Marketing Mix Response across Retail Formats - The Role of Shopping Trip Types

Online Appendix

November 2019

Marketing Mix Variables

RMS data

Price For each brand at a retail chain, market and week, we compute the price measure as a volume weighted index by dividing the revenue for that brand across all SKUs by the total volume-weighted units sold at the particular retail chain. If none of the SKUs of a particular brand are sold in a week at a retail chain, we replace the missing price index by the 4-week moving average price of that brand at the corresponding retail chain in that market.[1] Ailawadi et al. (2010) separately account for the effect of price changes and promotions. In the data, we only observe the retail prices, and thus, do not try and separate the effect of price variation due to quantity based price discrimination from variation due to discounts.

Line Length For any brand at a retailer in a week, we calculate the number of different SKUs available for the brand. The Nielsen RMS data only includes information on SKUs which are sold at a store in a week. Thus, to ensure that we do not under count the number of SKUs of infrequently purchased brands, or in slow moving categories, we smooth the SKU availability such that if an SKU is available in a five-week rolling window (two weeks before and two weeks after), we assume that the SKU is available at the particular chain in that week.

Non-Price Promotion We compute the non-price promotion variable as the percentage of SKUs of a brand associated with a feature at a retail chain in a week. The Nielsen RMS data includes feature information for only a subset of stores. However, as per the data guidelines, if a store features an SKU, it is safe to assume that all stores belonging to the same chain feature the same SKU. Using this information, we compute the percentage of SKUs on feature in a retail chain, and replace the missing values by a one month or three month moving average.

[1] Weeks with lower prices on particular SKUs are likely to see more sales of that SKU resulting in underestimating the price index, and possible endogeneity due to the correlation between the price index and the unobserved demand shocks. To account for this, we replicate the analysis with time invariant quantity weights corresponding to each SKU. We do not find any qualitative differences in our findings; the results are available from the authors.

Other Variables We also include in our analysis, measures to account for the competition faced by a retail chain. We do so by computing (i) number of stores belonging to the chain as a percentage of the total number of stores belonging to the particular format, (ii) number of stores belonging to the format as a percentage of the total number of stores in the DMA, and (iii) number of stores in the DMA. These variables account for within and cross format competition faced by a retail chain in a market. One potential issue with computing competition measures based on Nielsen data is that the data includes only a subset of stores in the United States. Our competition measures alleviate concerns relating to this because we compute percentage stores as opposed to including absolute numbers (first two measures). As the Nielsen RMS data set is widely used by consumer packaged goods manufacturers for making marketing decisions, we believe that it provides a reasonable representation for the purposes of our study.

Homescan panel data

Price The price variable is computed based on the average price of a brand at a retail chain in a market in a month. We first compute the average quantity weighted price of a brand at a retail chain in any given month, and then take the weighted average of the retail chain level price indices where the weights are based on the monthly quantity purchased on trips of that particular type. A potential issue with using month-specific quantities is that consumers are more likely to purchase a higher quantity when prices are lower thus leading to endogeneity in the weighted price index. To address this issue, we used time-invariant weights based on the average quantity sold at each retail chain. This results in a price index for a brand on a particular trip type in a market in a given month. Weighting the retail chain level indices by quantities belonging to a particular trip type ensures that the price index reflects the prices at stores which are most likely to be frequented for trips of the particular type. Thus, for example, if majority of the quantities purchased on unplanned trips are purchased at convenience stores, then the price index of the unplanned trips will be closer to the prices at convenience stores.

Line Length Unlike the price index, the line length measure at the trip type level does not lend itself to an easy interpretation. For instance, we know the price that was paid for any brand on a particular trip type; however, it is not clear what the line length of a brand on a trip type would be given that line length is typically associated with a store, retail chain or format, as opposed to a trip type. Thus, for any brand, we compute the line length measure on any trip type based on the line length at different retail chains and the quantity purchased on trips of that type to these retail chains. Specifically, for each brand, we compute the number of unique SKUs at a retail chain in a month, and then take a weighted average across these chains where the weights are based on the quantity of the brand purchased on trips of the particular type at these chains. Much like the price measure, we normalize the line length measure by the market-level averages, and use time-invariant quantity weights across retail chains. Computing the line length measure as such makes interpreting the findings based on the trip type level analysis more difficult and less intuitive. Having said this, we still include the line length measure in our analysis to understand whether changes in these measures affect brand shares on different trip types in a manner consistent with proposition 1.

Non-price Promotion The Homescan data, unlike the RMS data, does not include information on whether the chosen item was featured or not. Instead, the Homescan data provides information on whether the item was on a deal or not which could include a manufacturer or retailer coupon or a non-price promotion. To the extent that the value of any coupon is already accounted for in the price variable, the deal variable captures the effect of only the non-price part of the promotion. Thus, while different from features used in the RMS data, the deal variable in essence reflects the effect of non-price promotions and we would expect proposition 3 to hold for the deal variable as well.

