Appendix A

Do promotions benefit manufacturers, retailers, or both?

Shuba Srinivasan, Koen Pauwels, Dominique M. Hanssens, Marnik G. Dekimpe

A.1. Data: Variable Operationalization

In the first section, we provide information on the Dominick's data used in the paper, including the operationalization of manufacturer and retailer performance measures, holiday dummy variables and summary information on the categories in the data set.

A.1.1 Manufacturer and retailer performance measures

For the brands in a category, we consider brand sales as well as manufacturer revenues, defined as:
where $MS_{i,t}$ refers to market share of brand *i* at time *t*, Q_t is the category sales and $WP_{i,t}$ is the wholesale price of brand *i* at time *t*.

$$MR_{i,t} = MS_{i,t} \times Q_t \times WP_{i,t}$$

For the retailer, in addition to category sales, we also derive the total category revenue for the retailer as:

$$RR_t = \sum_{i=1}^{n} MS_{i,t} \times Q_t \times P_{i,t}$$

where $P_{i,t}$ refers to the price of brand *i* at time *t* and *n* is the total number of brands in a category. As both retailer and manufacturer revenues are expressed in dollars, the relative changes in $MR_{i,t}$ and RR_t due to a given price promotion yield insights into the division of promotional benefits between manufacturer and retailer. Additionally, we compute the retailer's total category margins (defined in dollars) as:

$$RM_t = \sum_{i=1}^{n} MS_{i,t} \times Q_t \times (P_{i,t} - WP_{i,t})$$

We note that the wholesale-price measure $WP_{i,t}$ does not capture the replacement cost of the item in a given week, but rather the average acquisition cost ($AAC_{i,t}$) of all items in inventory in that week. $AAC_{i,t}$ is obtained as a weighted average of the price paid by the retailer for brand *i* in week *t* and the retailer's average acquisition cost in *t-1* (for a description, see http://gsbwww.uchicago.edu/research/mkt/Databases/DFF/W.html). Even though this measure has been used repeatedly in the marketing literature (e.g. Besanko et al. 2001; Chevalier et al. 2000; Chintagunta 2002), there has been some concern that it may be sensitive to both sluggish adjustment and forward buying practices of the retailer. Moreover, we acknowledge that our manufacturer revenue measure for a given week does not perfectly reflect true manufacturer revenue for that same week since the timing of manufacturer trade deals and retail price discounts need not coincide. Still, manufacturer shipments ultimately increase to the extent that a consumer promotion is more or less successful in increasing end-user demand for the brand. Similarly, changes in the retailer's average acquisition cost do reflect changes in manufacturer revenue per unit. Therefore, the average acquisition cost seems an appropriate measure to study the total effects of price promotions on manufacturer revenues and retailer margins. Store revenue is captured by the total dollar sales summed over all Dominick's-defined departments for a given week. Store traffic is defined as the total number of customers visiting the store and buying at least one item in a given week.

