Are Food Price Promotions Predictable? The Hazard Function of Supermarket Discounts

Hao Lan, Tim Lloyd, Wyn Morgan, Paul W. Dobson

Online Appendix

Table of Contents

A.1 Scanner dataset details

The data period runs from September 8th 2001 to April 17th 2004. The time series are contiguous (in that there are no missing observations once the time series has begun) in 100% of cases, although some (10%) start later than 8th September 2001. All time series finish in the week ending 17th April 2004. The minimum number of observations for any product is 103 weeks. Data on products are at a highly disaggregate level at barcode level, so separate sizes of the same brand are recorded as different products with separate prices. Many of the products are national brands sold by all retail chains, so the dataset contains retailer-specific prices for such products. Each retailer-product combination is identified with a Unique Product Code (UPC), with their own time series of weekly prices. Thus, it represents a UPC-level price dataset of U.K. food retailing. Nielsen only supplied prices (not quantities) at the chain level.

Private label products with the same product profile, e.g. an 800-gram standard medium sliced white loaf of bread, are treated as one product and have the same product code in the dataset. Retailer-specific prices for these products (i.e. UPCs) represent the Tesco private-label 800-gram standard medium sliced white loaf, or the Sainsbury private-label 800-gram standard medium sliced white loaf, for example. Hence, private label products are treated analogously to the branded products stocked by multiple retailers in the scanner dataset.

This dimension of the dataset potentially offers insights into any differences between the pricing of manufacturer- and retailer-branded products.

Owing to the vast number of individual purchases actually made, most studies involving scanner data rely on price data that have been averaged in some way. The Nielsen Scantrack scanner prices used in this study are no exception and represent unit values (or "average revenue" prices), meaning that the price recorded in any given week for a specific (barcoded) product by a particular retailer represents the ratio of the total value of sales (price times quantity) to the total quantity. As such, prices in each of the retailers' outlets are weighted by the proportion of total sales transacted at each posted price. To see this, consider a product which sells at posted prices $p_1 + p_2 + \ldots + p_K$ with frequency $q_1 + q_2 + \ldots + q_K$ respectively (i.e. q_1 units are sold at p_1 and q_2 units are sold at p_2 and so forth) over a period of a week by a given retailer. The average revenue price ($\bar{p}$) is

$$\bar{p} = \frac{p_1 q_1 + p_2 q_2 + \ldots + p_K q_K}{q_1 + q_2 + \ldots + q_K} = \frac{\sum_{i=1}^{K} p_i q_i}{\sum_{i=1}^{K} q_i}$$

Since $\sum_{i=1}^{K} q_i = Q$ as the total number of units sold during the week, the average revenue price is

$$\overline{p} = \frac{\sum_{i=1}^{K} p_i q_i}{Q} = \sum_{i=1}^{K} p_i \frac{q_i}{Q}$$

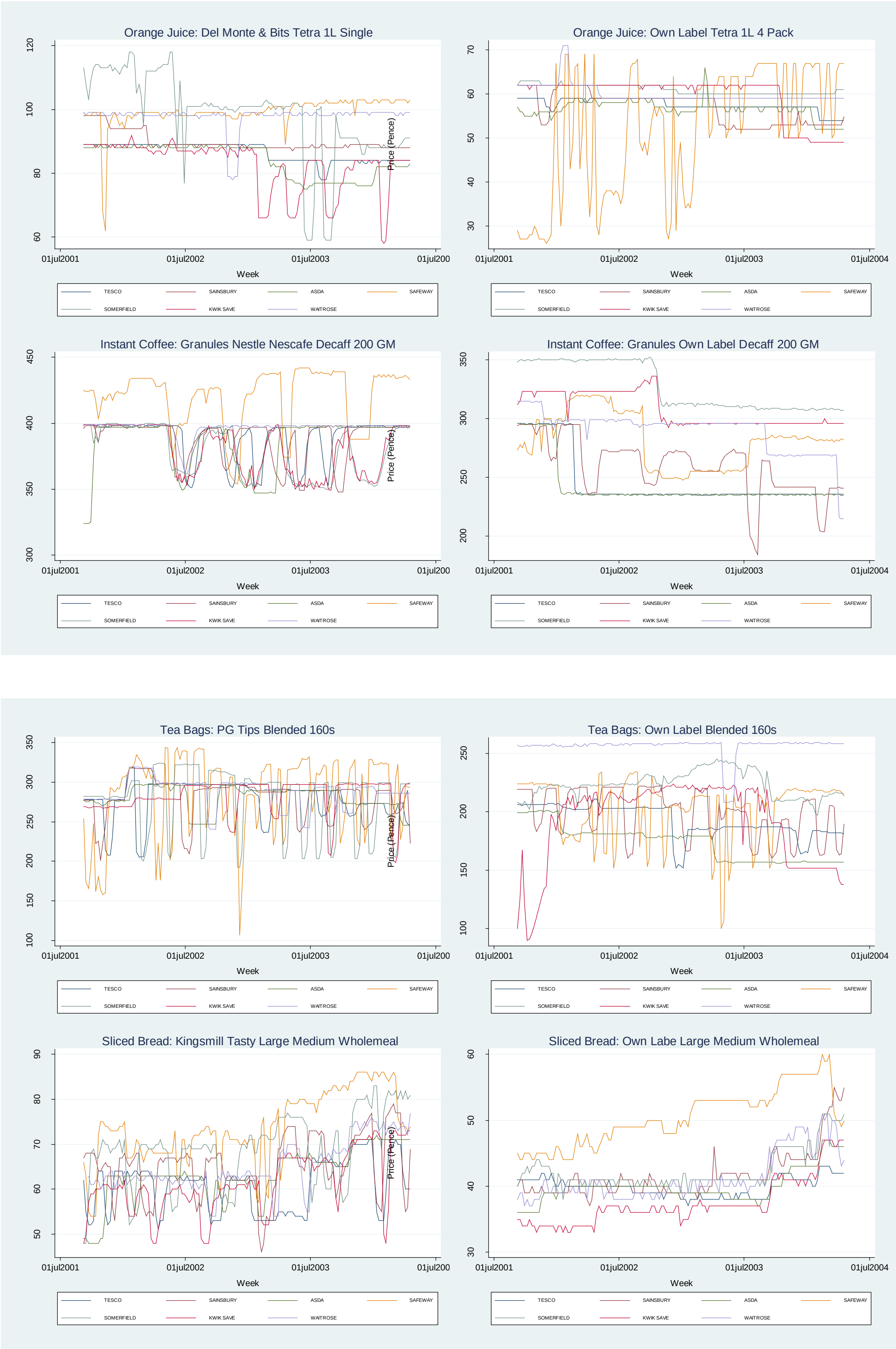
This is the average of each posted price weighted by the proportion of the units sold at that price. Hence, average revenue prices neatly reflect the relative importance of each of the prices at which the product was sold.[1] While ensuring that the average revenue price is representative, the weighting procedure does have implications that need to be borne in mind. Specifically, when using the frequency of changes of the average revenue price (and the duration that is implied by this frequency) to make inferences about the frequency of changes (and duration) of posted prices (i.e. those actually appearing on supermarket shelves).

