

## Chapter 2

# Sales Promotion Models

Harald J. van Heerde and Scott A. Neslin

Firms spend a significant part of their marketing budgets on sales promotions. Retail (2012) indicates that during 1997–2011, promotion accounted for roughly 75% of marketing expenditures for US packaged goods manufacturers; the other 25% was for advertising. In 2011, 58% of the budget was spent on promotion to the trade (i.e., from manufacturers to retailers), and 15% on manufacturer promotions to consumers. Since the impact of promotions on sales is usually immediate and strong (Blattberg et al. 1995), promotions are attractive to results-oriented managers seeking to increase sales in the short term (Neslin 2002). In a meta-analysis, Bijmolt et al. (2005) report that the average short-term sales promotion elasticity is  $-3.63$ , which implies that a 20% temporary price cut leads to a 73% rise in sales.<sup>1</sup> There are few, if any, other marketing instruments that are equally effective. Because of this, coupled with the availability of scanner data, marketing researchers have been very active in developing models for analyzing sales promotions. Most applications analyze promotions for consumer packaged goods, and this chapter reflects this practice. Nevertheless, many of the models could be applied to other settings as well.

This chapter discusses models for measuring sales promotion effects. Part I (Sects. 2.1–2.10) focuses on descriptive models, i.e., models that describe and explain sales promotion phenomena. We start by discussing promotions to consumers. Sections 2.1 through 2.5 focus on analyzing the direct impact of promotions on sales and decomposing that impact into a variety of sources. Section 2.6

---

<sup>1</sup>This figure holds for temporary price cuts without feature or display support. A feature or display may increase the sales effect up to a factor 9 (Narasimhan et al. 1996).

---

H.J. van Heerde (✉)  
Massey University, Auckland 0630, New Zealand  
e-mail: heerde@massey.ac.nz

S.A. Neslin  
Tuck School of Business, Dartmouth College, Hanover, USA

examines what happens after the sales promotion bump and describes models for measuring feedback effects, reference price effects, learning effects, permanent effects, and competitive reactions. Section 2.7 discusses endogeneity in sales promotion effect measurement and solutions to address it. Next we turn to descriptive models for promotions aimed at retailers (“trade promotions”) in Sect. 2.8, and discuss key issues and models concerning forward buying (Sect. 2.9) and pass-through (Sect. 2.10).

In Part II we discuss normative models, i.e. models that prescribe the best (profit maximizing) sales promotions. Section 2.11 covers models for planning promotions to consumers, Sect. 2.12 provides decision models on trade promotions for manufacturers, and Sect. 2.13 describes normative retailer models for optimal forward buying and pass-through. Part III concludes with a summary (Sect. 2.14), practical model guidelines (Sect. 2.15), and directions for future research (Sect. 2.16).

Part I: Descriptive Models

2.1 Promotions to the Consumer—Introduction

Promotions to the consumer include coupons, rebates, in-store temporary price cuts, feature advertising, and in-store displays. In analyzing promotion effects, we distinguish between *immediate effects* (the impact in the week  $t$  the promotion is implemented), *medium-term effects* (the weeks surrounding week  $t$ ) and *long-term effects* (that take place after the medium-term effects). Much of the work on sales promotion has focused on the sales promotion bump. The goal in modeling this bump is to allocate the increase in sales that occurs in period  $t$  to one or more of the sources listed in the second column of Table 2.1. The timing of the sources contributing the bump may be in week  $t$  itself (immediate effect) or in the surrounding weeks (medium-term effect). We discuss the immediate and medium-term effects in

Table 2.1 The impact of promotions to the consumer

The sales promotion bump (Sects. 2.1–2.5)		Long-term effects—beyond the sales promotion bump (Sect. 2.6)
Category growth	Increased consumption rate	Purchase-event feedback
	Outside industries	Reference prices
Within-category immediate effects	Cannibalization	Consumer learning
	Brand switching	Long-term effects
	Category switching	Competitive response
	Store switching	
Within-category medium-term effects	Acceleration	
	Deceleration	

Sects. 2.1–2.5. The effects of sales promotions on consumer behavior beyond the sales promotion bump are classified as long-term effects. These effects are list in Table 2.1 as well and discussed in more detail in Sect. 2.6.

We now discuss the decomposition of the sales promotion bump. We assume that the analysis is conducted at the level of an SKU (Stock Keeping Unit), i.e., a particular variety of a brand. The sales promotion sources contributing to the sales promotion bump for an SKU can be classified in three areas. The first is “Category Growth,” which means the increase in sales for the promoted SKU does not come at the expense of other products or stores within the category. Promotion-induced purchases may cause households to carry extra inventory, which is consumed at a higher rate. For example, a promotion on potato chips may increase household inventory and cause the household to consume more chips.

The second area is within-category immediate effects, i.e., the purchase draws from within the category in the same time period. This effect consists of:

- Cannibalization: the consumer switches from SKU  $j'$  of brand  $b$  to SKU  $j$  of brand  $b$ ;
- Brand switching: the consumer switches from brand  $b'$  to brand  $b$ ;
- Category switching: the consumer switches from a product from another category to brand  $b$ ;
- Store switching: the consumer switches from another store  $s'$  to store  $s$  to buy brand  $b$ ;

The third area consists of within-category medium-term effects, i.e., substitution from the period before the promotion (deceleration) or after (acceleration). Deceleration means the consumer postpones the category purchase to week  $t$ , because the consumer expects a promotion in week  $t$ . Deceleration implies purchase displacement from the past (before  $t$ ) to now. Acceleration refers to timing acceleration—the consumer buys the category in week  $t$  rather than later, or quantity acceleration—the consumer buys more of the category than usual. Both timing and quantity acceleration lead to higher household inventories, and both lead to a purchase displacement from the future ( $>t$ ) to now ( $t$ ). Note in defining acceleration we assume that the *postpromotion consumption rate does not increase*. If it does, we are in the case of Category Growth, and there is no purchase displacement.

Table 2.2 lists all possible combinations of effects derivable from Table 2.1. We first list Category Growth. Then, within-category substitution can come from four types of products (the item itself in another time period, other items within the same brand, other brands, and other categories), from two places (within the same store, from other stores) and from three time periods (before, during, and after the promotion). Hence in total there are  $4 \times 2 \times 3 = 24$  combinations, of which one drops out since a promoted product cannot substitute its own sales in the same period and store. These 23 combinations all imply some form of substitution, as we show in Table 2.2 (listings 2–24). For example, a brand- and store switch (source 7) implies that brand  $b'$  in store  $s'$  loses sales. The combination of store switching and acceleration (source 17) is sometimes referred to as indirect store switching

**Table 2.2** The 24 decomposition effects from manufacturer and retailer perspectives

#	Effect	Promotion of focal SKU of a brand in focal store in focal time period draws sales from...	Increased category consumption	Net unit sales effect for manufacturer	Net unit sales effect for retailer
<i>Category growth</i>					
1	Faster consumption	Household budget	Yes	+	+
<i>Within-category immediate</i>					
2	Cannibalization	Other SKUs from the same brand, in same store and same period	No	0	0
3	Brand switching	Other brands in same store in same period	No	+	0
4	Category switching	Other categories in same store in same period	Yes	+	0
5	Store switching	Same SKU in other stores in same period	No	0	+
6	Cannibalization and store switching	Other SKUs from the same brand, in other stores in same period	No	0	+
7	Brand switching and store switching	Other brands in other stores in same period	No	+	+
8	Category switching and store switching	Other categories in other stores in same period	Yes	+ or 0 <sup>a</sup>	+
<i>Within-category medium term</i>					
9	Acceleration	Same SKU in same store in future periods	No	0	0
10	Deceleration	Same SKU in same store in earlier periods	No	0	0
11	Cannibalization and acceleration	Other SKUs from the same brand, in same store and future periods	No	0	0
12	Cannibalization and deceleration	Other SKUs from the same brand in same store in earlier periods	No	0	0
13	Brand switching and acceleration	Other brands in same store in future periods	No	+	0
14	Brand switching and deceleration	Other brands in same store and earlier periods	No	+	0
15	Category switching and acceleration	Other categories in same store in future periods	Yes	+ or 0 <sup>a</sup>	0

(continued)

**Table 2.2** (continued)

#	Effect	Promotion of focal SKU of a brand in focal store in focal time period draws sales from...	Increased category consumption	Net unit sales effect for manufacturer	Net unit sales effect for retailer
16	Category switching and deceleration	Other categories in same store and earlier periods	Yes	+ or 0 <sup>a</sup>	0
17	Store switching and acceleration	Same SKU in other stores in future periods	No	0	+
18	Store switching and deceleration	Same SKU in other stores in earlier periods	No	0	+
19	Cannibalization, store switching and acceleration	Other SKUs from same brand in other stores in future periods	No	0	+
20	Cannibalization, store switching and deceleration	Other SKUs from same brand in other stores in earlier periods	No	0	+
21	Brand switching, store switching and acceleration	Other brands in other stores in future periods	No	+	+
22	Brand switching, store switching and deceleration	Other brands in other stores in earlier periods	No	+	+
23	Category switching, store switching and acceleration	Other categories in other stores in future periods	Yes	+ or 0 <sup>a</sup>	+
24	Category switching, store switching and deceleration	Other categories in other stores in earlier periods	Yes	+ or 0 <sup>a</sup>	+

<sup>a</sup>If the manufacturer produces the product in the other category that is being substituted, s/he does not benefit from the category switch (0); otherwise s/he does (+)

(Bucklin and Lattin 1992): the consumer visits both stores, but, because of the promotion in store  $s$ , she buys the promoted product in store  $s$  whereas otherwise she would have bought a product in store  $s'$ . Hence a current purchase in store  $s$  pre-empt a future purchase in store  $s'$ .

Table 2.2 shows where each of the 24 combinations draws from and also whether category consumption increases. This is certainly the case for Category Growth. Category consumption also increases when the contributing source involves category switching (#4, 8, 15, 16, 23, 24).

The effects listed in Table 2.2 are important because each has distinctive managerial implication in terms of unit sales. For example, brand switching benefits manufacturers but not retailers; store switching benefits retailers but not manufacturers; category growth benefits both manufacturers and retailers; category switching is neutral for retailers and may or may not be neutral for manufacturers, depending on whether the manufacturer has products in both categories. The rightmost two columns of Table 2.2 show that some combinations create a “+” for both manufacturers and retailers. So there is potential for conflict as well as “win-win” between manufacturers and retailers (Van Heerde and Gupta 2006). Further complicating matters is that both retailer and manufacturer profit margins may differ. A brand switch might actually benefit the retailer if the switched-to brand has higher margin. This motivates the retailer to demand a trade deal discount. The retailer can “pass through” some of the discount to promoting the brand, while at the same time increasing the margin on that brand. We discuss this in Sect. 2.9.

In sum, it is crucial to measure how the immediate impact of promotion is decomposed into the components shown in Table 2.2. Consequently, the decomposition of the sales promotion bump has gained considerable attention in the literature, and we summarize empirical generalizations in Sect. 2.5. First, however, we discuss the models necessary for measuring the decomposition. We start with individual-level incidence, choice, and quantity models in Sect. 2.2. In Sect. 2.3 we discuss individual-level models for store switching, category switching, cannibalization, and deceleration. In Sect. 2.4 we present aggregate (store-level) models.

## 2.2 Customer-Level Models—Incidence, Choice, and Quantity

Promotions can influence category incidence, brand choice, and purchase quantity, and historically, these decisions have received the most attention. The unit of analysis in this literature is typically the brand, not the SKU. The models for these household-level decisions are based largely on household panel scanner data, as are the models discussed in Sect. 2.3. However, these models have also utilized conjoint analysis (e.g., Eckert et al. 2012; Ailawadi et al. 2014). The models in Sect. 2.4 are based on aggregate (store- or higher) data, e.g., weekly store data.

The probability that household  $h$  buys  $q_{bt}^h$  units of brand  $b$  at shopping trip  $t$  is the product of three probabilities:

$$P(Q_{bt}^h = q_{bt}^h) = P(I_t^h = 1) \times P(C_t^h = b \mid I_t^h = 1) \times P(Q_{bt}^h \mid I_t^h = 1, C_t^h = b) \quad (2.1)$$

where

- $P(I_t^h = 1)$  is the probability that household  $h$  buys the category at trip  $t$  (incidence),
- $P(C_t^h = b \mid I_t^h = 1)$  is the probability that, conditional on incidence at  $t$ , household  $h$  buys brand  $b$ , and
- $P(Q_{bt}^h = q_{bt}^h \mid I_t^h = 1, C_t^h = b)$  is the probability that, conditional on a choice to buy brand  $b$  at trip  $t$ , the household buys  $q_{bt}^h$  units

### 2.2.1 Category Incidence Models

Category incidence is often modeled as a binary logit (e.g., Bucklin et al. 1998):

$$P(I_t^h = 1) = \frac{1}{1 + e^{-(\gamma_0 + \gamma_1 CV_t^h + \gamma_2 I_{t-1}^h + \gamma_3 \bar{C}^h + \gamma_5 INV_t^h)}} \quad (2.2)$$

where  $CV_t^h$  is the “inclusive value”, which in a nested logit framework is the expected maximum utility available to household  $h$  from buying a brand in the category at time  $t$ . It is given by the log of the denominator of the brand choice probability:  $CV_t^h = \ln \left( \sum_{b=1}^B \exp(u_{b'} + \beta X_{b't}^h) \right)$  (see Sect. 2.2.2).  $I_{t-1}^h$  is a lagged purchase incidence dummy (Ailawadi and Neslin 1998), and  $\bar{C}^h$  is the household’s average daily consumption computed from an initialization sample.  $INV_t^h$  is the inventory carried by household  $h$  at the beginning of shopping trip  $t$ . The standard approach to operationalize  $INV_t^h$  is:

$$INV_t^h = INV_{t-1}^h + PurQty_{t-1}^h - \bar{C}^h,$$

where  $PurQty_{t-1}^h$  is the quantity (in ounces) purchased by household  $h$  during trip  $t-1$ . Inventory should be mean-centered over time for a given household to remove household differences. The term  $\bar{C}^h$  assumes a constant purchase rate for household  $h$ . However, Ailawadi and Neslin (1998) propose that the consumption rate for household  $h$  at time  $t$  ( $Consumpt_t^h$ ) flexibly varies over time as a function of inventory:

$$Consumpt_t^h = INV_t^h \left[ \frac{\bar{C}^h}{\bar{C}^h + (INV_t^h)^f} \right], \quad (2.3)$$

where  $f$  is a parameter. A flexible consumption rate (smaller  $f$ ) means that promotion-induced stockpiling can increase category consumption (Effect #1 in Table 2.2; see also Sun 2005).

### 2.2.2 Brand Choice Model

The probability that household  $h$  buys brand  $b$  at time  $t$ , conditional on purchasing the category, is often given by a multinomial logit model<sup>2</sup> (Guadagni and Little 1983):

$$P(C_t^h = b | I_t^h = 1) = \frac{\exp(u_b + \beta X_{bt}^h)}{\sum_{b'=1}^B \exp(u_{b'} + \beta X_{b't}^h)}, \quad (2.4)$$

where  $B$  is the number of brands and  $V_{bt}^h$  is the “deterministic component” of the utility of household  $h$  for brand  $b$  at time  $t$  (Guadagni and Little 1983). A typical formulation would be:

$$V_{bt}^h = u_b + \beta X_{bt}^h = u_b + \beta_1 PRICE_{bt} + \beta_2 FEAT_{bt} + \beta_3 DISP_{bt} + \beta_4 BL_b^h + \beta_5 Last_{bt}^h, \quad (2.5)$$

where  $u_b$  is a brand-specific intercept,  $X_{bt}^h$  is a vector of marketing and household-specific covariates; and  $\beta$  is a vector of response coefficients. The components of  $X_{bt}^h$  might include  $PRICE_{bt}$ , the net price of brand  $b$  at time  $t$ ,  $FEAT_{bt}$  and  $DISP_{bt}$  as feature and display indicators for brand  $b$ , and  $BL_b^h$  is the intrinsic loyalty or preference for brand  $b$ , calculated as the within-household  $h$  market share of brand  $b$  in an initialization period and assumed constant over time (Bucklin et al. 1998). The BL term can be eliminated if differences in customer preference ( $\mu_b$ ) are modeled as unobserved heterogeneity (see Sect. 2.2.5). The term  $Last_{bt}^h$  is a dummy that is 1 in case brand  $b$  was bought last time by household  $h$ , and zero else. It captures purchase-event feedback or “state dependence” (see Sect. 2.6.1).

### 2.2.3 Purchase Quantity Model

Given purchase incidence and choice of brand  $b$ , the probability that household  $h$  buys  $q_{bt}^h = 1, 2, \dots, n$  units at time  $t$  is captured by a Poisson model with a truncation at the zero outcome (Bucklin et al. 1998). This can be written as:

$$P(Q_{bt}^h = q_{bt}^h | I_t^h = 1, C_t^h = b) = \frac{\exp(-\lambda_{bt}^h) (\lambda_{bt}^h)^{q_{bt}^h}}{[1 - \exp(-\lambda_{bt}^h)] q_{bt}^h!}, \quad (2.6)$$

---

<sup>2</sup>Multinomial probit is an alternative to the logit (e.g., Jedidi et al. 1999). The advantage of the probit model is that it avoids the independence of irrelevant alternatives (IIA) assumption of logit models (see Guadagni and Little 1983). However, it does not produce a closed form for the probability of consumer choice.



where  $\lambda_{bt}^h$  is the purchase rate of household  $h$  for brand  $b$  at time  $t$ . This parameter is a function of (mean-centered) inventory, the average number of units purchased by the household, and the size, price, and promotion status of the selected brand:

$$\lambda_{bt}^h = \exp\left(\theta_0 + \theta_1(Inv_t^h - \overline{Inv}^h) + \theta_2\overline{Q}^h + \theta_3SIZE_b + \theta_4PRICE_{bt} + \theta_5FEAT_{bt} + \theta_6DISP_{bt}\right)$$

### 2.2.4 Estimation

The likelihood function for incidence, choice, and quantity is given by:

$$L = \prod_{h=1}^H \prod_{t=1}^T \prod_{b=1}^B \left( P(I_t^h=1)^{Y_t^h} (1 - P(I_t^h=1))^{1-Y_t^h} P(C_t^h=b | I_t^h=1)^{Z_{bt}^h} P(Q_{bt}^h=q_{bt}^h | I_t^h=1, C_t^h=b)^{Z_{bt}^h} \right)$$

where

$Y_t^h$  Category purchase indicator, equals 1 if household  $h$  purchased the category on shopping trip  $t$ ; 0 otherwise

$Z_{bt}^h$  Brand purchase indicator, equals 1 if household  $h$  purchased brand  $b$  on shopping trip  $t$ ; 0 otherwise

Methods used to estimate the model include maximum likelihood, simulated maximum likelihood (Train 2003), and Bayesian “MCMC” estimation (Rossi et al. 2005). Once the incidence, choice, and quantity models have been estimated, we can calculate the incidence elasticity,  $\eta_I = \frac{\partial P(I)}{\partial PRICE} \frac{PRICE}{P(I)}$ ; the choice elasticity,  $\eta_{C|I} = \frac{\partial P(C|I)}{\partial PRICE} \frac{PRICE}{P(C|I)}$ ; and the quantity elasticity,  $\eta_{Q|I,C} = \frac{\partial E(Q)}{\partial PRICE} \frac{PRICE}{E(Q)}$ , where  $E(Q)$  is the expected purchase quantity (see Gupta 1988 for details). Gupta (1988) decomposes the total sales elasticity  $\eta_S$  into that due to purchase incidence ( $\eta_I$ ), brand switching ( $\eta_{C|I}$ ) and purchase quantity ( $\eta_{Q|I,C}$ ), so that  $\eta_S = \eta_I + \eta_{C|I} + \eta_{Q|I,C}$ . For example, a sales elasticity of  $-3$  with respect to promotion might be decomposed as  $-3 = -0.45 - 2.25 - 0.3$ , i.e., the brand switching elasticity comprises 75% of the total elasticity, whereas the incidence elasticity is 15% and the quantity elasticity is 10%. We refer to Sect. 2.5 for a more in-depth discussion on how (and how not) to interpret this result.

### 2.2.5 Heterogeneity

Consumers are naturally heterogeneous in their brand preferences, responsiveness to marketing actions, and how much they learn from the product usage experience. As a result, the parameters of the choice, incidence, and quantity models (Eqs. 2.2–2.6)

should be modeled to vary across consumers. For example, Eq. (2.5) could be written as:

$$V_{bt}^h = u_b^h + \beta^h X_{bt}^h = u_b^h + \beta_1^h PRICE_{bt} + \beta_2^h FEAT_{bt} + \beta_3^h DISP_{bt} + \beta_4^h BL_b^h + \beta_5^h LAST_{bt}^h \quad (2.7)$$

The brand-specific intercept from Eq. (2.5) is now household-specific ( $u_b^h$ ) meaning that households can differ in their preferences for various brands. The response coefficients for variable  $k$  ( $\beta_k^h$ ) also differ across households. This means that households can differ in their sensitivity to price, feature and display and the other variables in the utility model.

