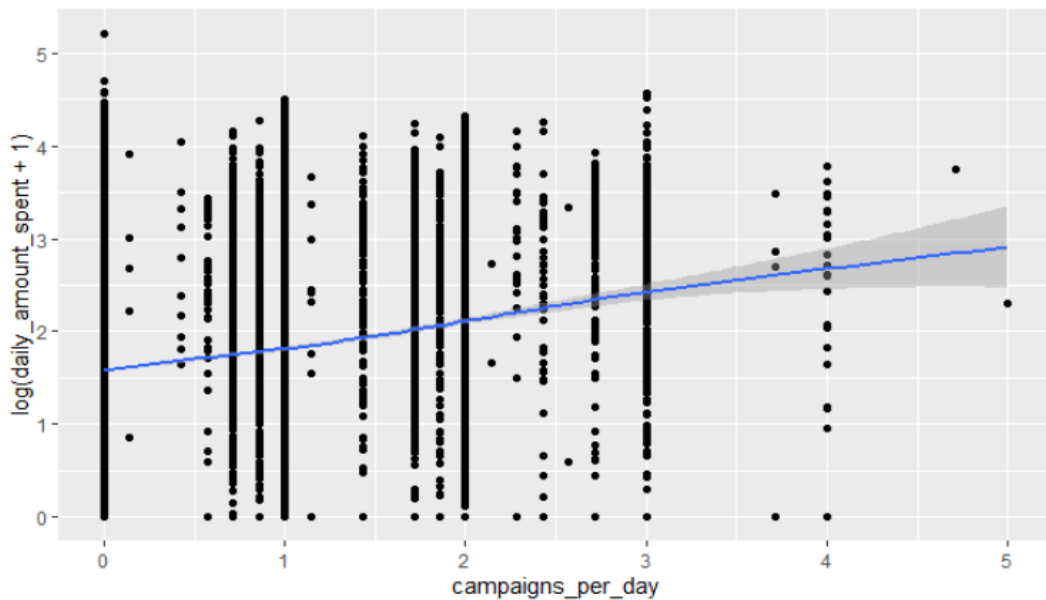


Figure 2: Relationship between average number of available campaigns and the week number (smoothing based on a general additive model)

A summary of models 1-4 is presented in Table 2. I have checked that using *week* as a factor variable has almost no effect on the resulting estimate of the effect of *campaigns_per_day*, which is why I simply account for the time trend to make the models more parsimonious. According to the OLS regression model (Model 1) the slope is 0.28, implying that when average daily number of campaigns increases by 1, the average daily amount spent increases by $(\exp(0.28)-1)*100\%=32\%$. However, as soon as I introduced random effects the promotional intensity became insignificant (Models 2-3). Model 3 allows both the intercept and the slope to vary across households, and the results are similar to those of Model 2. They indicate the insignificance of *campaigns_per_day* ($p>0.05$). Based on the standard deviations of random effects, it can be concluded that the largest source of random variation in $\log(\text{daily_amount_spent} + 1)$ is within-household variation of spendings in time, with less, yet substantial variation coming from differences in the conditional mean (intercept) across households and relatively small variation in the effect of *campaigns_per_day* across households.

Table 2: Parameter estimates of models 1-3 (dependent variable: $\log(\text{daily_amount_spent} + 1)$)

	Model 1		Model 2		Model 3	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Intercept	1.64***	0.012	1.66***	0.024	1.66***	0.024
Campaigns_per_day	0.28***	0.0089	-0.0066	0.0087	-0.0073	0.011
week	-0.0029***	0.00041	0.0022***	0.00036	0.0023***	0.00036
Random effects:	No		Std. Dev.		Std.Dev.	
Intercept			0.6204		0.6161	
Campaigns_per_day			-		0.1708	
Residual	1.161		0.9987		0.9931	
AIC	-		120568		120465	
BIC	-		120503.2		120396.9	
N	41652		41652		41652	

t-tests use Satterthwaite's method

*** - significant at the 1% level

** - significant at the 5% level

* - significant at the 10% level

Models 1-3 differ only with respect to random effects specification (no effects, random intercept, and random intercept and slope). The likelihood ratio test has shown the significance of the improvement of Model 3 compared to Model 2 ($\chi^2=106.83$, $df=2$, $p<0.001$).