When a retailer sells a product at the same price in all its outlets (i.e., in strictly following a national pricing policy) the average revenue price will only change if posted prices change. However, as soon as we leave the world of common pricing, a retailer's average revenue price reflects changes in both the posted price and the proportions sold at each of the posted prices in its outlets. In other words, the composition of purchases affects the (average revenue) price that is recorded; the upshot of which is that average revenue prices may change even if posted prices remain the same. The "fallacy of composition" inherent in average revenue prices has obvious implications when the frequency of a price change is the object of interest, since even a one penny change in price will change the frequency. In essence, average revenue prices such as those provided by Nielsen Scantrack will tend to overstate the frequency of price changes in any product that is offered for sale at different prices within any given retail chain because they vary with changes in posted prices and the composition of purchases. While the noise this creates in the price data (86% of all prices changes are 3% or less) is at odds with the precision of the EPOS monitoring (in which 100% of transactions are recorded) we do not expect it to affect the quality of the results in this paper simply because sales prices (by definition) are much lower than regular prices, and as such more easily distinguishable from noise. Nevertheless, in partly addressing this issue, the main results are repeated for 25% and 35% discount thresholds and reported in Table 4 in the main paper in respect of discount depth.

To give a flavor of the Nielsen scanner prices, **Figure A1-1** presents the prices of eight well-known branded and private label products in four categories across seven supermarkets in the dataset, where all the retail chains sell these items.

[1] A similar but distinct measure would weight posted prices by the number of stores selling the product at each price.

Figure A1-1. Weekly prices (pence) for a selection of products



As shown by the price lines for each of the different retailers in **Figure A1-1**, there are seven UPCs representing the national average prices in each of the retailers at weekly intervals. Without wishing to generalize, they display a number of interesting features, in particular the way in which sales punctuate the time series, albeit with a frequency and intensity that varies by product and retailer. When not on sale, prices tend to coalesce around particular levels, although this regular price is constant over the sample, particularly so for the products in tinned tomatoes and wrapped bread categories. It is also apparent that, despite representing the prices of identically barcoded products, there are persistent and substantial differences in the prices charged by retail chains.

In the scanner dataset, there are 1,704 UPCs with their distribution summarized in **Table A1-1**. Data are most prevalent in the bread (34%), soup (18%), coffee (8%) and orange juice (6%) product categories, each of which contains in excess of 100 UPCs. The least populated categories, such as frozen fish fingers (1%) and frozen pizza (1%), contain 20 UPCs each. However, the dataset is not a fully representative sample of consumer spending on food. In particular, fresh fruit and vegetables are not part of the dataset since they do not carry unique barcode indicators during the sample period. Even so, the range of categories covered is relatively broad, spanning beverages and foods including chilled, ambient, tinned and frozen foods. The most perishable category is wrapped bread.

Table A1-1. Distribution of UPCs by Product Category

Category	N	%	Format	N	%	Perishability	N	%
Breakfast Cereal	66	3.9	Ambient	348	20.0	Shelf storable	787	46.0
Corned Beef	30	1.8	Tinned	439	25.9	Shelf storable	787	46.0
Fish Fingers	20	1.2	Frozen	157	9.1	Perishable	917	54.0
Frozen Peas	34	2.0	Frozen	157	9.1	Perishable	917	54.0
Frozen Pizza	20	1.8	Frozen	157	9.1	Perishable	917	54.0
Instant Coffee	138	8.1	Ambient	348	20.0	Shelf storable	787	46.0
Jam	77	4.5	Ambient	348	20.0	Shelf storable	787	46.0
Orange Juice	108	6.3	Chilled	177	10.5	Perishable	917	54.0
Oven Chips	83	4.9	Frozen	157	9.1	Perishable	917	54.0
Tea Bags	67	3.9	Ambient	348	20.0	Shelf storable	787	46.0
Tinned Soup	308	18.1	Tinned	439	25.9	Shelf storable	787	46.0
Tinned Tomatoes	50	2.9	Tinned	439	25.9	Shelf storable	787	46.0
Tinned Tuna	51	3.0	Tinned	439	25.9	Shelf storable	787	46.0
Wrapped Bread	583	34.2	Fresh	583	34.5	Perishable	917	54.0
Yoghurt	69	4.1	Chilled	177	10.5	Perishable	917	54.0
Total	1704	100		1704	100		1704	100

Note: N refers to number of products in each classification

Table A1-2 summarizes the coverage of the dataset by retailer. The sample represents all seven national retailers well, where Tesco has the largest number of observations at 17%, while Waitrose has the least at 11%. The second and third highest are Sainsbury (at 16%) and Safeway (at 15%) respectively, both of which have more than 250 UPCs. One of the most interesting aspects of the dataset is that prices are available by retail chain, allowing time series analysis of identically barcoded products across retailers. Issues relating to the synchronization and staggering of prices (Lach and Tsiddon 1996; Berck et al. 2008) are most relevant here. In addition, the largest four retailers: Tesco, Sainsbury, Asda and Safeway (later replaced by Morrison) account for up to 63% of the total UPCs, which is consistent with the evidence that four major supermarkets dominate U.K. food retailing (Competition Commission 2000; 2003; 2008). There are 518 UPCs sold in all seven retail chains, accounting for 14% of the total.

