Full Length Article

Reward redemption effects in a loyalty program when customers choose how much and when to redeem

Matilda Dorotic [a,*], Peter C. Verhoef [a,b], Dennis Fok [c], Tammo H.A. Bijmolt [b]

[a] BI Norwegian Business School, Department of Marketing, Norway
[b] University of Groningen, Faculty of Economics and Business, Department of Marketing, The Netherlands
[c] Erasmus University Rotterdam, Erasmus School of Economics, The Netherlands

ARTICLE INFO

ABSTRACT

The redemption of loyalty program (LP) rewards has an important impact on LP members' behavior, particularly on purchase behavior before and after redeeming a reward. However, little is known about the interplay between members' purchase and redemption behavior when members are not pressured with point expiration and they choose for themselves when and how much to redeem. In this context, the effects of redemption are not straightforward, as little additional effort is required from an LP member to obtain the reward. Analyzing the behavior of 3094 members in such an LP, we find that the mere decision to redeem a reward significantly enhances purchase behavior before and after the redemption event, even when members redeem just a fraction of their accumulated points. Conceptually, we refer to this enhancement as the *redemption momentum*, which is an alternative and novel explanation of the existence of pre-reward effects that do not depend on points-pressure. In addition to the overall impact of redemption on purchases, prior purchase behavior also enhances redemption decisions. Finally, we find a number of moderating effects on purchase and redemption behavior that derive from the length of LP membership, age, income and direct mailings. Our study's most important managerial implication is that firms should avoid imposing point expiry and/or binding thresholds in order to enhance members' purchase behavior.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

In recent years, loyalty programs (LPs) have become the dominant tool for loyalty marketing worldwide. In the United States alone, the number of LP memberships exceeded 2.65 billion in 2012, increasing by 26.7% since 2010 (Berry, 2013). LPs aim to engage program members by rewarding their repeated purchases of a firm's product through (the redemption of) loyalty points that members collect on their purchases. Therefore, the benefits of an LP for a member become the most salient when redeeming a reward (Nunes & Drèze, 2006; Smith & Sparks, 2009a). Yet, as much as one-third of $48 billion worth of LP currency issued in 2010 remained unredeemed (Gordon & Hlavinka, 2011); likewise, *The Economist* estimated that "the total stock of unredeemed miles was worth more than all the dollar bills in circulation" (The Economist, 2005). To reduce liability, LPs introduced minimum thresholds and/or point expiration; however, this may undermine loyalty building efforts and engender customer frustration (Land, 2013; Stauss, Schmidt, & Schoeler, 2005). For example, point expiration is common in the airline industry where, due to restrictions on the

availability of "award seats," LP points often expire before members have an opportunity to cash in points (average award seat availability is only about 60% at major airlines (McCartney, 2012)). On the other hand, LPs are increasingly opting for a no-expiration (or long-term expiration) policy to avoid negative customer experiences. For instance, 96% of credit-card programs promote "no expiration" as their key sales feature (Land, 2013). On the other hand, without the expiration pressure to redeem points, firms fear that members' active engagement may decline and that their loyalty will fade in turn. Whether firms should encourage reward redemption and consider long-term expiration policies ranks among the least understood aspects of LPs (CRMtrends, 2012; Shugan, 2005).

Reward redemption may have an important impact on members' behavior, particularly on purchase behavior just before and after redeeming a reward. Having to reach a pre-specified threshold on time to obtain a reward motivates members to increase their expenditures—an effect known as *points pressure* (Taylor & Neslin, 2005). However, if a customer already has enough points or (s)he has too few points to be able to reach the threshold, the points pressure becomes negligible (Hartmann & Viard, 2008; Lewis, 2004). The question, then, is whether firms can expect redemption effects in LPs without significant binding deadlines that "require customers to jump through hoops to receive a reward" (Blattberg, Kim, & Neslin, 2008, p. 566).

* Corresponding author.
 E-mail address: matilda.dorotic@bi.no (M. Dorotic).

Unfortunately, the prevailing theoretical mechanisms to explain such effects are equivocal.

If firm-imposed motivators leading to points pressure are removed, then the presence of redemption effects depends on whether the redemption decision by itself impacts behavior. In LPs with continuous and linear rewarding schemes, members obtain a certain amount of LP currency for each dollar/euro spent and choose when to redeem (redemption timing) and what to redeem (redemption amount), based on their personal reward preferences and the collected balance of points (cf. Stourm, Bradlow, & Fader, 2013). Moreover, in continuous LPs, the program itself and/or its points typically do not expire for a longer period of time (e.g., retail LPs). This context allows us to investigate whether redemption effects on behavior in pre- and post-reward period can be evoked by the act of redeeming itself in the absence of firm-imposed thresholds. The decision to redeem points may precede the moment at which the reward is redeemed or it may occur at a point-of-sales without much prior planning, which has direct consequences on behavior.