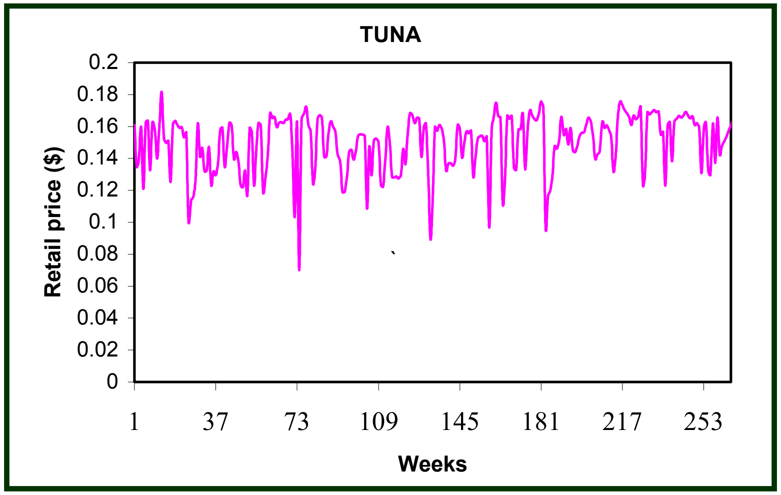
A.1.2 Holiday dummy variables

The total demand at most retail chains is volatile around major holidays. Hence, following Chevalier et al. (2000), we incorporate dummy variables that equal one in the shopping periods around the following holidays: Lent, Easter, Memorial Day, July 4th, Labor Day, Thanksgiving, the week following Thanksgiving, Christmas and the Superbowl. The database contains weekly data in which the weeks start on Thursday and end on Wednesday. We generate a set of dummy variables, one for each holiday. For Thursday holidays, the corresponding dummy variable is set to 1 for the two weeks prior to the holiday, but zero for the week including the holiday. For holidays taking place on all other days, the dummy variable is set to 1 for the week before the holiday and the week including the holiday. The Lent dummy variable takes the value one for the four weeks prior to the 2-week Easter shopping period (as e.g. tuna demand may increase during, and drop after Lent), the post-Thanksgiving variable has the value one for the week following Thanksgiving and the Christmas dummy is set to one for the week following Christmas to capture the shopping in anticipation of New Year's as in Chevalier et al. (2000). We incorporate a dummy variable corresponding to Halloween, since the demand for one of the categories we analyze -- front-end candies -- is likely to be much higher around this holiday. Since little candy is likely to be bought immediately after Halloween, we add an additional dummy variable that is set equal to 1 for the week following the holiday. For consistency, these eleven holiday dummy variables are incorporated in all categories analyzed.

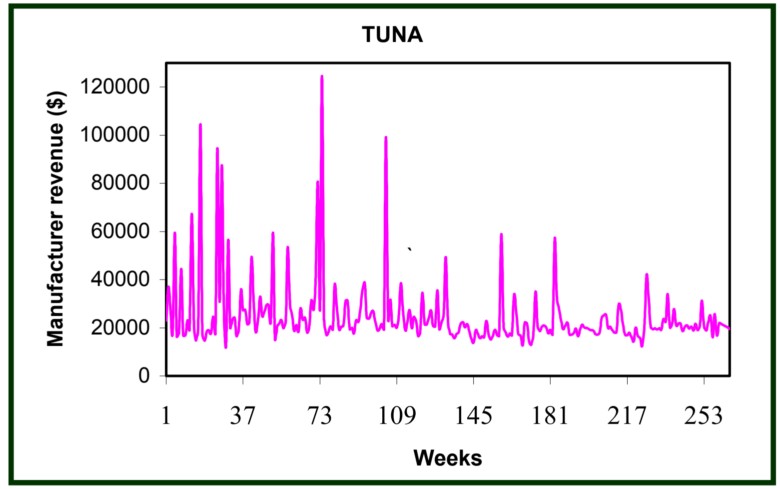
Figure A shows a plot of prices, manufacturer revenue, retailer revenue and retailer margin for a brand in the stationary canned-tuna market. Summary information on the average promotional frequency and depth in each of the categories in the data set is provided in Table A1. Table A2 shows the results of the unit-root tests. We excluded categories with structural breaks (due to new-product introductions) as identified by our unit-root tests, and thus confined our attention to 21 stationary categories in further analysis.