Table A1-2. Distribution of UPCs by Retailer

	All Products	
Retailer	UPCs	% of total
Asda	228	13.4
Kwik Save	221	13.0
Safeway	263	15.4
Sainsbury	275	16.1
Somerfield	242	14.2
Tesco	292	17.1
Waitrose	183	10.7
Total	1704	100

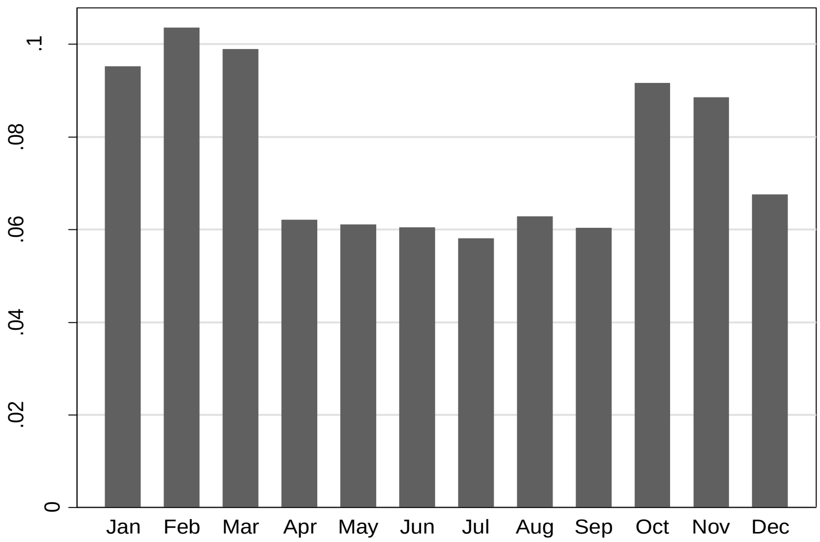
Table A1-3 reports the distribution of UPCs by brand status. The table shows that 82% of UPCs are branded products, with private label products therefore accounting for just under one-fifth of the UPCs in the dataset. This is mainly because private label products remained a relatively small proportion compared to brands during the sample period. As the table indicates, the composition of the subsample of common products is similar to that for all the products.

Table A1-3. Distribution of UPCs by Brand Status

	All Products		Common Products	
	UPCs	% of total	UPCs	% of total
National Brand	1399	82.1	546	84.8
Private Label	305	17.9	98	15.2
Total	1704	100.0	644	100.0

Figure A1-2 shows the distribution of sales by calendar month of the year. There is a clear seasonal split in our data, with promotional prices (discounts of 10% or more) being more prevalent in the autumn and winter, and correspondingly lower in spring and summer. In contrast to previous studies, there is dip in the proportion of sale prices in December. While this may simply reflect the composition of the dataset (fresh fruit and vegetables and festival foods such as a meat and alcohol are not included) monthly seasonal dummies augment all models in an attempt to shed some light on this issue in the following analysis.

Figure A1-2. The Seasonal Pattern of Food Sales in the U.K.



Note: Data represent the percentage of prices per month that are (10%+) promotional prices.

Finally, while our main analysis focuses on a discount threshold of >10%, we also consider deeper discounts with thresholds at >25% and >35%. **Table A1-4** shows the distribution of sales for the three different sales discount thresholds used, at >10%, >25% and >35% respectively, across the various classifications of the dataset. To visually highlight differences, we use traffic light colour coding. Those cells highlighted green (red) indicate where deeper sales are most common. From the table, it is apparent that deep sales are more highly concentrated than medium sales and are particularly apparent in two Hi-Lo retailers (Safeway and Somerfield) and in five categories (tinned tuna, oven chips, teabags, yoghurt and frozen pizza).

Table A1-4. The number of sales by discount depth threshold

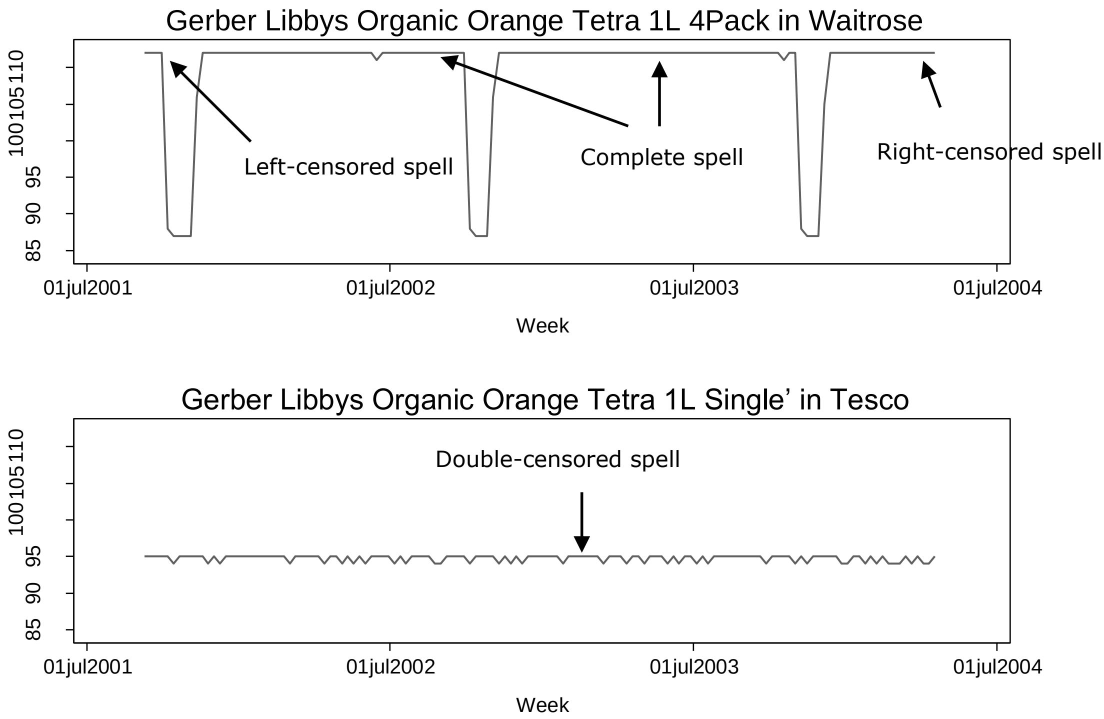
	All discounts (>10%)	Medium and deep discounts (>25%)	Medium and deep discounts share (%)	Deep discounts (>35%)	Deep discounts share (%)
TOTAL	3290	1429	43	498	15
Retailer					
Tesco	374	68	18	33	9
Sainsbury	418	115	28	22	5
Asda	33	4	12	1	3
Safeway	956	556	58	207	22
Somerfield	651	387	59	159	24
Kwik Save	598	215	36	68	11
Waitrose	260	84	32	8	3
Brand Type					
National Brands	2963	1301	44	458	15
Private Labels	327	128	39	40	12
Product Category					
Orange Juice	199	71	36	30	15
Instant Coffee	314	127	40	33	11
Tinned Tuna	139	104	75	63	45
Tinned Tomatoes	23	8	35	2	9
Tinned Soup	1007	547	54	147	15
Oven Chips	122	65	53	31	25
Corned Beef	53	25	47	2	4
Frozen Peas	30	8	27	0	0
Fish Fingers	7	1	14	0	0
Breakfast Cereal	115	40	35	14	12
Tea Bags	189	109	58	54	29
Yoghurt	220	154	70	60	27
Wrapped Bread	747	90	12	18	2
Jam	55	29	53	3	5
Frozen Pizza	199	51	73	41	59
Product Format					
Tinned	1222	684	56	214	18
Ambient	673	305	45	104	15
Frozen	229	125	55	72	31
Chilled	419	225	54	90	21
Fresh	747	90	12	18	2
Perishability					
Shelf-storable	1395	440	32	180	13
Perishable	1895	989	52	318	17

