A Game Theoretical Study of Cooperative Advertising with Multiple Retailers in a Distribution Channel

Jihua Zhang [1*], Jinxing Xie [2]

[1, 2] *Department of Mathematical Sciences, Tsinghua University. Beijing 100084, China*

Abstract

Extant studies of cooperative advertising mainly consider a single-manufacturer-single-retailer channel structure. This can provide limited insights, because a manufacturer, in real practices, would frequently deal with multiple retailers at the same time. In order to examine the impact of the retailer's multiplicity on channel members' decisions and total channel efficiencies, this paper develops a multiple-retailer model. In this model, the manufacturer and the retailers play the Stackelberg game to make optimal advertising decisions. Based on the quantitative results, it is observed that: 1) When there are multiple symmetric retailers, as the number of retailers scales up, the manufacturer's national advertising investment contributes increasingly to add to channel members' profits, but the total channel efficiency deteriorates; 2) When there are multiple asymmetric retailers, the distribution channel suffers from the manufacturer's uniform participation strategy due to the retailer's free-ride, and benefits with the manufacturer's retailer-specific participation strategy. This study derives equilibrium solutions for all games studied in closed form, and explicitly measures the gains/losses of channel efficiencies under different game settings.

Keywords: cooperative advertising, game theory, multiple retailers, scaling effect, free-riding

1. Introduction

In a typical distribution channel, the manufacturer usually makes its advertising via the national wide media so as to build up its long-term brand image and enlarge its potential client bases. In contrast, the retailer tends to make local advertising campaigns which focus more on the short-term sales effect. The discrepancy of two types of advertising is presumed of a magnitude sufficiency to create a role for cooperative (co-op) advertising. Co-op advertising is, practically, an interactive relationship between a manufacturer and a retailer in which the manufacturer pays a portion of the retailer's local advertising costs; The fraction shared by the manufacturer is commonly referred to as the manufacturer's participation rate. Co-op advertising offers consumers the information needed when they move through the final stages of purchase and a congruence of information and information needs that would be impossible if the manufacturer uses national advertising only (Huang & Li, 2001).

Co-op advertising prevails in today's marketing practices. It is reported that GE's budgets for local advertising are three times as high as its national advertising budget (Young & Greyser, 1983). Intel's expenditure on co-op advertising grows from $800 million in 1999 to $1.5 billion in 2001 (Elkin, 1999). In 1987, the estimated co-op advertising expenditure spent by the US companies amounted to $10 billion (Somers et al., 1990). In 1993, 20 billion is used for co-op advertising (Davis, 1994). According to Nagler (2006), the total US expenditure of co-op advertising in 2000 is estimated at $15 billion, nearly a four-fold increase in real terms in comparison with $900 million in 1970.

Increasing spending volumes and high levels of participation rate in co-op advertising have motivated theoretical studies to explore a most effective scheme for channel members to participate in. Lyon's (1932) discuss of the advertising allowances – money paid by a manufacturer to a dealer to cover a part of dealer's

[*] Corresponding author. *Tel.*: (+86-10) 6278-7812; *Fax*: (+86-10) 6278-5847;
E-mail address: zhang-jh02@mails.tsinghua.edu.cn

advertising expenses – provides what may have been the first formal, albeit qualitative treatment of the subject of co-op advertising. Following Lyon's seminal work, studies of co-op advertising, including both empirical investigations and analytical modeling works, never stop, thereafter, for the past eight decades. Our study follows the analytical modeling approach.

Regarding analytical modeling techniques, game theoretical models have been most widely used in co-op advertising researches. In particular, they are used to model the inherent inter-dependence and conflicts among channel members, and how the co-op advertising decisions, as well as channel performance is influenced by the various factors, such as advertising efforts at national and local levels, the participation rate, sales volumes, brand and store substitutions, pricing, among others. For an extensive survey of the game theoretical models in co-op advertising, one can refer to Taboubi & Zaccour (2005). We briefly review the co-op advertising literature related to the dynamic and static games, respectively. A dynamic game takes a long-term perspective over several (or infinite) time periods (e.g., Kim & Stealin, 1999; Jorgensen et al., 2000; Taboubi & Zaccour, 2002; Jorgensen & Zaccour, 2003; He et al., 2009). Customer demand is usually assumed to be based on a goodwill stock that is influenced by advertising in the current and previous periods (Nerlove & Arrow, 1962; Chintagunta & Jain, 1992). Investments on the goodwill stock are therefore considered to have the "carryover" effect from one period to the next. However, a common problem with the dynamic game models is that closed-form solutions are usually not available without significant model simplifications. On the other hand, static game analyzes the co-op advertising in a single-period (e.g., Berger, 1972; Dant & Berger, 1996; Berger & Magliozzi, 1992; Bergen & John, 1997; Huang & Li, 2001; Yue et al., 2006; Szmerekovsky & Zhang, 2009; Xie & Wei, 2009). One common assumption is that the game is played only once, and the brand-name investments, the local advertising spending, as well as player's participation strategies are all assumed to be time-invariant. Static game models are convenient tools to develop analytical solutions and to generate insights related to the key elements of the game structure, including advertising expenditures, participation rates, and channel efficiencies, so we choose to use static game models in this paper.

We note that all static analytical models mentioned above consider a single-retailer channel structure, which has several limitations: 1) It is unable to study the retailer's multiplicity and its impact, e.g., scaling effect; 2) It does not capture the *universal* impact of the manufacturer's brand-name advertising, which, once implemented, will benefit all resellers equally who sell products under the brand-name; 3) It fails to capture the impact of different retailers' attributes (i.e., profit margins, market sizes, etc.) on their co-op advertising decisions. Although marketing practices suggest that the manufacturer would often times deal with multiple retailers, existing literature of co-op advertising with multiple retailers is sparse. Karray & Zaccour (2007) considers a distribution channel consisting of two manufacturers and two retailers. The retailers choose their levels of marketing efforts (e.g., local advertising, displays, etc.) and the manufacturers control their participation rates for the retailers' marketing efforts if the co-op advertising program is an option. However, the national advertising efforts are specified as exogenous parameters instead of manufacturer's decision variables.

To the best of our knowledge, this study is among the first to develop a game theoretical model with multiple retailers. The main contribution is that it explicitly derives and quantitatively measures the impacts of the retailer's multiplicity on channel members' optimal decisions and on total channel efficiencies. Managerial insights can be summarized in the following two perspectives: When there are multiple symmetric retailers, we identify scaling effects on channel members' profits and the total channel efficiencies. Specifically, as the number of retailers scales up, the manufacturer's national advertising investment contributes increasingly to add to channel members' profits. Although the total channel efficiency may increase with the local advertising effectiveness and/or the manufacturer's relative channel power, it deteriorates quickly as the number of retailer scales up. When the retailers are asymmetric, we study how the different local market attributes and the manufacturer's uniform participation constraint affect the manufacturer's optimal participation strategy and the system's efficiency. It is observed that: 1) The manufacturer's uniform participation strategy can be inefficient due to the retailer's free-riding. Specifically, the manufacturer's uniform subsidy tends to be insufficient with retailers of larger market size and lower profit margins, and excessive with retailers of smaller market size and higher profit margin, giving the latter a free-ride; 2) In order to prevent the retailer's free-ride, the manufacturer can use the retailer-specific participation strategy, which suggests the manufacturer's participation rate with retailers of smaller market size and higher profit margin should be less generous than that of larger market size and lower profit margin.

This paper is organized as follows: Section 2 presents the model framework. Section 3-4 investigates impacts of the retailer's multiplicity when there are multiple symmetric/asymmetric retailers. Section 5 summarizes the findings and proposes directions for future research.

2. Model framework

Consider a distribution channel with one manufacturer M selling a certain product through n retailers R_1, R_2, …, R_n. Assume that different retailers are geographically *separated* and that cross-border products transferring is prohibited, so there are no intra-brand competitions. This assumption captures the real situation when a manufacturer's marketing channels are widespread, with each reseller being authorized as the sole representative of its local marketing area and making advertising and pricing decisions independently. As far as we know, some Chinese manufacturers of Electronic Instrument follow this way. They designate a first-class sales agent who is, headquartered in some first-tier city, fully in charge of the sales and serves in the city and the surrounding areas (e.g., Beijing for the *North-China* area, Shanghai/Guangzhou for the *East-China* area, ect.). Although geographically separated, the retailers benefit equally from the national advertising (e.g., brand-image advertising on national TV channels) because they carry the same products under the manufacturer's brand-name.