Figure A Plot of performance and price series for a leading brand

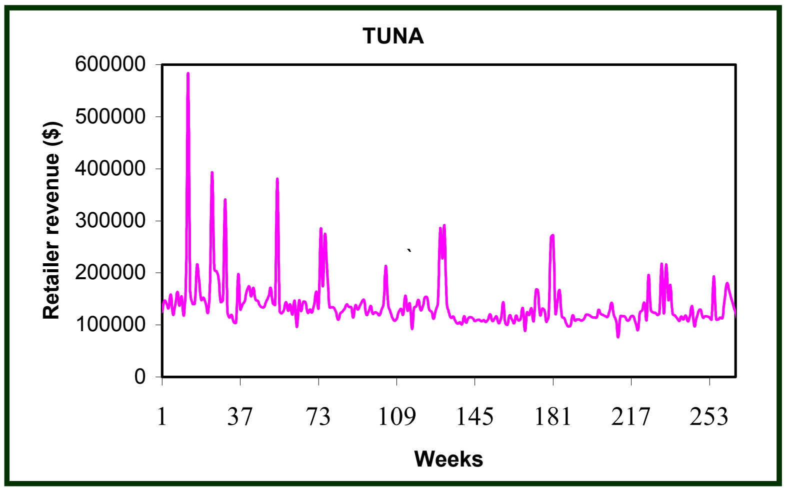
A: Plot of price series



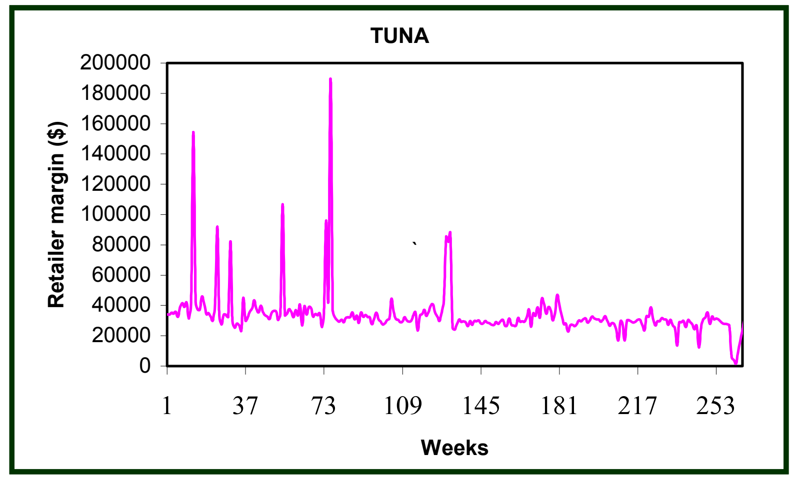
B: Plot of manufacturer revenue series



C: Plot of retailer revenue series



D: Plot of retailer margin series



A2 about here ---

Table A1 Average promotional frequency and promotional depth for the Dominick's data*

Category	Promotional frequency	Promotional depth
Analgesics	27%	9%
Beer	21%	7%
Bottled juice	23%	9%
Cereals	7%	10%
Cheese	32%	7%
Cookies	25%	10%
Crackers	36%	10%
Canned soup	19%	13%
Dish detergent	23%	8%
Frozen juice	30%	10%
Fabric softener	16%	8%
Front-end candies	13%	14%
Laundry detergent	26%	9%
Refrigerated juice	34%	13%
Soft drinks	38%	15%
Shampoo	27%	10%
Snack crackers	26%	10%
Soap	16%	21%
Toothpaste	26%	8%
Toilet tissue	36%	11%
Canned tuna	31%	10%
Average across categories	25%	11%

* Benchmark price = maximum price; average is taken across the top-three brands

Table A2 Unit-root tests[a]

	ADF unit root test		Evolving after exogenous break test[b]	Evolving after endogenous break test[c]
	Stationary	Evolving		
<u>Manufacturer Performance</u>				
Brand sales	71	3	1	-
Manufacturer revenue	70	5	-	-
<u>Retailer Performance</u>				
Category sales	22	3	-	-
Retailer revenue	22	3	-	-
Retailer margins	22	3	-	-
Store revenue	1	0	-	-
Store traffic	1	0	-	-
<u>Price Series</u>				
Retail price	64	11	-	-
Wholesale price	63	9	3	-

a- All the series in the table are stationary at the 5% levels.

b- Perron break test (1989, 1990).

c- Zivot and Andrews break test (1992).

A. 2 Validation

In this section, we present the results of various validation analyses, as summarized in Table B.

Table B Validation Comparison of median elasticities derived from alternative VAR specifications

	Focal model		**Extended model**	
	Median immediate elasticity	**Median total (cumulative) elasticity**	**Median immediate elasticity**	**Median total (cumulative) elasticity**
A1. Sensitivity to treatment of feature and display -- Added dynamics for exogenous feature and display				
<u>Manufacturer performance</u>				
Brand sales	3.20	3.72	3.14	3.62
Manufacturer revenue	2.35	2.30	2.26	2.18
<u>Retailer performance</u>				
Category sales	0.36	0.50	0.38	0.52
Retailer revenue	0.09	-0.05	0.08	-0.07
Retailer margin	-0.23	-0.70	-0.20	-0.76
Store revenue	0.00	0.01	0.00	0.00
Store traffic	0.01	0.00	0.01	0.00
A2. Sensitivity to treatment of feature and display -- Feature and display of promoted brand are treated as endogenous				
<u>Manufacturer performance</u>				
Brand sales	3.20	3.72	3.18	3.67
Manufacturer revenue	2.35	2.30	2.31	2.39
<u>Retailer performance</u>				
Category sales	0.36	0.50	0.39	0.47
Retailer revenue	0.09	-0.05	0.11	-0.09
Retailer margin	-0.23	-0.70	-0.25	-0.76
Store revenue	0.00	0.01	0.00	0.01
Store traffic	0.01	0.00	0.01	0.00
B. Sensitivity to aggregation across stores -- Log-log model				
<u>Manufacturer performance</u>				
Brand sales	3.20	3.72	3.02	3.42
Manufacturer revenue	2.35	2.30	2.19	2.15
<u>Retailer performance</u>				
Category sales	0.36	0.50	0.37	0.47
Retailer revenue	0.09	-0.05	0.11	-0.03
Retailer margin	-0.23	-0.70	-0.26	-0.80
Store revenue	0.00	0.01	0.01	0.01
Store traffic	0.01	0.00	0.01	0.01
C1. Sensitivity to inclusion of multiple performance measures -- Simultaneous inclusion of multiple sales or revenue variables				
<u>Manufacturer performance</u>				
Brand sales	3.20	3.72	3.11	3.94