Modeling heterogeneity adds much complexity to choice, incidence, and quantity models. It is worthwhile to make clear why modeling heterogeneity is important:

- *Spurious State Dependence*: In a homogeneous model, the state dependence parameter ( $\beta_5^h$ ) would be over-stated because it soaks up the variation due to heterogeneous preference as well as dynamically changing preference (see Keane 1997a, also Abramson et al. 2000).
- *Segmentation*: Marketing is about segmentation. By learning about heterogeneity, we make our models more useful because we can segment the market based on preference or response.
- *Avoid Independence of Irrelevant Alternatives (IIA)*: Logit models are open to the IIA criticism (see Guadagni and Little 1983). Modeling heterogeneity eliminates IIA problems at the aggregate level, although it still exists at the level of each household. Steenburgh (2008), Roederkerk et al. (2011), and Liu et al. (2015) develop choice models to overcome IIA.
- *Better Prediction*: Incorporating heterogeneity means that our models incorporate more information; hence they should be expected to predict better.

Researchers face a myriad of decisions in how to model heterogeneity:

- *Distribution of the Individual Parameters*: The distribution of individual-level parameters can be considered to be continuous (Chintagunta et al. 1991), discrete (Kamakura and Russell 1989), or finite mixture (e.g., Varki and Chintagunta 2004).
- *Parameters to be Considered Heterogeneous*: The parameters to model heterogeneously can include preference, most coefficients, or all coefficients.
- *Joint Distribution of the Parameters*: The distribution of the heterogeneous parameters can be considered to be uncorrelated, correlated, or no assumption made.
- *Incorporation of Observed Heterogeneity*: Heterogeneity can be thought of as “observed” versus “unobserved.” Observed heterogeneity means that the heterogeneity in any parameter can be captured by measurable variables such as demographics or preference ( $BL_b^h$ ). Observable heterogeneity is easy to

incorporate simply by including the observed variables in the model. The concern is that these measures do not capture all the heterogeneity, and so researchers often model unobserved heterogeneity in addition to observed heterogeneity.

- *Choice Set Heterogeneity*: Eq. (2.4) assumes that each household considers the same brands when making a choice. Researchers (e.g., Siddarth et al. 1995) have questioned this assumption and model heterogeneity in choice sets.
- *Estimation*: Maximum Likelihood (ML), Simulated Maximum Likelihood (SML), and Bayesian are possible ways to estimate the model.

The above choices give rise to  $3 \times 3 \times 3 \times 2 \times 2 \times 3 = 324$  possible ways to handle heterogeneity. No one route has emerged as most popular. Table 2.3 gives a summary of how a few papers have handled heterogeneity.

While no one single method has been shown to be superior, and sometimes the differences across various approaches are not crucial (e.g., Andrews et al. 2002a, b), the general conclusion is that it is crucial to include some form of unobserved heterogeneity in the model, at a minimum, in the brand-specific intercept. The reasons are (1) Heterogeneity improves fit and prediction (e.g., Chintagunta et al. 1991), (2) Heterogeneity changes the coefficients of other variables (e.g., Chintagunta et al. 1991), although Ailawadi et al. (1999) note that aggregate price elasticities may not change, (3) State dependence can be over-stated when preference heterogeneity is not included (e.g., Keane 1997a; see also Horsky et al. 2006).<sup>3</sup>

So far we have focused on parameter heterogeneity *across consumers*. Another form of heterogeneity is *over-time parameter variation*, even within the same consumer. For example, consumers may display cyclical purchase behavior, where periods of high consumption are alternated with periods of low consumption. Park and Gupta (2011) show this happens in the yogurt category. They use a hidden Markov Model (HMM) that distinguishes between high and low category purchase states. Transition between states is modelled probabilistically. Park and Gupta (2011) demonstrate that the model fits better than benchmark models and show that companies can benefit by targeting households that are in the high category purchase state.

### 2.2.6 Integrated Incidence, Choice, Quantity Models

Models have been developed that expressly integrate two or more consumer decisions. The integration takes place through correlations between error terms, the formulation of the utility function, or defining the set of decision options. For example, Krishnamurthi and Raj (1988) integrate brand choice and quantity decisions by correlating the error terms of the choice and quantity equations. Nested

---

<sup>3</sup>It is noteworthy that there is some evidence (Abramson et al. 2000; Chiang et al. 1999) that not including choice set heterogeneity significantly distorts parameters.

**Table 2.3** Ways to handle heterogeneity in household-level models (selected papers)

	Distribution of parameters	Which parameters heterogeneous	Joint distribution of parameters	Choice set heterogeneity	Including observed heterogeneity?	Estimation
Kamakura and Russel (1989)	Discrete	All	N/A	No	No	ML
Ailawadi et al. (2007b)	Continuous	Most	No stipulation	No	No	SML
Ansari et al. (2006)	Continuous	Most	Correlated	No	No	Bayesian
Gupta and Chintagunta (1994)	Discrete	All	N/A	No	Yes	ML
Seetharaman et al. (1999)	Continuous	All	Correlated	No	Yes	Bayesian
Chiang et al. (1999)	Continuous	All	Correlated	Yes	No	Bayesian

logit posits a utility function that integrates choice and incidence decisions (see Ben-Akiva and Lerman 1991). The incidence portion of the nested logit is what we describe in Sect. 2.2.1. Another set of models integrates choice and incidence by adding a “no-purchase” option, i.e., the consumer is assumed to choose among a set of alternatives,  $J - 1$  of which are brands; the  $J$ th is the no-purchase option. See Chib et al. (2004) and the papers to which they refer for examples. Chiang (1991) and Chintagunta (1993) develop integrated models of incidence, choice, and quantity. Bell and Hilber (2006) investigate the relationship between incidence and quantity. They find that consumers with greater storage constraints shop more often and purchase smaller quantities per visit. While modeling the decisions in an integrated way is attractive from econometric and behavioral perspectives, the overall decomposition results do not seem to differ much from separate models. For instance, for the coffee category Gupta (1988) uses non-integrated models whereas Chiang (1991) uses integrated models, but their results on the elasticity decomposition into brand switching, incidence and quantity are almost identical (see Table 2.5 in Sect. 2.5).

### 2.2.7 *Dynamic Structural Models*

Another approach to modeling consumer decisions is dynamic structural models. These models begin with the household utility function and include dynamic phenomena such as “forward-looking” consumer behavior, where consumers take into account future utility in making current-period decisions, and consumer learning of brand quality, often in the form of Bayesian learning (See Sect. 2.6.3). Erdem and Keane (1996) develop a dynamic structural model of brand choice that includes forward-looking and learning behavior. Gönül and Srinivasan (1996) develop a dynamic structural model of purchase incidence. Sun et al. (2003) develop a dynamic structural model of incidence and choice. Erdem et al. (2003), Sun (2005), and Chan et al. (2008) develop dynamic structural models of incidence, choice, and quantity. Using structural models versus nonstructural models seems to affect the elasticity decomposition. For instance, Sun et al. (2003) report a brand switching percentage of 56% for a dynamic structural model that accounts for forward-looking customers, whereas the percentage for the non-structural integrated model (nested logit) is 72%.

## 2.3 Customer-Level Models—Extensions

Researchers have developed models that extend the classical incidence-brand choice-purchase quantity set-up. The key extensions involve store switching (Sect. 2.3.1), cross-category effects (Sect. 2.3.2), SKU-level models (Sect. 2.3.3), and purchase deceleration (Sect. 2.3.4).

### 2.3.1 *Extension 1: Store Switching*

Bucklin and Lattin (1992) propose a model that can be used to capture store switching effects. Their store choice model is given by a multinomial logit model that includes store loyalty and store features as explanatory variables.

One challenge in measuring the effects of promotions on store choice is that consumers make store choice decisions based on a host of factors (e.g., location, produce quality, waiting lines) that have little to do with an individual brand's price and promotion. At the same time, price and promotion for brands and categories will affect store choice. Lourenço et al. (2015) offer a new approach to address the question of determining how individual promotions affect the store price image that consumers have.

Another challenge is that the standard logit model for store choice assumes that a consumer knows each store's prices and promotions, which seems unlikely. It seems more likely that there is indirect store switching: a promotional purchase in one store preempts a regular purchase in another store, as we discussed in Sect. 2.1. To tackle these complicating factors, increasingly sophisticated models of store choice are available (Bell and Lattin 1998; Bell et al. 1998; Rhee and Bell 2002; Bodapati and Srinivasan 2006; Singh et al. 2006; Guyt and Gijsbrechts 2014; Van Lin and Gijsbrechts 2014, 2015; Vroegrijk et al. 2013).

### 2.3.2 *Extension 2: Cross-Category Effects*

Cross-category effects means that a sales promotion for a brand in a category may either steal sales from brands in other categories (substitution effect) or it may enhance brand sales across categories (complementary effects). This is what retailers hope occurs (Lam et al. 2001). Ailawadi et al. (2007a) call these positive cross-category effects halo effects, and find empirical evidence for them in a major drugstore chain.

To capture cross-category effects, Manchanda et al. (1999) specify a model for shopping basket composition, i.e., what set of categories is bought on a specific shopping trip. They model this as a multivariate probit. Multivariate probits (or logits) differ from multinomial probits (or logits) in that more than one alternative can be chosen on the current purchase occasion. This is the case when we are modeling the *set* of categories purchased. Manchanda et al.'s model includes "complementarity" effects, whereas the price of one category influences sales in another category, and "coincidence" effects, where certain products are bought together.

Mehta and Ma (2012) propose a more elaborate model to capture multi-category brand choice decisions. While Manchanda, Ansari and Gupta only model purchase incidence decisions, Mehta and Ma (2012)'s model captures brand choice, incidence and quantity decisions. Their model allows for cross-category effects of

promotions both in the incidence and purchase quantity decisions. The model helps retailers in understanding how they should allocate promotional expenditures across brands within categories and coordinate timing of promotions of brands across categories.

### 2.3.3 Extension 3: SKU-Level Models

Most choice models are at the brand level. However, consumers buy specific SKUs. The advantage of an SKU-level model is specificity and allowing for cross-SKU elasticities, especially cross-SKU effects between SKUs for the same brand. E.g., does promoting Yoplait's 6-oz. vanilla yogurt take away from Yoplait's 6-oz. blueberry yogurt? The disadvantage of SKU-level models is there can be so many SKUs in a given category that the modeling becomes infeasible.

Fader and Hardie (1996) propose a parsimonious model for SKU choice that addresses this problem. Suppose there are  $N$  attributes and let  $L_n$  be the number of levels associated with the  $n$ th attribute. Define the set  $\{l_1, l_2, \dots, l_N\}$  as the unique set of attribute levels for brand  $b$ , SKU  $j$ . Fader and Hardie (1996) model the SKU-specific intercept in the logit model (Eq. 2.4) as:  $u_{bj} = \sum_{n=1}^N m_{bjn} \alpha_n$ , where  $m_{bjn}$  is an elementary row vector, the  $l_n$ th element of which equals 1, and  $\alpha_n$  is the vector of preferences over the  $L_n$  levels of attribute  $n$ . A similar approach was followed by Ho and Chong (2003) and Chintagunta and Dubé (2005). For the covariates  $X_{bjt}^h$  (Eq. 2.4) we may use SKU-level versions of the variables in the brand choice model.

### 2.3.4 Extension 4: Deceleration

Deceleration means that consumers anticipate promotions, and consequently they may postpone purchases until a promotion is offered. To capture deceleration, we need a model component for the effect of an expected future promotion on current purchase behavior. Van Heerde et al. (2000) and Macé and Neslin (2004) use actual future prices in a model of brand sales to capture deceleration in a model of brand sales (more about this in Sect. 2.4.2). For household data, Sun et al. (2003) present a structural model for the promotion expectation process. They assume that consumers expect future promotions according to a first-order Markov model. The authors find that the estimated expectations conform rather well with the actual promotion schedule (see also Erdem et al. 2003).

Sun et al. (2003) propose measuring deceleration by adding a variable  $PromTime_t^h$  to the category incidence model (Eq. 2.2). This represents the time since the last promotion, and is meant to capture that consumers may hold out until the next promotion. To obtain  $PromTime_t^h$  they calculate the average time between

promotions in the category. If the time since the last promotion in the category seen by the consumer is greater than this average,  $PromTime_t^h$  equals 1; otherwise it equals 0. If  $PromTime_t^h$  is 1, a consumer may expect a promotion soon, and defers the current purchase. As a result, they expect (and find) the estimated coefficient for  $PromTime_t^h$  to be negative.

### 2.3.5 Discussion

While the literature provides models for each of the key consumer responses to sales promotion, there are no papers yet that combine all possible responses that are listed in Table 2.2. Van Heerde and Gupta (2006) come close by combining store switching, category incidence, brand and SKU choice, purchase quantity, increased consumption effects and deceleration effects. This allows them to identify 18 of the 24 possible sources of the sales promotion bump from Table 2.2. Since Van Heerde and Gupta (2006) do not model category choice (Sect. 2.3.2), they do not measure effects related to category switching.

To estimate all 24 decomposition sources, one would have to estimate a model for all pertinent consumer decisions. This would require specifying the probability for the decision to choose store  $s$ , category  $k$ , brand  $b$ , SKU  $j$  and quantity  $q_{skbt}^h$  as:

$$P(Q_{skbt}^h = q_{skbt}^h) = P(S_t^h = s) \times P(Cat_{skt}^h = 1 | S_t^h = s) \times P(C_{skt}^h = \{b, j\} | Cat_{skt}^h = 1, S_t^h = s) \\ \times P(Q_{skbt}^h = q_{skbt}^h | S_t^h = s, Cat_{skt}^h = 1, C_{skt}^h = \{b, j\}) \quad (2.8)$$

The first three components at the right hand side can be modeled using store, category, and SKU models reviewed in this section. The quantity decision could be modeled using Eq. (2.6). However, it isn't clear whether one could derive an analytical formula to derive the 24-state decomposition in Table 2.2, or whether one would have to use simulation. For example, it would be difficult to distinguish between a category switch and a new category purchase that represents additional expenditures from the household's total budget.

## 2.4 Store-Level Models of Sales Promotion Effects

A large body of research has developed store-level models for sales promotion effects. These models draw on weekly store-level scanner data that are more readily available, more representative, and easier to process than household-level scanner data (Van Heerde et al. 2004). Table 2.4 shows that the phenomena that are studied with store-level vs household-level data are similar, but the terminology can differ.



**Table 2.4** Related terms in models of household- and store data

Household data	Store data
Purchase (Sect. 2.1)	Sales (Sect. 2.4.1)
Category incidence (Sect. 2.2.1)	Number of buyers in category
Brand switching (Sect. 2.2.2)	Cross-brand effects (Sects. 2.4.1 and 2.4.3)
Acceleration (Sect. 2.2.3)	Postpromotion dips (Sect. 2.4.2)
Deceleration (Sect. 2.3.4)	Prepromotion dips (Sect. 2.4.2)

### 2.4.1 Scan\*Pro Model

Perhaps the most well-known store-level model for sales promotion effects is Scan\*Pro (Wittink et al. 1988).<sup>4</sup> It is a multiplicative regression model for brand sales, explained by own-brand and cross-brand prices and promotions:

$$S_{bst} = \lambda_{bs} \mu_{bt} \prod_{b'=1}^B \left\{ PI_{b'st}^{\beta_{b'b}} \cdot \gamma_{1b'b}^{FEATONLY_{b'st}} \cdot \gamma_{2b'b}^{DISPONLY_{b'st}} \cdot \gamma_{3b'b}^{FEAT \& DISP_{b'st}} \right\} e^{\mu_{bst}}, \quad (2.9)$$

where

$S_{bst}$	sales (in units) of brand $b$ in store $s$ in week $t$
$PI_{bst}$	price index (ratio of current to regular price) of brand $b$ in store $s$ in week $t$
$FEATONLY_{bst}$	indicator for feature only: 1 if there is a feature without a display for brand $b$ in store $s$ in week $t$ , 0 else
$DISPONLY_{bst}$	indicator for display only: 1 if there is a display without feature for brand $b$ in store $s$ in week $t$ , 0 else
$FEAT \& DISP_{bst}$	indicator for feature and display: 1 if there is a feature and display for brand $b$ in store $s$ in week $t$ , 0 else

The model includes a brand-store specific intercept  $\lambda_{bs}$ , a brand-week specific intercept  $\mu_{bt}$ , own- ( $\beta_{bb}$ ) and cross-price ( $\beta_{b'b}$ ,  $b' \neq b$ ) elasticities, and multipliers for own- ( $\gamma_{1bb}$ ,  $\gamma_{2bb}$ ,  $\gamma_{3bb}$ ) and cross-brand ( $\gamma_{1b'b}$ ,  $\gamma_{2b'b}$ ,  $\gamma_{3b'b}$ ,  $b' \neq b$ ) effects for feature, display, and feature & display. Note by including these latter three variables, one can investigate a possible interaction between Feature and Display. The model is linearized by taking logs and estimated by ordinary least squares.<sup>5</sup>

Van Heerde et al. (2001) propose a semiparametric version of Scan\*Pro, estimated by nonparametric techniques. Their results show that the response of sales to the percentage price discount is S-shaped, i.e., there are threshold and saturation effects.

<sup>4</sup>This working paper was reprinted as Chap. 12 of Wieringa et al. (2011).

<sup>5</sup>Another important method is “PromotionScan” (Abraham and Lodish 1993). PromotionScan is a time series approach to estimating the short-term promotion bump. It is based on the “Promoter” method (See Sect. 2.8).

A key independent variable in the Scan\*Pro model is the price index (the ratio of actual to regular price). In store data, both actual and regular prices are typically available whereas household data tend to include price paid only. The price index captures the effects of promotional price changes, which may be quite different from regular price effects (Mulhern and Leone 1991; Bijmolt et al. 2005). If there is sufficient variation in regular price, it can be included as a separate predictor variable as in Mulhern and Leone (1991).

Note that Eq. (2.9) allows for asymmetric cross-effects between brands, i.e., the impact of promoting Brand A on sales of Brand B is not the same as the impact of promoting Brand B on sales of Brand A. This is an important feature because asymmetric brand switching has been consistently found in the promotions literature, although its causes are not yet completely explained (Neslin 2002). A downside of modeling all cross effects is that for  $N$  brands one needs  $N^2$  parameters per marketing mix instrument. As a result, parameter estimates can be unreliable and lack face validity. Roederkerk et al. (2013) use a parsimonious specification where the cross effects between SKUs are a function of the similarity of these SKUs.

Aggregate logit models (Sect. 2.4.4) overcome this problem by estimating only one parameter per instrument, from which the  $N^2$  own- and cross effects are derived. In case a modeler wants to consider not only cross-brand effects within categories but also across category effects, the number of cross-effects may become really problematic. To handle this case, Kamakura and Kang (2007) present a parsimonious solution based on a principal-component representation of response parameters.

While the Scan\*Pro Model specifies a model for brand sales, an alternative approach is to model both category sales and market share, as brand sales is the product of the two components. Breugelmans and Campo (2011) use this approach to study the effectiveness of in-store displays in a virtual store environment. They use an attraction specification for the market share model, which ensures that market shares add up to one across brands and are bound between zero and one. A difference between this approach and Scan\*Pro is that market share models tend to impose a structure on the own and cross effects (e.g., cross effects are a mathematical function of own effects) whereas the Scan\*Pro model estimates the cross effects freely. Freely estimated cross effects allow the data to “speak for themselves,” but they do not impose any constraint on the estimates, and hence cross effects with counterintuitive signs are possible in the Scan\*Pro model. Incorrect signs for cross effects are less likely in market share models.

### 2.4.2 Models for Pre- and Postpromotion Dips

Van Heerde et al. (2000) and Macé and Neslin (2004) have used store-level data to measure the aggregate effects of acceleration (i.e., postpromotion dips) and deceleration effects (i.e., prepromotion dips). The (simplified) model is of the form:

$$\ln S_t = \alpha_0 + \alpha_1 \ln PI_t + \sum_{u=1}^S \beta_u \ln PI_{t-u} + \sum_{v=1}^{S'} \gamma_v \ln PI_{t+v} + \varepsilon_t, \quad (2.10)$$

where

$\ln S_t$  natural log of sales in week  $t$

$\ln PI_t$  natural log of price index (ratio of current to regular price) in week  $t$

$\beta_u$   $u$ -week lagged effect of promotion, corresponding to post-promotion dips (acceleration). A positive  $\beta_u$  indicates that a price decrease is followed by a sales dip  $u$  weeks later

$\gamma_v$   $v$ -week lead effect of promotion, corresponding to pre-promotion dips (deceleration). A positive  $\gamma_v$  indicates that a price decrease in  $v$  weeks from now is preceded by a decrease in sales now

Macé and Neslin (2004) estimate pre- and postpromotion dips on data spanning 83 stores, 399 weeks, and 30,861 SKUs from 10 product categories. They find that 22% of the sales promotion bump is attributable to postpromotion dips, and 11% to prepromotion dips. Hence, pre- and postpromotion together are one-third of the sales promotion bump, which is remarkably close to the 32% cross-period effect reported by Van Heerde et al. (2004). Macé and Neslin (2004) find that SKU, category, and store-trading area customer characteristics explain significant variation in pre- and postpromotion elasticities.