Point estimates of random effects as well as the corresponding conditional standard deviations were extracted from Model 3. Random effects larger than 0 imply higher than average responsiveness, smaller than zero – lower than average responsiveness of amount spent to coupon promotions. Top-10 random effects of *campaign_per_day* are visualized in Figure 3 and top-10 lowest random effects for the slope – in Figure 4.

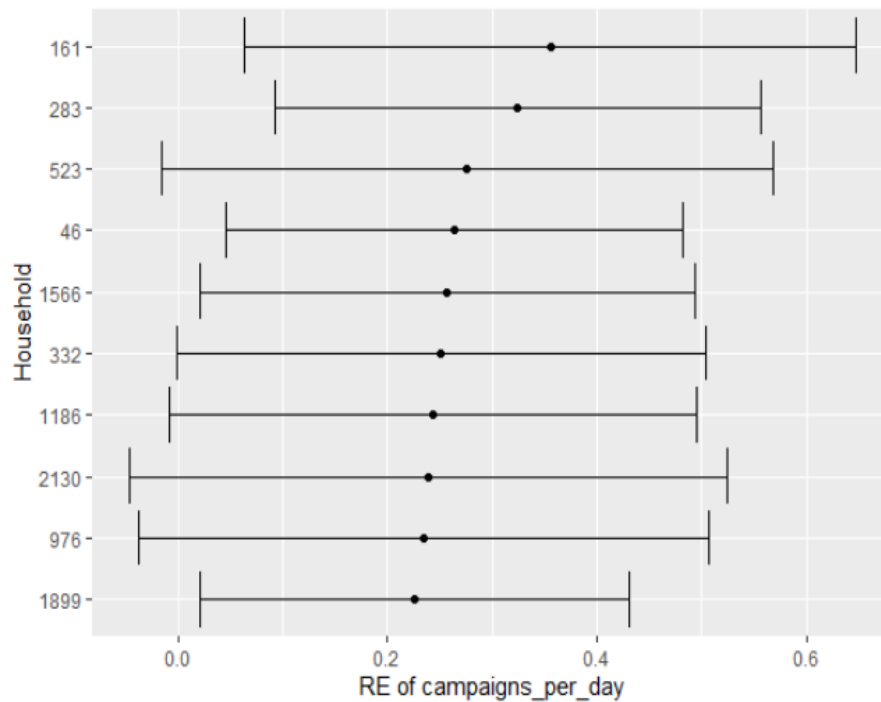


Figure 3: Top-10 households having positive random effect of *campaigns_per_day*

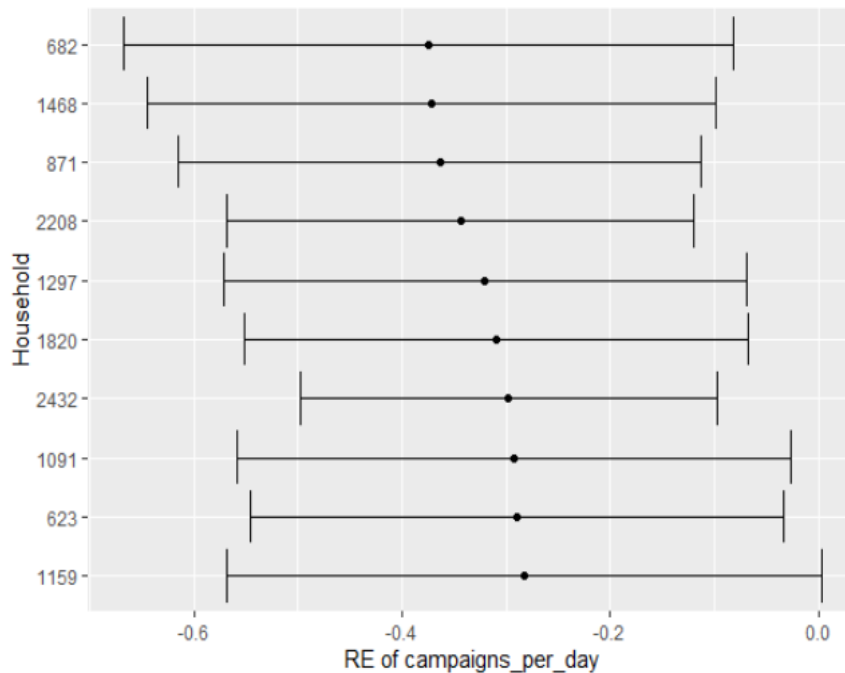


Figure 4: Top-10 households having negative random effect of campaigns_per_day

Among top-10 positive effects only a half are significantly different from zero. At the same time, among top-10 negative effects 9 significantly differ from zero. These estimates imply that either coupon promotions tend to be ineffective or that there is a correlation between individual effects and the exposure to coupon promotions (e.g. if coupon promotions were sent more often to buyers with low propensity to purchase). The latter can be tested by introducing fixed individual effects of households (800 dummy variables). Fixed effects are allowed to be correlated with the treatment, but it would make the problem computationally too complex to be handled using a standard computer. However, some exploratory analysis has not revealed any substantial differences between households that were and were not treated, which is why I can assume the result implies that coupon promotions did not have a significantly positive short-term effect.