Our study mainly focuses on channel members' advertising decisions. So we do not take directly pricing decisions into account, instead, we assume a fixed gross unit margin (advertising costs excluded) for channel members. Let M's gross profit margin be ρ_0 and R_i's be ρ_i, i=1, 2, …, n. Fixed margins can be justified with a short-run planning horizon. Similar assumptions are also used by Jorgensen et al. (2000) and Sigue & Chintagunta (2008).

Let a_0 and a_i, i=1, 2, …, n, be M's and R_i's advertising spendings, respectively. We assume the resulting customer demand for retailer R_i, $S_i = S_i(a_0, a_i)$, often called the *sales response* function, is jointly determined by both the national advertising spending a_0 and local advertising spending a_i. There is a substantial literature on the estimation of the sales response function with respect to advertising investments. Following (Xie & Wei, 2009) we use

$$S_i(a_0, a_i) = b_i(\alpha_i + \sqrt{a_0} + k_i\sqrt{a_i}),\ i=1,\ 2,\ \ldots,\ n, \quad (1)$$

where $\alpha_i \geq 0$ is the demand base with respect to zero advertising input, $k_i \geq 0$ is the relative effectiveness of the local advertising compared with the national advertising, and $b_i \geq 0$ is the market size. To obtain a better understanding of $S_i(a_0, a_i)$, we suggest take $(\alpha_i + \sqrt{a_0} + k_i\sqrt{a_i})$ as the average demand of a typical customer within R_i's marketing area and b_i the population size. We assume further $\alpha_1 = \ldots = \alpha_n = \alpha$ and $k_1 = \ldots = k_n = k$ to imply that a typical customer's average demand level is much the same in all local markets. Then the n local markets are primarily differentiated by the population size within the marketing areas. Technically, there are no difficulties to extend the model by considering α_i and k_i being dependent with the retailer's index i. However, the assumption of a uniform α and k can help us to better focus on the key parameter of b_i, which we will use in this paper to distinguish *large* retailers from *small* retailers. Finally, for simplicity of our discussion, we normalize $\alpha = 0$. Note that α is not the retailer's decision variable; Instead, it is only a constant parameter, not coupling with any decision variables. So normalizing $\alpha = 0$ will not affect the equilibrium outcome regarding channel members' optimal decisions. Then R_i's sales response function is specified as

$$S_i(a_0, a_i) = b_i(\sqrt{a_0} + k\sqrt{a_i}). \quad (2)$$

Obviously, $S_i(a_0, a_i)$ is continuously differentiable, strictly increasing, and strictly (joint) concave with respect to (a_0, a_i). The square root formulation also reflects the commonly observed "advertising saturation effect," i.e., additional advertising spending generates continuously diminishing returns, which has been verified and supported by Simon & Arndt (1980) based on their review of over 100 case studies.

To implement co-op advertising, let M shares a portion $t_i \in [0, 1]$ of R_i's local advertising cost a_i.

Let $\boldsymbol{a} := (a_1, a_2, ..., a_n)$ and $\boldsymbol{t} := (t_1, t_2, ..., t_n)$, then profits for M, R_i, i=1, 2, …, n, and the whole distribution channel are:

$$\Pi_0(a_0, \boldsymbol{a}, \boldsymbol{t}) = \rho_0 \sum_{i=1}^{n} b_i(\sqrt{a_0} + k\sqrt{a_i}) - \sum_{i=1}^{n} t_i a_i - a_0\,, \tag{3}$$

$$\Pi_i(a_0, a_i) = \rho_i b_i(\sqrt{a_0} + k\sqrt{a_i}) - (1 - t_i)a_i\,,\ i=1,\ 2,\ ...,\ n. \tag{4}$$

$$\Pi_S(a_0, \boldsymbol{a}) = \Pi_0 + \sum_{i=1}^{n} \Pi_i = \sum_{i=1}^{n} (\rho_0 + \rho_i) b_i(\sqrt{a_0} + k\sqrt{a_i}) - \sum_{i=1}^{n} a_i - a_0\,. \tag{5}$$

All throughout this paper, we use subscripts “0”, “i” (i=1, 2, …, n), and “S” to mark the parameters corresponding to M, R_i, and the whole channel system.

3. Scaling effect with multiple homogeneous retailers

Suppose the n retailers in the distribution channel are homogeneous. Without loss of generality, let all retailers be identical with R_1, then $b_i = b_1$, $\rho_i = \rho_1$, for all i=2, 3, …, n.

3.1. Benchmark case — centralized channel

Following Equation (5), the total channel profit can be written as

$$\Pi_S(a_0, \boldsymbol{a}) = (\rho_0 + \rho_1) b_1 \sum_{i=1}^{n} (\sqrt{a_0} + k\sqrt{a_i}) - \sum_{i=1}^{n} a_i - a_0\,. \tag{6}$$

Please note that $\Pi_S(a_0, \boldsymbol{a})$ is a strict (jointly) concave function with respect to $(a_0, \boldsymbol{a})$. To see this point, one can simply check the Hessian of $\Pi_S(a_0, \boldsymbol{a})$ being negative-defined. By definition, $\nabla^2(\Pi_S) = (A_{jl})$, where $A_{jl} = \partial^2 \Pi_S / \partial a_j \partial a_l$, $j, l = 0, 1, ..., n$. If $j \neq l$, $A_{jl} = 0$; Otherwise, if j=l=0, $A_{jl} = (\rho_1 + \rho_0) \sum_{i=1}^{n} \frac{\partial^2 S_i}{\partial a_0^2} = -\frac{n}{4}(\rho_1 + \rho_0) b_1 a_0^{-3/2} < 0$, if $j = l \geq 1$, $A_{jl} = -\frac{n}{4}(\rho_1 + \rho_0) k b_1 a_j^{-\frac{3}{2}} < 0$. Thus, $\nabla^2(\Pi_S)$ is negative, so $\Pi_S(a_0, \boldsymbol{a})$ has a unique global optimal point, which can be solved from the first order condition $\nabla \Pi_S = 0$. Denote the optimal point by ($\bar{a}_0^C$, $\bar{\boldsymbol{a}}^C$) := ($\bar{a}_0^C$, $\bar{a}_1^C$, …, $\bar{a}_n^C$).

Proposition 1. *With n symmetric retailers in the distribution channel the centralized solution that maximizes the total channel profit is given by*

$$\bar{a}_0^C = \frac{1}{4} b_1^2 n^2 (\rho_0 + \rho_1)^2\,, \tag{7}$$

$$\bar{a}_i^C = \frac{1}{4} b_1^2 k^2 (\rho_0 + \rho_1)^2\,,\ i=1,\ 2,\ …,\ n. \tag{8}$$

Hereafter in this paper, we will use “C” to mark centralized solutions, and the upper-bar “ ¯ ” to mark equilibrium outcomes corresponding to the n-symmetric-retailer case.

Note that $(\rho_0 + \rho_1)$ represents the whole channel’s gross margin for each unit sales. Then, Equations (7)-(8) suggest that the optimal levels for national/local advertising expenditures should be proportional to the squared value of the channel’s gross margin scaled by the market size, i.e., $[b_1(\rho_0 + \rho_1)]^2$. What’s different, M’s optimal advertising level $\bar{a}_0^C$ is increasing in n, whereas, R_i’s optimal advertising level $\bar{a}_i^C$ stays constant with n. Substitute ($\bar{a}_0^C$, $\bar{\boldsymbol{a}}^C$) into Equation (6), we obtain the optimal channel profit by

$$\overline{\Pi}_S^C = \frac{1}{4} n(k^2+n) b_1^2 (\rho_0+\rho_1)^2 . \tag{9}$$

Referring to Equations (7)-(8), we can actually interpret Equation (9) in the following way: $\overline{\Pi}_S^C = \frac{1}{4} n(k^2+n) b_1^2 (\rho_0+\rho_1)^2 = n\left(\frac{1}{4} k^2 b_1^2 (\rho_0+\rho_1)^2\right) + \frac{1}{4} n^2 b_1^2 (\rho_0+\rho_1)^2 = n\overline{a}_1^C + \overline{a}_0^C$, which reveals that, the optimal centralized channel profit $\overline{\Pi}_S^C$ consists of two parts: one is M's optimal national advertising expense $\overline{a}_0^C$, the other is obtained by adding up all retailers' optimal local advertising expenses $n\overline{a}_1^C$. Since $\overline{a}_0^C / n\overline{a}_1^C = n/k^2$ increases in n, we conclude that the national advertising's relative contribution to total channel profit increases with n.