Note that models for household- and store-level data deal differently with purchase acceleration. Since typical household-level models do not incorporate a store choice model, acceleration effects manifest both *within* the same store (source 9 in Table 2.2) and *across* stores (source 17 in Table 2.2). In store-level models such as (2.10), the aggregate outcome of acceleration (postpromotion dips) is only captured *within* the same store, which is source 9 in Table 2.2. As a result, one may expect larger acceleration effects in household-level than in store-level models.

### 2.4.3 Store-Level Decomposition Model

Van Heerde et al. (2004) propose a regression-based method for decomposing own-brand effects into cross-brand (brand switching), cross-period (acceleration & deceleration), and category expansion effects. The method uses the identity that total category sales (TCS) during periods  $t - S'$  through  $t + S$  equals sales of the target brand in period  $t$  ("own-brand sales" or OBS) plus sales of other brands in period  $t$  ("cross-brand sales" or CBS) plus total category sales in period  $t - S'$  through  $t + S$ , excluding period  $t$  ("pre- and post-period category sales" or PPCS). Therefore,  $TCS = OBS + CBS + PPCS$ , or  $-OBS = CBS + PPBC - TCS$ . The method regresses these four variables on the same set of regressors:

$$\begin{aligned}
\text{--OBS (own-brand sales)} \quad & -S_{bt} = \alpha^{ob} + \beta^{ob} PI_{bt} + \sum_{k=1}^K \gamma_k^{ob} X_{kt} + \varepsilon_{bt}^{ob} \\
\text{CBS (cross-brand sales)} \quad & \sum_{b'=1}^B S_{b't} = \alpha^{cb} + \beta^{cb} PI_{bt} + \sum_{k=1}^K \gamma_k^{cb} X_{kt} + \varepsilon_{bt}^{cb} \\
& b' \neq b \\
\text{PPCS (cross-period sales)} \quad & \sum_{u=-S'}^S \sum_{\substack{b'=1 \\ b \neq b}}^B S_{b't+u} = \alpha^{cp} + \beta^{cp} PI_{bt} \\
& + \sum_{k=1}^K \gamma_k^{cp} X_{kt} + \varepsilon_{bt}^{cp} \\
\text{--TCS (total category sales)} \quad & -\sum_{u=-S'}^S \sum_{b=1}^B S_{bt+u} = \alpha^{ce} + \beta^{ce} PI_{bt} \\
& + \sum_{k=1}^K \gamma_k^{ce} X_{kt} + \varepsilon_{bt}^{ce}
\end{aligned}$$

where  $\sum_{k=1}^K \gamma_k X_{kt}$  captures the effects of covariates such as cross-brand instruments, store dummies, and week dummies. Since  $\text{--OBS} = \text{CBS} + \text{PPBC} - \text{TCS}$ , the parameters for the price indices ( $PI$ ) add up in the following way:

$$\beta^{ob} = \beta^{cb} + \beta^{cp} + \beta^{ce}, \quad (2.11)$$

Equation (2.11) decomposes the own-brand promotion effect ( $\beta^{ob}$ ) into cross-brand switching ( $\beta^{cb}$ ), acceleration and decoration ( $\beta^{cp}$ ), and category expansion ( $\beta^{ce}$ ). Equation 2.11 can be divided through by  $\beta^{ob}$  to provide a percentage decomposition.

All parameters in Eq. (2.11) are expected to be positive. If there is a promotional price discount for brand  $b$ ,  $PI_{bt}$  decreases, own brand sales increases (presumably), and hence minus own brand sales decreases. Consequently, the regression coefficient  $\beta^{ob}$  will be positive. Similarly, a price discount for brand  $b$  decreases cross-brand sales (presumably), which implies  $\beta^{cb} > 0$ . Furthermore, if a decrease in  $PI_{bt}$  leads to a decrease in cross-period sales (i.e., pre- and postpromotion dips)  $\beta^{cp} > 0$ . Finally, if the price discount for brand  $b$  manages to increase category sales, then total category sales increase and the negative decreases, and  $\beta^{ce} > 0$ .

Van Heerde et al. (2004) obtain the decomposition for four types of promotional price discounts: without feature- or display support, with feature-only support, with display-only support, and with feature- and display support. To accomplish this, they use a specific set of independent variables, discussed in the appendix to this chapter.

Van Heerde et al. (2004) provide two extensions of (2.11). One of them defines the model at the SKU level and splits the “cross-brand effect” (or cross-item effect) into within-brand cannibalization ( $\beta^{cbw}$ ) and between-brand switching ( $\beta^{cbb}$ ), i.e.,  $\beta^{cb} = \beta^{cbw} + \beta^{cbb}$ . The other extension splits the category expansion effect into a cross-store effect ( $\beta^{cs}$ ) and a market-expansion effect ( $\beta^{me}$ ) (the category-growth effect in Tables 2.1 and 2.2), i.e.,  $\beta^{ce} = \beta^{cs} + \beta^{me}$ . This allows the model to quantify within-brand SKU switching as well as store switching.

Leeftang et al. (2008) extend the model in yet another direction by accounting for cross-category effects. The model allows for positive (complementary) and negative (substitution) cross-category effects. The method uses pairs of categories between which complementary or substitution effects can be expected (e.g., canned

beer and bottled beer). Leeflang and Parreña Selva (2012) also study cross-category effects but focus on the moderating factors of these effects, such as the physical distance between categories in a store.

#### ***2.4.4 Heterogeneity Across Stores***

The original Scan\*Pro paper (Wittink et al. 1988) reports strong differences in promotional responses across US regions. Brand managers may exploit these differences by tailoring promotions at the regional level. Several studies have since examined how promotional effectiveness varies across stores. Hoch et al. (1995) and Montgomery (1997) use hierarchical Bayesian methods to allow price response parameters to differ across stores. Hoch et al. (1994) and Montgomery (1997) use store and trading characteristics to explain variation in price sensitivity across stores. Montgomery (1997) shows how the model can be used to adjust prices at the store level for enhanced profitability.

Haans and Gijsbrechts (2011) study how the effectiveness of promotions on category sales varies with the size of the store. They use a hierarchical linear model where log category sales is explained by log price, discount depth, feature and display and a quantity discount variable. The response parameters are store-specific, and they are explained by the size of the store in a second layer. Haans and Gijsbrechts (2011) find that while the percentage lift due to promotions is smaller in larger stores, the absolute effect are larger.

#### ***2.4.5 Aggregate Logit Model***

A frequent criticism of regression models of promotion is they are not rooted in economic theory. The aggregate logit model overcomes this issue, which is one reason it is increasingly popular in marketing science. Another reason is that it accommodates own- and cross effects with an economy of parameters, which is something that does not hold for the aggregate models discussed in Sects. 2.4.1–2.4.3. The aggregate logit model was introduced by Berry et al. (1995). Its logic is that individual consumers maximize utility and choose brands according to a multinomial logit model.

Estimation of aggregate logit models can take into account price endogeneity (Besanko et al. 1998; Villas-Boas and Winer 1999). For instance in times of a positive demand shock, managers may increase price. To account for this, researchers have correlated price with the error term (Besanko et al. 1998; Villas-Boas and Winer 1999). If the endogenous nature of price is ignored, its coefficient  $\beta_2$  may be underestimated quite severely as shown in a meta-analysis of price elasticity (Bijmolt et al. 2005).

To complement the demand model (2.17), Besanko et al. (1998) assume a certain model of competitive conduct, and derive a supply model from that. Next,

**Table 2.5** Sales bump decomposition results reported in the literature

Study	Category	% secondary demand (=brand switching effect)	% primary demand (=own sales effect not due to brand switching)
Elasticity decomposition			
Gupta (1988)	Coffee	84	16
Chiang (1991)	Coffee (feature)	81	19
	Coffee (display)	85	15
Chintagunta (1993)	Yogurt	40	60
Bucklin et al. (1998)	Yogurt	58	42
Bell et al. (1999)	Margarine	94	6
	Soft drinks	86	14
	Sugar	84	16
	Paper towels	83	17
	Bathroom tissue	81	19
	Dryer softeners	79	21
	Yogurt	78	22
	Ice cream	77	23
	Potato chips	72	28
	Bacon	72	28
	Liquid detergents	70	30
	Coffee	53	47
	Butter	49	51
Chib et al. (2004)	Cola (price)	78	22
	Cola (display)	68	32
	Cola (feature)	64	36
Van Heerde et al. (2003)	Sugar	65	35
	Yogurt	58	42
	Tuna	49	51
Nair et al. (2005)	Orange juice	65	35
Average elasticity decomposition		71	29
Unit-sales decomposition			
Pauwels et al. (2002)	Soup	11	89
	Yogurt	39	61
Van Heerde et al. (2003)	Sugar	45	55
	Yogurt	33	67
	Tuna	22	78

(continued)

**Table 2.5** (continued)

Study	Category	% secondary demand (=brand switching effect)	% primary demand (=own sales effect not due to brand switching)
Sun et al. (2003)	Ketchup	56	44
Van Heerde et al. (2004)	Peanut butter	43	57
	Shampoo	31	69
	Tuna	31	69
	Bathroom tissue	21	79
Nair et al. (2005)	Orange juice	8	92
Sun (2005)	Tuna	42	58
	Yogurt	39	61
Chan et al. (2008)	Tuna	28	72
	Paper towels	14	86
Leeflang et al. (2008)	Bottled beer	18	82
	Canned beer	11	89
	Concentrated fabric softeners	13	87
	Non-concentrated fabric softeners	29	71
	Concentrated dish detergents	4	96
	Non-concentrated dish detergents	28	72
	Detergents	39	61
Average unit sales effect decomposition		28	72

the demand and supply models are estimated simultaneously. Whereas the original aggregate logit model assumed homogenous consumers, several papers in marketing have relaxed this assumption (e.g., Chintagunta 2001; Dubé et al. 2002). Nevo (2000) provides guidelines how to estimate a heterogeneous aggregate logit model. Moreover, Nair et al. (2005) propose aggregate models that not only capture underlying brand choice decisions, but also incidence and quantity. Their demand elasticity breakdown shows that brand choice accounts for 65% and incidence and quantity for 35%, which is in the same ballpark as the breakdowns obtained from individual-level data (see Table 2.5 in Sect. 2.5).

Park and Gupta (2009) develop a simulated maximum likelihood estimation method for the random coefficient aggregate logit model. The method is especially suitable when there are relatively small samples of shoppers, leading to measurement error. The estimation accounts for endogeneity and it yields unbiased and efficient estimates of the demand parameters.

One drawback of aggregate logit models is that it is difficult to identify unobserved heterogeneity with aggregate data (Bodapati and Gupta 2004). Another one is that it requires the specification of an outside good to account for non-incidence.

This outside good has to be based on assumptions on population size and category consumption, which may be questionable. Yet another drawback of current aggregate logit models is that they typically ignore dynamic and quantity effects such as acceleration, deceleration, and purchase-event feedback. An issue that also needs further development is that the assumed competitive conduct refers to long-term stable prices instead of to Hi-Lo pricing that is the essence of promotional pricing (Neslin 2002, p. 45).

## 2.5 Generalizations About the Decomposition

Table 2.5 summarizes the findings of the literature on the decomposition of the sales promotion bump. For each study we indicate the product category and the percentage attributed to secondary demand effects (effects that cause substitution from other brands, i.e., brand switching) and the percentage due to primary demand effects (the part of the sales promotion bump that is not due to brand switching). There are two fundamental approaches to calculating the decomposition, the “elasticity” approach and the “unit sales” approach. The elasticity approach is explained in Sect. 2.2.4, and was originated by Gupta (1988). It is based on the mathematical relationship that the elasticity of the probability of buying brand  $b$  at time  $t$  with respect to (an assumed continuous measure of) promotion equals the sum of the elasticities of brand choice, purchase incidence, and purchase quantity with respect to promotion. The unit sales composition looks at changes in actual sales of the promoted brand as well as other brands in the category. The upper part of Table 2.5 shows the elasticity decomposition, the lower part shows the unit sales decomposition. There are two important findings in Table 2.5:

1. The unit sales decomposition yields lower secondary demand effects compared to the elasticity decomposition, generally in the 10–40% range.
2. Using the elasticity decomposition, there is a general downward trend in the percentage allocated to secondary demand, starting from about 80–85% and now at 45–70%.

The difference in findings for unit sales versus elasticity decompositions is detailed in Van Heerde et al. (2003). The brand choice component of the elasticity decomposition represents the change in the brand choice probability conditional on making a category purchase. The cross-brand component focuses on how many sales other brands lose when the focal brand is promoted (Van Heerde et al. 2003). The key difference is the way they treat the increase in purchase incidence due to the promotion. This increase is ignored in the elasticity decomposition, but in the unit sales decomposition, it is included in the calculation of the net sales loss for the other brands (Van Heerde et al. 2003). As a result, the actual loss in cross-brand sales is much less than what the elasticity-based secondary demand fraction suggests. This difference holds within the same category, as illustrated by the entries



for orange juice (Nair et al. 2005) and sugar, yogurt and tuna (Van Heerde et al. 2003) in the top and lower parts of Table 2.5.

One possible reason for the general downward trend in secondary demand effects in elasticity is that more recent models allow for unobserved household heterogeneity. The argument is this (see also Keane 1997b; Sun et al. 2003). Suppose there are two brands, A and B. A large segment of customers lies in wait for brand A to be on promotion, while the other segment lies in wait for brand B's promotions. When brand A is promoted, it is almost exclusively bought by the first segment. As a result, the conditional choice probability for brand A increases spectacularly when it is promoted, while the conditional choice probability for brand B approaches zero. When brand B is promoted, the reverse occurs. Such a phenomenon leads to a very strong conditional brand choice elasticity. However, actually there is very little switching between A and B (promotion is influencing incidence, not switching), and hence there is small cross-brand sales loss in the unit sales decomposition. If this explanation holds, then models that allow for unobserved household heterogeneity should show a lower percentage brand switching than models that assume homogeneity. This could be the reason why the more recent elasticity decomposition results, which tend to be derived from heterogeneous models, show less brand switching. This seems a worthwhile direction for further research.

## 2.6 Long-Term Impact—Beyond the Immediate Sales Bump

Promotions affect consumers beyond the immediate sales bump. Promotions may lead to purchase-event feedback (Sect. 2.6.1). Promotions may also affect reference prices (Sect. 2.6.2). Over time, consumers learn price promotion patterns (Sect. 2.6.3). Promotions may affect long-term consumer behavior (Sect. 2.6.4). Finally, competitors may react (Sect. 2.6.5).

### 2.6.1 Purchase-Event Feedback

Purchase event feedback is the degree to which current purchases affect future brand preferences. This is known as “state dependence” in the economics literature (see Roy et al. 1996), and is due to consumer learning from the product purchase and usage experience.

Researchers have captured purchase event feedback by including a lagged purchase indicator such as  $Last_{bt}^h$  in Eq. (2.5). However, Blattberg and Neslin (1990) distinguish between the purchase effect and the promotion usage effect—the purchase-event feedback from a purchase on promotion may be different than the feedback from a regular purchase. For example, self-perception theory suggests that if the consumer concludes he or she bought the brand because of the promotion

rather than brand preference, purchase event feedback will be weakened (Dodson et al. 1978). Behavioral learning theory (Rothschild and Gaidis 1981) suggests promotion purchasing could enhance or detract from purchase event feedback. The effect could be positive if the promotion serves as a reward and thus encourages future purchasing, or negative if promotion merely trains consumers to buy on promotion. To investigate this, Gedenk and Neslin (1999) distinguish whether or not the last purchase was made on promotion. They find that price promotions detract from feedback. This finding is the same as originally reported by Guadagni and Little (1983)—a promotion purchase is less reinforcing than a non-promotion purchase, but better than no purchase at all. It is also the same as found by Seetharaman (2004).

Ailawadi et al. (2007b) propose yet another mechanism for feedback effects of promotions. They postulate that acceleration in the form of larger purchase quantity enhances purchase-event feedback because the household consumes more of the brand over a continuous period of time. In an empirical study of yogurt and ketchup, Ailawadi et al. (2007a, b) find that larger purchase quantity is associated with an increase in repeat purchase rates.

The measurement of purchase event feedback is quite challenging because of its potential confound with customer heterogeneity. Failure to account adequately for customer heterogeneity produces spurious state dependence findings (see Sects. 2.2.6 and 2.6.3).

## 2.6.2 Reference Prices

The reference price is the standard to which consumers compare an observed price in order to assess the latter's attractiveness (Kalyanaram and Winer 1995). Although there are many ways to operationalize reference price (Winer 1986), a significant body of literature supports the notion that individuals make brand choices based on this comparison. Briesch et al. (1997) conclude that a brand-specific exponentially smoothed reference price provides the best fit and prediction:

$$RP_{bt}^h = \alpha RP_{bt-1}^h + (1 - \alpha) \text{Price}_{bt-1}, \quad (2.12)$$

where

$RP_{bt}^h$  household  $h$ 's reference price of brand  $b$  at purchase occasion  $t$   
 $\alpha$  carryover parameter,  $0 \leq \alpha \leq 1$

Prospect theory (Kahneman and Tversky 1979) predicts that consumers react more strongly to price increases than to price decreases (Kalyanaram and Winer 1995). To operationalize this, Erdem et al. (2001) define *LOSS* as the difference between the actual price and the reference price, given that the reference price is

lower than the actual price. Similarly, *GAIN* is the difference given that the reference price is higher than the actual price:

$$\begin{aligned} LOSS_{bt}^h &= \max \{ \text{Price}_{bt-1} - RP_{bt}^h, 0 \} \\ GAIN_{bt}^h &= \max \{ RP_{bt}^h - \text{Price}_{bt-1}, 0 \} \end{aligned}$$

To capture the direct effect of price as well as the effects of losses and gains, the utility function in the brand choice model (2.5) can be specified as (Briesch et al. 1997):

$$\beta X_{bt}^h = \beta_1 PRICE_{bt} + \beta_2 FEAT_{bt} + \beta_3 DISP_{bt} + \beta_4 BL_b^h + \beta_5 Last_{bt}^h + \beta_6 LOSS_{bt}^h + \beta_7 GAIN_{bt}^h. \quad (2.13)$$

In (2.13), we expect  $\beta_1 < 0$ ,  $\beta_6 < 0$ , and  $\beta_7 > 0$ . If losses loom larger than gains,  $|\beta_6| > |\beta_7|$ .

Recent papers question the findings regarding loss-aversion (Bell and Lattin 2000), and whether the reference price effect itself has been significantly over-estimated (Chang et al. 1999). The Chang et al. argument is that price sensitive consumers time their purchases to promotions, so observations of purchases with low prices over-represent price sensitive consumers and over-estimate both the loss and gain aspects of reference prices. Further work is needed to take into account these points. From a modeling standpoint, these papers illustrate the subtle but important challenges in modeling household-level data.

### 2.6.3 Consumer Learning

Frequent exposure to sales promotions may affect consumer perceptions of promotional activity (Krishna et al. 1991) and change their response to promotion. Mela et al. (1997) study the long-term effects of promotion and advertising on consumers' brand choice behavior. They use 8 1/4 years of panel data for a frequently packaged good. Their results indicate that consumers become more price and promotion sensitive over time because of reduced advertising and increased promotions. Mela et al. (1998) conclude that increased long-term exposure of households to promotions reduces their category purchase rate. However, when households do decide to buy, they buy more of the product. Such behavior is indicative of an increasing tendency to "lie in wait" for especially good promotions. This study was among the first to provide evidence of purchase deceleration (see Sect. 2.3.4).

Bijmolt et al. (2005) provide a meta-analysis of 1851 price elasticities reported in four decades of academic research in marketing. A salient finding is that in the period 1956–1999, the average (ceteris paribus) elasticity of sales to price went from  $-1.8$  to  $-3.5$ . The relative elasticities (i.e., choice and market share) are quite

stable (i.e., no significant change). Thus, the primary demand part of the sales elasticity is increasing over time, whereas the secondary demand part is stable. This finding is consistent with “lie-in-wait” behavior reported by Mela et al. (1998), but inconsistent with an increased sensitivity of the brand choice decision to price reported by Mela et al. (1997).