The insignificance of the promotional effect makes it especially relevant to explore whether some demographic groups are responsive to coupon promotions despite the overall insignificance of such promotions. Model 4 allows testing the hypotheses that both baseline level of household spending (the intercept) and the sensitivity of coupon promotions (the slope) vary by demographic characteristics of households. None of the demographic variables or their interactions with *campaigns_per_day* were significant. The output of the model is rather long and can be found in the appendix.

The insignificance of demographic variables may be related to the nonlinearity of

the effects. To get additional insights, I built a regression tree of the random effect's estimate for the slope with demographic variables as explanatory variables using the *rpart* package (Figure 5).

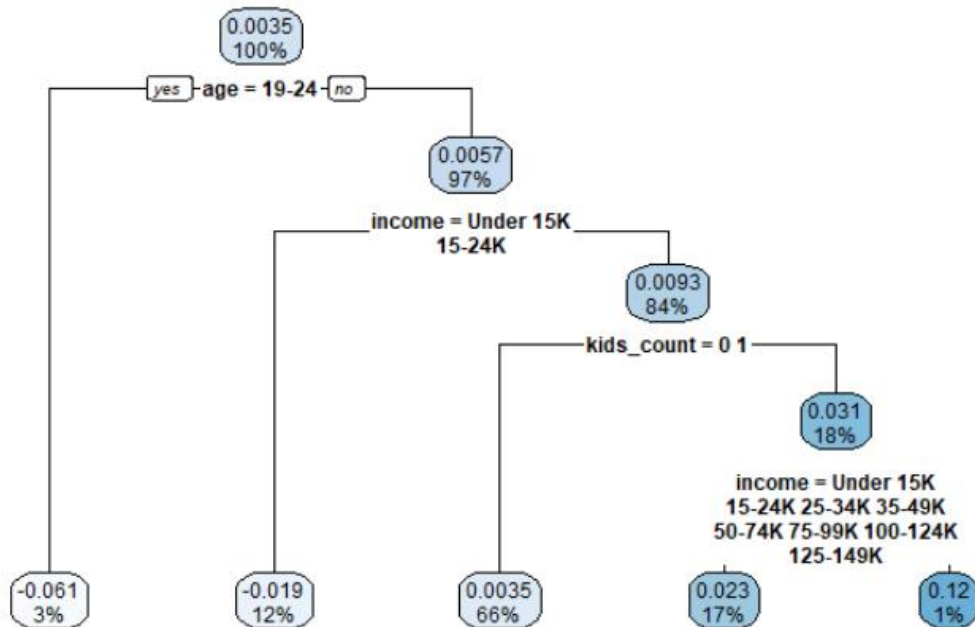


Figure 5: Regression tree of the random effect's for campaigns_per_day heterogeneity across demographic groups

Even though the differences across groups do not look very substantial, the tree is still insightful. Buyers aged 19-24 (the youngest age group in the dataset) are the least responsive (mean random effect=-0.061). The higher the income and the number of kids the more responsive the household is. More specifically, the highest differentiation is between households with 0 or 1 kid and others, as well as between those making under \$24K annually and those making more. The highest mean random effect is among households where the householder's age is above 24 years old, having more than 1 kid and making more than \$150K annually, but the segment is too small to be accounted for in the model (only 1% of the sample). Therefore, a more parsimonious model can include binary indicators of *age* (19-24 vs. others), *kids_count* (less than 2 vs. 2+) and *income* (\$25K+ vs. others). Model 5 (Table 3) is such a parsimonious version of Model 4.