Manufacturer revenue	2.35	2.30	2.32	2.26
C2. Sensitivity to inclusion of multiple performance measures -- Separate models on brand sales and wholesale price				
Manufacturer performance Manufacturer revenue	2.35	2.30	2.30	2.17
C3. Sensitivity to inclusion of multiple performance measures -- Store traffic as additional endogenous variable				
Manufacturer performance Brand sales Manufacturer revenue	 3.20 2.35	 3.72 2.30	 3.11 2.36	 3.75 2.40
Retailer performance Category sales Retailer revenue Retailer margin Store revenue	 0.36 0.09 -0.23 0.00	 0.50 -0.05 -0.70 0.01	 0.39 0.07 -0.34 0.00	 0.54 -0.07 -0.79 0.01

Table B Validation Comparison of median elasticities derived from alternative VAR specifications

	Focal model		Extended model	
	Median immediate elasticity	**Median total (cumulative) elasticity**	**Median immediate elasticity**	**Median total (cumulative) elasticity**
C4. Sensitivity to inclusion of multiple performance measures -- Simultaneous inclusion of both manufacturer and retailer revenue				
Manufacturer revenue - retailer revenue (significant instances)	--	--	97%	98%
Manufacturer revenue - retailer revenue (mean estimate)	2.36	1.96	2.39	1.99
D1. Sensitivity to the nature of the price-setting mechanism -- WP of the promoted brand as an additional endogenous variables				
Manufacturer performance Brand sales Manufacturer revenue	 3.20 2.35	 3.72 2.30	 3.28 2.34	 3.74 2.32
Retailer performance Category sales Retailer revenue Retailer margin Store revenue Store traffic	 0.36 0.09 -0.23 0.00 0.01	 0.50 -0.05 -0.70 0.01 0.00	 0.38 0.10 -0.28 0.00 0.00	 0.56 -0.07 -0.75 0.01 0.01
D2. Sensitivity to the nature of the price-setting mechanism -- Weighted price for other categories as additional endogenous variable in store-traffic model				
Retailer performance Store traffic	 0.01	 0.00	 0.01	 0.00
E. Sensitivity to potential cross-category influence -- Price in complementary/substitute category as additional endogenous variable*				
Manufacturer performance				

Brand sales	3.12	3.64	3.19	3.68
Manufacturer revenue	2.22	2.17	2.24	2.19
Retailer performance				
Category sales	0.50	0.54	0.51	0.63
Retailer revenue	0.11	-0.05	0.11	-0.08
Retailer margin	-0.21	-0.65	-0.28	-0.69
Store revenue	0.01	0.01	0.01	0.01
Store traffic	0.00	0.00	0.00	0.00
F. Sensitivity to the adopted wholesale price operationalization -- Alternative wholesale price measure**				
Manufacturer revenue	2.10	2.01	2.12	1.91
Retailer margin	-0.10	-0.18	-0.17	-0.23
G. Sensitivity to over-parameterization (Different lags per equation)				
Manufacturer performance				
Brand sales	3.20	3.72	3.18	3.79
Manufacturer revenue	2.35	2.30	2.33	2.30
Retailer performance				
Category sales	0.36	0.50	0.37	0.59
Retailer revenue	0.09	-0.05	0.10	-0.08
Retailer margin	-0.23	-0.70	-0.26	-0.80
Store revenue	0.00	0.01	0.00	0.01
Store traffic	0.01	0.00	0.00	0.00
H. Sensitivity to over-parameterization (Retaining only significant |t >1| estimates				
Manufacturer performance				
Brand sales	3.20	3.72	3.15	3.81
Manufacturer revenue	2.35	2.30	2.43	2.29
Retailer performance				
Category sales	0.36	0.50	0.39	0.59
Retailer revenue	0.09	-0.05	0.11	-0.04
Retailer margin	-0.23	-0.70	-0.28	-0.82
Store revenue	0.00	0.01	0.00	0.00
Store traffic	0.01	0.00	0.01	0.00

* Computed on the following subset of categories: fabric softener and laundry detergents, crackers, snack crackers and cheese, bottled, refrigerated and frozen juice, shampoo and soap.