Insight 1. *In a centralized channel, as the retailers' number scales up, the national advertising plays an increasingly prominent role in terms of generating profits for the whole distribution channel.*

3.2. Decentralized channel

For the decentralized channel, we model channel members' decision process as a two-stage Stackelberg game. In the first stage, M decides its national advertising expenditure a_0 and participation rates t_i, i=1, 2, …, n, and in the second stage, R_i simultaneously chooses its local advertising level a_i, i=1, 2, …, n, accordingly. Due to the symmetry of the model, let M's participation rates t_i with R_i, i=1, 2, …, n all be identical with t_1, then players' profit functions are reduced to

$$\Pi_0(a_0, \boldsymbol{a}, t_1) = \rho_0 b_1 \sum_{i=1}^{n} (\sqrt{a_0} + k\sqrt{a_i}) - t_1 \sum_{i=1}^{n} a_i - a_0 , \tag{10}$$

$$\Pi_i(a_0, a_i, t_1) = \rho_1 b_1 (\sqrt{a_0} + k\sqrt{a_i}) - (1-t_1) a_i , \quad i=1,\ 2,\ \ldots,\ n. \tag{11}$$

Note that at the second stage of the game, given a_0 and t_1, R_i's simultaneous move generates a Nash Equilibrium. To deal with the Nash Equilibrium, one should always discuss its *existence* and *uniqueness*. Fortunately, for the current model studied, both the existence and the uniqueness of the Nash Equilibrium are guaranteed. For existence, we use Topkis's (1979) argument of the *supermodular game* as the sufficient condition (Because $\partial^2 \Pi_i / \partial a_i \partial a_j = 0$ for all $j \neq i$, i=1, 2, …, n, then the game is always a supermodular game). For uniqueness, we refer to Moulin's (1986) statement, which requires that the slope of each player's best responses never exceed 1 in absolute value. In our model, the condition $\left| \dfrac{\partial^2 \Pi_i / \partial a_i \partial a_j}{\partial^2 \Pi_i / \partial a_i^2} \right| = 0 < 1$, $j \neq i$, i=1, 2, …, n, is always valid.

Lemma 1. *Given M's decision of a_0 and t_1, the retailers' simultaneous move generates a unique Nash equilibrium, which can be explicitly expressed as*

$$a_i^*(a_1^*, \ldots a_{i-1}^*, a_{i+1}^* \ldots, a_n^* \mid a_0, t_1) = \frac{1}{4} \frac{\rho_1^2 b_1^2 k^2}{(1-t_1)^2} , \quad i=1,\ 2,\ \ldots,\ n. \tag{12}$$

Because $\partial a_i^* / \partial t_1 > 0$, then a higher level of M's participation rate t_1 induces a higher level of R_i's local advertising input in equilibrium.

Proposition 2. *On condition that $\rho_0 \geq \frac{1}{2}\rho_1$*[†]*, the unique sub-game perfect equilibrium solution for the Stackelberg game is given by*

[†] When $\rho_0 < \frac{1}{2}\rho_1$, M's optimal participation rate will be zero, i.e., M chooses not to participation into the retailers' local advertising. Similar arguments apply to Proposition 4 and 5 analogously.

$$\bar{a}_0^D = \frac{1}{4}\rho_0^2 n^2 b_1^2, \tag{13}$$

$$\bar{a}_i^D = \frac{1}{16} b_1^2 k^2 (2\rho_0 + \rho_1)^2,\ i=1,\ 2,\ \ldots,\ n, \tag{14}$$

$$\bar{t}_1^D = \frac{2\rho_0 - \rho_1}{2\rho_0 + \rho_1}. \tag{15}$$

Hereafter in this paper, we will use the superscript "*D*" to notify all equilibrium solutions and the associated items with a decentralized channel.

Equation (15) reveals that M's optimal participation rate $\bar{t}_1^D \in [0,1]$ is only dependent with channel members' gross margins: It is increasing in ρ_0 and decreasing in ρ_1. Please note that the ratio $\rho_0 / \rho_{\mathrm{i}}$ represents M's relative channel power against R$_i$, and the kink point 1/2 is the retailer's local advertising-demand elasticity in absolute value ($\left|\partial \log(S_i / a_i) / \partial a_i\right| = \frac{1}{2}$), therefore, condition $\rho_0 \geq \frac{1}{2}\rho_1$ is equivalent to saying that M's relative channel power is larger than or equals to the retailer's advertising-demand elasticity in absolute value. We assume this condition hold all through this section.

Compare Equations (13)-(14) with (7)-(8), we obtain similar observations as in Section 3.1: 1) R$_i$'s optimal advertising cost $\bar{a}_1^D$ is independent with *n*; 2) M's optimal advertising investment $\bar{a}_0^D$ is proportional to n^2, therefore, will increase quickly as the number of retailers scales up. Substitute (13)-(14) into (10)-(11), we get M's and R$_i$'s profits as:

$$\overline{\Pi}_0^D = \frac{1}{16} b_1^2 n\left[4n\rho_0^2 + k^2(2\rho_0 + \rho_1)^2\right], \tag{16}$$

$$\overline{\Pi}_i^D = \frac{1}{8} b_1^2 \rho_1\left[4n\rho_0 + k^2(2\rho_0 + \rho_1)\right],\ i=1,\ 2,\ \ldots,\ n. \tag{17}$$

Equations (16) can be reformed as $\overline{\Pi}_0^D = n\left[\frac{1}{16} b_1^2 k^2 (2\rho_0 + \rho_1)^2\right] + \frac{1}{4}\rho_0^2 n^2 b_1^2 = n\bar{a}_1^D + \bar{a}_0^D$, which suggests us to count M's optimal profit by two parts: one is M's own advertising investment $\bar{a}_0^D$ at the national level, the other is the sum of all its retailers' advertising expenses $n\bar{a}_1^D$ at the local level. Since $\bar{a}_0^D / n\bar{a}_1^D$ increases in *n*, we conclude that the national advertising's relative contribution to M's profit increases with *n*. Similarly, Equation (17) can be reformed as $\overline{\Pi}_i^D = \gamma \cdot \bar{a}_1^D + \phi \cdot \left(\bar{a}_0^D / n\right)$, *i*=1, 2, …, *n*, being a linear combination of R$_i$'s optimal local advertising spending $\bar{a}_1^D$ and the *average spending* of M's optimal national advertising, $\left(\bar{a}_0^D / n\right)$, where $\gamma = 2\rho_1 / (2\rho_0 + \rho_1)$, $\phi = 2\rho_1 / \rho_0$. While $\bar{a}_1^D$ is constant with *n*, $\left(\bar{a}_0^D / n\right) = \frac{1}{4} n \rho_0^2 b_1^2$ is increasing with *n*. Therefore, M's national advertising spending increasingly contributes to add to R$_i$'s profit as the retailer's number scales up.

Insight 2. *The manufacturer's national advertising conducts a universal impact in generating channel members' profits: It helps generate sales in all local markets that carry products under its brand-name. As the retailer's number scales up, it contributes increasingly to add profits to both the manufacturer and the retailers.*

In general, the channel profit based on decentralized decisions (13)-(15) will be less than that based on centralized decisions (7)-(8). To quantitatively measure the system efficiency under the current decentralized system, we introduce the ratio $\eta \in [0,1]$ as the (optimal) channel profit under the decentralized game setting over that under the centralized game setting. For example, we will use $\bar{\eta} := \overline{\Pi}_S^D / \overline{\Pi}_S^C$ to measure the channel efficiency for the decentralized game setting with multiple symmetric retailers, in which $\overline{\Pi}_S^D$ is obtained by adding all members' profits:

$$\overline{\Pi}_S^D = n\overline{\Pi}_1^D + \overline{\Pi}_0^D = \frac{1}{16} b_1^2 n\left[4n\rho_0(\rho_0 + 2\rho_1) + k^2(2\rho_0 + \rho_1)(2\rho_0 + 3\rho_1)\right]. \quad (18)$$

Proposition 3. *The system efficiency of the current decentralized game setting is*

$$\overline{\eta} = 1 - \left(\frac{n + \frac{k^2}{4}}{n + k^2}\right) \cdot \frac{\rho_1^2}{(\rho_0 + \rho_1)^2}, \quad (19)$$

which has the following properties:

(i) For any given ρ_0, ρ_1 *and k,* $\overline{\eta}$ *is a convex decreasing function of n, and* $\lim_{n \to +\infty} \overline{\eta} = \rho_0(\rho_0 + 2\rho_1)/(\rho_0 + \rho_1)^2$

(ii) For any given n, $\overline{\eta}$ *increases with k and* ρ_0/ρ_1*. The efficiency bound is estimated by* $\overline{\eta} \in \left[\frac{5}{9}, 1\right]$*. When k*=0 *and* $\rho_0 = \frac{1}{2}\rho_1$*,* $\overline{\eta}$ *gets its lower bound 5/9; when* $\rho_0 \to +\infty$*,* $\overline{\eta}$ *gets its upper bound* 1.