Analyzing the purchase and redemption behavior of 3094 members in a Dutch continuous LP, we find that in as much as 70% of redemptions, the decision to redeem is made a short time ahead of the redemption. Having made the decision motivates customers within the LP, resulting in an increase in purchase behavior prior to the redemption event, even when customers subsequently redeem just a small fraction of their overall point balance. We label this effect *redemption momentum* and note that this effect complements the *points pressure* effect, which may occur for members who have an insufficient amount of points in the weeks before a redemption.

In the post-reward period, the redemption enhances feelings of gratitude, importance, satisfaction or obliged reciprocity, which may in turn spur purchase behavior (Palmatier, Jarvis, Bechkoff, & Kardes, 2009). However, empirical findings on the post-reward effects on members' behavior are scarce and the results are mixed in the literature. In some cases, points pressure shifts purchases in time and creates post-redemption dips due to stockpiling. This is not expected to occur when members can choose timing and redemption amounts. Our study provides support for positive post-reward effects when customers do not face binding deadlines and can choose the redemption timing and amount.

Finally, redemption effects on purchase behavior may vary across LP members (Kopalle, Sun, Neslin, Sun, & Swaminathan, 2012; Stourm et al., 2013; Zhang & Breugelmans, 2012). In particular, the effects may be moderated by members' prior experience with the LP (length of LP membership) and various socio-demographic aspects (age, income, etc.), as well as the amount of direct mailing promotions that members obtain (Lewis, 2004). Yet, those interaction effects have not been extensively investigated. In response, we provide an integrated analysis of the main and interaction effects.

In summary, the contribution of this paper is threefold. First, we explore whether LPs can foster redemption effects without imposing restrictive deadlines. To this end, we examine alternative mechanisms that drive (pre-)redemption effects and propose the novel *redemption momentum* mechanism, which goes beyond the traditional points pressure explanations. Second, this study tackles the interrelatedness of purchase and redemption decision-making by simultaneously modeling purchase incidence, purchase amount, redemption decision and redemption amount. Moreover, our model studies the interplay between redemption and purchases, accounting both for endogeneity of redemption and endogeneity of personalized mailings to LP members. Third, this study provides an integrated analysis of potential moderating effects, such as relationship length, socio-demographics and direct mailings, on the relationship between redemption and purchases. In this way, our paper answers the call to simultaneously model diverse LP mechanisms to better understand the underlying processes and sources of incremental sales in LPs (Blattberg et al., 2008; Kopalle et al., 2012).

The paper proceeds by discussing the theoretical background and existing studies on the effects of reward redemption. It then continues with the model formulation, a description of the data, the empirical analyses and the results. We conclude with a discussion of key findings and managerial implications.

2. Prior literature

Marketing literature has extensively studied the effects of LPs on customer behavior (Leenheer, van Heerde, Bijmolt, & Smidts, 2007; Liu, 2007). A synthesis of available evidence indicates that, overall, LPs enhance LP members' behavior (Dorotic, Bijmolt, & Verhoef, 2012) through increases in purchase volume/frequency (Drèze & Hoch, 1998; Lewis, 2004; Liu, 2007; Taylor & Neslin, 2005) and share of wallet at the LP provider (Leenheer et al., 2007; Verhoef, 2003). However, the role that reward redemption itself plays in this increase is not clear. Existing research on LP rewards has mainly focused on the attractiveness of different reward types and their impact on profitability (Kim, Shi, & Srinivasan, 2001; Kivetz & Simonson, 2002; Zhang, Krishna, & Dhar, 2000), while reward redemption effects themselves have received relatively less attention (Dorotic et al., 2012; Smith & Sparks, 2009a).

Below we separately review the literature on three key aspects: pre-reward effects, post-reward effects, and the impact of mailings and other main moderators. Table 1 provides an overview of (selected) prior research, summarizes their main findings, and positions our study.

2.1. Pre-reward effects

Literature to date almost exclusively links pre-reward effects to the goal-pursuit theory and the *points pressure mechanism* (Kivetz et al., 2006; Kopalle et al., 2012; Taylor & Neslin, 2005). Points pressure suggests that pre-reward effects are driven by members' anticipation of obtaining future rewards and/or by switching costs, which together constitute the pressure to collect a sufficient amount of points for a reward (Hartmann & Viard, 2008; Kopalle et al., 2012; Lewis, 2004). Researchers provide evidence of pre-reward effects in short-term LPs, in which members must reach a spending threshold during a time-limited period to obtain a pre-specified reward (e.g., "Spend X on groceries within 3 months, get a free turkey" or "Buy 10, get 1 free") (Kivetz et al., 2006; Lal & Bell, 2003; Taylor & Neslin, 2005). In such sales promotion-like LPs, the points pressure is high due to the high potential sunk costs and saliency of explicit goals.

In continuous LPs, empirical support for pre-reward effects is found for those LPs with distinctive customer tiers (Drèze & Nunes, 2011; Kopalle et al., 2012) and for retailers with specific, firm-defined redemption thresholds (Lewis, 2004; Zhang & Breugelmans, 2012). These studies reaffirm that pre-reward effects occur through explicit threshold reward structures set by a firm (e.g., LP tiers or "for each 500 collected points that customers obtain a voucher/discount"). Such a known external threshold may induce pressure to build up purchases to reach the threshold, thereby spurring the points pressure.