Consumers’ learning of product quality has been examined within the purview of Bayesian learning models. See Sect. 2.2.6 for a discussion of dynamic structural models, several of which incorporate Bayesian learning. Erdem et al. (2008) develop a Bayesian learning model particularly relevant to promotions. They investigate the extent to which consumers learn the quality of a brand through experience, price/promotions, advertising frequency, and advertising content. They find experience is the most important source of learning, while price/promotions and the combined impact of advertising are equally important. They find that price promotions induce negative learning, reducing the total sales impact of promotion by 27% (p. 1122). The authors note however that these effects have to be disentangled from stockpiling (p. 1124).

This highlights the difficulty in measuring learning in structural models. Shin et al. (2012) show that failing to account correctly for customer preference heterogeneity can bias the estimate of learning. Their study is particularly interesting because it combines survey data of brand preferences with the usual consumer panel data. They find that without survey data, the extent of customer learning is significantly over-stated.

DeVecchio et al. (2006) shed further light on learning from promotions by conducting a meta-analysis of 42 studies generating 132 inferred correlations between “the use of sales promotion and post-promotion brand preference” (p. 207). They consider studies that measure choice as well as brand perceptions, based on field or laboratory data. They find that on average promotion does not have a statistically significant association with brand preference. However, there are moderators in both negative and positive directions. Promotions have a negative effect when the discount is 20% or more, when it is an unannounced price reduction, and when it is for a durable good. Promotions have a positive effect when the promotion is a coupon or a premium, when it is a packaged good, and when the brand is competing against several brands. These findings suggest that researchers modeling what consumers learn from promotion need to take into account the form of promotion and the steepness of the promotion discount.

#### ***2.6.4 Long-Term Effects***

Researchers have also started to investigate the long-term effects of promotions. If sales promotions are successful in attracting (new) consumers to the brand or increase their consumption rate permanently, the sales impact of the promotion should be observed beyond the immediate sales promotion bump. For aggregate (sales) data, there are two primary ways of modeling the long-term effects of sales

promotion: Vector Autoregressive Models with X-variables (VARX) and Vector Error-Correction Models. A two-brand VARX model for sales promotion effects could be specified as follows (cf. Nijs et al. 2001):

$$\begin{aligned}
 \begin{bmatrix} \Delta \ln S_{1t} \\ \Delta \ln S_{2t} \\ \Delta \ln \text{Price}_{1t} \\ \Delta \ln \text{Price}_{2t} \end{bmatrix} &= \begin{bmatrix} c_{0,S1} + \sum_{s=2}^{13} c_{s,S1} SD_{st} + \delta_{S1t} \\ c_{0,S1} + \sum_{s=2}^{13} c_{s,S1} SD_{st} + \delta_{S2t} \\ c_{0,P1} + \sum_{s=2}^{13} c_{s,P1} SD_{st} + \delta_{P1t} \\ c_{0,P1} + \sum_{s=2}^{13} c_{s,P1} SD_{st} + \delta_{P2t} \end{bmatrix} + \sum_{i=1}^8 \begin{bmatrix} \phi_{11}^i & \phi_{12}^i & \phi_{13}^i & \phi_{14}^i \\ \phi_{21}^i & \phi_{22}^i & \phi_{23}^i & \phi_{24}^i \\ \phi_{31}^i & \phi_{32}^i & \phi_{33}^i & \phi_{34}^i \\ \phi_{41}^i & \phi_{42}^i & \phi_{43}^i & \phi_{44}^i \end{bmatrix} \begin{bmatrix} \Delta \ln S_{1t-i} \\ \Delta \ln S_{2t-i} \\ \Delta \ln \text{Price}_{1t-i} \\ \Delta \ln \text{Price}_{2t-i} \end{bmatrix} \\
 &+ \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} & \gamma_{14} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} & \gamma_{24} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} & \gamma_{34} \\ \gamma_{41} & \gamma_{42} & \gamma_{43} & \gamma_{44} \end{bmatrix} \begin{bmatrix} \Delta FEAT_{1t} \\ \Delta DISP_{1t} \\ \Delta FEAT_{2t} \\ \Delta DISP_{2t} \end{bmatrix} + \begin{bmatrix} u_{S1t} \\ u_{S2t} \\ u_{P1t} \\ u_{P2t} \end{bmatrix}
 \end{aligned} \tag{2.14}$$

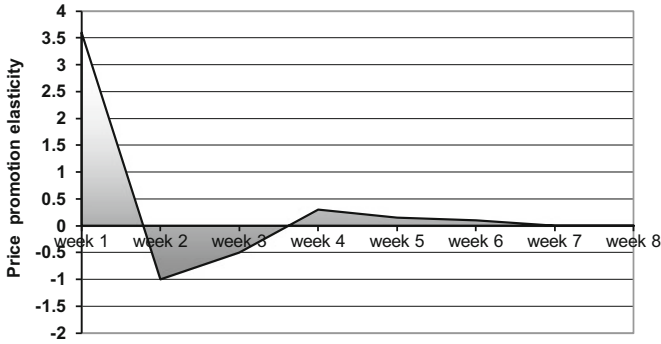
where

- $\Delta \ln S_{bt}$   $\ln S_{bt} - \ln S_{bt-1}$ , i.e., current minus lagged sales of brand  $b$ .
- $\Delta \ln \text{Price}_{bt}$   $\ln \text{Price}_{bt} - \ln \text{Price}_{bt-1}$ , i.e., current minus lagged price of brand  $b$
- $SD_{st}$  a 4-weekly seasonal dummy variable (1 during 4-week period  $s$ , 0 else)
- $t$  deterministic trend variable
- $\Delta FEAT_{bt}$   $FEAT_{bt} - FEAT_{bt-1}$ , i.e., the current minus lagged feature dummy for brand  $b$
- $\Delta DISP_{bt}$   $DISP_{bt} - DISP_{bt-1}$ , i.e., the current minus lagged display dummy for brand  $b$

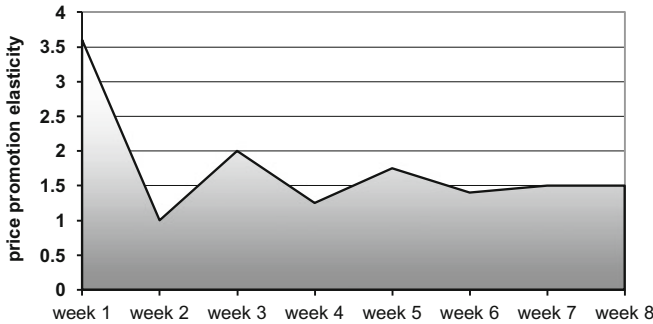
Equation (2.14) is estimated by OLS or SUR (Seemingly Unrelated Regression). Once the parameters have been estimated, researchers calculate Impulse Response Functions (IRF) to track the incremental impact of a one standard deviation price promotion shock on sales in periods  $t, t+1, t+2, \dots$ . The permanent effect of a promotion is the asymptotic value of log sales when  $t \rightarrow \infty$ . Figure 2.1 shows a hypothetical Impulse Response Function. The area under the curve in grey in the figures is called the “long-term effect” or “cumulative effect”. In Fig. 2.1, there is a zero permanent effect of a promotion in week 1. Such a pattern corresponds to no unit root in the sales series (Dekimpe and Hanssens 1995). This means that the long-term (=cumulative) effect is finite. When sales have a unit root, there can be a nonzero permanent effect of a one-time promotion as illustrated in Fig. 2.2. In this case, the long-term effect (area under the curve) is infinite.

Fok et al. (2006) propose a Vector-Error Correction model that directly captures the short- and long-term effects of sales promotions on *brand sales*:

$$\Delta \ln S_t = \beta_0 + \sum_{k=1}^K A_k^{sr} \Delta X_{kt} + \Pi \left( \ln S_{t-1} - \sum_{k=1}^K A_k^{lr} X_{k,t-1} \right) + \nu_t, \quad \nu_t \sim N(0, V), \tag{2.15}$$



**Fig. 2.1** Promotion in week 1 with a zero permanent effect



**Fig. 2.2** Promotion in week 1 with a positive permanent effect

where

- $\Delta$  first difference operator:  $\Delta X_t = X_t - X_{t-1}$
- $S_t$  Vector ( $B \times 1$ ) with sales (in kilo) of brands  $b = 1, \dots, B$  in week  $t$
- $X_{kt}$  Vector ( $B \times 1$ ) with marketing-mix variable  $k$  ( $k = 1, \dots, K$ ) of brands  $b = 1, \dots, B$  in week  $t$
- $\beta_0$  Vector ( $B \times 1$ ) with intercepts of brands  $b = 1, \dots, B$
- $A_k^{sr}$  Matrix ( $B \times B$ ) with short-term effects of marketing-mix variable  $k$
- $A_k^{lr}$  Matrix ( $B \times B$ ) with long-term effects of marketing-mix variable  $k$
- $\Pi$  Diagonal matrix ( $B \times B$ ) with adjustment effects
- $\nu_t$  Vector ( $B \times 1$ ) of error terms of brands  $b = 1, \dots, B$  in week  $t$
- $V$  Variance-Covariance matrix ( $B \times B$ ) of the error term  $\nu_t$

The diagonal elements of  $A_k^{sr}$  and  $A_k^{lr}$  measure respectively, the short- and long-run effects of the  $k$ -th marketing-mix variable of each brand, while the off-diagonal elements capture cross effects. The  $\Pi$  parameters reflect the speed of adjustment towards the underlying long-term equilibrium. We refer to Fok et al.

**Table 2.6** Category-demand effects of price promotions across 460 Categories

	Short term effects (%)	Permanent effects (%)
Positive	58	2
Negative	5	0
Zero	37	98

This table is based on Nijs et al. (2001)

**Table 2.7** Own-brand sales effects across 442 categories

	Non-significant (%)	Positive own-sales elasticity (%)	Negative own-sales elasticity (%)	Mean own-sales elasticity
<i>Short term effects</i>				
Price promotions	30.96	63.54	5.50	3.989
Advertising	67.00	20.45	12.55	0.014
<i>Permanent effects</i>				
Price promotions	94.99	4.15	0.86	0.046
Advertising	98.23	1.28	0.49	0.000

This table is based on Steenkamp et al. (2005)

(2006) for a formal proof of these properties, and to Van Heerde et al. (2007, 2010, 2013) for applications. Horváth and Franses (2003) provide an in-depth discussion about testing for whether the Error Correction model is appropriate, and the cost of estimating an Error Correction model when it is not appropriate, and vice versa.

Using a VARX model, Nijs et al. (2001) study the effects of consumer price promotions on *category sales* across 460 consumer product categories over a 4-year period. The data describe national sales in Dutch supermarkets and cover virtually the entire marketing mix. The key results are in Table 2.6. In 98% of the cases, there is no permanent effect of promotions on category sales, i.e., the Impulse Response Function resembles Fig. 2.1. This is a sensible result, in that one would not expect permanent category sales effects in the mature categories carried by most supermarkets. An interesting extension of this work would be to look at new categories such as tablet computers or HDTVs.

Steenkamp et al. (2005) use a VARX model to study the permanent effects of promotions and advertising on brand sales for the top three brands from 442 frequently purchased consumer product categories in the Netherlands. Their major results are displayed in Table 2.7. A key conclusion is that in the far majority of cases, there are no permanent effects of promotion and advertising on own-brand sales. In the short term, these effects do exist, and they are more prevalent and stronger for promotions than for advertising.<sup>6</sup>

<sup>6</sup>Note that the final version of this paper (Steenkamp et al. 2005) does not contain these results anymore, since the journal requested the authors to focus on competitive reactions.

Slotegraaf and Pauwels (2008) use persistence (time series) modeling to measure the long-term promotion effects for 100 brands. A key distinguishing feature of the study is that it does not limit itself to just the top three or top five brands per category, but studies 9–25 brands per category, including small brands. Slotegraaf and Pauwels (2008) find that permanent effects are fairly common, as 14% of the brands have a positive sales evolution. On average, price promotions yield a higher permanent elasticity (0.06) than feature (0.003) and display (0.002). Slotegraaf and Pauwels (2008) also document substantial variation in these elasticities across brands. Both permanent and cumulative sales effects from promotions are larger for brands with higher brand equity and more product introductions.

Interestingly, Pauwels et al. (2002) show, based on data from the canned soup and yogurt categories, that permanent promotion effects are virtually absent for brand choice, category incidence, and purchase quantity.

Fok et al. (2006) apply their VEC model (2.20) to seven years of U.S. data on weekly *brand sales* of 100 different brands in 25 product categories. On average, the cumulative promotional price elasticity (−1.91) tends to be smaller in absolute value than the immediate promotional price elasticity (−2.29). Hence, some of the positive effects of a price promotion are compensated in the periods following the promotion by, for example, the effects of acceleration. Actually, the implied size of the post-promotion dip is 17% ( $100\% * ((2.29 - 1.91)/2.29)$ ), which is quite close to 22%, which is what we calculate for the postpromotion dip in Macé and Neslin (2004).<sup>7</sup> It is also interesting to note that the short-term regular price elasticity (−0.54) is much smaller in magnitude than the short-term price promotion elasticity (−2.29).

Ataman et al. (2008) study the long-term sales effects of promotion and other marketing variables. They apply Bayesian time-varying parameter models to analyze the sales for 225 *newly introduced* CPG brands observed across five years. While distribution is the strongest driver of long-term sales, feature and display ranked as the second-most effective instrument. For *mature* CPG brands, Ataman et al. (2010) show that price discounting lead to lower baseline sales in the long run and a to a heightened price sensitivity.

Datta et al. (2015) study the long-term effects of free-trial promotions. Many service providers offer consumers a free use of the service for a limited time in the hopes of retaining these customers. Datta et al. (2015) use a logit model for the monthly decision to keep the service. They show that the acquisition through a free trial affects the baseline retention rate even after the free trial is over. Free-trial customers, on the other hand, are more sensitive to their own usage rates and to marketing activities, creating opportunities for companies to retain these customers.

---

<sup>7</sup>The 22% cannot directly be calculated from Table 2.4 in Macé and Neslin (2004) because the elasticities reported in that table are point elasticities and therefore do not exactly correspond to the 20% price cut effects calculated in that table. However, calculations using the detailed results summarized in Macé and Neslin's Table 2.4, reveal that on average, 66.2% of the combined pre and post effect is due to post effects. Since the combined effect reported in Macé and Neslin's Table 2.4 is 33.3% of the bump ( $1 - 0.667$  from the last column in the table), the percentage due to postpromotion dips is  $0.662 \times 0.333 = 0.2204 = 22.0\%$ .



### 2.6.5 Competitive Reactions<sup>8</sup>

Since promotions affect cross-brand sales and market shares, and they are relatively easy to replicate, competitive reactions are likely. Competitors may either retaliate or accommodate a promotion initiated by a rival brand. Moreover, they may respond in-kind with the same instrument (e.g., price cut followed by price cut) or with another instrument (e.g., price cut followed by volume-plus promotion). Leeflang and Wittink (1992, 1996) specify reaction functions that allow for the measurement of the degree and nature of competitive reactions:

$$\ln\left(\frac{P_{bt}}{P_{b,t-1}}\right) = \alpha_b + \sum_{\substack{b'=1 \\ b' \neq b}}^B \sum_{t^*=1}^{T^*+1} \beta_{bb't^*} \ln\left(\frac{P_{b',t-t^*+1}}{P_{b',t-t^*}}\right) + \sum_{t^*=2}^{T^*+1} \beta_{bb't^*} \ln\left(\frac{P_{b,t-t^*+1}}{P_{b,t-t^*}}\right) + \sum_{b'=1}^B \sum_{t^*=1}^{T^*+1} \sum_{x=1}^3 \tau_{xbbt^*} (wx_{b',t-t^*+1} - wx_{b',t-t^*}) + \varepsilon_{bt} \quad (2.16)$$

$\frac{P_{bt}}{P_{b,t-1}}$	ratio of successive prices for brand $b$ in period $t$
$wx_{b,t} - wx_{b,t-1}$	first difference for the three types of promotions for brand $b$ : $x = 1$ (feature), $x = 2$ (display), and $x = 3$ (feature and display)

The parameter  $\beta_{bb't^*}$  represents competitive reactions with the same instrument: the price response by brand  $b$  to a price change by brand  $b'$  that took place  $t^*$  periods ago. Parameter  $\tau_{xbbt^*}$  captures competitive reactions with different instruments: the price response by brand  $b$  to a promotion of type  $x$  by brand  $b'$  that took place  $t^*$  periods ago.

For the grocery category under study, Leeflang and Wittink (1992) find that competitor reactions occur quite frequently, especially using the same marketing instrument as the initiator. By studying competitive reactions based on over 400 consumer product categories over a four-year time span, Steenkamp et al. (2005) test the empirical generalizability of Leeflang and Wittink (1992, 1996). They use VARX models similar to Eq. (2.14). Table 2.8 shows that the predominant reaction to a price promotion attack is no reaction at all. Indeed, for 54% of the brands under price promotion attack, the average short-term promotion reaction is not significantly different from zero. Furthermore, the significant short-term promotion reactions are twice more as likely to be retaliatory than accommodating (30% vs. 16%). Table 2.8 also shows that long-term reactions are very rare. In over 90% of the instances, price promotion attacks do not elicit a persistent or long-term price promotion on the part of the defending brand.

<sup>8</sup>Leeflang (2008) provides an in-depth discussion on models for competitive reactions, including structural models.

**Table 2.8** Competitive reactions to price promotions

Reaction with price promotion	Short-term effect (%)	Long-term effect (%)
No reaction	54	92
Competitive reaction	30	5
Cooperative reaction	16	3

This table is based on Steenkamp et al. (2005)

Steenkamp et al. (2005) find that absence of reaction corresponds primarily to the absence of harmful cross-sales effects. Only 118 out of 954 brands miss an opportunity in that they could have defended their position, but chose not to. When managers do opt to retaliate, effective retaliation is prevalent (63%). In 56% of these cases the response neutralizes the competitive attack, whereas in 36% of these cases the net effect is positive for the defending brand.

An interesting perspective is provided by Pauwels (2007). He finds that competitive response to promotions plays a relatively minor role in post-promotion effects. The major factor in post-promotion effects is the company’s own “inertia” to continue promoting in subsequent weeks. This is a very interesting finding in that it says companies are highly myopic when it comes to formulating promotion policy, basing the frequency of future promotions on the frequency of past promotions, rather than considering the competitive implications.

The competitive reactions studies discussed so far are all in “business as usual” situations. During price wars, which are an extreme form of price competition, competitive reactions may become fiercer, partly fueled by media coverage (Van Heerde et al. 2015).

**2.7 Endogeneity**

**2.7.1 What Is Endogeneity?**

Managers may plan sales promotions based on demand factors that are not included in the model. For example, a manager may anticipate a demand increase due to external factors (e.g., favorable weather, events) and decide to cut back on sales promotions or discounts because the product will be in high demand regardless. Alternatively, a manager may capitalize on these demand shocks and offer more or better sales promotion deals than usual to seize the moment. Because these demand shocks are not observed by the researcher (the researcher does not have the data to capture them), they are not explicitly in the model and they become part of the error term. This produces a correlation between the sales promotion variable in the model and the error term. When the model is estimated, the estimate for the effect of the sales promotion variable on utility will be incorrect, because it will partly include

the effect of the unobserved demand shock that is correlated with the sales promotion variable. This is the endogeneity problem.

More formally, the endogeneity problem occurs when an independent variable is correlated with the error term. This leads to a biased and inconsistent estimate for the effect of the independent variable on the dependent variable. Endogeneity has become a major theme in the academic marketing literature, especially the endogeneity of pricing. The concern is that managers may set prices strategically and that therefore the observed variation in prices is not random but rather correlated with the error term. Other types of sales promotions (e.g., features, displays) may also be set strategically.

### 2.7.2 Addressing Endogeneity Through Control Variables

Rather than right away jumping to more technical approaches to address endogeneity, it is essential to have a good understanding of the “easy” fixes (Rossi 2014; Germann et al. 2015). Perhaps the demand factors that managers use to set pricing and sales promotion may be measurable. If so they need to be included in the model. Examples include seasonal dummies (holidays, Christmas, Mother’s Day), temperature, and event indicators. Moreover, there are likely to be cross-sectional differences between brands leading to differences in marketing variables. For example, a high-quality, popular brand may be able to charge higher prices than a low-quality, unpopular brand. If the model does not include a brand intercept, the higher utility for the first brand will be attributed to its higher price—which is an incorrect inference. Hence fixed effects for brands (dummies for brands) are required to avoid biases due to parameter inferences based on cross-sectional differences. If this is not feasible, then quality ratings for brands (based on, e.g., consumer reports or online reviews) can be included in the model to make the previously unobservable quality differences observable, which means they are no longer part of the error term or a cause of endogeneity. Including these variables will help the researcher to obtain more correct estimates for sales promotion effects.

The Scan\*Pro model includes weekly dummies  $\mu_{bt}$  to capture market-wide marketing activities that are unobserved by the researcher such as coupon drops or advertising. It also includes store dummies to filter out any cross-store form of endogeneity (e.g., bigger stores charging lower prices). It estimates the model by brand, which means that cross-brand differences (which are potentially endogenous) are not reflected in parameter estimates. Other demand shocks can be captured by including variables for holidays, events and days of the week.