Note: traffic light colour coding: (i) yellow indicates depth threshold average (+/- 2%); red indicates value less (-2%+) than the average; green indicates value greater (+2%+) than the average.

A.2 Censored spells in chain level prices

As is standard in duration econometrics, the main analysis uses a dataset of complete and right-censored spells. In the original spell dataset, we identify 6,007 regular price spells, of which 55% are complete, 18% left-censored, 17% right-censored and 10% double-censored. A left-censored spell is a regular price spell that starts by a sale before the sampling frame begins and is terminated by a sale within the sample; a right-censored spell denotes a spell that begins at a point of sale within the sample but continues beyond the end of the sample frame. If a spell is both left- and right-censored simultaneously, it is commonly called a double-censored spell (and not on sale duration the sample). There are examples of complete, left-censored and right-censored regular price spells in **Figure A2-1**. In the figure, the top price series is Gerber Libbys Organic Orange juice (Tetra 1L 4Pack) in Waitrose and the bottom one is Gerber Libbys Organic Orange Juice (Tetra 1L Single) in Tesco. A left-censored, complete, and right-censored spells are created from the top price series respectively; a double-censored spell is created from the bottom price series.

Figure A2-1.- Examples of types of censored spells



Using Maximum Likelihood methods, it is possible to handle the likelihood of right-censored spells using the survival function (i.e. $S(t_c) = 1 - F(t_c)$ where t_c is a fixed known censoring time). Double-censored spells measure UPC items that are never on sale during the sample period. Since our main interest is in the “time to sale” and the factors that affect this duration, double-censored spells are discarded since this represents the regular price spells of products that have never been on sale. It is also common to discard left censored spells, since to include them requires modification of the likelihood function which is both technically challenging requires regarding the rate at which regular spells occur across the various classifications of the data (see D'Addio and Rosholmand 2002; Cameron and Trivedi 2005).

To investigate the impact of left censoring we begin by reviewing the distribution of left censored cells (**Table A2-1)**. While they appear to be quite evenly distributed across the key classifications in the sample, they are relatively numerous in Asda and tinned tomatoes (although few in absolute number in the sample as a whole). Results from the re-estimation of hazard models by category and by retailer including left censored cells are reported in **Table A2-2** and **Table A2-3** and suggest that the shape of the hazard function is typically invariant to the treatment of left censored spells. Specifically, estimates of time dependence are positive in all categories except orange juice and wrapped bread and in all retailers except Safeway suggesting that discarding left censored data has a limited impact on the estimation. Nakamura and Steinsson (2008) arrive at a similar conclusion.

Table A2-1. Distribution of censored spells by retailer and category

	Left-censored spell	
	Frequency	% in group
ASDA	85	25.8
KWIK SAVE	149	15.5
SAFEWAY	206	14.6
SAINSBURY	193	22.1
SOMERFIELD	177	16.7
TESCO	162	19.8
WAITROSE	122	22.3
Breakfast Cereal	39	17.8
Corned Beef	20	19.4
Fish Fingers	8	22.9
Frozen Peas	14	17.9
Frozen Pizza	17	15.9
Instant Coffee	99	18.1
Jam	29	18.0
Orange Juice	66	17.7
Oven Chips	54	22.5
Tea Bags	48	15.9
Tinned Soup	269	17.3
Tinned Tomatoes	31	31.6
Tinned Tuna	31	14.2
Wrapped Bread	311	19.2
Yoghurt	58	16.7
Total	1094	18.2

Table A2-2. Hazard functions including left-censored spells by category

	Orange Juice	Instant Coffee	Tinned Tuna	Tinned Tomatoes	Tinned Soup	Oven Chips	Corned Beef	Breakfast Cereal	Tea Bags	Yoghurt	Wrapped Bread	Frozen Pizza
Baseline (p)	0.963	1.101**	1.167**	1.156	0.999	1.141**	1.226**	1.147**	1.181***	1.631***	0.982	1.708***
	(0.050)	(0.045)	(0.076)	(0.139)	(0.021)	(0.067)	(0.118)	(0.077)	(0.064)	(0.072)	(0.026)	(0.142)
label	1.894***	0.616	1.000	1.000	0.717***	1.000	2.952*	1.000	1.022	0.314**	0.701*	1.000
	(0.420)	(0.186)	(.)	(.)	(0.092)	(.)	(1.697)	(.)	(0.484)	(0.176)	(0.146)	(.)
rival	3.199***	1.072	1.484*	7.675***	1.827***	0.984	7.889***	2.191***	1.199	2.824***	1.158*	0.464**
	(0.704)	(0.143)	(0.324)	(5.714)	(0.134)	(0.163)	(2.814)	(0.450)	(0.242)	(0.508)	(0.093)	(0.150)
label $\times$ *rival*	0.520**	1.885	1.000	1.000	1.370	1.000	0.294	1.000	0.931	1.000	12.209**	1.000
	(0.166)	(0.738)	(.)	(.)	(0.299)	(.)	(0.277)	(.)	(0.443)	(.)	(13.572)	(.)
Month dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
UPC Random Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LR test of σ_α^2=0 (P value)	0.001	0.000	0.003	0.092	1.000	1.000	1.000	0.010	0.000	0.000	0.000	0.000
N	331	507	199	79	1518	211	93	192	282	336	1351	104

Notes: standard error is reported in brackets. ***, ** and * denote that the null hypothesis (of unity in the hazard ratio) is rejected at the 1%, 5% and 10% level. . denote categories with insufficient data. LR test refers to likelihood –ratio test.