Since the ratio ρ_i / ρ_0, *i*=1, 2 represents retailer R_i's relative *channel power*. Proposition 3 states that a strict Pareto improvement bilateral participation scheme is guaranteed as long as the retailer's *average* channel power is greater or equals to the manufacturer's. Recall that big manufacturers used to have more channel power than retailers before the early 1980's. Manufacturers such as Procter&Gamble (P&G) limited the quantities of high-demand products they would deliver to a given supermarket chain and insisted that the supermarket carry all sizes of a certain product (Huang et al. 2002). However, things changed ever since the middle 1980's with a market trend in which the retailers began to retain equal or even more powers (Achenbaum and Mitchel 1987; Buzzell et al. 1990; Kumar 1996;).

According to (ii) of Proposition 3, $\overline{\eta}$ increases as the local advertising's relative effectiveness (*k*) gets higher and/or the manufacturer's relative channel power (ρ_0 / ρ_1) gets larger. However, as shown by (i) of Proposition 3, the channel efficiency suffers from a increasing number of retailers. Figure 1 shows how the channel efficiency changes with the retailer's number. In particular, if *k*=1, $\rho_0 = 0.5$ and $\rho_1 = 1$, the channel efficiency is below 60% when there are more than 7 retailers in the distribution channel.

Insight 3. *With multiple symmetric retailers in the distribution channel, the channel efficiency increases with the local advertising effectiveness and/or the manufacturer's relative channel power, however, it deteriorates quickly as the number of retailer scales up.*

4. Optimal participation strategies with asymmetric retailers

According to Berger (1972), co-op advertising ventures in practice are frequently specified on a 50-50 basis, with each participant paying half of the expenses. Nagler (2006) conducted a large-scale study of 2,286 brands in the US, finding 61.5% of the retailers enjoy a participation rate of 50% from the manufacturers. Does it make sense for different retailers to enjoy a uniform level of participation rate (for instance, 50%)? Intuitively, if the manufacturer's uniform participation rate is derived from a sense of fairness rather than systematic profit maximization, it can be inefficient. One important reason can be the retailer's free-riding. It is commonly believed that a well-designed channel contract should prevent horizontal free-riding among retailers (Lal, 1990).

In this section, we consider co-op advertising with multiple asymmetric retailers. We will examine how different parameters affect the manufacturer's co-op advertising decisions, and will quantify the efficiency losses due to retailers' free-riding under the manufacturer's uniform participation strategy. Without loss of generality, let *n*=2.

4.1. The retailer's free-riding under the manufacturer's uniform participation strategy

Consider M, the Stackelberg game leader, offers a uniform co-op advertising participation rates to R_1

and R_2 ($t_1 = t_2 = t$). Then channel members' profits are:

$$\Pi_0(a_0,a_1,a_2,t) = \rho_0\left(\sum_{i=1}^{2} b_i(\sqrt{a_0} + k\sqrt{a_i})\right) - t\sum_{i=1}^{2} a_i - a_0, \tag{20}$$

$$\Pi_1(a_0,a_1,t) = \rho_1 b_1(\sqrt{a_0} + k\sqrt{a_1}) - (1-t)a_1, \tag{21}$$

$$\Pi_2(a_0,a_2;t) = \rho_2 b_2(\sqrt{a_0} + k\sqrt{a_2}) - (1-t)a_2. \tag{22}$$

The game sequence is the same as in Section 3. Without loss of generality, we assume $\rho_1 \le \rho_2$ all through this paper. Before moving on, let's define $\rho_R := \theta_1\rho_1 + \theta_2\rho_2$, where $\theta_i = b_i^2\rho_i \big/ (b_1^2\rho_1 + b_2^2\rho_2)$. By definition, θ_i, i=1, 2, and ρ_R have the following properties:

Lemma 2.

(i) $\theta_i \in [0,1]$, i=1, 2, $\theta_1 + \theta_2 = 1$.

(ii) θ_i *increases in* b_i *and decreases in* b_j, $j \ne i$, i=1, 2. *When* $b_1/b_2 \to +\infty$, *we get* $\theta_1 \to 1$, *and* $\rho_R \to \rho_1$; *Conversely, when* $b_1/b_2 \to 0$, *we get* $\theta_1 \to 0$, *and* $\rho_R \to \rho_2$.

According to Lemma 2, ρ_R is a convex combination of ρ_1 and ρ_2, weighted by θ_1 and θ_2, so $\rho_1 \le \rho_R \le \rho_2$. We call ρ_R the retailer's *average profit margin*.

Proposition 4. *On condition that* $\rho_0 \ge \frac{1}{2}\rho_R$, *the unique sub-game perfect equilibrium solution for the Stackelberg game will be given by*

$$\hat{a}_0^D = \frac{1}{4}(b_1 + b_2)^2\rho_0^2, \tag{23}$$

$$\hat{a}_1^D = \frac{1}{16}\frac{\rho_1^2 b_1^2 k^2 (2\rho_0 + \rho_R)^2}{\rho_R^2}, \tag{24}$$

$$\hat{a}_2^D = \frac{1}{16}\frac{\rho_2^2 b_2^2 k^2 (2\rho_0 + \rho_R)^2}{\rho_R^2}. \tag{25}$$

The manufacturer's optimal participation rate will be

$$\hat{t}^D = \frac{2\rho_0 - \rho_R}{2\rho_0 + \rho_R}. \tag{26}$$

Hereafter in this paper, we will use the upper-cap "$\wedge$" to notify all the equilibrium solutions and the associated items with M's uniform participation strategy.

With Proposition 4 we are ready to observe the retailer's free-riding when M uses a uniform participation rate $\hat{t}^D$ with differentiated retailers. According to Equation (26), M's optimal participation rate increases in ρ_0 and decreases in ρ_R. Let's consider the extreme case when R_1's market size b_1 is larger than R_2's market size b_2 by far, i.e., $b_1/b_2 \to +\infty$. By property (ii) of Lemma 2, we have $\theta_1 \to 1$, and consequently ρ_R goes down to ρ_1, then M's optimal participation rate will be raised up to $(2\rho_0 - \rho_1)/(2\rho_0 + \rho_1)$, as if R_1 is the only downstream reseller. For this situation, we say that R_2 gets a free-ride, because R_2 enjoys the same high level of participation rate as R_1. Extreme situations like $b_1/b_2 \to +\infty$ are unusual in real scenarios, but situations with $b_1 \ge b_2$ can be *reasonably* expected given $\rho_1 \le \rho_2$ [‡]; So long as $b_1 \ge b_2$, one can use similar argument to show that the retailer with smaller market size gets the free-ride.

Insight 4. *With multiple asymmetric retailers, the manufacturer's uniform participation strategy results in the retailer's free-riding. The manufacturer's optimal uniform participation rate tends to be in favor of retailers with larger market sizes; therefore, retailers with smaller market sizes (usually accompanied by higher profit margins) get the free-ride.*

4.2. The manufacturer's retailer-specific participation strategy

Because of the retailers' free-riding, the manufacturer's uniform participation strategy can be inefficient. In order to quantify the inefficiencies and disclose the fact that M's uniform participation strategy indeed imposes inappropriate incentives in co-op advertising, we introduce the manufacturer's *retailer-specific* participation strategy. Under the retailer-specific participation strategy, M is allowed to use differentiated participation rates towards different retailers. Denote M's participation rate with R_i by t_i, i=1, 2. Then channel members profit functions are

$$\Pi_0(a_0, a_1, a_2, t_1, t_2) = \rho_0\left(\sum_{i=1}^{2} b_i(\sqrt{a_0} + k\sqrt{a_i})\right) - \sum_{i=1}^{2} t_i a_i - a_0, \tag{27}$$

$$\Pi_1(a_0, a_1, t_1) = \rho_1 b_1(\sqrt{a_0} + k\sqrt{a_1}) - (1-t_1)a_1, \tag{28}$$

$$\Pi_2(a_0, a_2, t_2) = \rho_2 b_2(\sqrt{a_0} + k\sqrt{a_2}) - (1-t_2)a_2. \tag{29}$$

The game sequence is the same as been specified in the uniform participation case.