** Computed on the toothpaste category.

A.2.1 Sensitivity to the treatment of feature and display

The exogenous treatment of feature and display is relaxed in two ways. First, we extend our model by allowing for two lags of feature and display variables.[1] For all performance measures, the elasticity estimates closely match those obtained from the focal model (panel A1). Second, we treat the feature and display activity of the brand experiencing a price promotion as endogenous. This results in a model with six endogenous and four exogenous variables (the feature and display activity of the two competing brands). Again, results very similar to our focal model are obtained (panel A2).

A.2.2 Sensitivity to aggregation across stores

Our aggregation across stores with heterogeneous marketing-mix activities may have caused some biases (see Christen et al. 1997 for an in-depth discussion). Previous work has asserted that DFF aligns its promotions across stores (see e.g. Hoch et al. 1995); still, there may have been some deviations in terms of individual-store compliance. The linear (focal) model has been shown to be least sensitive to the store aggregation issue (Christen et al. 1997). In contrast, the log-log model is more susceptible to store aggregation bias, provided aggregation issues are a concern. If the results are very similar for the two specifications, this implies that the store aggregation bias is not a serious issue. Panel B shows that the log-log model indeed yields results that are very similar to those obtained with the linear model.

A.2.3 Sensitivity to inclusion of multiple performance measures

Simultaneous incorporation of all performance measures of both retailer and manufacturers into one giant VAR model would put too much strain on an already heavily parameterized model (see also Pesaran and Smith 1998). To examine whether incorporation of only one performance variable affects our substantive insights, four additional analyses are implemented. First, we incorporate simultaneously all three brand sales (manufacturer revenue) variables, along with their corresponding price variables, in a six-equation model. As shown in Panel C1, very similar elasticity estimates are again obtained for both performance indicators. Second, the financial metric of manufacturer revenue that is a focus of our research is a composite of unit sales and prices. Therefore we conduct validation analyses by running separate models on unit sales and wholesale prices, and comparing the computed additional manufacturer revenues (evaluated at the series' sample mean) with the ones obtained through the focal model's IRFs. These new results (panel C2) confirm our findings on manufacturer revenue. Third, we capture retailer pricing considerations beyond the focal category by including the store-traffic variable as an additional endogenous variable in all model specifications, resulting in a set of five-equation VAR models. Once more, comparable elasticity estimates are obtained (panel C3). Fourth, we verify whether the impulse responses for manufacturers and retailers are indeed statistically distinct by including manufacturer and retailer revenue, along with the three price variables, in the same VAR model. Using the aforementioned bootstrap procedure, we computed 250 times an estimate of the difference between the manufacturer and retailer elasticities. A subsequent test through the sample standard errors obtained from the empirical distribution on these elasticity differences revealed that in 97% (98%) of the cases, the immediate (cumulative) manufacturer revenue elasticity is significantly different from the immediate (cumulative) retailer revenue elasticity (Panel C4), indicating that our statistical inference is robust to this issue.

A.2.4 Sensitivity to the nature of the price-setting mechanism

The focal four-equation VAR model contains three price equations (see Equation 1), which capture previous research's arguments that retail prices are based on (1) the current prices of all (major) brands in the category (e.g. Chintagunta 2002), (2) past prices and performance and, (3) demand seasonality and special shopping periods. Moreover, retail price setting considerations may also include (1) the wholesale price or acquisition cost (e.g. Kim and Staelin 1996), (2) store-traffic implications (Drèze 1995), and (3) cross-category effects, if any. As for the store traffic, we already demonstrated the robustness of the model with respect to this variable in section *A.2.3*. In addition, we estimated an augmented (5-equation) model, in which we add the wholesale price of the shocked brand (panel D1); the summary statistics are very similar to the ones derived from our focal model. Next, the store-traffic model is augmented with a weighted price variable for the other categories, resulting in a five-equation model with the store-traffic variable, three price series from category i and a weighted price for the j other categories ($i \neq j$) as endogenous variables (panel D2). No substantial differences in the summary statistics for store traffic are observed.