Table 3: Parameter estimates of Model 5

Fixed effects:	Estimate	Std. Error
Intercept	1.66***	0.024
Campaigns_per_day	-0.0073***	0.011
Kids_count2	0.10*	0.062
Income25	0.12**	0.061
Age25	0.042	0.098
Campaigns_per_day: kids_count2	0.035	0.028
Campaigns_per_day: income25	0.050 *	0.028
Campaigns_per_day: age25	0.097**	0.042
Week	0.0023***	0.00036
Random effects:	Std.Dev.	
Intercept	0.6144	
Campaigns_per_day	0.1674	
Residual	0.9932	
AIC	120480	
BIC	120348.7	
N	41652	

Model 5 is significantly preferred to model 3 according to the likelihood ratio test ($\chi^2=41.538$, $df=8$, $p<0.001$). According to it, buyers older than 24 years with household income exceeding 25 thousand dollars annually are more responsive to coupon promotions. However, a linear hypothesis test did not allow us to reject the hypothesis that the sum of parameter estimates of `campaigns_per_day`, `campaigns_per_day:income25`, and `campaigns_per_day:age25` significantly differ from zero ($p>0.05$). While the distribution of residuals is very close to normal as indicated by the normal quantiles plot (Figure 6), there is an undesirable pattern in the residuals vs. fitted plot (Figure 7). There is a clear diagonal line of dots, where residuals equal the negative of the fitted value. This line corresponds to observations that are zero but were predicted to be non-zero. Such a pattern is not surprising as in the presence of excess zeros, a one-step model will often overestimate the amount spent. Otherwise, the plot shows only moderate heteroscedasticity (decrease in residual variance as fitted values increase). The problem cannot be solved easily (omitting zeros is undesirable as they contain useful information), but a possible solution is outlined in the discussion section.