### 2.7.3 Addressing Endogeneity Through Instrumental Variables

If there is reason to believe that there is still a potential endogeneity problem, a different estimation method than Ordinary Least Squares is required. For aggregate data, the most common approach is Two-Stage Least Squares (2SLS). The idea is to isolate exogenous variation in the endogenous regressor of interest, and use this variable instead of the original endogeneity-plagued variable. 2SLS first predicts the endogenous variable by regressing it on instrumental variables (IVs) and the exogenous variables from the demand model. IVs are variables that are correlated with the endogenous variable but uncorrelated with the error term of the demand model. In the context of sales promotion models, it means that IVs are not correlated with the error term of demand but do explain variation in the endogenous variable of interest (e.g., price, feature, display). For example, ingredient costs are often a good instrument because costs do not directly drive sales yet they are correlated with consumer prices.

In 2SLS, the predicted endogenous variable is included in the focal demand equation instead of the original endogenous variable, and this is the second stage regression. This step ensures that only the exogenous variation in the endogenous variable is considered. The equation can then be estimated using OLS, yielding consistent estimates. The standard errors need to be corrected for the fact that the model includes an estimated regressor; standard packages such as SPSS or STATA do this automatically.

For discrete choice models, the control function approach *should be used rather than 2SLS* (Petrin and Train 2010). The method *also requires instrumental variables (IVs)*. The control function approach estimates a first-stage regression for the endogenous variable of interest, e.g., price, explained by the instrumental variable (s) and the exogenous variables in the model. The residual from this regression represents variation in price that cannot be explained by observables, hence represents the endogenous component of price. This residual is added to the model, e.g., in the choice utility function (Eq. 2.3) and is called a control function because it “controls” for the endogeneity of price. The original price variable is kept in the model; i.e., it is not replaced by the forecast from the first stage regression, as is done in Two-Stage Least Squares (2SLS). Next, the model that includes the control function is estimated with Maximum Likelihood. Because this step involves the estimated control function rather than the true value, standard errors from standard ML are incorrect. Bootstrapping or asymptotic formulas need to be used to obtain the correct standard errors (Petrin and Train 2010).

### 2.7.4 *Problems in Addressing Endogeneity*

The IVs for price or promotion need to be valid (uncorrelated with the error term) and sufficiently strong (have significant explanatory power in the first-stage regression). If the IVs are not valid, the estimator can become even more biased than before the endogeneity correction. If the IVs are not sufficiently strong, the estimator will have large standard errors, leading to insignificance and sometimes counterintuitive signs (Rossi 2014). In practice, these two requirements are hard to meet simultaneously. Strong IVs are often invalid because they may be correlated with current demand. For example, lagged prices will typically explain current prices well, but may also drive current demand due to post-promotion effects, and hence they are invalid IVs. Valid IVs are often weak, because they are too far removed from the endogenous regressor.

Another problem with endogeneity corrections is that they lead to worse model fit (in- and out of sample) of the sales promotion model, and hence predictive performance cannot be used to establish the success of an endogeneity correction (Ebbes et al. 2011). Hence, all in all we caution against a mechanical use of endogeneity correction estimation methods. Much thought needs to be given to whether there is actually a problem once all key control variables have been included.

### 2.7.5 *IV-Free Methods to Address Endogeneity*

IVs need to be strong (sufficiently correlated with the endogenous regressor) and valid (uncorrelated with the error term). If proper IVs cannot be found, instrument-free methods can be used. One approach is through Gaussian copulas (Park and Gupta 2012). Copulas are statistical functions that capture the joint distribution between two stochastic variables that each may have different types of marginal distributions. The approach directly models the correlation between the endogenous regressor ( $X$ ) and the error term ( $\epsilon$ ), and by doing so, it eliminates the endogeneity problem. Suppose the marginal distribution of the endogenous regressor is  $H(X)$  and the marginal distribution of the error term is  $G(\epsilon)$ . Their joint density is captured through a bivariate Gaussian copula (Park and Gupta 2012, Eq 2.6). This leads to a likelihood function that can be optimized with Maximum Likelihood to obtain consistent estimates of the model parameters.

However, there is also a simpler, equivalent way to estimate the model if the error term has a normal distribution (Park and Gupta 2012, Eq. 2.10). First, the researcher estimates the empirical cumulative distribution of the endogenous variable:  $H(X)$ . This essentially means sorting the observations from low to high, and then calculating for each observation the proportion of the observations that are less than or equal to the focal observation. Next, the researcher calculates the inverse standard normal CDF for each observation:  $\Phi^{-1}(H(X))$ . Finally, this term is

added to the demand equation of interest, where the original endogenous regressor (X) is kept as is, and this equation is estimated with OLS. Now X will be estimated consistently. Bootstrapping needs to be applied to obtain correct standard errors. A key requirement is that the endogenous variable has a non-normal distribution, which needs to be established first through a normality test.

Gaussian copulas were recently applied by Burmester et al. (2015) and Datta et al. (2015). The copula method by Park and Gupta (2012) can also be applied for discrete choice models (household-level data), for discrete endogenous regressors and for “slope endogeneity,” which is an endogeneity problem that arises when the manager decides on marketing actions based on the response parameter heterogeneity.

An alternative instrument-free method is offered by Ebbes et al. (2005). It uses “latent instrumental variables.” The idea is that underlying the distribution of the endogenous variables are latent discrete support points, which are the latent IVs. The method then splits the variation of the endogenous regressor into an exogenous part (captured by the discrete support points) and an endogenous part, which is assumed to be distributed bivariate normal together with the error term of the demand equation.

### ***2.7.6 VAR Models and Endogeneity***

Finally, we note that VAR models and other vector-based time series models (VARX, VEC) treat multiple variables as endogenous. For example, sales, price and promotion may be part of a three-variate vector that constitutes the dependent (endogenous) variable in these models. These models allow for the measurement of feedback effects of sales on price and promotion and for inertia (e.g., price and promotion depending on own lagged values). However, treating variables as endogenous is not the same as correcting for an endogeneity bias. The VAR model for example does not try to infer a causal effect of one variable on another. In a VAR model, the immediate effect is captured through the covariance matrix of the error term, which means that the effect is bidirectional (Y1 affecting Y2 as much as Y2 affecting Y1). In other words, there is nothing in a VAR model that tries to correct for endogeneity bias when inferring the effect of one variable on another.

## **2.8 Promotions to the Trade—Introduction**

Manufacturers use promotional discounts to the trade as an incentive for the trade to buy more of the brand, and to sell more to consumers by passing through at least part of the discount. The key two phenomena that determine the effectiveness of trade promotions are forward buying (Sect. 2.8) and pass-through (Sect. 2.9). In

Sect. 2.11 we present decision models for manufacturers who want to optimize their trade promotions, and in Sect. 2.12 we discuss models for retailers who want to optimize pass-through and forward-buying.

## 2.9 Forward Buying

Trade promotions offered by manufacturers often lead to forward buying by retailers. Forward buying is essentially purchase acceleration by retailers in that retailers purchase during the promotion period to satisfy demand in future periods (Neslin 2002, p. 36). While a retailer may sell part of the extra stock to consumers at a discount, their key incentive to forward buy is to sell the other part at regular price. We show an example in Fig. 2.3a, b.

Suppose a manufacturer offers a trade deal in period  $t - 1$  in order to stimulate a retailer to promote in period  $t$ . Figure 2.3a shows how a retailer may order higher quantities in period  $t - 1$  than usually to benefit from the lower wholesale price offered by the manufacturer in period  $t$ .<sup>9</sup> Forward buying implies that the stock bought in period  $t - 1$  not only satisfies the extra demand during the consumer promotion in period  $t$  (see Fig. 2.3b), but also demand sold at regular price in period  $t + 1$  and beyond. Figure 2.3b shows that not only the retailer forward buys but also consumers: the bump in period  $t$  is followed by a postpromotion dip in period  $t + 1$ .

To measure the effectiveness and profitability of trade promotions, Blattberg and Levin (1987) model the interplay between factory shipments, retailer promotions, retailer sales, and retailer inventories:

$$\text{FactoryShipments}_t = f_1(\text{Inventories}_{t-1}, \text{Trade Promotions}_t) \quad (2.17)$$

$$\text{Consumer Promotions}_t = f_2(\text{Trade Promotions}_t, \text{Trade Promotions}_{t-1}, \text{Inventories}_{t-1}) \quad (2.18)$$

$$\text{Retailer Sales}_t = f_3(\text{Consumer Promotions}_{t-1}) \quad (2.19)$$

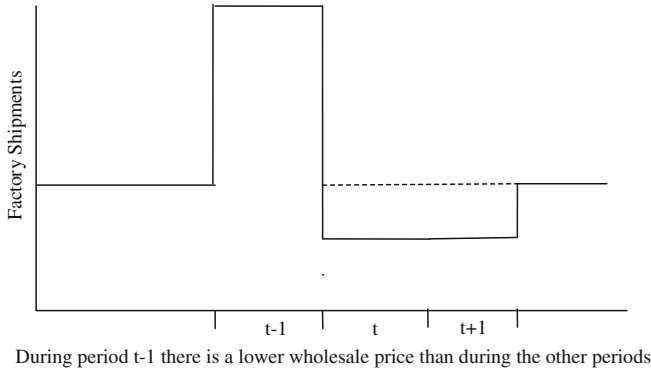
$$\text{Inventories}_t = f_4(\text{Inventories}_{t-1}, \text{FactoryShipments}_t, \text{Retailer Sales}_{t-1}) \quad (2.20)$$

Equation (2.17) captures the effect of trade promotions and inventories on factory shipments, i.e., how much the retailer orders (and therefore gets shipped). The willingness of retailers to run a consumer promotion depends on the availability of trade promotions and its own inventories (Eq. 2.18). Retail sales are driven by

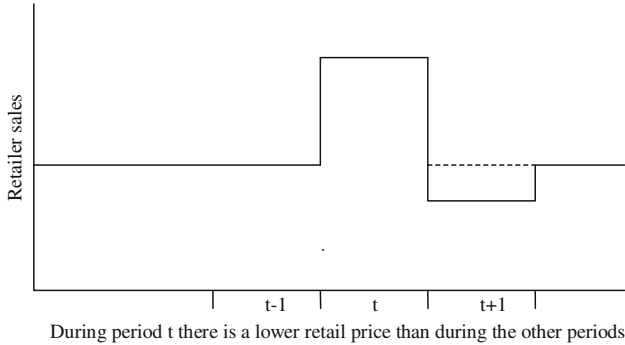
---

<sup>9</sup>“Factory Shipments” are shipments from the manufacturer to the retailer and reflect retailer orders or manufacturer sales. Blattberg and Levin (1987) use factory shipments data. Abraham and Lodish (1987) note their method can be applied to factory shipment data as well as other data such as warehouse withdrawals.

(a)



(b)



**Fig. 2.3** **a** Response of factory shipments to trade promotion, **b** response of retailer sales to consumer promotion

consumer promotions (Eq. 2.19), and Eq. (2.20) shows that retail inventories are a function of its own lag, inflow (Factory shipments) and outflow (Retailer sales).

Another approach to evaluate the effectiveness of the trade promotion is to estimate what factory shipment “sales” (Eq. 2.17) would have been in the absence of the promotion (Abraham and Lodish 1987). Once we estimate baseline sales, the size of the bump can be quantified as the actual factory shipments in the period with the wholesale discount (period  $t - 1$  in Fig. 2.3a) minus predicted baseline factory shipments for the same period. To estimate baseline factory shipments, Abraham and Lodish (1987) develop PROMOTER, a time series approach that tries to identify a “base” sales level by extrapolating the sales level during “normal” periods



(see Abraham and Lodish 1993 for an application to retail promotions). The PROMOTER model decomposes sales into three components:

$$S_t = B_t + P_t + E_t \quad (2.21)$$

where

- $S_t$  Factory shipments at time  $t$
- $B_t$  Baseline at time  $t$  to be estimated
- $P_t$  Promotion effect at time  $t$  if any
- $E_t$  Noise term

## 2.10 Pass-Through

The American Marketing Association defines pass-through as: “The number or percentage of sales promotion incentives offered to wholesalers or retailers by manufacturers that are extended to consumers by those channel members” (American Marketing Association 2015). For trade deals, it is the percentage of the discount that is passed on to the consumer in the form of a price reduction. Trade deals constitute close to 60% of manufacturers’ marketing budgets (Retail 2012). Manufacturers believe that only 52–66% of their trade spending is passed through to the consumer (Retail 2012). Though published numbers on pass-through range from 0 to 200%, pass-through is often less than 100% (Neslin 2002, p. 34).

Calculating pass-through fundamentally involves the relationship between retailer costs, which decrease when the retailer receives a trade deal, and retailer selling price. Questions have emerged as to how to calculate retailer costs. For example, should the researcher use “average acquisition cost” (AAC) or “transaction cost” (see Nijs et al. 2010)? Transaction cost is the unit cost the retailer paid for product in the current week. AAC can be calculated as a weighted average of the AAC of retailer inventory at the end of the previous week and cost of product the retailer purchased in the current week (Besanko et al. 2005, p. 129). The estimate of pass-through can differ depending on whether one uses AAC or transaction cost.

Table 2.9 shows an example. We assume normal wholesale price (transaction cost for the retailer) is \$4 per case; a trade deal lowers this to \$3 in period 4. The retailer starts with 400 cases in inventory and we assume purchases enough inventory each week exactly to cover demand. The table shows the cases purchased by the retailer, cases purchased by the consumer, and ending inventory. This is used to calculate AAC. For example, in period 3 AAC is  $(150 \times \$4 + (400 - 150) \times \$4)/(400) = \$4$ .<sup>10</sup> However, in period 4, AAC is  $(250 \times \$3 + (400 -$

<sup>10</sup>We use the formula provided on page 129 of Besanko et al. (2005):  $AAC(t) = [(Retailer\ Purchases\ in\ t) \times Wholesale\ price\ in\ period\ t] + [(Inventory\ at\ end\ of\ t - 1) - (Retail\ sales\ in\ t)] \times AAC(t - 1)] \times (Inventory\ at\ end\ of\ t)^{-1}$ .

**Table 2.9** Calculating pass-through for hypothetical data

Period	Wholesale price (transaction cost)	Starting inventory	Cases bought by retailer	Cases bought by consumer	Ending inventory	Average acquisition cost (AAC)	Retail price
1	\$4	400	150	150	400	\$4	\$5
2	\$4	400	150	150	400	\$4	\$5
3	\$4	400	150	150	400	\$4	\$5
4	\$3	400	250	250	400	\$3.38	\$4.50
5	\$4	400	150	150	400	\$3.61	\$5
6	\$4	400	150	150	400	\$3.76	\$5
7	\$4	400	150	150	400	\$3.85	\$5
8	\$4	400	150	150	400	\$3.90	\$5

$250) \times \$4)/(400) = \$3.38$ . So in period 3, the transaction cost is \$3 while the average AAC is \$3.38. We assume the retailer lowers the retail price from \$5 to \$4.50 in period 4 when the trade deal is offered.

Following BDG, we calculate pass-through by regressing retail price versus AAC (see pages 128–129 of BDG). We obtain:

$$\widehat{\text{Retail\_Price}}_t = 2.57 + 0.62 \times \text{Acquisition\_Cost}_t$$

According to this linear model, a \$1 decrease in acquisition cost leads to a \$0.62 decrease in retail price. Hence the pass-through rate is 62%. However, from a transaction cost perspective, the \$1 reduction in transaction cost resulted in a \$0.50 pass-through, or 50%. Nijs et al. (2010) assert that using AAC results in inflated estimates of pass-through. The reason is apparent in Table 2.9. AAC is essentially a moving average of previous AAC and current transaction costs. Therefore the full reduction of \$1 in period 4 is cushioned by previous AAC, so period 4 AAC decreases to \$3.38, not \$3. The regression sees a retail price of \$4.50 associated with a cost of \$3.38, so infers that the retail price reduction per dollar cost reduction is larger than if the transaction cost of \$3 were used in period 4. Our purpose is not to adjudicate whether transaction costs or AAC is more appropriate. Our point is that the calculation of relevant costs is crucial for estimating the pass-through rate. The above is a simple example, and does not include carrying costs, forward buying, and other factors that can make cost accounting of inventory a challenge.

The above treats trade deals as discrete incentives applied to a focal product. This may be appropriate for off invoice or “scanback” trade deals (see Gómez et al. 2007 for descriptions). However, much trade promotion spending involves lump sum payments for activities such as advertising programs. The retailer may use lump sum payments in various ways for different SKUs at different times. It thus is difficult to map a particular trade deal payment to a particular retailer promotion action. Ailawadi and Harlam (2009) offer a fresh perspective on pass-through in this environment. They propose to measure pass-through by “adding up the promotion

funding, in all its forms, that the retailer receives from a manufacturer in a year and comparing it to the retailer's total spending on price promotions for the manufacturer's products" (p. 783). Using this approach, they find interesting results. For example, pass-through is over 100% in the two years they examined, meaning that the retailer spends more on its promotions than it receives in funding. However, this finding is across all manufacturers, even those who provide no funding yet receive promotions. The pass-through rate among manufacturers that provide funding was 20% (Table 2.3 in Ailawadi and Harlam).

Factors such as price elasticity of the promoted brand, retailer size, retail price, category importance to the retailer, and ability of the promoted brand to take away from others, have been shown to influence pass-through (Besanko et al. 2005; Pauwels 2007; Nijs et al. 2010; Ailawadi and Harlam 2009). One crucial question is the role of market share. Intuition suggests higher share brands should command higher pass-through. However, a simple retailer profit model shows why a high-share brand might not receive high pass-through. The argument is that higher share brands have higher baseline sales, which means the retailer sacrifices more margin by putting it on sale. Let:

- B baseline sales for promoted brand
- $B_0$  baseline sales for rest of category
- M profit margin for promoted brand
- $M_0$  profit margin for rest of category
- D trade deal discount
- $\delta$  pass through
- $\eta$  gain in sales for promoted brand

Then: Retailer Profit with no pass through	$B(M + D) + B_0M_0$
Retailer profit with $\delta$ pass through	$B(M + D - \delta) + \eta(M + D - \delta) + (B_0 - \eta)M_0$
Difference	$-\delta B + \eta(M + D - \delta) - \eta M_0$

A high share brand has higher B and this exerts a force for decreasing  $\delta$ . One could argue that higher share brands have higher  $\eta$ 's, but that isn't good if  $M_0 > M + D - \delta$ , which could be the case if high share brands have lower regular margins for the retailer (they are stronger brands). This leads the retailer to decrease  $\delta$  to minimize the baseline loss and make  $M + D - \delta$  larger than  $M_0$ . The manufacturer might argue that promoting its high share brand grows the category through store switching, so that the  $\eta$  incremental units for the brand do not come completely from other brands in the category. In any case, the above analysis shows the challenges the high share brand must overcome to obtain higher pass-through.

Recent studies suggest that despite the pass-through handicap discussed above, higher share brands do command higher pass-through. Although early work by Walters (1989) found market share had no impact on pass-through, more recent studies by BDG, Pauwels (2007), Ailawadi and Harlam (2009), and Nijs et al.

(2010) find that it does. An interesting avenue for future research would be to learn why this occurs despite the baseline handicap discussed above.

Another issue that has received attention is cross-brand pass-through. BDG find that retailers adjust the prices of other brands in the category when passing through a promotion for the focal brand. They find for example that trade deals for large brands are less likely than small brands to generate positive cross-brand pass-through, i.e., large brands do not induce the retailer to reduce the retail price of competing smaller products. In a critique of BDG, McAlister (2007) argues that BDG find so many significant coefficients for cross-brand pass-through because they inadvertently inflate the number of independent observations by a factor of 15 (15 is the number of price zones). When she corrects for this, she finds “the number of stable, significant coefficients for other brands’ wholesale prices is lower than one would expect by chance.” (p. 876). See Dubé and Gupta (2008) and Duan et al. (2011) for further discourse on this issue. It is also noteworthy that McAlister (2007) argues that the use of average acquisition cost can also distort estimates of own and cross-brand pass-through.

## Part II: Normative Models

In Part I we have focused on descriptive models for sales promotions, i.e. models that describe, analyze, and explain sales promotion phenomena. In Part II we discuss normative (decision) models for sales promotions, i.e., models that tell the decision maker what is the best (profit maximizing) decision on sales promotion activities. Section 2.10 focuses on models for promotions to consumers, whereas Sect. 2.11 zooms in on trade promotion models for manufacturers. Section 2.12 takes the perspective of a retailer who tries to optimize forward buying and pass-through in response to trade promotions offered by manufacturers.