Proposition 5. *If* $\rho_0 \ge \frac{1}{2}\rho_i$, i=1, 2, *then under the manufacturer's retailer-specific participation strategy, channel members' equilibrium advertising levels are*

$$\tilde{a}_0^D = \frac{1}{4}(b_1 + b_2)^2 \rho_0^2, \tag{30}$$

$$\tilde{a}_1^D = \frac{1}{16} b_1^2 k^2 (2\rho_0 + \rho_1)^2, \tag{31}$$

$$\tilde{a}_2^D = \frac{1}{16} b_2^2 k^2 (2\rho_0 + \rho_2)^2. \tag{32}$$

The manufacturer's optimal participation rates are

$$\tilde{t}_1^D = \frac{2\rho_0 - \rho_1}{2\rho_0 + \rho_1}, \tag{33}$$

$$\tilde{t}_2^D = \frac{2\rho_0 - \rho_2}{2\rho_0 + \rho_2}. \tag{34}$$

Proof of Proposition 5 is almost the same as Proposition 4. So we omit it here to save space. Hereafter in this paper, we use the tilde "~" to represent equilibrium solutions and the associated items for cases with M's retailer-specific participation strategy.

[‡] Marketing practices suggest that a retailer who possesses a giant market size usually sets its unit margins being comparatively low, for example, the Wal-Mart, being famous for its *Everyday Low Prices*, operates with a low gross margin of no more than 4% over decades. A range of studies has found that prices at Wal-Mart are anywhere from 8 to 39 percent less than its major competitors, and that even a very small increase in its costs, without a corresponding increase in revenues, would wipe out all Wal-Mart's profits entirely (Furman, 2005). In contrast, a retailer with limited market size would often times impose a high unit margin. This is well explained by the old marketing discipline: *low profits go along with good sales*.

Making comparisons between Equations (33)-(34) and (26), and also by Lemma 2, we know that $\tilde{t}_2^D \le \hat{t}^D \le \tilde{t}_1^D$. This observation can be intuitively explained as

Insight 5. *In contrast with the optimal retailer-specific participation strategy, M's optimal participation rate under the uniform participating strategy imposes inappropriate incentives to the retailers: It is insufficient with the retailer of larger market sizes (i.e., R_1 in our model), and is excessive with the retailer of smaller market sizes (i.e., R_2 in our model).*

Regarding how to make a difference between asymmetric retailers, Equations (33)-(34) suggest that M's optimal participation rate towards the retailer with smaller market sizes should be less than that with larger sizes. The following proposition, along with Figure 2, quantitatively compares M's uniform participation strategy with M's retailer-specific participation strategy. Results of Proposition 6 are natural consequences of Proposition 4 and 5, so we omit the proof to save space.

Proposition 6.

(i) *If $\rho_0 \le \frac{1}{2}\rho_1$, then M's optimal policy is to take no participations in the retailers' local advertising, either under the uniform participation strategy or under the retailer-specific participation strategy (i.e. $\tilde{t}_1^D = \hat{t}^D = \tilde{t}_2^D = 0$).*

(ii) *If $\frac{1}{2}\rho_1 \le \rho_0 \le \frac{1}{2}\rho_R$, then M will participate into R_1's local advertising with rate $\tilde{t}_1^D = (2\rho_0 - \rho_1)/(2\rho_0 + \rho_1)$ and participate into R_2's local advertising with rate $\tilde{t}_2^D = 0$ under the retailer-specific participation strategy; But will take no participations with either retailers under the uniform participation strategy (i.e. $\hat{t}^D = 0$).*

(iii) *If $\frac{1}{2}\rho_R \le \rho_0 \le \frac{1}{2}\rho_2$, then M will participate into R_1's local advertising with rate $\tilde{t}_1^D = (2\rho_0 - \rho_1)/(2\rho_0 + \rho_1)$ and participate into R_2's with rate $\tilde{t}_2^D = 0$ under the retailer-specific participation strategy; And will participation into both retailers' local advertising with rate $\hat{t}^D = (2\rho_0 - \rho_R)/(2\rho_0 + \rho_R)$ under the uniform participation strategy.*

(iv) *If $\frac{1}{2}\rho_2 \le \rho_0$, then M will participate into R_1's local advertising with rate $\tilde{t}_1^D = (2\rho_0 - \rho_1)/(2\rho_0 + \rho_1)$ and participate into R_2's local advertising with rate $\tilde{t}_2^D = (2\rho_0 - \rho_2)/(2\rho_0 + \rho_2)$ under the retailer-specific participation strategy; And will participations into both retailers' local advertising with rate $\hat{t}^D = (2\rho_0 - \rho_R)/(2\rho_0 + \rho_R)$ under the uniform participation strategy.*

4.3. System efficiency estimations

For all cases discussed above, we can measure the channel efficiencies by comparing the optimal channel profit under a specific game setting against the optimal profit of the centralized channel. We use two measurements: One is the *absolute* measure, denoted by $\delta \ge 0$, which describes the distance of the equilibrium channel profit under decentralized solution and the optimal total channel profit under centralized solution. For example, when M is allowed to use the retailer-specific participation rates, we define $\tilde{\delta} := \Pi_S^C - \tilde{\Pi}_S^D$; when M is confined to the uniform participation strategy, $\hat{\delta} := \Pi_S^C - \hat{\Pi}_S^D$; The other is the *relative* measure, denoted by $\eta \in [0,1]$, which is the ratio of the equilibrium channel profit under decentralized solution over the optimal total channel profit under centralized solution (η has been introduced in Section 3.2). For example, when M is allowed to use the retailer-specific participation rates, we define $\tilde{\eta} := \tilde{\Pi}_S^D / \Pi_S^C$; when M is confined to the uniform participation strategy, $\hat{\eta} := \hat{\Pi}_S^D / \Pi_S^C$. By definition, the absolute measure and the relative measure are consistent; a lower (higher) value of δ (η)

indicates a higher level of system efficiency.

We list all equilibrium solutions in Table 1 for a reference, where the total channel profit under with centralized solution is $\Pi_S^C = \frac{1}{4}\sum_{i=1}^{2} k^2 b_i^2 (\rho_0 + \rho_i)^2 + \frac{1}{4}\left[\sum_{i=1}^{2} b_2(\rho_0 + \rho_i)\right]^2$. The 2 rows at the bottom gives the efficiency estimation by both absolute measure and relative measure. We first focus on the absolute measure. Compare $\hat{\delta}$ with $\tilde{\delta}$, we quantify the efficiency losses due to M's uniform participation strategy by $\hat{\delta} - \tilde{\delta} = \tilde{\Pi}_S^D - \hat{\Pi}_S^D = \frac{1}{4}\rho_0 b_1^2 b_2^2 k^2 (\rho_1 - \rho_2)^2 \big/ (\rho_1^2 b_1^2 + \rho_2^2 b_2^2)$, which is greater than zero *as long as* $\rho_1 \neq \rho_2$. The intuitions here are quite clear: 1) M's uniform participation strategy incurs efficiency losses as long as the retailers' gross margins *do not coincide*; 2) The system's inefficiency gets worse when the distance of ρ_1 and ρ_2 enlarges. Since the uniform participation strategy can be taken as a special case of the retailer-specific participation strategy with $t_1 = t_2$, so formally, we may consider the item $\hat{\delta} - \tilde{\delta} > 0$ as the *added value* to the system by switching to the retailer-specific participation strategy.

We then focus on the relative measures for channel efficiencies. We do not fully display η in its original form, instead we estimate its upper bound and lower bound (Detailed derivations in the appendix). All bound estimations are conducted with $k \geq 0$, $\rho_0 \geq \frac{1}{2}\rho_i \geq 0$ and $b_i \geq 0$, i=1, 2. The lower bounds (upper bounds) are corresponding to the worst (best) situation of the channel efficiencies. For example, the lower bound 5/9 of $\tilde{\eta}$ suggests that, when the retailer-specific participation strategy is adopted, the distribution channel is, theoretically speaking, at risk of attaining a low efficiency of mere 55.6%. For another example, the lower bound 0 of $\hat{\eta}$ suggests that the uniform participation strategy can be extremely inefficient if $\rho_1 = 0$ and $b_1 = \rho_2 = 2\rho_0 = k \to +\infty$. The explanation is quite straight forward. When the above conditions are met, it is reasonable to postulate that: 1) R_1's gross margin is smaller than R_2 by far, and 2) R_1's market size is larger than R_2 by far. According to property (ii) of Lemma 2, M's uniform participation rate $\hat{t}^D$ tends to be one, which will give R_2 a *big* free-ride, because M's participation rate with R_2 *would have been* zero under M's retailer-specific participation strategy. It is, indeed, the negative impact of R_2's free-riding that drives the system's efficiency to zero.