Similar to the price variable, we observe whether an individual item purchased was associated with a deal or not. We, thus, compute the non-price promotion variable based on the average proportion of products of a brand which were purchased on a deal at a retail chain in a month. Much like the price variable, we first compute the average quantity weighted proportion of products purchased on a deal of a brand at a retail chain in any given month, and then take the weighted average of these indices where the weights are based on the monthly quantity purchased on trips of that particular type. For any trip type, the resulting non-price promotion variable thus reflects the proportion of products which were on a deal at the format where most trips of that type were made.

Instrumental Variable Regression

Table 1: Instrumental Variable Regression Coefficients by Format

	Price	**Non-Price Promotion**	**Line Length**
Convenience	-0.471	0.295	1.666
	(0.008)	(0.079)	(0.0069)
Drug	-1.827	1.028	0.795
	(0.001)	(0.002)	(0.0007)
Supermarkets	-2.101	0.787	0.349
	(0.001)	(0.001)	(0.0001)
Mass Merchandiser	-1.532	0.821	0.593
	(0.001)	(0.002)	(0.0002)

Note. The Table reports the format-specific coefficients from the regressions run by pooling data across all product categories. In addition to fixed effects, we instrument for prices, features and line length using Hausman-style instruments. Specifically, we instrument any marketing mix variable using the average value of the same variable for the same brand in the same retail chain and week across all other markets. Standard errors are reported in parenthesis. All estimates are significantly different from 0 at the 99% confidence level. Further, all estimates are significantly different from each other at the 99% confidence level.

Profitability Analysis based on Bootstrapping

Table 2: Annual Profit Changes from a Change in Marketing Mix Variables

	Absolute Change			**Percentage Change**		
	Drug	**Supermkt**	**MM**	**Drug**	**Supermkt**	**MM**
Price						
Dry Dog Food	146	1453	621	4.24%	4.05%	3.62%
	(0)	(1)	(1)	(0.00%)	(0.00%)	(0.00%)
Liq. HD Detergent	553	1529	974	3.89%	3.84%	4.02%
	(1)	(2)	(1)	(0.00%)	(0.00%)	(0.00%)
Toilet Tissue	882	3148	1508	3.54%	3.47%	3.32%
	(1)	(12)	(8)	(0.00%)	(0.01%)	(0.00%)
Paper Towel	675	52	171	2.75%	1.75%	1.94%
	(10)	(15)	(10)	(0.01%)	(0.00%)	(0.00%)
Non-Price Promotion						
Dry Dog Food	1 (0)	20	13	0.02%	0.02%	0.02%
		(1)	(0)	(0.00%)	(0.00%)	(0.00%)
Liq. HD Detergent	26	117	16	0.05%	0.05%	0.01%
	(0)	(1)	(0)	(0.00%)	(0.00%)	(0.00%)
Toilet Tissue	21	298	124	0.09%	0.11%	0.07%
	(0)	(3)	(2)	(0.00%)	(0.00%)	(0.01%)
Paper Towel	87	604	132	0.17%	0.16%	0.10%
	(1)	(7)	(3)	(0.01%)	(0.01%)	(0.00%)
Line Length						
Dry Dog Food	439	4803	1935	11.74%	6.32%	4.41%
	(3)	(21)	(3)	(0.04%)	(0.03%)	(0.01%)
Liq. HD Detergent	3532	9791	3799	7.16%	6.77%	5.09%
	(11)	(12)	(5)	(0.01%)	(0.01%)	(0.00%)
Toilet Tissue	2174	30092	13404	7.65%	14.32%	11.45%
	(12)	(57)	(30)	(0.04%)	(0.02%)	(0.02%)
Paper Towel	11976	65043	19284	23.20%	27.11%	18.59%
	(108)	(133)	(55)	(0.10%)	(0.04%)	(0.03%)

Note. The top, middle and bottom panels of the Table report the changes in profit for P&G corresponding to changes in prices, non-price promotions and line length, respectively. The left (right) panel reports the absolute (percentage) change in profits. For prices, the profit changes are calculated based on a 5% increase in price and a 4% increase in input costs. For non-price promotionsand line length, profit changes are calculated based on a 1% increase in either variable, respectively. The reported numbers are calculated by bootstrapping over the first-stage standard errors of the estimated coefficients. We bootstrap over 500 iterations and report the standard errors in parenthesis. All estimates are statistically significant at the 99% confidence level.

References

Ailawadi, Kusum L, Jie Zhang, Aradhna Krishna and Michael W Kruger (2010), 'When wal-mart enters: How incumbent retailers react and how this affects their sales outcomes', *Journal of Marketing Research* **47**(4), 577–593.