## 2.11 Decision Models for Promotions to the Consumer

### 2.11.1 Retailer Promotion Optimization

Tellis and Zufryden (1995) formulate a model to maximize cumulative *retailer* category profits over a finite horizon, by optimizing the depth and timing of discounts, and order quantities, for multiple brands. The model is based on an integration of consumer decisions in purchase incidence, brand choice and quantity. The retailer profit objective function is given by:

$$\max_{\{Disc_{bt}, O_{bt}, \delta_{bt}, \xi_{bt}\}} \left\{ \sum_{b,t} (M \cdot S_{bt} \cdot (\text{Price}_{bt} m_{bt} - Disc_{bt})) - \sum_{b,t} (\xi_{bt} F_{bt} + h_{bt} \cdot (I_{bt} - I_{bt-1}) / 2 + \delta_{bt} Tag_{bt}) \right\}, \quad (2.22)$$

where

$Disc_{bt}$	retailer discount level for brand $b$ , during period $t$ ( $\geq 0$ )
$O_{bt}$	retailer order quantity for brand $b$ , made at beginning of period $t$ ( $\geq 0$ )
$\delta_{bt}$	integer price-change indicator (=1 if a price change was made for brand $b$ during $t$ relative to $t - 1$ ; 0 otherwise)
$\xi_{bt}$	integer order time indicator (=1 if an order for brand $b$ is placed during period $t$ ; 0 otherwise)
$M$	total household market size
$S_{bt}$	average sales of brand $b$ per customer during period $t$ , computed as a function of causal variables (including, $Disc_{bt}$ ), and obtained via models for incidence, brand choice, and quantity
$Price_{bt}$	regular price of brand $b$ during period $t$ ,
$m_{bt}$	regular retailer profit margin (excluding inventory costs) of brand $b$ during period $t$ ,
$F_{bt}$	fixed costs of ordering brand $b$ during period $t$
$h_{bt}$	cost per unit of holding inventory of brand $b$ during period $t$
$I_{bt}$	retailer inventory for brand $b$ during period $t$ (this depends on orders $O_{bt}$ , sales $S_{bt}$ , and market size $M$ )
$Tag_{bt}$	cost of retagging shelves if a price change of brand $b$ occurs during period $t$

The retailer profit function (Eq. 2.22) equals the profit margin before inventory costs less inventory cost for the product category. Inventory costs include the fixed costs of placing an order, the average cost of holding inventory, and the costs for changing retail price such as retagging shelves. This optimization includes constraints that ensure that (1) inventories are updated appropriately and (2) demand is always met (see Tellis and Zufryden 1995 for details).

Natter et al. (2007) present a decision support system for dynamic retail pricing and promotion planning. Their weekly demand model incorporates price, reference price effects, seasonality, article availability information, features and discounts. They quantify demand interdependencies (complements and substitutes) and integrate the net impact of these interdependencies into an optimal pricing model. The methodology was developed and implemented at BauMax, an Austrian Do-It-Yourself (DIY) retailer. Eight pricing rounds with thousands of different Stock Keeping Units (SKUs) served as a testing ground for the approach. The final marketing decision-support system implemented in rounds six through eight increased gross profit on average by 8.1% and sales by 2.1%, relative to predicted baselines.

Divakar et al. (2005) develop a sales forecasting model and decision support system that has been implemented at a major consumer packaged goods company. Managers are able to track forecast versus actual sales in a user-friendly “marketing dashboard” computing environment and drill down to understand the reasons for potential discrepancies. Based on that understanding, managers can adjust price and promotion accordingly. Divakar et al. report the company estimated that the DSS resulted in savings of \$11 million on an investment of less than \$1 million.

The authors emphasize the importance of organization “buy-in,” relevance, and diagnostics for successful real-world adoption of promotion models for decision-making.

An interesting promotion optimization issue is whether the manufacturer should strive for competing stores to alternate promoting its brand, or promote in the same week. The concern is that alternating weeks may allow consumers to switch stores and virtually always buy the brand on deal. Guyt and Gijsbrechts (2014) examine this using a model of store, category, and brand choice. They examine the sales impact of “out-of-phase” (alternating) versus “in-phase” (same-week) promotions. They show how store switching plays a key role in determining whether out-of-phase or in-phase promotion calendars are better for the manufacturer. These conclusions come from simulations, not formal optimization. Formal optimization would be a promising next step. This is a very real issue for manufacturers negotiating the timing of trade deal pass-through with retailers.

### 2.11.2 Targeting Promotions to Consumers

Since the beginning of the “Age of Addressability” (Blattberg and Deighton 1991), companies have increased their ability to target promotions to individual customers. The promotion can be delivered via email, retailer apps, website customization, or “old fashioned” direct mail. From a modeling perspective, the impetus for promotion targeting comes from parameter heterogeneity. Rossi et al. (1996) were the first to show how individual-level parameters could be used to devise customer-level promotions.

Zhang and Krishnamurthi (2004) provide a method for customizing promotions in online stores. Their approach provides recommendations on when to promote how much to whom. They take the perspective of a manufacturer who wants to optimize its expected gross profit from a household over three shopping trips with respect to a brand’s price promotions. The corresponding objective function is:

$$\max_{\{Disc_{bt+s}, s=0,1,2\}} \left\{ \sum_{s=0}^2 P(I_{t+s}^h = 1, C_{t+s}^h = b) \cdot E(Q_{bt+s}^h | I_{t+s}^h = 1, C_{t+s}^h = b) \cdot (m_{bt+s} - Disc_{bt+s}) \right\} \quad (2.23)$$

(all symbols have been defined previously; see Eqs. (2.1) and (2.22)). The optimization is subject to the constraint that discounts are nonnegative and that they do not exceed a fixed fraction of the regular price, to prevent brand equity erosion. The authors demonstrate that their approach may lead to much higher profits, especially because it prevents wasting money on price discounts that are too steep.

Zhang and Wedel (2009) followed up this work by studying targeted promotions in online versus offline stores. They estimated response models (incidence, choice, quantity) separately for the online and offline channels of a focal retailer. A key

difference between online and offline targeting is the authors assume that online redemption rate is 100% since the delivery would be through customer-level modification of the website. Redemption for offline promotions was assumed to be 15% because delivery would be through coupons made available at checkout. The authors also distinguished between “loyalty” promotions offered to customers who purchased the focal brand last time, and “competitive” promotions offered to customers who purchase a different brand last. The authors find that loyalty promotions were more profitable in online stores, while competitive promotions were more profitable in offline stores. The reason is the difference in customer behavior. Customers in the online store were strongly state dependent. This meant it was difficult to get them to switch brands and easier to get them to stick with the current brand, so loyalty promotions were more profitable. Customers in offline stores were less state dependent hence it was easier to get them to switch brands with a promotion.

Zhang and Wedel’s optimization used a three-period “rolling horizon”. The advantage of this approach is that decisions are made on a periodic basis using the exact data from the customer’s recent purchase history, not their estimated data (see also Neslin et al. 2009). However, the optimization has to be re-run each decision period. Another approach is dynamic programming. Dynamic programming produces a policy function that specifies, given the “state variables” that characterize the customer, whether the firm should or should not offer a promotion at a particular time to that customer. Khan et al. (2009) used finite horizon dynamic programming to determine when to offer which customers free shipping or coupons for an online grocery retailer. Neslin et al. (2013) used an infinite horizon dynamic program to determine when to offer which customers email or direct mail promotions for a meal preparation service.

There are several modeling issues to resolve for optimal promotion targeting. These include the optimization method as well as the particulars of the response model. One implementation issue is what should be the level of customization—optimize the timing but do not consider cross-sectional heterogeneity in response (mass targeting), optimize timing and consider segment-level heterogeneity, and optimize timing and consider customer-level heterogeneity. If model estimates are precise and computation time is not an issue, customer-level targeting is preferred. However, in real world applications, customer-specific parameters may be imprecise, and running optimization at the customer level can be computationally difficult. Zhang and Wedel (2009) as well as Khan et al. (2009) investigated this issue. Khan et al. found that in comparison to baseline, targeting based on timing alone increased profits by 7.8%, targeting based on timing at the segment level increased profits by 10.9%, and targeting based on timing at the individual level increased profits by 13.2%. Zhang and Wedel’s results are similar in that mass targeting does quite well, and there are decreasing returns to segment-level and then customer-level customization. Everything however depends on the response function—the nature of heterogeneity and the precision with which individual-level parameters can be measured. Future research is needed to decide what level of optimization is best.

The above research uses estimated response models and formal optimization to target promotions. Another approach is to generate a targeting policy without the benefit of modeling or optimization and to test the policy with a field experiment. The concern of course is the firm may be “leaving money on the table”. See Venkatesan and Farris (2012) for an analysis of a supermarket targeted coupon strategy, and Luo et al. (2014) for an analysis of mobile-delivered coupon.

## 2.12 Manufacturer Decision Models for Trade Promotions

Silva-Risso et al. (1999) present a decision support system that permits to search for a manufacturer’s optimal trade promotion calendar. By modeling the purchase incidence (timing), choice and quantity decisions they decompose total sales into incremental and non-incremental. The manufacturer’s objective function is given by:

$$\begin{aligned}
 \max_{\kappa_{bt}, FEAT_{bt}, DISP_{bt}} \left\{ \sum_{t=1}^T \rho^t \cdot M \cdot E(\Delta S_{bt}) \cdot (\text{Price}_{bt} \cdot (1 - \kappa_{bt} \cdot DSTEP) - MCOST_{bt}) \right. \\
 - \sum_{t=1}^T \rho^t \cdot M \cdot E(B_{bt}) \cdot (\text{Price}_{bt} \cdot \kappa_{bt} \cdot DSTEP) \\
 - \sum_{t=1}^T (\rho^t \cdot \delta_{bt} \cdot Tag_{bt} + FEAT_{bt} \cdot FCOST_{bt} + DISP_{bt} \cdot DCOST_{bt}) \\
 \left. + \sum_{t=T+1}^{T+13} \rho^t \cdot M \cdot E(\Delta S_{bt} |_{\text{no promotion}}) \cdot (\overline{\text{Price}_b} - \overline{MCOST_b}) \right\}
 \end{aligned} \tag{2.24}$$

where  $\delta_{bt}$  and  $Tag_{bt}$ , have been defined previously (Sect. 2.10), and

$\kappa_{bt}$	0, 1, 2, ..., 10. This is the discount multiplier for brand-size $b$ in week $t$ . If $\kappa_{bt}=0$ , brand-size $b$ is sold at the base price in week $t$ . When the manufacturer offers a discount, it is computed as a multiple of a discount step level, e.g., 5%
$\rho$	discount rate
$M$	average number of category consumers that shop in the store or chain
$E(\Delta S_{bt})$	expected number of incremental units of brand-size $b$ in week $t$ due to the promotion
$\text{Price}_{bt}$	wholesale base price of brand-size $b$ in week $t$
$DSTEP$	base discount step, e.g., 5%
$MCOST_{bt}$	manufacturer’s marginal cost of brand-size $b$ in week $t$
$E(B_{bt})$	expected number of baseline plus borrowed units of brand-size $b$ in week $t$



$FCOST_{bt}$	fixed cost charged by retailer to run a feature ad for brand-size $b$ in week $t$
$DCOST_{bt}$	fixed cost charged by retailer to set up a display for brand-size $b$ in week $t$
$E(\Delta S_{bt}  _{\text{no promotion}})$	expected number of incremental units of brand-size $b$ in week $t = T+1, \dots, T+13$ due to carry-over effects from promotions in the period $t = 1, \dots, T$
$\overline{\text{Price}}_b$	average wholesale base price of brand-size $b$
$\overline{MCOST}_b$	average manufacturer's marginal cost of brand-size $b$

The objective function has four components: (1) the expected contribution from incremental units, (2) the expected opportunity cost of selling at a discount to consumers who would have bought the brand at the regular price, (3) the fixed costs associated with promotion decisions, and (4) the carry-over effects from consumption and purchase feedback over a 13-week period subsequent to the planning horizon. The objective function is maximized subject to constraints on the minimal and maximal number of promotions. Furthermore, the retailer may insist on a minimum level of category profits.

Neslin et al. (1995) develop a model that optimizes trade promotions and advertising. Their dynamic optimization model considers the actions of manufacturers, retailers, and consumers. The manufacturer attempts to maximize its profits by advertising directly to consumers and offering periodic trade deal discounts in the hope that the retailer will in turn pass through a retailer promotion to the consumer. Neslin et al. (1995) specify a multi-equation model for the retailer order and pass-through decisions, and for the effects of advertising and promotion on aggregate consumer demand. One of their findings is an intrinsic negative relationship between optimal levels of advertising and promotion. Higher advertising begets higher baseline sales, which increases the cost of promotion in the form of lost margin on baseline sales. Higher levels of promotion erode margin, thereby decreasing the incremental contribution of advertising. The result is that forces that tend to increase promotion tend to decrease advertising. This could be offset if advertising enhances consumer response to promotions, but the point is, *absent this interaction*, this research suggests optimal promotion and advertising expenditures are negatively related.

### 2.13 Retailer Decision Models for Forward Buying and Pass-Through

Retailers may benefit from trade promotions by forward buying. Blattberg and Neslin (1990, p. 460) derive the optimal amount of forward buying:

$$W^* = \frac{52 \cdot (G \cdot P - HC)}{(PC \cdot P \cdot CC + 13 \cdot SC)} \quad (2.25)$$

where

- $W^*$  optimal number of weeks supply to forward buy
- $G$  increase in profit margin per case from trade deal (trade deal discount in dollars)
- $P$  number of cases per pallet
- $HC$  handling cost per pallet
- $PC$  purchase cost per case (regular wholesale price –  $G$ )
- $CC$  cost of capital (borrowing costs to finance the forward buy)
- $SC$  storage cost per pallet per month

Equation (2.25) shows the rich interplay among variables that determine the amount of forward buying. For example, a larger trade deal discount ( $G$ ) directly encourages more forward buying. It also decreases the retailer's purchase cost ( $PC$ ), which decreases the retailer's financing costs ( $PC \cdot P \cdot CC$ ). This encourages even more forward buying. Equation (2.25) also shows that all else equal, larger costs of capital ( $CC$ ), storage ( $SC$ ), and handling ( $HC$ ) discourage forward buying.

Equation (2.25) is an inventory management model that does not take into account the increase in demand that results from the retailer passing through a portion of the trade deal discount to decrease retail price. There have been a few studies that have derived the optimal level of pass-through for a retailer. While not taking into account all the inventory factors in Eq. (2.25), Tyagi (1999) found that pass-through depends on the following function:

$$\varphi = \frac{S(\text{Price}^*)S''(\text{Price}^*)}{S'(\text{Price}^*)}, \quad (2.26)$$

where  $S$  is the demand function at the retail level,  $\text{Price}^*$  is the optimal retail price, and the primes stand for the first or second derivatives of the demand function. Specifically, if  $\varphi < 1$ , retailer pass-through is less than 100%; if  $\varphi = 1$ , retailer pass-through is 100%, and if  $\varphi > 1$ , retailer pass-through is greater than 100%. Tyagi (1999) shows that for the linear and all concave consumer demand functions, optimal pass-through is less than 100%. However, for commonly used demand functions such as the constant elasticity demand function (e.g., the Scan\*Pro model from Sect. 2.4.1), a rational retailer engages in greater than 100% pass-through. Moorthy (2005) generalizes Tyagi's formulation in several directions. First, besides whole price changes, Moorthy (2005) considers other cost components as well, such as inventory and labor costs. Second, Moorthy (2005) considers multiple retailers and multiple brands. Moorthy finds for example that cross-brand pass-through may be optimal, i.e., when decreasing the price of Brand A, it may be optimal to increase the price of Brand B.

As noted earlier, off-invoice trade promotions imply that retailers obtain a discount on the wholesale price for every unit *bought* on promotion. However, manufacturers often lose money on these deals as a result of forward-buying by retailers (Drèze and Bell 2003). Current trade promotion practice often shuns off-invoice trade deals in favor of trade deals that compensate retailers based on how much they *sell*, not *buy* (see Gomez et al. 2007). An example is the scanback deal, which gives retailers a discount on units *sold* during the promotion. Drèze and Bell (2003) develop a theory to compare retailer pricing decisions and profitability under scan-back and traditional off-invoice trade deals. They derive that for a given set of deal parameters (regular price, deal size, and deal duration), the retailer always prefers an off-invoice deal (because of the benefits of forward buying), whereas the manufacturer always prefers a scan-back. However, manufacturers can redesign the scanback to replicate the retailer profits generated by the off-invoice deal. The redesign makes the retailer indifferent between the off-invoice and the scan-back and makes the manufacturer strictly better off. The benefit of scan-back deals for retailers is that they economize on excess inventory costs, since scan-back deals do not lead to forward buying. For a redesigned scan-back in which the deal length is the same as off-invoice, but the deal depth is increased, consumer demand increases.

### Part III: Conclusions

This part concludes this chapter on sales promotion models. Section 2.13 presents a summary of the key empirical findings on sales promotion effectiveness. Next, Sect. 2.14 offers practical guidelines in model implementation, and Sect. 2.15 elaborates on avenues of further research in the sales promotion realm.

#### 2.14 Summary

This chapter has presented several models for the effects of sales promotions. In order to determine which of these models may be most relevant, we now summarize their key findings:

- Promotions to consumers lead to very strong sales promotion bumps in the short term. Hence it is essential that a model captures short-term effects.
- As a generalized finding across many categories and brands, brand switching effects expressed in unit sales are about 1/3 of the bump (Table 2.5), acceleration and deceleration effects are also 1/3 (Macé and Neslin 2004), and the remaining 1/3 is sometimes labeled “category expansion” (Van Heerde et al. 2004). Therefore it is important that any short-term model distinguishes at least among these main sources. See Table 2.2 for deeper insights of the sources for the short-term promotion bump.

- There are significant purchase-event feedback effects of promotions. In the first couple of weeks buying on promotion, a consumer's purchase behavior is affected by that promotion. Hence it is important to accommodate feedback effects in models (Sect. 2.6.1).
- Permanent effects of promotions on brand and category sales are rare (Sect. 2.6.4). That is, the effect of a promotion typically dies out after a number of weeks. Hence it may not be necessary to model these permanent effects.
- Two key factors that drive the profitability of trade promotions are pass-through and forward buying (Sects. 2.8 and 2.9). Incremental sales at retail are driven by pass-through combined with consumer response to promotions (Sect. 2.11). Any optimization model for trade dealing needs at least to include these phenomena.
- The rise of scanback deals may call for new models for pass-through (Sect. 2.12).

## 2.15 Practical Modeling Guidelines

In this section we provide a number of practical guidelines for building sales promotion models. Irrespective of whether the aim is to build a descriptive model (Part I) or normative model (Part II), it is important first to evaluate the available data. Do they match the modeling objective? If the goal is to learn about consumer heterogeneity, data at the consumer level are required. If the goal is to understand aggregate promotion effects, data at the store-level or at a higher aggregation level are sufficient. Once the data have been collected, it is important that the most important causal drivers of the performance measure of interest are available. In other words, the independent variables in the dataset should be able to explain a significant proportion of the variation in the dependent variable. A next step is to check descriptive statistics (means, variances, time series plots) and identify (and possibly correct or delete) outliers.

The subsequent step is to specify a descriptive model. A descriptive model may be the end goal in itself or constitute the building block of a normative model. We provide in Table 2.10 a number of descriptive models, with a few (admittedly subjective) pros and cons of each. As for the individual-level models, our view is that the minimum requirements include heterogeneity (see Sect. 2.2.5) and purchase-event feedback (Sect. 2.6.1). In aggregate regression- and time series models, it is important to include dynamic effects (Sect. 2.4.2). While aggregate logit models offer the benefits of (1) consistency with individual-level utility maximization and (2) parsimony in modeling cross-effects, they currently lack dynamic effects.