Table A2-3. Hazard functions including left-censored spells by retailer

	TESCO	SAINSBURY	ASDA	SAFEWAY	SOMERFIELD	KWIK SAVE	WAITROSE
Baseline (p)	1.023	1.135***	1.240***	0.917***	1.254***	1.152***	1.284***
	(0.039)	(0.040)	(0.097)	(0.020)	(0.037)	(0.035)	(0.057)
label	0.879	0.638**	0.206**	0.844	0.509**	1.064	1.217
	(0.217)	(0.131)	(0.136)	(0.154)	(0.157)	(0.325)	(0.277)
rivals	1.564***	1.569***	0.955	1.635***	1.261**	1.288**	1.826***
	(0.156)	(0.181)	(0.251)	(0.132)	(0.125)	(0.134)	(0.294)
label $\times$ *rival*	0.712	1.062	5.499	0.931	1.926	0.660	1.202
	(0.302)	(0.329)	(6.937)	(0.234)	(0.809)	(0.242)	(0.608)
Category dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
UPC Random Effects	Yes	Yes	No	Yes	Yes	Yes	Yes
LR test of σ_α^2=0 (P value)	0.427	0.000	0.496	0.000	0.000	0.000	0.000
N	689	793	186	1357	997	888	487

Notes: Coefficients minus one represent the proportional change in the occurrence of a sale. Standard error is reported in brackets. ***, ** and * denote that the null hypothesis of no effect (i.e. unity in the hazard ratio) is rejected at the 1%, 5% and 10% level respectively. LR test refers to likelihood –ratio test.

A.3 Testing Heterogeneity

In non-parametric analysis, it is common to test the equality of hazard functions across two or more groups. **Table A3-1** reports the log rank test proposed by Mantel and Haenszel (1959) is employed and shows that, in all cases, the null hypothesis of equality is rejected at the 1% significance level using $\chi^2(n)$ statistic indicating that the significant heterogeneity across each of the (n) retailers, categories, formats, shelf storability and brand status in the sample.

Table A3-1. Log-rank test results of heterogeneity

Heterogeneous group	Log-rank test
Retailer	χ^2 (6) = 646.90***
Category	χ^2(14) = 124.87***
Format	χ^2(4) = 39.33***
Storability	χ^2(1) = 24.64***
Brand Status	χ^2 (1) = 79.21***

Note: ***, ** and * denote that the null hypothesis of equality is rejected at the 1%, 5% and 10% level respectively.

A.4 Specification Testing

In the paper, we report four model specifications in Table 2 to illustrate the effect of UPC-level heterogeneity on the time dependence of sales, finding that the sign of the hazard changes when sales in rival retailers ($rival$) and UPC-specific random effects (UPC-RE) are accounted for. Here we check the robustness of this result using alternative model specifications, as summarized in **Table A4-1**. At the head of the table is the hazard ratio for the Model 2 from the main paper (UPC-RE and rival). Model A1 uses UPC-level random effects (UPC-RE) only without sales in rival retailers ($rival$). Model 2 uses UPC-FE with rival. Model A3 uses UPC-FE alone. Model A4 has neither panel estimator but pools the data like Model 1 in main paper although it does includes sales in rival retailers ($rival$).

Positive time dependence (rising hazard) is invariant to whether we use fixed or random effects (alone or in combination with $rival$) and produces similar estimates of the hazard. Random effects is preferred given that UPC level fixed effects vastly reduces the degrees of freedom and wipes out all the time invariant variables. Model 4 suffers from the same aggregation bias present in Model 1 in the main paper underlining the important role the panel estimator performs in taking account of the microeconomic heterogeneities.

Table A4-1. Specification

Time dependence parameter (p)		
	Hazard Ratio	Standard Error
Model 2 : UPC-RE, Rival	**1.135*** **	**(0.017)**
Model A1: UPC-RE	1.126***	(0.017)
Model A2: UPC-FE; rival	1.304***	(0.018)
Model A3: UPC-FE	1.290***	(0.018)
Model A4: Pooled	0.898***	(0.012)

Notes: ***, ** and * denote that the null hypothesis (of unity in the hazard ratio) is rejected at the 1%, 5% and 10% level. All RE models contain label and monthly seasonal dummies. “rival” refers to rival price variable. See text for details.

A.5 Hazard models by retailer (ASDA specification check)

Table 3 in the main paper reports hazard functions by retailer. Mindful of brevity and the implications for the degrees of freedom that Model 2 implies results are presented in the main paper for Model 4 which collapse the 15 categories of food into the binary classification 'shelf-storable'. Even with this most parsimonious specification it is not possible to estimate the Asda model with random effects. To assess whether the absence of random effects has a bearing on the time dependence for Asda, **Table A5-1** reports alternative models that exclude some of the other variables but that permit estimation of the Asda model with random effects. All variations support the finding that Asda's baseline hazard is time invariant (flat). In addition, no rejection of the likelihood-ratio test of $\sigma_\alpha^2 = 0$ also supports that UPC level random effects play little role in the Asda model.

Table A5-1. Various Simplifications of the Hazard function (Model 4) for Asda

	(1)	(2)	(3)	(4)
Baseline (p)	0.852	0.851	0.841	0.924
	(0.141)	(0.141)	(0.139)	(0.165)
label		0.849		
		(0.916)		
rival	1.461	1.445	1.489	1.571
	(0.583)	(0.585)	(0.590)	(0.726)
Shelf storable			0.599	
			(0.226)	
UPC Random Effects	Yes	Yes	Yes	Yes
LR test of σ_α^2=0 (P value)	0.192	0.195	0.236	0.252
Observations	101	101	101	101

Notes: Specification (1) includes rival sales variable only; (2) includes both label and rival sales variable; (3) includes rival sales and storable dummy; (4) includes rival sales and monthly dummies. Other combinations of effects do not estimate owing to insufficient observations. Notes from previous table apply here.

A.6 Hazard models by category

Here we explore any additional heterogeneity in the regular price spells. At issue is the effect of category level heterogeneity on the estimation of the model specifications in Table 2 in the main paper. By estimating hazard function at the category level, we are able to examine different shapes of hazard function across categories. For this robustness check, **Table A6-1** reports the estimation of Model 2 in Table 2 in the paper for each category of food in turn. Three product categories (fish fingers, frozen peas, and jam) could not be estimated due to insufficient data. In all categories the estimated time dependence is greater than unity; significantly so in all but wrapped bread. The estimates of p vary in accordance with the category hazard ratios in Table 2 of the main paper, with the occurrence of sales being highest in frozen pizza and yoghurt. The category breakdown reveals some differences: the private label effect is not common to all categories, private labels being promoted more frequently than brands in these two categories. Other results are similar to those implied by the hazard ratios in Table 2 of the paper, so qualitative conclusions remain unchanged. Importantly, the effect of staggering of prices is common in most categories and where significant it is found regardless of brand status in the manner reported in the paper.