5. Concluding remarks

This study is strongly motivated by the scarcity of the modeling works which explicitly consider co-op advertising in a multiple-retailer's environment. So in this paper, as our main contribution, a game theoretical model with multiple retailers is developed to explore the impacts of the retailer's multiplicity on channel members' optimal decisions and on the total channel efficiency. We examine and quantify the impacts in two different aspects. It is observed that: 1) With multiple symmetric retailers, when the number of retailers scales up, the manufacturer's national advertising investment contributes increasingly to add to channel members' profits, but the total channel efficiency goes down quickly and converges to a certain value; 2) With multiple asymmetric retailers (differentiated by gross margins and market sizes), the manufacturer's uniform participation strategy can be inefficient due to the retailer's free-riding. We show that the manufacturer's optimal decision under the uniform participation constraint tends to impose inappropriate incentives to retailers, and then discuss how the detrimental impacts of this distortion can be mitigated and corrected by adopting the retailer-specific participation strategy. For both the symmetric and asymmetric-retailer cases, we solve the games all in closed form and explicitly estimate efficiency bounds.

For future researches, we would first like to extend the current study to consider retailers' advertising/pricing competitions. Another interesting attempt would be to design some mechanism to coordinate the distribution channel. Finally, all models proposed in this paper are deterministic in nature. It would also be interesting to investigate how uncertainties and dynamics could change the conclusions of this paper.

Figures and Tables

Figure 1. Scaling effect on the system's efficiency with multiple symmetric retailers when parameters are specified as $\rho_0 = 0.5$, $\rho_1 = 1$, and k=1

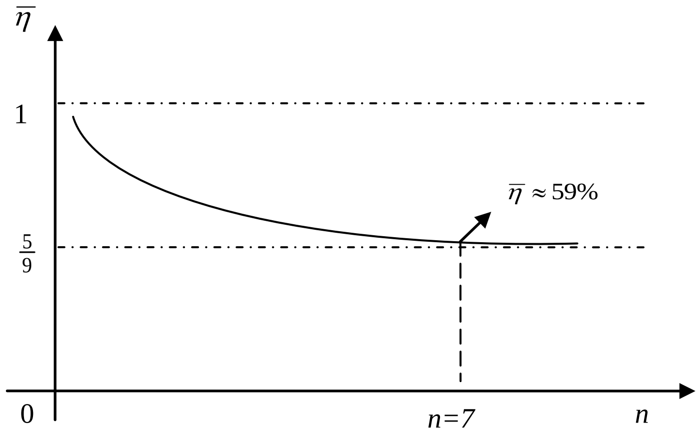

Figure 2. Comparison of M's optimal participation rates with R_1 and R_2 under both the uniform and retailer-specific participation strategies

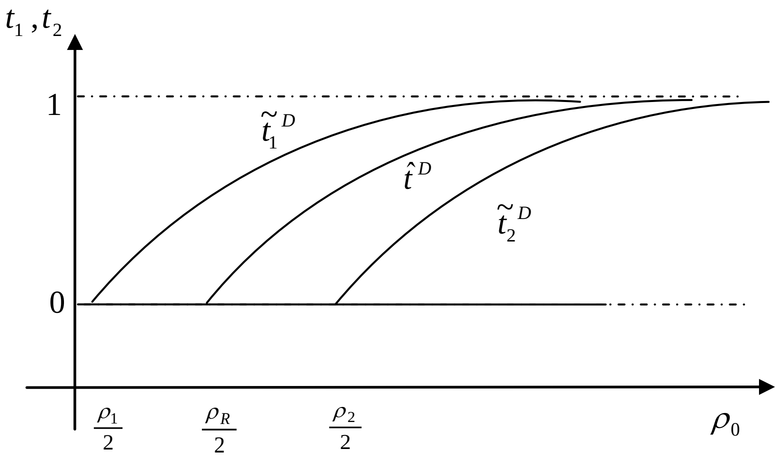

Table 1. Channel performance and system efficiency estimation with two asymmetric retailers

	Retailer-specific participation strategy	Uniform participation strategy
Participation rates	$\tilde{t}_1^D = \frac{2\rho_0 - \rho_1}{2\rho_0 + \rho_1}$ $\tilde{t}_2^D = \frac{2\rho_0 - \rho_2}{2\rho_0 + \rho_2}$	$\hat{t}^D = \frac{2\rho_0 - \rho_R}{2\rho_0 + \rho_R}$
Local advertising	$\tilde{a}_1^D = \frac{1}{16} b_1^2 k^2 (2\rho_0 + \rho_1)^2$	$\hat{a}_1^D = \frac{1}{16} \frac{\rho_1^2 b_1^2 k^2 (2\rho_0 + \rho_R)^2}{\rho_R^2}$
	$\tilde{a}_2^D = \frac{1}{16} b_2^2 k^2 (2\rho_0 + \rho_2)^2$	$\hat{a}_2^D = \frac{1}{16} \frac{\rho_2^2 b_2^2 k^2 (2\rho_0 + \rho_R)^2}{\rho_R^2}$

National advertising	$\tilde{a}_0^D = \frac{1}{4}(b_1 + b_2)^2 \rho_0^2$	$\hat{a}_0^D = \frac{1}{4}(b_1 + b_2)^2 \rho_0^2$
Efficiency by absolute measure	$\tilde{\delta} = \frac{1}{16}\sum_{i=1}^{2} k^2 \rho_i^2 b_i^2 + \frac{1}{4}\left(\sum_{i=1}^{2} \rho_i b_i\right)^2$	$\hat{\delta} = \frac{1}{16}\sum_{i=1}^{2} k^2 \rho_i^2 b_i^2 + \frac{1}{4}\left(\sum_{i=1}^{2} \rho_i b_i\right)^2 + \frac{1}{4}\frac{\rho_0 b_1^2 b_2^2 k^2 (\rho_1 - \rho_2)^2}{\rho_1^2 b_1^2 + \rho_2^2 b_2^2}$
Efficiency by relative measure	$\tilde{\eta} \in [\frac{5}{9}, 1]$	$\hat{\eta} \in [0, 1]$

Appendix

Proof of Proposition 1. The first order condition $\nabla\Pi_S(a_0,\boldsymbol{a})=0$ is equivalent to

$$\frac{\partial \Pi_S}{\partial a_0}=\frac{1}{2}b_1 n(\rho_0+\rho_1)a_0^{-\frac{1}{2}}-1=0 \tag{A1}$$

$$\frac{\partial \Pi_S}{\partial a_i}=\frac{1}{2}b_1 k(\rho_0+\rho_1)a_i^{-\frac{1}{2}}-1=0,\ i=1,\ 2,\ \ldots,\ n. \tag{A2}$$

Solving the above two Equations, we obtain (7)-(8) as the unique centralized solution.

Proof of Lemma 1. Given M's decision of a_0 and t_1, the n retailers will move simultaneously to choose their local advertising levels that optimize their individual profits Π_i, i=1, 2, …, n. According to Equation (11), Π_i is independent with a_j, $j\neq i$, so $a_i^*=\frac{1}{4}\rho_1^2 b_1^2 k^2/(1-t_1)^2$ which maximizes Π_i should be the dominant strategy for R$_i$. a_i^* is also unique because Π_i is strictly concave with a_i, so we obtain the unique Nash equilibrium as shown by Equation (12).

Proof of Proposition 2. According to Lemma 1, the retailers' Nash equilibrium is given by $a_i^*(a_1^*,\ldots a_{i-1}^*,a_{i+1}^*\ldots,a_n^*\mid t_1)=\frac{1}{4}\rho_1^2 b_1^2 k^2/(1-t_1)^2$, i=1, 2, …, n. Substitute a_i^*, i=1, 2, …, n, into M's profit function, we get

$$\Pi_0(a_0,t_1)=n\rho_0 b_1\left(\sqrt{a_0}+\frac{\rho_1 b_1 k^2}{2(1-t_1)^2}\right)-n\frac{\rho_1^2 b_1^2 k^2 t_1}{4(1-t_1)^2}-a_0, \tag{A3}$$

The first order conditions of $\Pi_0(a_0,t_1)$ in (A3) on a_0 and t_1 are, respectively,

$$\frac{\partial \Pi_0}{\partial a_0}=\frac{1}{2}\rho_0 n b_1 a_0^{-\frac{1}{2}}-1=0, \tag{A4}$$

$$\frac{\partial \Pi_0}{\partial t_1}=-\frac{1}{4}\frac{n b_1^2 k^2\rho_1(\rho_1-2\rho_0+t_1\rho_1+2t_1\rho_0)}{(1-t_1)^3}=0, \tag{A5}$$

Solving (A4)-(A5), we get (13) and (15). Obviously, Π_0 in Equation (A3) is strictly concave with a_0 for any given t_1, so $\bar{a}_0^D$ in (13) is the unique optimal national advertising level that maximizes M's profit. According Equation (A5), if $2\rho_0-\rho_1>0$, then Π_0 will increase with t_1 when $0<t_1<\frac{2\rho_0-\rho_1}{2\rho_0+\rho_1}$ and decreases with t_1 when $\frac{2\rho_0-\rho_1}{2\rho_0+\rho_1}<t_1<1$; However, if $2\rho_0-\rho_1<0$, then Π_0 will decrease with all $0<t_1<1$. So M's optimal participation rate will be $\bar{t}_1^D=\frac{2\rho_0-\rho_1}{2\rho_0+\rho_1}$ if $2\rho_0-\rho_1>0$ and $\bar{t}_1^D=0$ if $2\rho_0-\rho_1<0$. Here we are only interested in the former case where M's optimal participation rate takes positive values. When $\rho_0>\frac{1}{2}\rho_1$, by substituting (15) into (12) we get Equation (14).