Software is increasingly available to estimate the models included in Table 2.10. There is no need to program maximum likelihood estimation for logit models, regression models, or Poisson models, as these are readily available in SPSS,

**Table 2.10** A selection of models for sales promotion effects

Model	Dependent variable	Model	Advantage	Disadvantage	Key studies
<i>Household-level data</i>					
Individual-level purchase behavior models	Brand choice	Multinomial logit model	<ul style="list-style-type: none"><li>• Consistent with utility theory</li></ul>	IIA-assumption (can be avoided by using a probit model)	<ul style="list-style-type: none"><li>• Guadagni and Little (1983)</li><li>• Gupta (1988)</li></ul>
	Purchase incidence	Binomial logit model	<ul style="list-style-type: none"><li>• Proper model</li></ul>	–	<ul style="list-style-type: none"><li>• Gupta (1988)</li></ul>
	Purchase quantity	Poisson model	<ul style="list-style-type: none"><li>• Easy to estimate</li></ul>	Mean-variance equality	Bucklin et al. (1998)
	Consumption	Consumption model	<ul style="list-style-type: none"><li>• Allows for flexible consumption</li></ul>		Ailawadi and Neslin (1998)
	Store choice	Multinomial logit model	<ul style="list-style-type: none"><li>• Consistent with utility theory</li></ul>	IIA-assumption	Bucklin and Lattin (1992)
	Category choice	Multivariate Probit	<ul style="list-style-type: none"><li>• Proper model</li></ul>		Manchanda et al. (1999)
	SKU choice	Multinomial logit model	<ul style="list-style-type: none"><li>• Parsimonious</li></ul>	IIA-assumption	Fader and Hardie (1996)
	Incidence, choice, and quantity	Dynamic structural model	<ul style="list-style-type: none"><li>• Complete, integrated</li></ul>	Difficult to estimate	Erdem et al. (2003), Sun (2005)
<i>Weekly store-level data</i>					
Baseline model	Brand sales to consumers or to retailers	Time series smoothing procedure to distinguish baseline and promotional sales	<ul style="list-style-type: none"><li>• Intuitively appealing</li><li>• Easy to implement</li></ul>	<ul style="list-style-type: none"><li>• Difficult to correct for all confounds</li><li>• No parameterized model for running policy simulations</li></ul>	<ul style="list-style-type: none"><li>• Abraham and Lodish (1987, 1993)</li></ul> (continued)

Table 2.10 (continued)

Model	Dependent variable	Model	Advantage	Disadvantage	Key studies
Regression model	Brand sales at weekly store level	Scan*Pro model: multiplicative regression model	<ul style="list-style-type: none"> <li>• Fits data well</li> <li>• Tested in many applications</li> </ul>	<ul style="list-style-type: none"> <li>• No dynamic effects</li> <li>• Many independent variables</li> </ul>	Witink et al. (1988)
		Scan*Pro model with lead and lagged effects	<ul style="list-style-type: none"> <li>• Allows to measure pre- and postpromotion dips</li> </ul>	<ul style="list-style-type: none"> <li>• Many independent variables</li> </ul>	Van Heerde et al. (2000)
		System of linear additive models	<ul style="list-style-type: none"> <li>• Gives decomposition effects</li> </ul>	<ul style="list-style-type: none"> <li>• Linearity assumption</li> <li>• Many independent variables</li> </ul>	Van Heerde et al. (2004)
Time series model	Brand sales at market level, or category sales at market level	Vector autoregressive model	<ul style="list-style-type: none"> <li>• Allows for distinction between long- and short-term effects</li> </ul>	<ul style="list-style-type: none"> <li>• Many parameters</li> </ul>	<ul style="list-style-type: none"> <li>• Nijs et al. (2001)</li> <li>• Steenkamp et al. (2005)</li> </ul>
	Brand sales at market level, or category sales at market level	Vector error correction model	<ul style="list-style-type: none"> <li>• Separate parameters for long- and short-term effects</li> </ul>		<ul style="list-style-type: none"> <li>• Fok et al. (2006)</li> </ul>
Aggregate logit model	Brand sales at store level	Aggregate logit model, derived from aggregating individual logit models	<ul style="list-style-type: none"> <li>• Consistent with utility theory</li> <li>• Parsimonious</li> </ul>	<ul style="list-style-type: none"> <li>• Hard to implement</li> <li>• No dynamic effects</li> <li>• Identification of heterogeneity</li> <li>• Specification of outside good</li> </ul>	<ul style="list-style-type: none"> <li>• Berry et al. (1995)</li> <li>• Dubé et al. (2002)</li> </ul>

EvIEWS, STATA, SAS, Limdep and other statistical packages. STATA also provides multinomial probit estimation. For time series models, EvIEWS, Stata and SAS/ETS are good choices. Some models (such as multivariate probit models, consumption models, dynamic structural models, aggregate logit models) are not (yet) available in the major commercial statistical programs. Custom programs are required to estimate these models, which can be accomplished in Gauss, Matlab, Stata, SAS, and the free statistical platform R (see Rossi et al. 2005 for several Bayesian models in R).

## 2.16 Future Research

We hope this chapter enables researchers to implement models that lead to more effective and profitable promotions. We also hope this chapter stimulates new research in areas that have not yet obtained sufficient attention. One such area is the effects of sales promotions for non-grocery products. While most of the models discussed in this chapter are for grocery products, it is unclear whether they are applicable to promotion effects for other items such as durables or services. Though some headway has been made (e.g., Van Heerde et al. (2005) present a promotion model for clothing stores), there is ample room for additional model development. We expect that deceleration effects are stronger in categories that are more expensive (per item) than grocery products. For example, consumers anticipate last-minute holiday deals, many consumers postpone purchasing clothes until the sales season starts (Van Heerde et al. 2005), and the same may apply to car and furniture purchases.

Promotions involve dynamic effects, whose effects are not yet fully captured by current models. For example, we lack optimal retailer models that take into account consumer learning and expectations. One could take the model of Sun et al. (2003) as a starting point and next optimize profit from the firm perspective. Another gap in the literature is store-level models that disentangle state dependence, reference prices, and purchase timing effects.

The decision models for consumer promotions and trade promotions are typically based on descriptive models of demand responses to promotions. However, these decision models tend to exclude competitor responses. If competitors respond to a player's "optimal promotion plan", the outcome may become suboptimal for this player. It seems worthwhile to develop decision models that explicitly account for competitive reactions.

While the literature provides many insights into the effects of price promotions, features and displays, relatively little is known about other promotion types. For example, it is not clear whether promotions that offer more value for the same price (20% extra volume, buy-one-get-one free) are more or less effective than equivalent price promotions (20% lower price, 50% lower price).

The field of online promotions is virtually untapped. For example: are the effects of emailed coupons different from the effects of traditional paper coupons? What

are the impacts of free shipping promotions or price discounts communicated in banner ads or on a retail website? Can we design an optimal contact model for email promotions for frequent shoppers that maximizes both retailer and manufacturer profit?

It seems that a key input that is difficult to collect is promotion activities and customer behavior for the *non-focal* store. This is in contrast to scanner data. Online stores and manufacturers have lots data as they pertain to their store. These can and are being used to tailor promotions, (see Zhang and Krishnamurthi 2004; Zhang and Wedel 2009) but ideally data will become available for non-focal stores. In short, promotion models have focused largely on frequently purchased products and estimated using scanner data. There are several issues that need to be resolved in that industry, and the digitized environment opens up an entirely new opportunity for promotions research.

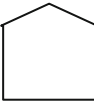
### **Appendix: Variable Definition for the Decomposition in Sect. 2.4.3**

Van Heerde et al. (2004) obtain decomposition (2.11) for price index variables with four types of support (with/without feature, with/without display). To achieve this, they transform the original four promotion variables (PI, FEATONLY, DISPONLY, FEAT&DISP) from the Scan\*Pro model into seven new variables: price index with feature-support (PF), price index with display-only support (PD), price index with feature and display support (PFD), price index without support (PWO), plus FWO (Feature without price cut), DWO (Display without price cut), and FDWO (Feature&Display without price cut). Regular price is indicated by a price index with value “1”. A 20% discount would be indicated by 0.8. The PWO, PF, PD, and PFD variables are defaulted to “1” if there is no price discount, but change depending on whether there is a discount and if so how it is supported. The FWO, DWO, and FDWO variables default to “0” and can equal “1” only if there is a feature, display, or both, without a price cut.

To illustrate the transformation, Table 2.11 contains the four original and seven new variables. In case # 1 there is no promotion, and the original price index (PI) equals 1 while the FEATONLY, DISPONLY, FEAT&DISP are zero, as defined in the Scan\*Pro model. Since there is no (supported or unsupported) price discount in case #1, the four new price index variables (PWO, PF, PD, PFD) are all at their nonpromotional value of 1. The FWO, DWO, and FDWO variables are zero since there is no feature or display without a price cut. In case # 2 there is a twenty percent price discount without any support, which shows up in the original variables as a price index of 0.8 while FEATONLY, DISPONLY, FEAT&DISP remain zero. Since this is a price cut without support, among the new price indices only the price index without support (PWO) variable is decreased to 0.8. The other three price indices PF, PD, and PFD stay at their nonpromotional level of 1, while the FWO, DWO, and FDWO variables stay at their default value of 0. Case # 3



**Table 2.11** Transforming the four Scan\*Pro variables into seven new variables

Case #	Scan*Pro variables					Seven new variables (Van Heerde et al. 2000, 2004)						
	Price index	Feature-only	Display-only	Feature and display		Price index without support	Price index with feature-only support	Price index with display only support	Price index with feature and display support	Feature-only without price cut	Display-only without price cut	Feature and display without price cut
	PI	FEAT	DISP	FEAT & DISP		PWO	PF	PD	PFD	FWO	DWO	FDWO
1	1	0	0	0		1	1	1	1	0	0	0
2	0.8	0	0	0		0.8	1	1	1	0	0	0
3	0.8	1	0	0		1	0.8	1	1	0	0	0
4	0.8	0	1	0		1	1	0.8	1	0	0	0
5	0.8	0	0	1		1	1	1	0.8	0	0	0
6	1	1	0	0		1	1	1	1	1	0	0
7	1	0	1	0		1	1	1	1	0	1	0
8	1	0	0	1		1	1	1	1	0	0	1

represents a twenty percent price cut with a feature-only, and hence price index with feature-only support (PF) is lowered to 0.8 (and again, all other variables are at their nonpromotional levels). Analogously, in cases # 4 and 5 the variables PD and PFD become 0.8, in turn. In case # 6 there is a feature without a price cut, which can be seen from the original variables since FEATONLY becomes 1 while PI remains 1. Consequently, among the new variables, FWO becomes 1, while DWO and FDWO remain 0, and PWO, PF, PD, and PFD stay 1 since there is no price cut. Cases #7 and 8 show how DWO and FDWO are defined.

Since price cuts tend to be communicated with feature and or display, the four Scan\*Pro Variables tend to be highly correlated. The seven new variables, in contrast, describe seven mutually exclusive promotion situations, and they tend to be much less correlated. While the seven new variables are larger in number than the four Scan\*Pro variables, a few of them typically do not vary and can therefore be excluded from models (especially the FDWO variable). Researchers who are concerned about multicollinearity in their (store-level) model may consider using the new set of seven variables proposed in this appendix.

## References

- Abraham, Magid M., and Leonard M. Lodish. 1987. Promoter: An automated promotion evaluation system. *Marketing Science* 6 (2): 101–123.
- Abraham, Magid M., and Leonard M. Lodish. 1993. An implemented system for improving promotion productivity using store scanner data. *Marketing Science* 12 (3): 248–269.
- Abramson, Charles, Rick L. Andrews, Imran S. Currim, and Morgan Jones. 2000. Parameter bias from unobserved effects in the multinomial logit model of consumer choice. *Journal of Marketing Research* 37 (4): 410–426.
- Ailawadi, Kusum L., and Bari A. Harlam. 2009. Findings—Retailer promotion pass-through: A measure, its magnitude, and its determinants. *Marketing Science* 28 (4): 782–791.
- Ailawadi, Kusum L., and Scott A. Neslin. 1998. The effect of promotion on consumption: Buying more and consuming it faster. *Journal of Marketing Research* 35 (August): 390–398.
- Ailawadi, Kusum L., Karen Gedenk, and Scott A. Neslin. 1999. Heterogeneity and purchase event feedback in choice models: An empirical analysis with implications for model building. *International Journal of Research in Marketing* 16 (3): 177–198.
- Ailawadi, Kusum L., Jacques Cesar, Bari Harlam, and David Trounce. 2007a. Quantifying and improving promotion profitability at CVS. *Marketing Science* 26 (4): 566–575.
- Ailawadi, Kusum L., Karen Gedenk, Christian Lutzky, and Scott A. Neslin. 2007b. Decomposition of the sales impact of promotion-induced stockpiling. *Journal of Marketing Research* 44 (3): 450–467.
- Ailawadi, Kusum L., Karen Gedenk, Tobias Langer, Yu Ma, and Scott A. Neslin. 2014. Consumer response to uncertain promotions: An empirical analysis of conditional rebates. *International Journal of Research in Marketing* 31 (1): 94–106.
- American Marketing Association. 2015. <https://www.ama.org/resources/Pages/Dictionary.aspx?dLetter=P>. Accessed 13 Aug 2015.
- Andrews, Rick L., Andrew Ainslie, and Imran S. Currim. 2002a. An empirical comparison of logit choice models with discrete versus continuous representations of heterogeneity. *Journal of Marketing Research* 39 (November): 479–487.

- Andrews, Rick L., Asim Ansari, and Imran S. Currim. 2002b. Hierarchical Bayes versus finite mixture conjoint analysis models: A comparison of fit, prediction, and partworth recovery. *Journal of Marketing Research* 39 (1): 87–98.
- Ansari, Asim, Carl F. Mela, and Scott A. Neslin. 2006. Customer channel migration. Working paper, Columbia Business School, New York.
- Ataman, Berk, Carl F. Mela, and Harald J. van Heerde. 2008. Building brands. *Marketing Science* 27 (November–December): 1036–1054.
- Ataman, Berk, Harald J. van Heerde, and Carl F. Mela. 2010. The long-term effect of marketing strategy on brand sales. *Journal of Marketing Research* 47 (5): 866–882.
- Bell, David R., Jeongweng Chiang, and V. Padmanabhan. 1999. The decomposition of promotional response: An empirical generalization. *Marketing Science* 18 (4): 504–526.
- Bell, David R., and Christian A.L. Hilber. 2006. An empirical test of the theory of sales: Do household storage constraints affect consumer and store behavior? *Quantitative Marketing and Economics* 4 (2): 87–117.
- Bell, David R., and James M. Lattin. 1998. “Shopping behavior and consumer preference for store price format: Why “large basket” shoppers prefer EDLP.” *Marketing Science* 17 (1): 66–88.
- Bell, David R., and James M. Lattin. 2000. Looking for loss aversion in scanner panel data: The confounding effect of price response heterogeneity. *Marketing Science* 19 (2): 185–200.
- Bell, David R., Teck-Hua Ho, and Christopher S. Tang. 1998. Determining where to shop: Fixed and variable costs of shopping. *Journal of Marketing Research* 35 (August): 352–369.
- Bell, David R., and James M. Lattin. 1998. “Shopping Behavior and Consumer Preference for Store Price Format: Why “Large Basket” Shoppers Prefer EDLP.” *Marketing Science* 17 (1): 66–88.
- Ben-Akiva, Moshe, and Steven R. Lerman. 1991. *Discrete choice analysis*. Cambridge, MA: MIT Press.
- Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* 63 (4): 841–890.
- Besanko, David, Sachin Gupta, and Dipak Jain. 1998. Logit demand estimation under competitive pricing behavior: An equilibrium framework. *Management Science* 44 (11, Part 1 of 2): 1533–1547.
- Besanko, David, Jean-Pierre Dubé, and Sachin Gupta. 2005. Own-brand and cross-brand retail pass-through. *Marketing Science* 24 (1): 123–137.
- Bijmolt, Tammo H.A., Harald J. Van Heerde, and Rik G.M. Pieters. 2005. New empirical generalizations on the determinants of price elasticity. *Journal of Marketing Research* 42 (May): 141–156.
- Blattberg, Robert C., and John Deighton. 1991. Interactive marketing: Exploiting the age of addressability. *Sloan Management Review* 33 (1): 21–32.
- Blattberg, Robert C., and Alan Levin. 1987. Modelling the effectiveness and profitability of trade promotions. *Marketing Science* 6 (2): 124–146.
- Blattberg, Robert C., and Scott A. Neslin. 1990. *Sales promotion, concepts, methods and strategies*. Englewood Cliffs, New Jersey: Prentice Hall.
- Blattberg, Robert C., Richard Briesch, and Edward J. Fox. 1995. How promotions work. *Marketing Science* 14 (3, Part 2 of 2): g122–g132.
- Bodapati, Anand V., and Sachin Gupta. 2004. The recoverability of segmentation structure from store-level aggregate data. *Journal of Marketing Research* 41 (August): 351–364.
- Bodapati, Anand V., and V. Srinivasan. 2006. The impact of feature advertising on customer store choice. Working paper, Stanford Graduate School of Business.
- Breugelmans, Els, and Katia Campo. 2011. Effectiveness of in-store displays in a virtual store environment. *Journal of Retailing* 87 (1): 75–89.
- Briesch, Richard A., Lakshman Krishnamurthi, Tridib Mazumdar, and S.P. Raj. 1997. A comparative analysis of reference price models. *Journal of Consumer Research* 24 (September): 202–214.
- Bucklin, Randolph E., and James M. Lattin. 1992. A model of product category competition among grocery retailers. *Journal of Retailing* 68 (3): 271–293.

- Bucklin, Randolph E., Sunil Gupta, and S. Siddarth. 1998. Determining segmentation in sales response across consumer purchase behaviors. *Journal of Marketing Research* 35 (May): 189–197.
- Burmester, Alexa B., Jan U. Becker, Harald J. van Heerde, and Michel Clement. 2015. Pre- and post-launch effects of publicity and advertising on sales. *International Journal of Research in Marketing*.
- Chan, Tat, Chakravarthi Narasimhan, and Qin Zhang. 2008. Decomposing promotional effects with a dynamic structural model of flexible consumption. *Journal of Marketing Research* 45 (August): 487–498.
- Chang, Kwangpil, S. Siddarth, and Charles B. Weinberg. 1999. The impact of heterogeneity in purchase timing and price responsiveness on estimates of sticker shock effects. *Marketing Science* 18 (2): 178–192.
- Chiang, Jeongwen. 1991. A simultaneous approach to the whether, what and how much to buy questions. *Marketing Science* 10 (4): 297–315.
- Chiang, Jeongwen, Siddhartha Chib, and Chakravarthi Narasimhan. 1999. Markov Chain Monte Carlo and models of consideration set and parameter heterogeneity. *Journal of Econometrics* 89 (1–2): 223–248.
- Chib, Siddhartha, P.B. Seetharaman, and Andrei Strijnev. 2004. Model of brand choice with a no-purchase option calibrated to scanner-panel data. *Journal of Marketing Research* 41 (May): 184–196.
- Chintagunta, Pradeep K. 1993. Investigating purchase incidence, brand choice, and purchase quantity decisions of households. *Marketing Science* 12 (2): 184–208.
- Chintagunta, Pradeep K. 2001. Endogeneity and heterogeneity in a probit demand model: Estimation using aggregate data. *Marketing Science* 20 (4): 442–456.
- Chintagunta, Pradeep K., Dipak C. Jain, and Naufel J. Vilcassim. 1991. “Investigating heterogeneity in brand preferences in logit models for panel data.” *Journal of Marketing Research* 28 (4): 417–428.
- Chintagunta, Pradeep K., and Jean-Pierre Dubé. 2005. Estimating a stockkeeping-unit-level brand choice model that combines household panel data and store data. *Journal of Marketing Research* 42 (August): 368–379.
- Datta, Hannes, Bram Foubert, and Harald J. van Heerde. 2015. The challenge of retaining customers acquired by free trials. *Journal of Marketing Research* 52 (April): 217–234.
- Dekimpe, Marnik G., and Dominique M. Hanssens. 1995. The persistence of marketing effects on sales. *Marketing Science* 14 (1): 1–21.
- DeVecchio, Devon, David H. Henard, and Traci H. Freling. 2006. the effect of sales promotion on post-promotion brand preference: A meta-analysis. *Journal of Retailing* 82 (3): 203–213.
- Divakar, Suresh, Brian T. Ratchford, and Venkatesh Shankar. 2005. CHAN4CAST: A multichannel multiregion forecasting model for consumer packaged goods. *Marketing Science* 24 (3): 333–350.
- Dodson, Joe A., Alice M. Tybout, and Brian Sternthal. 1978. Impact of deal and deal retraction on brand switching. *Journal of Marketing Research* 15 (May): 72–81.
- Drèze, Xavier, and David R. Bell. 2003. Creating win-win trade promotions: Theory and empirical analysis of scan-back trade deals. *Marketing Science* 22 (1): 16–39.
- Duan, Jason A., Leigh McAlister, and Shameek Sinha. 2011. Commentary-reexamining Bayesian model-comparison evidence of cross-brand pass-through. *Marketing Science* 30 (3): 550–561.
- Dubé, Jean-Pierre, Pradeep K. Chintagunta, Amil Petrin, Bart J. Bronnenberg, P.B. Ron Goettler, K.Sudhir Seetharaman, Raphael Thomadsen, and Ying Zhao. 2002. Structural applications of the discrete choice model. *Marketing Letters* 13 (3): 207–220.
- Dubé, Jean-Pierre, and Sachin Gupta. 2008. Cross-brand pass-through in supermarket pricing. *Marketing Science* 27 (3): 324–333.
- Eckert, Christine, Jordan J. Louviere, and Towhidul Islam. 2012. Seeing the forest despite the trees: Brand effects on choice uncertainty. *International Journal of Research in Marketing* 29 (3): 256–264.