Table A6-1. Hazard model by category (Model 2)

	Orange Juice	Instant Coffee	Tinned Tuna	Tinned Tomatoes	Tinned Soup	Oven Chips	Corned Beef	Breakfast Cereal	Tea Bags	Yoghurt	Wrapped Bread	Frozen Pizza
Baseline (p)	1.040	1.108**	1.273***	1.095	1.504***	1.398***	1.469***	1.122	1.368***	1.782***	0.999	2.174***
	(0.062)	(0.052)	(0.092)	(0.198)	(0.037)	(0.107)	(0.171)	(0.092)	(0.084)	(0.084)	(0.032)	(0.192)
label	2.348***	0.579	1.000	1.000	0.451***	1.000	5.185**	1.000	1.044	0.528	0.582*	1.000
	(0.698)	(0.211)	(.)	(.)	(0.122)	(.)	(3.747)	(.)	(0.639)	(0.358)	(0.176)	(.)
rival	8.426***	1.489**	2.078***	61.365***	2.384***	0.843	37.182***	3.047***	1.711**	3.639***	1.411***	0.961
	(2.525)	(0.232)	(0.504)	(83.436)	(0.268)	(0.207)	(21.321)	(0.760)	(0.406)	(0.684)	(0.143)	(0.365)
label × *rival*	0.243***	1.377	1.000	1.000	2.128*	1.000	0.081*	1.000	1.347	1.000	1.000	1.000
	(0.100)	(0.667)	(.)	(.)	(0.863)	(.)	(0.117)	(.)	(0.795)	(.)	(.)	(.)
SAINSBURY	0.943	1.027	0.115**	0.747	0.497***	1.049	0.219*	0.583	1.310	0.648	0.932	10.118**
	(0.559)	(0.271)	(0.101)	(0.845)	(0.116)	(0.581)	(0.194)	(0.241)	(0.777)	(0.385)	(0.244)	(11.117)
ASDA	1.158	0.000	0.000	0.710	0.145***	1.654	0.600	0.126***	0.000	0.000	0.491*	0.000
	(1.474)	(0.000)	(0.000)	(1.161)	(0.056)	(1.584)	(0.725)	(0.084)	(0.000)	(0.000)	(0.211)	(0.000)
SAFEWAY	2.988**	1.435	1.878	0.235	14.549***	7.557***	1.276	1.106	5.658***	2.501*	1.587*	5.484*
	(1.619)	(0.400)	(0.968)	(0.357)	(3.155)	(3.597)	(0.656)	(0.433)	(3.143)	(1.211)	(0.399)	(5.023)
SOMERFIELD	1.915	0.456*	0.715	1.209	1.221	3.116**	1.334	0.646	2.456	14.068***	2.198***	4.949
	(1.086)	(0.205)	(0.330)	(1.389)	(0.263)	(1.668)	(0.661)	(0.262)	(1.361)	(6.495)	(0.546)	(4.939)
KWIK SAVE	6.666***	0.460**	0.505	0.000	1.977***	1.233	3.028**	0.442*	1.563	9.710***	1.792**	0.821
	(3.693)	(0.167)	(0.276)	(0.000)	(0.453)	(0.700)	(1.638)	(0.212)	(0.878)	(4.531)	(0.458)	(0.807)
WAITROSE	1.033	0.322***	0.296	0.697	1.969**	3.629**	5.610**	0.370**	0.971	0.562	1.504	0.163
	(0.634)	(0.136)	(0.227)	(0.972)	(0.539)	(1.996)	(4.632)	(0.180)	(0.609)	(0.396)	(0.443)	(0.182)
Month dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
UPC Random Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LR test of $\sigma_\alpha^2=0$ (P value)	0.000	0.000	0.021	1.000	0.000	0.056	1.000	0.177	0.000	0.000	0.000	0.006
N	265	408	168	48	1249	157	73	153	234	278	1040	87

Notes: standard error is reported in brackets. ***, ** and * denote that the null hypothesis (of unity in the hazard ratio) is rejected at the 1%, 5% and 10% level. . denote categories with insufficient data. LR test refers to likelihood–ratio test.

A.7 Hazard model by perishability and format

Here we present additional estimation results regarding perishability in the data in **Table A7-1**. We replicate Model 3 and Model 4 in Table 2 by storability and format dummies. The format dummy is created to be Tinned (Tinned Tuna, Tinned Tomatoes, Tinned Soup, Corned Beef), Ambient (Instant Coffee, Breakfast Cereal, Tea Bags, Jam), Frozen (Oven Chips, Frozen Peas, Fish Fingers, Frozen Pizza), Chilled (Orange Juice, Yogurt), and Fresh (Wrapped Bread). Furthermore, the Storability dummy is defined as Perishable (fresh, chilled and frozen) and Shelf storable (ambient and tinned).