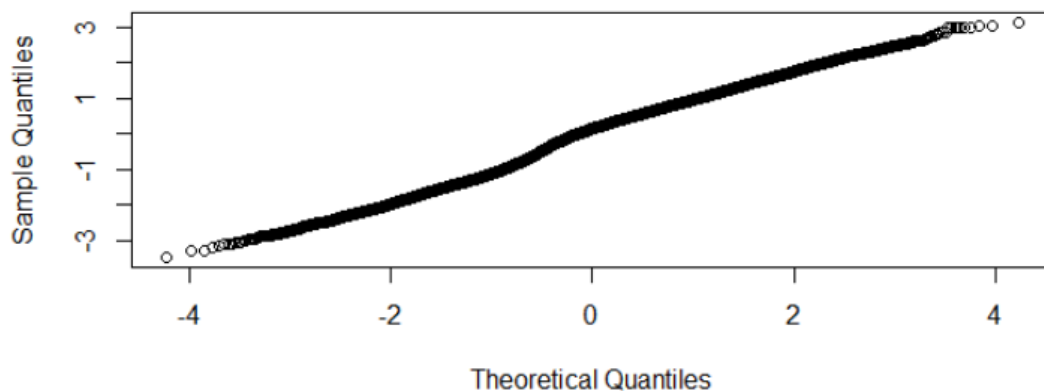


Figure 6: Normal quantiles plot for Model 5 residuals

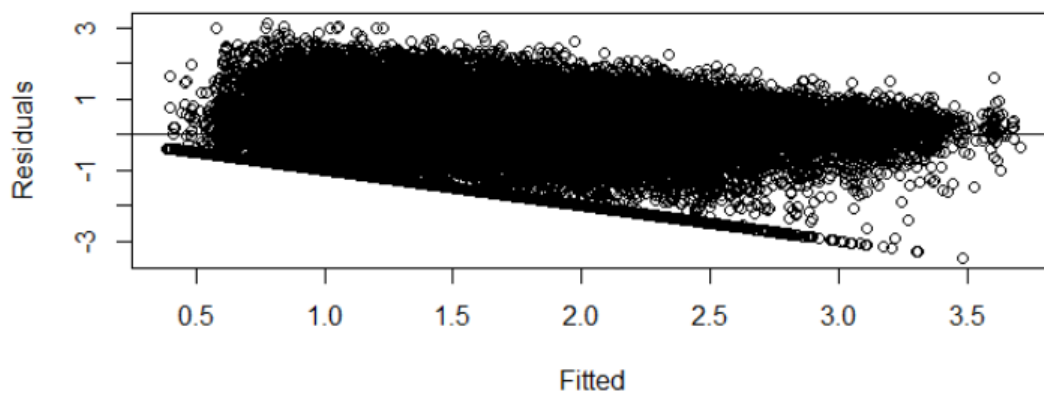


Figure 7: Residuals vs. fitted plot based on Model 5

5. Conclusion

This study proposes a mixed linear modeling framework for revealing the average treatment effects of coupon promotions as well as for figuring out what kind of customers are more responsive to coupon promotions and, therefore, should be targeted to increase returns on coupon promotions. It turned out that receiving more coupon promotions does not lead to higher spending among households. None of the available demographic characteristics was significantly linearly related to the responsiveness to coupon promotions according to the initial model with cross-level interaction terms. However, exploratory analysis of random slopes with the help of a regression tree allowed to correct the specification, which allowed confirming that

people aged 25+ with annual household income exceeding \$25,000 are more responsive, yet their response is insignificantly different from zero. Despite being highly uncertain (relatively high conditional variance), random effects for the slope representing individual estimates of the sensitivity to coupon promotions, may be useful segmentation/targeting variables for identifying specific households that are the most/the least responsive to coupon promotions.

While I illustrated the approach by inferring the effect of the total daily average number of campaigns, the framework allows for the assessment of multiple promotional effects (corresponding to, for instance, different types of campaigns or different types of promoted products).

The model clearly indicated the insignificance of coupon campaigns effect and minor dependency of this effect on demographic characteristics of households, which is unlikely to change if some other estimation method is used. However, one limitation of the model is that because the dependent variable is zero-inflated (contains many zeros) the residuals turned out to be heteroscedastic as indicated by the residuals-versus fitted plot. To account for the fact that 0 usually means “no shopping trip occurred that week” rather than “went to the store and bought nothing”, I need a model that has two components: one for explaining the probability of going to the store and for explaining the amount spent if the household goes to the store. A multilevel version of the tobit model would be appropriate (8). To the best of my knowledge, such a model is not available as a preprogrammed routine in any multilevel modeling software and requires custom programming using Bayesian framework such as Stan.

Another limitation of the model is that it reflects only short-term effects of coupon promotions. While I do not expect large stockpiling effects in the case of coupon promotions, it is still not clear whether being exposed to coupons increases overall annual household spending. In order to address this issue, I would analyze how spending changed after coupon promotions started in two groups of customers: those who were not exposed to any coupons during the whole year and those who were exposed to some coupon promotions. Basically, a significant change in the difference between the two groups (compared to the baseline pre-promotion difference) will indicate the significance of coupon promotions.

References

- [1] Guadagni PM, Little JDC. (1982). A logit model of brand choice calibrated on scanner data. *Mark Sci.* 2(3):203–38.
- [2] Jain DC, Vilcassim NJ, Chintagunta PK. (1994). A random-coefficients logit brand-choice model applied to panel data. *J Bus Econ Stat.* 12(3):317–28.
- [3] Fader PS, Hardie BGS. (1996). Modeling consumer choice among SKUs. *J Mark Res.* 33(4):442–52.
- [4] Thunström L. (2010) Preference heterogeneity and habit persistence: The case of breakfast cereal consumption. *J Agric Econ.* 61(1):76–96.
- [5] Hoffmann J, Bronnmann J. (2019). Bottle size matters: Heterogeneity in the German carbonated soft drink market. *Agribusiness.* 35(4):556–73.
- [6] Bates D, Mächler M, Bolker BM, Walker SC. (2015). Fitting linear mixed-effects models using lme4. *J Stat Softw.* 67(1).
- [7] Kuznetsova A, Brockhoff PB, Christensen RHB. (2017). lmerTest package: tests in linear mixed effects models. *J Stat Softw.* 82(13).
- [8] Zeng Q, Wen H, Huang H, Abdel-Aty M. (2017). A Bayesian spatial random parameters Tobit model for analyzing crash rates on roadway segments. *Accid Anal Prev.* 100:37–43.