Proof of Proposition 3.

(i) Take the first order and second order derivatives of $\overline{\eta}$ with respect to n, we have

$$\frac{\partial\overline{\eta}}{\partial n}=-\frac{3}{4}\frac{k^2\rho_1^2}{(\rho_1+\rho_0)^2(k^2+n)^2}<0, \tag{A6}$$

$$\frac{\partial^2\overline{\eta}}{\partial n^2}=\frac{3}{2}\frac{k^2\rho_1^2}{(\rho_1+\rho_0)^2(k^2+n)^3}>0. \tag{A7}$$

So, for any given ρ_0, ρ_1 and *k*, $\bar{\eta}$ is a convex decreasing function of *n*. Let *n* goes to infinity, we have $\lim_{n\to\infty}\bar{\eta} = \rho_0(\rho_0+2\rho_1)/(\rho_0+\rho_1)^2$.

(ii) For any given *n*, ρ_0 and ρ_1, we have

$$\frac{\partial\bar{\eta}}{\partial k} = \frac{3}{2}\frac{nk\rho_1^2}{(\rho_1+\rho_0)^2(k^2+n)^2} > 0\,. \tag{A8}$$

Let $r = \rho_0/\rho_1$, then for any given *n* and *k*, we also have

$$\frac{\partial\bar{\eta}}{\partial r} = \frac{1}{2}\frac{k^2+4n}{(1+\rho_0)^3(k^2+n)} > 0\,. \tag{A9}$$

Since $\bar{\eta}$ increases with *k* for all *n*, ρ_0 and ρ_1, so $\bar{\eta}$ attains its lower bound when *k* reaches down to 0, that is $\bar{\eta}\big|_{k=0} = \rho_0(\rho_0+2\rho_1)/(\rho_0+\rho_1)^2$, which has a minimum value of 5/9 when $\rho_0 = \frac{1}{2}\rho_1$. On the other hand, $\bar{\eta}$ increases with *r* for all *n* and *k*, so when $\rho_0 \to \infty$, $\bar{\eta}$ touches its upper bound 1.

Proof of Proposition 4. We use backward-induction. At the second stage of the game, for any given a_0 and t_1, R_i's profit function is strictly concave with a_i, so R_i's optimal response $a_i^*(a_0,t_1)$ is uniquely obtained as:

$$a_i^*(a_0,t_1) = \frac{1}{4}\frac{\rho_i^2 b_i^2 k^2}{(1-t_1)^2}\,,\ i=1,\ 2. \tag{A10}$$

Substituting $a_i^*(a_0,t_1)$, *i*=1, 2, into Equation (20), we get M's profit function as

$$\Pi_0(a_0, a_1^*(a_0,t), a_2^*(a_0,t), t_1) = \rho_0\sum_{i=1}^{2} b_i\left(\sqrt{a_0} + \frac{1}{2}\frac{\rho_i\, b_i\, k^2}{(1-t_1)}\right) - \frac{t_1}{4}\sum_{i=1}^{2}\frac{\rho_i^2 b_i^2 k^2}{(1-t_1)^2} - a_0\,. \tag{A11}$$

Take first order conditions of (A11) on a_0 and t_1, respectively, we have

$$\frac{\partial\Pi_0}{\partial a_0} = \frac{1}{2}\rho_0(b_1+b_2)a_0^{-\frac{1}{2}} - 1 = 0\,, \tag{A12}$$

$$\frac{\partial\Pi_0}{\partial t_1} = -\frac{k^2}{4}\frac{\left(2\rho_0(b_1^2\rho_1+b_2^2\rho_2)+b_1^2\rho_1^2+b_2^2\rho_2^2\right)\cdot t_1 - 2\rho_0(b_1^2\rho_1+b_2^2\rho_2) - b_1^2\rho_1^2 - b_2^2\rho_2^2}{(1-t_1)^3} = 0\,. \tag{A13}$$

Solving (A12)-(A13), we get (23) and (26). Next, we will confirm that (23) and (26) gives the unique solution that maximizes M's profit. Π_0 in Equation (A11) is additive separable with respect to a_0 and t_1, so we may deal with the two variables separately. Obviously, Π_0 in Equation (A11) is strictly concave with a_0 for any given t_1, so $\bar{a}_0^D$ in Equation (23) is the unique optimal national advertising level. Then we check how Π_0 changes with t_1. According Equation (A13), if $2\rho_0 - \rho_R \ge 0$, then Π_0 will increase with t_1 when $0 \le t_1 \le \frac{2\rho_0-\rho_R}{2\rho_0+\rho_R}$ and decreases with t_1 when $\frac{2\rho_0-\rho_R}{2\rho_0+\rho_R} \le t_1 \le 1$; However, if $2\rho_0 - \rho_R \le 0$, then Π_0 will decrease with all $0 \le t_1 \le 1$. So M's unique optimal participation rate will be $\hat{t}_1^D = \frac{2\rho_0-\rho_R}{2\rho_0+\rho_R}$ if $\rho_0 \ge \frac{1}{2}\rho_R$ and $\hat{t}_1^D = 0$ if $\rho_0 \le \frac{1}{2}\rho_R$. Here we are only interested in the former case where M's optimal participation rate takes positive values. When $\rho_0 \ge \frac{1}{2}\rho_R$, by substituting (26) into (A10) we obtain (24)-(25).

Bound estimation for channel efficiencies by relative measure in Table 1

Without loss of generality, we assume $\rho_1 \le \rho_2$. Then the condition $\rho_0 \ge \frac{1}{2}\rho_i$, *i*=1, 2 is equivalent to

$\rho_0 \geq \frac{1}{2}\rho_2$, and the parameter ranges will be $k \geq 0$, $\rho_0 \geq \frac{1}{2}\rho_2 \geq \frac{1}{2}\rho_1 \geq 0$, and $b_i \geq 0$, i=1, 2. First of all, we consider the case when M adopts the retailer-specific participation strategy. The channel efficiency by relative measure is

$$\tilde{\eta} = \tilde{\Pi}_S^D / \Pi_S^C = 1 - \frac{1}{4} \cdot \left[(\rho_1^2 b_1^2 + \rho_2^2 b_2^2)k^2 + 4(\rho_1 b_1 + \rho_2 b_2)^2\right] \cdot \Big[\rho_0^2\left((b_1^2 + b_2^2)k^2 + (b_1 + b_2)^2\right) + 2\rho_0\left((\rho_1 b_1^2 + \rho_2 b_2^2)k^2 + (b_1 + b_2)(\rho_1 b_1 + \rho_2 b_2)\right) + (\rho_1^2 b_1^2 + \rho_2^2 b_2^2)k^2 + (\rho_1 b_1 + \rho_2 b_2)^2\Big]^{-1}. \quad \text{(A14)}$$