Table A7-1. Hazard model by perishability and format

	Perishable	Shelf storable	Tinned	Ambient	Frozen	Chilled	Fresh
Baseline (p)	1.037	1.235***	1.370***	1.135***	1.250***	1.177***	0.999
	(0.024)	(0.024)	(0.032)	(0.037)	(0.072)	(0.048)	(0.032)
label	0.969	0.715**	0.499***	1.116	1.000	1.364	0.582*
	(0.165)	(0.118)	(0.120)	(0.259)	(.)	(0.322)	(0.176)
rival	1.684***	2.200***	2.645***	1.714***	1.153	3.957***	1.411***
	(0.124)	(0.159)	(0.250)	(0.197)	(0.207)	(0.596)	(0.143)
label×rival	1.354	0.713	1.900*	0.449***	1.000	0.660	1.000
	(0.397)	(0.153)	(0.729)	(0.122)	(.)	(0.210)	(.)
SAINSBURY	1.012	0.696**	0.396***	1.014	1.154	0.959	0.932
	(0.200)	(0.108)	(0.082)	(0.230)	(0.494)	(0.399)	(0.244)
ASDA	0.442**	0.108***	0.124***	0.063***	0.192**	0.303	0.491*
	(0.152)	(0.033)	(0.044)	(0.036)	(0.158)	(0.348)	(0.211)
SAFEWAY	2.145***	3.733***	7.240***	1.500*	4.252***	2.150**	1.587*
	(0.394)	(0.562)	(1.389)	(0.329)	(1.577)	(0.796)	(0.399)
SOMERFIELD	2.869***	0.980	1.062	0.797	2.481**	4.737***	2.198***
	(0.537)	(0.154)	(0.203)	(0.208)	(1.034)	(1.754)	(0.546)
KWIK SAVE	2.775***	1.034	1.496*	0.542**	1.573	6.033***	1.792**
	(0.528)	(0.171)	(0.313)	(0.138)	(0.676)	(2.231)	(0.458)
WAITROSE	1.257	0.996	1.279	0.585*	1.699	0.668	1.504
	(0.266)	(0.187)	(0.314)	(0.162)	(0.782)	(0.298)	(0.443)
Month dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
UPC random effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
LR test of σ_α^2=0 (P value)	0.000	0.000	0.000	0.000	0.000	0.000	0.000
N	1886	2417	1538	879	303	543	1040

Notes from previous table apply here.

A.8 Definition of rival sales variable

Here we report the estimation results of Model 4 in Table 2 from main analysis using different definitions of rival sales. As we mentioned in footnote 7 in the main paper, we also estimate models for sales in the previous week and previous fortnight. **Table A8-1** reports the robustness estimation results. It shows that these models produce qualitatively similar results.

Table A8-1. Estimation results using different rival sale definitions (Model 4)

	Previous week		Previous fortnight	
	Hazard	S.E.	Hazard	S.E.
Baseline (p)	1.135***	(0.017)	1.131***	(0.017)
label	0.871	(0.102)	0.856	(0.101)
rival	1.908***	(0.094)	1.815***	(0.091)
$label \times rival$	0.614***	(0.109)	0.752	(0.131)
shelf storable	1.353***	(0.098)	1.340***	(0.097)
SAINSBURY	0.794*	(0.099)	0.789*	(0.099)
ASDA	0.191***	(0.044)	0.185***	(0.042)
SAFEWAY	2.657***	(0.316)	2.631***	(0.312)
SOMERFIELD	1.661***	(0.207)	1.676***	(0.209)
KWIK SAVE	1.796***	(0.231)	1.742***	(0.224)
WAITROSE	1.052	(0.150)	1.033	(0.147)
Month dummy	Yes		Yes	
UPC Random Effects	Yes		Yes	
LR test of σ_α^2=0 (P value)	0.000		0.000	
N	4303		4303	

Notes: Coefficients minus one represent the proportional change in the occurrence of a sale. Standard error is reported in brackets. ***, ** and * denote that the null hypothesis of no effect (i.e. unity in the hazard ratio) is rejected at the 1%, 5% and 10% level respectively. LR test refers to likelihood –ratio test.

A.9 Weekly dummies

Here we replicate the estimation in Table 2 in the main analysis using weekly dummies. The results are present in **Table A9-1**. The key estimates are similar to the ones in Table 2 in the main analysis. Owing to the fact that degree of freedom are significantly reduced by including weekly dummies we report monthly dummies, which appear to be effective in controlling for time heterogeneity.

Table A9-1. Proportional hazard models of sales in U.K. food retailing

	Model 1		**Model 2**		**Model 3**		**Model 4**	
	Hazard	S.E.	Hazard	S.E.	Hazard	S.E.	Hazard	S.E.
Baseline (p)	0.887***	(0.012)	1.174***	(0.017)	1.174***	(0.017)	1.174***	(0.017)
Instant Coffee	1.299***	(0.123)	1.227	(0.229)				
Tinned Tuna	1.642***	(0.191)	1.662**	(0.403)				
Tinned Tomatoes	0.352***	(0.079)	0.345***	(0.112)				
Tinned Soup	1.596***	(0.133)	1.502**	(0.243)				
Oven Chips	0.822	(0.098)	0.742	(0.165)				
Corned Beef	0.941	(0.148)	0.953	(0.277)				
Frozen Peas	0.717*	(0.143)	0.652	(0.236)				
Fish Fingers	0.691	(0.269)	0.642	(0.368)				
Breakfast Cereal	1.396***	(0.170)	1.415	(0.336)				
Tea Bags	1.309**	(0.137)	1.214	(0.255)				
Yoghurt	1.627***	(0.173)	1.459*	(0.298)				
Wrapped Bread	0.890	(0.075)	0.935	(0.151)				
Jam	0.662***	(0.106)	0.620*	(0.173)				
Frozen Pizza	1.515***	(0.219)	1.372	(0.410)				
Ambient					1.124	(0.138)		
Frozen					0.963	(0.127)		
Chilled					0.683**	(0.114)		
Fresh					0.766**	(0.096)		
Shelf Storable							1.317***	(0.098)
SAINSBURY	0.979	(0.074)	0.987	(0.126)	0.955	(0.124)	0.950	(0.124)
ASDA	0.256***	(0.048)	0.217***	(0.051)	0.211***	(0.050)	0.213***	(0.050)
SAFEWAY	2.403***	(0.159)	3.275***	(0.400)	3.169***	(0.391)	3.198***	(0.394)
SOMERFIELD	1.705***	(0.121)	1.896***	(0.242)	1.801***	(0.233)	1.856***	(0.240)
KWIK SAVE	1.708***	(0.120)	1.932***	(0.254)	1.890***	(0.252)	1.984***	(0.265)
WAITROSE	1.127	(0.096)	1.164	(0.171)	1.098	(0.162)	1.091	(0.161)
Label	0.755***	(0.048)	0.869	(0.114)	0.793*	(0.100)	0.839	(0.103)
Rival			1.914***	(0.101)	1.980***	(0.104)	1.987***	(0.105)

label×rival		0.970 (0.171)	0.867 (0.149)	0.839 (0.144)
Weekly dummy	Yes	Yes	Yes	Yes
UPC Random Effects	No	Yes	Yes	Yes
Likelihood-ratio test of σ_{α}^{2}=0 (P value)		0.000	0.000	0.000
Observations	4303	4303	4303	4303
UPCs	1703	1703	1703	1703

Notes: Coefficients minus one represent the proportional change in the occurrence of a sale. Results are relative to the base group, orange juice, branded products stocked by the market leader, Tesco in Model 2 (chilled in Model 3 and Perishable Model 3). ***, ** and * denote that the null hypothesis of no effect (i.e. unity in the hazard ratio) is rejected at the 1%, 5% and 10% level respectively.