[1]Comparable results were obtained for one and three lags.

A.2.5 Sensitivity to potential cross-category influences

Previous research has also suggested that price changes in one category typically do not affect demand in other categories, *unless* these categories are obvious demand complements/substitutes. Our dataset contains a number of such obvious candidates: fabric softener and laundry detergents, crackers, snack crackers and cheese, bottled, refrigerated and frozen juice, shampoo and soap. For those category pairs, we estimate an augmented model, in which the market-share weighted price of the top 3 brands in category j is added as an endogenous variable to the focal model in category i (rotating i and j, 16 such analyses are performed, resulting in the additional estimation of 176 VAR models). Similar summary statistics on these analyses are presented in Table B (panel E). Once again, no substantial differences in results are observed.

A.2.6 Sensitivity to the adopted wholesale price operationalization

Two of our performance measures, manufacturer revenue and retailer margin, depend on the adopted wholesale price (WP) definition. The retailer herself uses this measure as the relevant acquisition cost to compute profit margins, and has been applied in previous literature (see e.g. Besanko et al. 2001; Chevalier et al. 2001). However, the operationalization has been criticized on two grounds: (i) it may be subject to sluggish adjustment, as the older, higher-priced inventory needs to be sold off first, and (ii) it may be affected by forward-buying practices on the part of the retailer. Therefore, we obtained an additional dataset that features the base manufacturer wholesale price to the retailer and the starting and ending date of manufacturer promotions to retailers for a period from 1991 to 1994. This data allows us to compute an alternative wholesale price measure for a category also present in the Dominick's data: toothpaste. Because it records the period in which manufacturer trade deals are offered, this measure is not affected by retailer inventory management and thus not subject to sluggish adjustment nor forward buying. The alternative wholesale price measure is, however, subject to some other issues: it only features a subset of SKUs per brand and a subset of the Dominick's product categories, and it only informs us whether a manufacturer offered a deal, not whether the retailer accepted it. Despite these differences between the two wholesale price measures, they lead to similar time series of manufacturer revenue and retailer margin and result in comparable IRFs and elasticity estimates (Table B, Panel F).

A.2.7 Sensitivity to VAR parameterization

Finally, we assess the sensitivity to potential over-parameterization of the VAR. An examination of the frequency counts of the number of lags per type of model (Table C1) shows that the maximum number of lags is three.

Table C1 Summary of number of lags in the VAR model

Performance measure	1 lag	2 lags	3 lags
Manufacturer Performance			
Brand sales	92%	6%	2%
Manufacturer revenue	92%	6%	2%
Retailer performance			
Category sales	94%	6%	0%
Retailer revenue	92%	8%	0%
Retailer margins	94%	6%	0%
Store revenue	94%	6%	0%
Store traffic	94%	6%	0%

And, the majority of the estimated models (over 90%) has one lag, in line with earlier VAR-based promotional studies (see e.g. Dekimpe and Hanssens 1999), suggesting that over-parameterization is not a major concern. Nevertheless, we assess the sensitivity of our findings to the VAR parameterization by implementing validation runs on the number of lags in the model. These focus on three issues. First, instead of selecting the same number of lags for the model, we let the maximum number of lags vary with each endogenous variable in each equation following the procedure advocated by Hsiao (1981, 1982). A comparison of the elasticity estimates obtained using this approach shows that the results are very similar to those from our focal model [Panel G]. Second, we start with a complete VAR system and constrain to zero those parameters with |t-value| less than 1, as advocated by Pesaran, Pierse and Lee (1993) and used in marketing by Dekimpe and Hanssens (1999) among others. Once again, the results are similar to our focal model results [please see Panel H].

Third, to assess the sensitivity to model parameterization, we compare the mean square forecast errors (MSFE) of our model with those of simpler model specifications. Specifically, we obtained a sequence of one-step ahead (static) forecasts, using actual rather than forecasted values for lagged dependent variables, as is customary for static forecasts. We obtain out-of-sample forecasts for 8 weeks of data corresponding to the period October 1994-November 1994, given that our estimation period is from September 1989 to September 1994. To assess the sensitivity of the forecasting performance to model parameterization, we compare the mean square forecast errors (MSFE) of our model with those of simpler model specifications. Specifically:

a. Instead of imposing a common lag structure on every endogenous variable, we let the maximum number of lags vary with each endogenous variable in each equation using the procedure indicated in point 1 above, re-estimate the VAR, and derive the MSFE;
b. we start with a complete VAR system and constrain to zero those parameters with |t-value| less than 1;
c. we start with a complete VAR system and constrain to zero some off-diagonal parameters. Specifically, we first constrain carry-over effects to zero (i.e. we set $\beta^i_{12} = \beta^i_{13} = \beta^i_{14} = 0$). Second, we constrain demand-feedback effects to zero (we set $\beta^i_{21} = \beta^i_{31} = \beta^i_{41} = 0$). Third, we constrain competitive reaction effects to zero (we set $\beta^i_{23} = \beta^i_{24} = \beta^i_{32} = \beta^i_{34} = \beta^i_{42} = \beta^i_{43} = 0$).
d. Finally, we constrain the seasonal dummy variables to be common across brands.

Table C2 Forecasting performance: MSFE of focal model relative to restricted model (median values)

	Restricted model specification					
	Model a: Different lags per equation	**Model b: Estimates with t-value less than 1 set to zero**	**Model c: Off-diagonal parameters set to zero**			**Model d: Seasonals common across brands**
			No Carry-over effects	**No Demand Feedback**	**No Competitive Reaction**	
<u>Manufacturer Performance</u>						
Brand sales	0.99	0.99	0.90	0.94	1.00	1.04
Manufacturer revenue	0.98	0.99	0.91	0.95	1.01	0.98
<u>Retailer performance</u>						
Category sales	1.01	1.01	0.93	0.97	0.96	1.02
Retailer revenue	0.99	1.05	0.95	0.98	0.96	1.05
Retailer margins	1.03	0.98	0.94	0.99	0.98	1.09
Store revenue	1.03	0.98	0.99	0.94	0.97	1.13
Store traffic	1.04	0.97	0.97	0.93	0.96	1.15

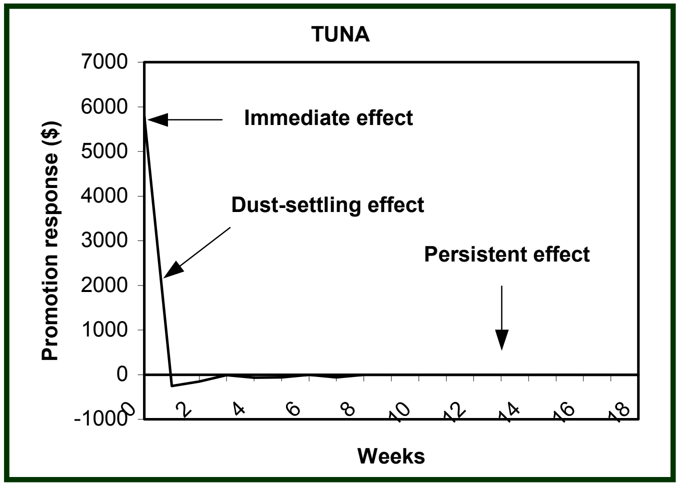
In all instances, the MSFE of the focal model relative to the restricted model, based on out-of-sample forecasts for eight periods, shows that the relative performance of the restricted models is not much improved (versus the focal model), as shown in Table C2. In sum, through the adopted VAR specification, we achieved a balance between efficiency, flexibility and parsimony.

Figure B Impulse Response Functions

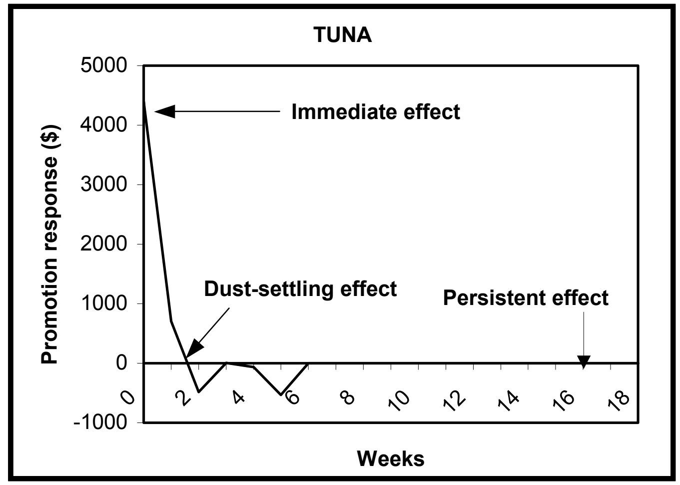
Figure B shows an example of the incremental effect over time of a price promotion of one cent per ounce for one of the leading brands in the tuna market on the manufacturer's (Panel A) and the retailer's (Panel B) revenues. Both parties experience a significant and immediate revenue increase in the promotional period, and a post-promotional dip around period 2. However, given the level/trend stationarity of the performance series, neither player experiences an enduring revenue gain (i.e. the incremental revenue impact converges to zero). Furthermore, both the immediate effect (\$5,790 versus \$4,400) and the cumulative impact (\$5,180 versus \$4,030) prior to convergence are more pronounced for the manufacturer than for the retailer. This is also the case in Panel C and Panel D, which trace the over-time impact of a one-cent price promotion in the stationary cheese market, where only the manufacturer (Panel C) enjoys an immediate revenue increase (\$8,200), while both the immediate and cumulative effects (-\$10,430 and –\$18,010, respectively) for the retailer are negative (Panel D). Hence, in the former case, the retailer's and the manufacturer's financial interests are aligned, while this is clearly not the case in the latter example.