Since

$$\partial\tilde{\eta} / \partial\rho_0 = \frac{1}{2} \cdot \left[(b_1^2(\rho_0 + \rho_1) + b_2^2(\rho_0 + \rho_2))k^2 + (b_1 + b_2)(b_1(\rho_0 + \rho_1) + b_2(\rho_0 + \rho_2))\right] \cdot \left[(\rho_1^2 b_1^2 + \rho_2^2 b_2^2)k^2 + 4(\rho_2 b_2 + \rho_1 b_1)^2\right] \cdot \left[(b_1^2(\rho_0 + \rho_1)^2 + b_2^2(\rho_0 + \rho_2)^2)k^2 + (b_1(\rho_0 + \rho_1) + b_2(\rho_0 + \rho_2))^2\right]^{-2} > 0, \quad \text{(A15)}$$

so $\tilde{\eta}$ increases in ρ_0 for all $\rho_2 \geq \rho_1 \geq 0$, $b_1, b_2 \geq 0$ and $k \geq 0$. Then the upper bound of $\tilde{\eta}$ must be 1, because $\tilde{\eta}\big|_{\rho_0 \to +\infty} = 1$. On the other hand, the lower bound of $\tilde{\eta}$ must be attained at $\rho_0 = \frac{1}{2}\rho_2$, where Equation (A14) will be reduced to

$$\tilde{\eta}\big|_{\rho_0 = \frac{1}{2}\rho_2} = \left[(4\rho_1\rho_2 b_1^2 + 3\rho_1^2 b_1^2 + \rho_1^2 b_2^2 + 8\rho_2^2 b_2^2)k^2 + \rho_2(b_1 + b_2)(5\rho_2 b_2 + \rho_2 b_1 + 4\rho_1 b_1)\right] \cdot \left[(4\rho_1\rho_2 b_1^2 + \rho_2^2 b_1^2 + 4\rho_1^2 b_1^2 + 9\rho_2^2 b_2^2)k^2 + (3\rho_2 b_2 + \rho_2 b_1 + 2\rho_1 b_1)^2\right]^{-1}. \quad \text{(A16)}$$

Take the first order derivative of (A16) with k, we find that it will be increasing with k for all $\rho_2 \geq \rho_1 \geq 0$ and $b_1, b_2 \geq 0$, then we know that the lower bound of $\tilde{\eta}$ must be attained at $\rho_0 = \frac{1}{2}\rho_2$ and k = 0, where Equation (A16) will be further reduced to

$$\tilde{\eta}\big|_{\rho_0 = \frac{1}{2}\rho_2, k=0} = \frac{\rho_2(b_1 + b_2)(5\rho_2 b_2 + b_1\rho_2 + 4b_1\rho_1)}{(b_1\rho_2 + 3b_2\rho_2 + 2b_1\rho_1)^2}. \quad \text{(A17)}$$

Again, Equation (A17) is increasing in b_1 for all $\rho_1 \geq 0$ and $b_2 \geq 0$, so we can simply take b_1 down to zero to obtain the lower bound of (A17) being 5/9, which should be minimal possible value of $\tilde{\eta}$ with $k \geq 0$, $\rho_0 \geq \frac{1}{2}\rho_2 \geq \frac{1}{2}\rho_1 \geq 0$ and $b_1, b_2 \geq 0$. So the efficiency bounds are estimated as $\tilde{\eta} \in [\frac{5}{9}, 1]$.

Then we consider the case when M adopts the uniform participation strategy. The channel efficiency by relative measure is presented as

$$\hat{\eta} = \tilde{\Pi}_S^D / \Pi_S^C = 1 - \frac{1}{4} \cdot \left[(\rho_1^2 b_1^2 + \rho_2^2 b_2^2 + 4b_1^2 b_2^2 \rho_0^2(\rho_2 - \rho_1)(\rho_1^2 b_1^2 + \rho_2^2 b_2^2)^{-1})k^2 + 4(\rho_1 b_1 + \rho_2 b_2)^2\right] \cdot \left[(b_1^2(\rho_0 + \rho_1)^2 + b_2^2(\rho_0 + \rho_2)^2)k^2 + (b_1(\rho_0 + \rho_1) + b_2(\rho_0 + \rho_2))^2\right]^{-2}. \quad \text{(A18)}$$

One can check that when $\rho_1 = 0$ and $\rho_2 = k = b_1 = 2\rho_0 \to +\infty$, $\hat{\eta}$ can reach down to 0; When $\rho_1 = \rho_2$ and $\rho_0 \to \infty$, $\hat{\eta}$ can rise up to 1. So the efficiency bound is estimated by $\hat{\eta} \in [0,1]$.

References

[1] Bergen, M. & John, G. (1997). Understanding cooperative advertising participation rates in conventional channels. Journal of Marketing Research, 34(3): 357-369

[2] Berger, P.D. (1972). Vertical cooperative advertising ventures. Journal of Marketing Research, 9(3): 309-312

[3] Berger, P.D. & Magliozzi, T. (1992). Optimal co-operative advertising decisions in direct-mail operations. Journal of the Operational Research Society, 43(11): 1079-1086

[4] Chintagunta, P.K. & Jain, D. (1992). A dynamic model of channel member strategies for marketing expenditures. Marketing Science, 11(2): 168-188

[5] Dant, R.P. & Berger, P.D. (1996). Modeling cooperative advertising decisions in franchising. Journal of the Operational Research Society, 47(9): 1120-1136

[6] Davis, R.A. (1994). Retailers open doors wide for co-op. Advertising Age 1, August 30

[7] Elkin, T. (1999). Co-op crossroads. Advertising Age 70: 1 and 24-26

[8] Furman, J. (2005). Wal-Mart: A Progressive Success Story. Washington DC: Center for American Progress, http://www.americanprogress.org/kf/walmart_progressive.pdf

[9] He, X., Prasad, A. & Sethi, S.P. (2009). Cooperative and advertising in a dynamic stochastic supply chain: feedback Stackelberg strategies. Production and Operations Management, 18(1): 78-94

[10] Huang, Z. & Li, S.X. (2001). Co-op advertising models in a manufacturer-retailer supply chain: a game theory approach. European Journal of Operational Research, 135(3): 527-544

[11] Jorgensen, S., Sigue, S.P. & Zaccour, G. (2000). Dynamic cooperative advertising in a channel. Journal of Retailing, 76(1): 71-92

[12] Jorgensen, S. & Zaccour, G. (2003). A differential game of retailer promotions. Automatica, 39(7): 1145-1155

[13] Karray, S. & Zaccour, G. (2007). Effectiveness of coop advertising programs in competitive distribution channels. International Game Theory Review, 9(2): 151-167

[14] Kim, S.Y. & Staelin, R. (1999). Manufacturer allowances and retailer pass-through rates in a competitive environment. Management Science, 18(1): 59-76

[15] Lal, R. (1990). Improving channel coordination trough franchising. Marketing Science, 9(4): 299-318

[16] Lyon, L. (1932). Advertising Allowances. Washington DC: The Brookings Institution

[17] Moulin, H. (1986). Game Theory for the Social Sciences. New York: New York University Press

[18] Nagler, M.G. (2006). An exploratory analysis of the determinants of cooperative advertising participation rates. Marketing Letters, 17(2): 91-102

[19] Nerlove, M. & Arrow, K.J. (1962). Optimal advertising policy under dynamic conditions. Economica, 29(114): 129-142

[20] Simon, J.L. & Arndt, J. (1980). The shape of the advertising function. Journal of Advertising Research, 20(4): 11-28

[21] Sigue, S.P. & Chintagunta, P. (2008). Advertising strategies in a franchise system. European Journal of Operational Research, 198(2): 655-665

[22] Somers, T.M., Gupta, Y.P. & Herriott, S.R. (1990). Analysis of cooperative advertising expenditures: a transfer-function modeling approach. Journal of Advertising Research, 30(5): 35-45

[23] Szmerekovsky, J.G. & Zhang, J. (2009). Pricing and two-tier advertising with one manufacturer and one retailer. European Journal of Operational Research, 192(3): 904-917

[24] Taboubi, S. & Zaccour, G. (2002). Impact of retailer's myopia on channel's strategies. Optimal

Control and Differential Games: Essays in Honor of Steffen Jorgensen, G. Zaccour ed., Kluwer Academic Publishers: 179–192

[25] Taboubi, S. & Zaccour, G. (2005). Coordination Mechanisms in Marketing Channels: a Survey of Game Theory Models. Les Cahiers du GERAD, G-2005-36

[26] Topkis, D.M. (1979). Equilibrium points in nonzero-sum n-person submodular games. SIAM Journal on Control and Optimization, 17(6): 773-787

[27] Xie, J. & Wei, J.C. (2009). Coordinating advertising and pricing in a manufacturer-retailer channel. European Journal of Operational Research, 197(2): 785-791

[28] Young, R.F. & Greyser, S.A. (1983). Managing Cooperative Advertising: A strategic Approach. MA: Lexington Books

[29] Yue, J., Austin, J., Wang, M.C. & Huang, Z. (2006). Coordination of cooperative advertising in a two-level supply chain when manufacturer offers discount. European Journal of Operational Research, 168(1): 65-85