A.10 Results reporting estimates of monthly dummies

To save space in the main paper, coefficients on the monthly dummies in the models in Table 2 were not reported. **Table A10-1** reports these estimates, in particular the hazard ratio of 1.274 for December commented on in the paper.

Table A10-1. Proportional hazard models of sales in U.K. food retailing

	Model 1		**Model 2**		**Model 3**		**Model 4**	
	Hazard	S.E.	Hazard	S.E.	Hazard	S.E.	Hazard	S.E.
Baseline (p)	0.868***	(0.012)	1.133***	(0.017)	1.133***	(0.017)	1.133***	(0.017)
Breakfast Cereal	1.267**	(0.152)	1.347	(0.311)				
Corned Beef	0.946	(0.147)	0.981	(0.278)				
Fish Fingers	0.666	(0.258)	0.593	(0.330)				
Frozen Peas	0.733	(0.145)	0.725	(0.252)				
Frozen Pizza	1.388**	(0.197)	1.288	(0.371)				
Instant Coffee	1.203**	(0.112)	1.205	(0.217)				
Jam	0.644***	(0.100)	0.599*	(0.161)				
Oven Chips	0.840	(0.099)	0.773	(0.166)				
Tea Bags	1.292**	(0.134)	1.218	(0.248)				
Tinned Soup	1.463***	(0.118)	1.405**	(0.219)				
Tinned Tomatoes	0.323***	(0.072)	0.344***	(0.109)				
Tinned Tuna	1.567***	(0.178)	1.688**	(0.398)				
Wrapped Bread	0.848**	(0.070)	0.889	(0.139)				
Yoghurt	1.304***	(0.133)	1.255	(0.247)				
Ambient					1.014	(0.129)		
Fresh					0.783**	(0.094)		
Frozen					0.747*	(0.120)		
Tinned					1.151	(0.135)		
Shelf Storable							1.324***	(0.095)
ASDA	0.224***	(0.041)	0.194***	(0.044)	0.187***	(0.043)	0.188***	(0.043)
KWIK SAVE	1.550***	(0.104)	1.701***	(0.215)	1.657***	(0.212)	1.721***	(0.220)
SAFEWAY	2.031***	(0.127)	2.658***	(0.311)	2.576***	(0.304)	2.602***	(0.306)
SAINSBURY	0.859**	(0.062)	0.825	(0.101)	0.796*	(0.099)	0.789*	(0.098)
SOMERFIELD	1.503***	(0.100)	1.708***	(0.210)	1.623***	(0.201)	1.662***	(0.206)
WAITROSE	1.031	(0.085)	1.112	(0.156)	1.048	(0.148)	1.042	(0.148)
Rival	0.753***	(0.048)	0.887	(0.113)	0.823	(0.100)	0.861	(0.103)
Label			1.873***	(0.096)	1.936***	(0.099)	1.941***	(0.100)
label×rival			0.934	(0.163)	0.830	(0.142)	0.812	(0.139)
February	0.990	(0.083)	1.026	(0.098)	1.022	(0.098)	1.021	(0.098)
March	0.706***	(0.058)	0.733***	(0.070)	0.733***	(0.070)	0.742***	(0.071)
April	0.978	(0.089)	0.891	(0.093)	0.876	(0.091)	0.874	(0.091)

May	1.010	(0.096)	0.726***	(0.079)	0.718***	(0.078)	0.721***	(0.079)
June	1.020	(0.098)	0.919	(0.104)	0.900	(0.102)	0.896	(0.101)
July	1.481***	(0.157)	1.375**	(0.171)	1.338**	(0.165)	1.324**	(0.164)
August	1.305***	(0.126)	1.076	(0.119)	1.047	(0.116)	1.035	(0.114)
September	1.217**	(0.115)	1.177	(0.129)	1.156	(0.127)	1.144	(0.126)
October	1.348***	(0.114)	1.242**	(0.123)	1.225**	(0.122)	1.226**	(0.122)
November	1.082	(0.088)	1.097	(0.102)	1.084	(0.101)	1.082	(0.100)
December	1.122	(0.093)	1.274**	(0.125)	1.265**	(0.124)	1.259**	(0.124)
UPC Random Effects	No		Yes		Yes		Yes	
Likelihood-ratio test of σ_{α}^{2}=0 (P value)			0.000		0.000		0.000	
Observations	4303		4303		4303		4303	
UPCs	1703		1703		1703		1703	

Notes: Coefficients minus one represent the proportional change in the occurrence of a sale. Results are relative to the base group, orange juice, branded products stocked by the market leader, Tesco in Model 2 (chilled in Model 3 and Perishable Model 3). ***, ** and * denote that the null hypothesis of no effect (i.e. unity in the hazard ratio) is rejected at the 1%, 5% and 10% level respectively.

References

Berck, P., J. Brown, J.M. Perloff, and S.B. Villas-Boas. 2008. “Sales: tests of theories on causality and timing.” *International Journal of Industrial Organization*, 26(6): 1257-1273.

Cameron, A. and P. Trivedi. 2005. *Microeconometrics: methods and applications*. Cambridge University Press.

Competition Commission. 2000. *Supermarkets: A Report on the Supply of Groceries from Multiple Stores in the United Kingdom*. London: HMSO.

Competition Commission. 2003. *Safeway plc and Asda Group Limited (owned by Wal-Mart Stores Inc); Wm Morrison Supermarkets PLC; J Sainsbury plc; and Tesco plc: A report on the mergers in contemplation*. London: HMSO.

Competition Commission. 2008. *The Supply of Groceries in the UK*, London: HMSO.

D'Addio, A. and Rosholm, M. 2006, “Left-Censoring in Duration Data: Theory and Applications,” Economics Working Papers 2002-5, Department of Economics and Business Economics, Aarhus University.

Lach, S. and D. Tsiddon. 1996. “Staggering and synchronization in price-setting: evidence from multiproduct firms.” *American Economic Review*, 86(5): 1175-1196.

Mantel, N. and W. Haenszel. 1959. “Statistical aspects of the analysis of data from retrospective studies of disease.” *Journal of National Cancer Institute*, 22, 719–748.

Nakamura, E. and J. Steinsson. 2008. “Five facts about prices: A reevaluation of menu cost models.” *Quarterly Journal of Economics*, 123(4): 1415-1464.