- Ebbes, Peter, Michel Wedel, Ulf Böckenholt, and Ton Steerneman. 2005. Solving and testing for regressor-error (in)dependence when no instrumental variables are available: With new evidence for the effect of education on income. *Quantitative Marketing & Economics* 3 (4): 365–392.
- Ebbes, Peter, Dominik Papies, and Harald J. van Heerde. 2011. The sense and non-sense of holdout sample validation in the presence of endogeneity. *Marketing Science* 30 (6): 1115–1122.
- Erdem, Tülin, and Michael P. Keane. 1996. Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing Science* 15 (1): 1–20.
- Erdem, Tülin, Susumu Imai, and Michael P. Keane. 2003. Brand and quantity choice dynamics under price uncertainty. *Quantitative Marketing and Economics* 1 (1): 5–64.
- Erdem, Tülin, Glenn Mayhew, and Baohong Sun. 2001. Understanding reference-price shoppers: A within- and cross-category analysis. *Journal of Marketing Research* 38 (November): 445–457.
- Erdem, Tülin, Michael P. Keane, and Baohong Sun. 2008. A dynamic model of brand choice when price and advertising signal product quality. *Marketing Science* 27 (6): 1111–1125.
- Fader, Peter S., and Bruce G.S. Hardie. 1996. Modeling consumer choice among SKUs. *Journal of Marketing Research* 33 (November): 442–452.
- Fok, Dennis, Csilla Horváth, Richard Paap, and Philip Hans Franses. 2006. A hierarchical Bayes error correction model to explain dynamic effects of price changes. *Journal of Marketing Research* 43 (August): 443–461.
- Gedenk, Karen, and Scott A. Neslin. 1999. The role of retail promotion in determining future brand loyalty: Its effect on future purchase event feedback. *Journal of Retailing* 75 (4): 433–459.
- Germann, Frank, Peter Ebbes, and Rajdeep Grewal. 2015. The chief marketing officer matters! *Journal of Marketing* 79 (3): 1–22.
- Gómez, Miguel I., Vithala R. Rao, and Edward W. McLaughlin. 2007. Empirical analysis of budget and allocation of trade promotions in the U.S. supermarket industry. *Journal of Marketing Research* 44 (3): 410–424.
- Gönül, Füsün, and Kannan Srinivasan. 1996. Estimating the impact of consumer expectations on purchase behavior: A dynamic structural model. *Marketing Science* 15 (3): 262–279.
- Guadagni, Peter M., and John D.C. Little. 1983. A logit model calibrated on scanner data. *Marketing Science* 2 (3): 203–238.
- Gupta, Sunil. 1988. Impact of sales promotion on when, what and how much to buy. *Journal of Marketing Research* 25 (November): 342–355.
- Gupta, Sachin, and Pradeep K. Chintagunta. 1994. On using demographic variables to determine segment membership in logit mixture models. *Journal of Marketing Research* 31 (1): 128–136.
- Guyt, Jonne Y., and Els Gijsbrechts. 2014. Take turns or March in sync? The impact of the national brand promotion calendar on manufacturer and retailer performance. *Journal of Marketing Research* 51 (6): 753–772.
- Haans, Hans, and Els Gijsbrechts. 2011. One-deal-fits-all? On category sales promotion effectiveness in smaller versus larger supermarkets. *Journal of Retailing* 87 (4): 420–436.
- Ho, Teck-Hua, and Juin-Kuan Chong. 2003. A parsimonious model of stockkeeping-unit choice. *Journal of Marketing Research* 40 (August): 351–365.
- Hoch, Stephen J., Byung-Do Kim, Alan L. Montgomery, and Peter E. Rossi. 1995. Determinants of store-level price elasticity. *Journal of Marketing Research* 32 (1): 17–29.
- Hoch, Stephen J., Xavier Dreze, and Mary E. Purk. 1994. “EDLP, Hi-Lo, and Margin Arithmetic.” *Journal of Marketing* 58 (4): 16–27.
- Horsky, Dan, Sanjog Misra, and Paul Nelson. 2006. Observed and unobserved preference heterogeneity in brand-choice models. *Marketing Science* 25 (4): 322–335.
- Horváth, Csilla, and Philip Hans B.F. Franses. 2003. Deriving dynamic marketing effectiveness from econometric time series models. In *ERIM report series research in management*. Rotterdam: Erasmus Research Institute of Management.

- Jedidi, Kamel, Carl F. Mela, and Sunil Gupta. 1999. Managing advertising and promotion for long-run profitability. *Marketing Science* 18 (1): 1–22.
- Kahneman, Daniel, and Amos Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47 (March): 263–291.
- Kalyanaram, Gurumurthy, and Russell S. Winer. 1995. Empirical generalizations from reference price research. *Marketing Science* 14 (3): G161–G169.
- Kamakura, Wagner A., and Woosong Kang. 2007. Chain-wide and store-level analysis for cross-category management. *Journal of Retailing* 83 (2): 159–170.
- Kamakura, Wagner A., and Gary J. Russell. 1989. A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research* 26 (November): 379–390.
- Kantar Retail. 2012. *2012 trade promotion industry benchmarking study*. Wilton, CT: Kantar Retail.
- Keane, Michael P. 1997a. Modeling heterogeneity and state dependence in consumer choice behavior. *Journal of Business and Economic Statistics* 15 (3): 310–327.
- Keane, Michael P. 1997b. Current issues in discrete choice modeling. *Marketing Letters* 8 (3): 307–322.
- Khan, Romana, Michael Lewis, and Vishal Singh. 2009. customer management and the value of one-to-one marketing. *Marketing Science* 28 (6): 1073–1079.
- Krishna, Aradhna, Imran S. Currim, and Robert W. Shoemaker. 1991. Consumer perceptions of promotional activity. *Journal of Marketing* 55: 4–16.
- Krishnamurthi, Lakshman, and S.P. Raj. 1988. A model of brand choice and purchase quantity price sensitivities. *Marketing Science* 7 (1): 1–20.
- Lam, Shun Yin, Mark Vandebosch, John Hulland, and Michael Pearce. 2001. Evaluating promotions in shopping environments: Decomposing sales response into attraction, conversion, and spending effects. *Marketing Science* 20 (2): 194–215.
- Leeflang, Peter S.H. 2008. Modeling competitive responsiveness. In *Chapter 7 in handbook of marketing decision models*, ed. Berend Wierenga, 211–251. International Series in Operational Research & Management Science. New York (USA): Springer Science and Business Media.
- Leeflang, Peter S.H., and Josefa Parreña Selva. 2012. Cross-category demand effects of price promotions. *Journal of the Academy of Marketing Science* 40 (4): 572–586.
- Leeflang, Peter S.H., and Dick R. Wittink. 1992. Diagnosing competitive reactions using (aggregated) scanner data. *International Journal of Research in Marketing* 9 (1): 39–57.
- Leeflang, Peter S.H., and Dick R. Wittink. 1996. Competitive reaction versus consumer response: Do managers overreact? *International Journal of Research in Marketing* 13 (2): 103–120.
- Leeflang, Peter S.H., Josefa Parreña Selva, Albert van Dijk, and Dick R. Wittink. 2008. Decomposing the sales promotion bump accounting for cross-category effects. *International Journal of Research in Marketing* 25 (3): 201–214.
- Liu, Qiang, Thomas J. Steenburgh, and Sachin Gupta. 2015. The cross attributes flexible substitution logit: Uncovering category expansion and share impacts of marketing instruments. *Marketing Science* 34 (1): 144–159.
- Lourenço, Carlos J.S., Els Gijsbrechts, and Richard Paap. 2015. The impact of category prices on store price image formation: An empirical analysis. *Journal of Marketing Research* 52 (2): 200–216.
- Luo, Xueming, Michelle Andrews, Zheng Fang, and Chee Wei Phang. 2014. Mobile targeting. *Management Science* 60 (7): 1738–1756.
- Macé, Sandrine, and Scott A. Neslin. 2004. The determinants of pre- and postpromotion dips in sales of frequently purchased goods. *Journal of Marketing Research* 41 (August): 339–350.
- Manchanda, Puneet, Asim Ansari, and Sunil Gupta. 1999. The “shopping basket”: A model for multicategory purchase incidence decisions. *Marketing Science* 18 (2): 95–114.
- McAlister, Leigh. 2007. Comment—cross-brand pass-through: Fact or artifact? *Marketing Science* 26 (6): 876–898.

- Mehta, Nitin, and Yu Ma. 2012. A multicategory model of consumers' purchase incidence, quantity, and brand choice decisions: Methodological issues and implications on promotional decisions. *Journal of Marketing Research* 49 (August): 435–451.
- Mela, Carl F., Sunil Gupta, and Donald R. Lehmann. 1997. The long-term impact of promotion and advertising on consumer brand choice. *Journal of Marketing Research* 34 (May): 248–261.
- Mela, Carl F., Kamel Jedidi, and Douglas Bowman. 1998. The long-term impact of promotions on consumer stockpiling behavior. *Journal of Marketing Research* 35 (May): 250–262.
- Montgomery, Alan L. 1997. Creating micro-marketing pricing strategies using supermarket scanner data. *Marketing Science* 16 (4): 315–337.
- Moorthy, Shridhar. 2005. A general theory of pass-through in channels with category management and retail competition. *Marketing Science* 24 (1): 110–122.
- Mulhern, Francis J., and Robert P. Leone. 1991. Implicit price bundling of retail products: A multiproduct approach to maximizing store profitability. *Journal of Marketing* 55 (October): 63–76.
- Nair, Harikesh, Jean-Pierre Dubé, and Pradeep Chintagunta. 2005. Accounting for primary and secondary demand effects with aggregate data. *Marketing Science* 24 (3): 444–460.
- Narasimhan, Chakravathi, Scott A. Neslin, and Subrata K. Sen. 1996. promotional elasticities and category characteristics. *Journal of Marketing* 60 (April): 17–30.
- Natter, Martin, Andreas Mild, Thomas Reutterer, and Alfred Taudes. 2007. An assortment-wide decision-support system for dynamic pricing and promotion planning. *Marketing Science* 26 (4): 576–583.
- Neslin, Scott A. 2002. *Sales promotion*. Relevant knowledge series. Cambridge, MA: Marketing Science Institute.
- Neslin, Scott A., Stephen G. Powell, and Linda G. Schneider Stone. 1995. The effects of retailer and consumer response on optimal manufacturer advertising and trade promotion strategies. *Management Science* 41 (5): 749–766.
- Neslin, Scott A., Thomas P. Novak, Kenneth R. Baker, and Donna L. Hoffman. 2009. An optimal contact model for maximizing online panel response rates. *Management Science* 55 (5): 727–737.
- Neslin, Scott A., Gail Ayala Taylor, Kimberly D. Grantham, and Kimberly R. McNeil. 2013. Overcoming the 'recency trap' in customer relationship management. *Journal of the Academy of Marketing Science* 41 (3): 320–337.
- Nevo, Aviv. 2000. A practitioner's guide to estimation of random-coefficient logit models of demand. *Journal of Economics & Management Strategy* 9 (4): 513–548.
- Nijs, Vincent, Marnik G. Dekimpe, Jan Benedict Steenkamp, and Dominique M. Hanssens. 2001. The category demand effects of price promotions. *Marketing Science* 20 (1): 1–22.
- Nijs, Vincent, Kanishka Misra, Eric T. Anderson, Karsten Hansen, and Lakshman Krishnamurthi. 2010. Channel pass-through of trade promotions. *Marketing Science* 29 (2): 250–267.
- Park, Sungho, and Sachin Gupta. 2009. Simulated maximum likelihood estimator for the random coefficient logit model using aggregate data. *Journal of Marketing Research* 46 (August): 531–542.
- Park, Sungho, and Sachin Gupta. 2011. A regime-switching model of cyclical category buying. *Marketing Science* 30 (3): 469–480.
- Park, Sungho, and Sachin Gupta. 2012. Handling endogenous regressors by joint estimation using copulas. *Marketing Science* 31 (4): 567–586.
- Pauwels, Koen. 2007. How retailer and competitor decisions drive the long-term effectiveness of manufacturer promotions for fast moving consumer goods. *Journal of Retailing* 83 (3): 297–308.
- Pauwels, Koen, Dominique M. Hanssens, and S. Siddarth. 2002. The long-term effects of price promotions on category incidence, brand choice and purchase quantity. *Journal of Marketing Research* 39 (November): 421–439.
- Petrin, Amil, and Kenneth Train. 2010. A control function approach to endogeneity in consumer choice models. *Journal of Marketing Research* 47 (1): 3–13.

- Rhee, Hongjai, and David R. Bell. 2002. The inter-store mobility of supermarket shoppers. *Journal of Retailing* 78 (4): 225–237.
- Roederkerk, Robert P., Harald J. van Heerde, and Tammo H.A. Bijmolt. 2011. Incorporating context effects into a choice model. *Journal of Marketing Research* 48 (August): 767–780.
- Roederkerk, Robert P., Harald J. van Heerde, and Tammo H.A. Bijmolt. 2013. Optimizing retail assortments. *Marketing Science* 32 (5): 699–715.
- Rossi, Peter. 2014. Even the rich can make themselves poor: A critical examination of IV methods in marketing applications. *Marketing Science* 33 (5): 655–672.
- Rossi, Peter E., Robert E. McCulloch, and Greg M. Allenby. 1996. The value of purchase history data in target marketing. *Marketing Science* 15 (4): 321–340.
- Rossi, Peter E., Greg M. Allenby, and Robert McCulloch. 2005. *Bayesian statistics and marketing*. Chichester, West Sussex, England: Wiley.
- Rothschild, Michael L., and William C. Gaidis. 1981. Behavioral learning theory: Its relevance to marketing and promotions. *Journal of Marketing* 45 (2): 70–78.
- Roy, Rishin, Pradeep K. Chintagunta, and Sudeep Haldar. 1996. A framework for investigating habits, ‘the hand of the past,’ and heterogeneity in dynamic brand choice. *Marketing Science* 15 (3): 280–299.
- Seetharaman, P.B. 2004. Modeling multiple sources of state dependence in random utility models: A distributed lag approach. *Marketing Science* 23 (2): 263–271.
- Seetharaman, P.B., Andrew Ainslie, and Pradeep K. Chintagunta. 1999. Investigating household state dependence effects across categories. *Journal of Marketing Research* 36 (4): 488–500.
- Shin, Sangwoo, Sanjog Misra, and Dan Horsky. 2012. Disentangling preferences and learning in brand choice models. *Marketing Science* 31 (1): 115–137.
- Siddarth, S., Randolph E. Bucklin, and Donald G. Morrison. 1995. Making the cut: Modeling and analyzing choice set restriction in scanner panel data. *Journal of Marketing Research* 32 (3): 255–266.
- Silva-Risso, Jorge M., Randolph E. Bucklin, and Donald G. Morrison. 1999. A decision support system for planning manufacturers’ sales promotion calendars. *Marketing Science* 18 (3): 274–300.
- Singh, Vishal, Karsten Hansen, and Robert Blattberg. 2006. Market entry and consumer behavior: An investigation of a Wal-Mart supercenter. *Marketing Science* 25 (2): 457–476.
- Slotegraaf, Rebecca J., and Koen Pauwels. 2008. The impact of brand equity and innovation on the long-term effectiveness of promotions. *Journal of Marketing Research* 45 (3): 293–306.
- Steenburgh, Thomas J. 2008. The invariant proportion of substitution property (IPS) of discrete-choice models. *Marketing Science* 27 (2): 300–307.
- Steenkamp, Jan Benedict E.M., Vincent R. Nijs, Dominique M. Hanssens, and Marnik G. Dekimpe. 2005. Competitive reactions to advertising and promotion attacks. *Marketing Science* 24 (1): 35–54.
- Sun, Baohong. 2005. Promotion effects on endogenous consumption. *Marketing Science* 24 (3): 430–443.
- Sun, Baohong, Scott A. Neslin, and Kannan Srinivasan. 2003. Measuring the impact of promotions on brand switching when consumers are forward looking. *Journal of Marketing Research* 40 (November): 389–405.
- Tellis, Gerard J., and Fred S. Zufryden. 1995. Tackling the retailer decision maze: Which brands to discount, how much, when and why? *Marketing Science* 14 (3, Part 1 of 2): 271–299.
- Train, Kenneth. 2003. *Discrete choice methods with simulation*. Cambridge, UK: Cambridge University Press. <http://elsa.berkeley.edu/books/choice2.html>.
- Tyagi, Rajeev K. 1999. A characterization of retailer response to manufacturer trade deals. *Journal of Marketing Research* 36 (November): 510–516.
- Van Heerde, Harald J., and Tammo HA Bijmolt. 2005. “Decomposing the promotional revenue bump for loyalty program members versus non-members.” *Journal of Marketing Research* 42 (November): 443–457.
- Van Heerde, Harald J., Sachin Gupta. 2006. The origin of demand: A system to classify the sources of the sales promotion bump. Working paper, University of Waikato.



- Van Heerde, Harald J., Peter S.H. Leeflang, and Dick R. Wittink. 2000. The estimation of pre- and postpromotion dips with store-level scanner data. *Journal of Marketing Research* 37 (August): 383–395.
- Van Heerde, Harald J., Peter S.H. Leeflang, and Dick R. Wittink. 2001. Semiparametric analysis to estimate the deal effect curve. *Journal of Marketing Research* 38 (May): 197–215.
- Van Heerde, Harald J., Sachin Gupta, and Dick R. Wittink. 2003. Is 75% of the sales promotion bump due to brand switching? No, only 33% is. *Journal of Marketing Research* 40 (November): 481–491.
- Van Heerde, J. Harald, Peter S.H. Leeflang, and Dick R. Wittink. 2004. Decomposing the sales promotion bump with store data. *Marketing Science* 23 (3): 317–334.
- Van Heerde, J. Harald, Kristiaan Helsen, and Marnik Dekimpe. 2007. The impact of a product-harm crisis on marketing effectiveness. *Marketing Science* 26 (2): 230–245.
- Van Heerde, J. Harald, Shuba Srinivasan, and Marnik G. Dekimpe. 2010. Estimating cannibalization rates for a pioneering innovation. *Marketing Science* 29 (6): 1024–1039.
- Van Heerde, Harald J., Maarten Gijsenberg, Marnik G. Dekimpe, and Jan-Benedict E.M. Steenkamp. 2013. Advertising and price effectiveness over the business cycle. *Journal of Marketing Research* 50 (April): 177–193.
- Van Heerde, Harald J., Els Gijsbrechts, and Koen Pauwels. 2015. Fanning the flames? How media coverage of a price war impacts retailers, consumers and investors. *Journal of Marketing Research*.
- Van Lin, Arjen, and Els Gijsbrechts. 2014. Shopper loyalty to whom? Chain versus outlet loyalty in the context of store acquisitions. *Journal of Marketing Research* 51 (3): 352–370.
- Van Lin, Arjen, and Els Gijsbrechts. 2015. The battle for health and beauty: What drives supermarket and drugstore category-promotion lifts? *International Journal of Research in Marketing*. <http://dx.doi.org/10.1016/j.ijresmar.2015.09.002>.
- Varki, Sajeev, and Pradeep K. Chintagunta. 2004. The augmented latent class model: Incorporating additional heterogeneity in the latent class model for panel data. *Journal of Marketing Research* 41 (2): 226–233.
- Venkatesan, Rajkumar, and Paul W. Farris. 2012. Measuring and managing returns from retailer-customized coupon campaigns. *Journal of Marketing* 76 (1): 76–94.
- Villas-Boas, J. Miguel, and Russell S. Winer. 1999. “Endogeneity in brand choice models.” *Management Science* 45 (10): 1324–1338.
- Vroegrijk, Mark, Els Gijsbrechts, and Katia Campo. 2013. Close encounter with the hard discounter: A multiple-store shopping perspective on the impact of local hard-discounter entry. *Journal of Marketing Research* 50 (5): 606–626.
- Walters, Rockney G. 1989. An empirical investigation into retailer response to manufacturer trade promotions. *Journal of Retailing* 65 (2): 253–272.
- Wieringa, Jaap E., Peter C. Verhoef, and Janny C. Hoekstra. 2011. *Liber Amicorum in honor of Peter S.H. Leeflang*. Rijksuniversiteit Groningen.
- Winer, Russell S. 1986. A reference price model of brand choice for frequently purchased products. *Journal of Consumer Research* 13 (September): 250–256.
- Wittink, Dick R., Michael J. Addona, William J. Hawkes, and John C. Porter. 1988. The estimation, validation, and use of promotional effects based on scanner data. Working paper, Cornell University.
- Zhang, Jie, and Lakshman Krishnamurthi. 2004. Customizing promotions in online stores. *Marketing Science* 24 (4): 561–578.
- Zhang, Jie, and Michel Wedel. 2009. The effectiveness of customized promotions in online and offline stores. *Journal of Marketing Research* 48 (2): 190–206.

Handbook of Marketing Decision Models

Wierenga, B.; van der Lans, R. (Eds.)

2017, XI, 598 p. 35 illus., 19 illus. in color., Hardcover

ISBN: 978-3-319-56939-0