Fig. B Impulse-Response functions

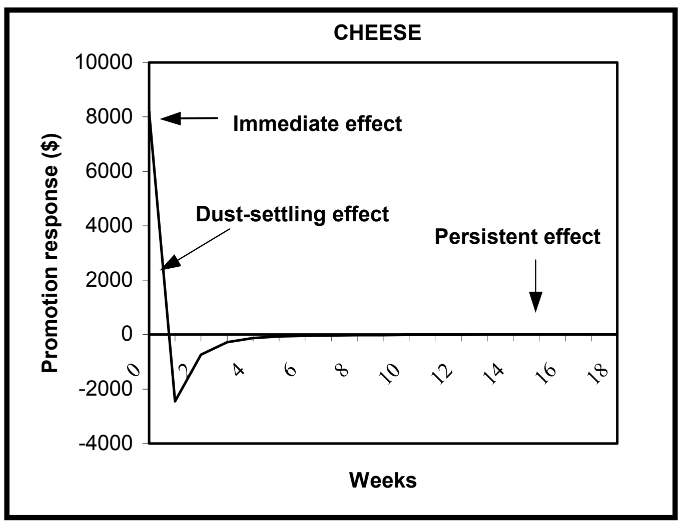
A: Impulse response function of a price promotion of one cent per ounce on manufacturer revenue



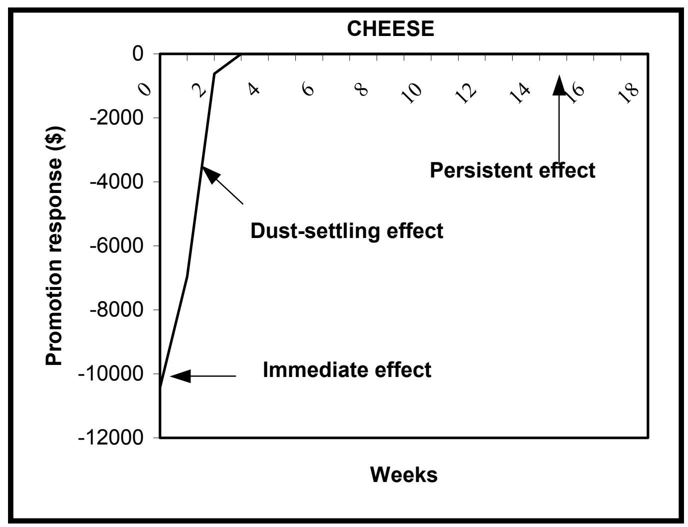
B: Impulse response function of a price promotion of one cent per ounce on retailer revenue



C: Impulse response function of a price promotion of one cent per ounce on manufacturer revenue



D: Impulse response function of a price promotion of one cent per ounce on retailer revenue



Additional references

Hsiao, Cheng (1981), "Autoregressive modeling and money-income causality detection," *Journal of Monetary Economics*, 7, 85-106.

Hsiao, Cheng (1982), *Time series and causal ordering of Canadian Money, Income and Interest Rates*, Elsevier Science Publishers B.V., North-Holland, Amsterdam.

Pesaran M. Hashem, Pierse Richard and K.C. Lee (1993), "Persistence, cointegration and aggregation: A disaggregated analysis of output fluctuations in the US Economy", *Journal of Econometrics*, 1993, (56), 57-88