Nonetheless, Smith and Sparks (2009a) found that in a typical continuous retail LP, where customers endogenously choose how much and when to redeem, only the smallest group of analyzed redeemers (approximately 10%) demonstrated a planning behavior of saving points in order to reach a higher-value reward. The majority of redemptions seemed to be driven by the notion of rewarding and treating oneself from the accumulated balance, sometimes on impulse (Smith & Sparks, 2009a,b). Moreover, recent psychological insights indicate that goal-pursuit may not be the only mechanism driving LP behavior (Henderson, Beck, & Palmatier, 2011; Wiebenga & Fennis, 2014). The findings of Stourm et al. (2013) indicate that in the absence of firm-driven restrictions on the amount and timing of redemption, members may form latent thresholds of redemption based on their subjective perceptions of their points' value relative to cash. Therefore, the points-pressure mechanism alone may not be sufficient in explaining

Table 1
Effects of reward redemption: Overview of findings and study positioning.

	Study	Design	Purchase incidence	Spending	Pre-reward effects	Post-reward effects	Personalized mailings	Moderators of rewarding effects
Short-term LPs	Kivetz et al. (2006)	Experimental	✔	×	+	0	+	×
	Nunes and Drèze (2006)	Experimental	✔	×	+	×	×	Effect stronger if presented in earned points (vs. purchases)
	Lal and Bell (2003)	Empirical	×	✔	+	+	×	Effects strongest for low baseline spenders
	Taylor and Neslin (2005)	Empirical	×	✔	+	+	×	Effects strongest for low baseline spenders
Continuous LPs	Lewis (2004) *Reward thresholds observed*	Empirical (Online retailing)	✔	✔	+	×	+	Effects strongest for high spenders Positive impact on purchase behavior of mailing a discount voucher
	Kopalle et al. (2012) *Reward thresholds observed, points expire*	Empirical (Hotel chain)	✔	✔	+	0/+ (only for price sensitive segment)	×	2 redemption options (free stay vs. free upgrade)
	Drèze and Nunes (2011) *Reward thresholds observed, points expire*	Empirical (Airlines)	✔	✔	+	0 (resetting but not to the initial level)	×	×
	This study *Reward thresholds unobserved and heterogeneous, points do not expire*	Empirical (Retailing)	✔	✔	✔	✔	✔	Income age Relationship duration

Notes: ✔ analyzed effect; × effect not analyzed; + positive effect; 0 post-reward dip (resetting).
The study by Zhang and Breugelmans (2012) is not included in this table because the design of the analyzed LP is not directly comparable.

the impact of redemption on pre-reward purchase behavior. We posit that in the absence of external thresholds (points pressure), members form internal, latent states that affect their behavior before and after redemption, as explained in the subsequent sections.

2.2. Post-reward effects

Post-reward effects are mostly attributed to the *rewarded-behavior mechanism* (Blattberg et al., 2008; Taylor & Neslin, 2005). Reward redemption enhances subsequent purchase frequency and volume through behavioral learning that ties repurchases to rewards (Rothschild & Gaidis, 1981). Furthermore, a reward obtained through an LP can evoke the belief of a windfall gain or good deal (Arkes et al., 1994; Smith & Sparks, 2009b), a sense of appreciation from the firm that evokes reciprocal feelings (e.g., gratitude, indebtedness) in customers (Gwinner, Gremler, & Bitner, 1998; Palmatier et al., 2009), a sense of belongingness (Dowling & Uncles, 1997), or an elevated sense of status (Drèze & Nunes, 2009). Therefore, reward redemption may induce positive post-reward effects by reinforcing attitudinal attachment, which then affects purchase behavior (Haisley & Loewenstein, 2011; Taylor & Neslin, 2005). This post-reward effect is instrumental for building long-term relationships with LP members (Kumar & Shah, 2004; Palmatier et al., 2009).

However, the empirical support for post-reward effects is mixed. Some studies reveal positive post-reward effects on purchase behavior in short-term LPs, albeit mainly among light users or for particular types of rewards (Lal & Bell, 2003; Roehm, Pullins, & Roehm, 2002; Taylor & Neslin, 2005). Kivetz et al. (2006) found no support for such effects in an experimental study; they instead found evidence for post-reward resetting (i.e., a dip in the purchase behavior after redeeming a reward when purchases return to their pre-reward baseline levels). In a continuous LP setting, Drèze and Nunes (2011) also found post-reward resetting in an airline LP, but not to the initial level, which implies some positive post-reward effect. However, they studied a customer tier program, where reaching a higher tier entitles members to preferential treatment and higher status. It is therefore hard to judge whether increased baseline behavior after redemption is due to the new benefits or the redemption itself. In a similar setting, Kopalle et al. (2012) did not find the rewarded behavior effect for a customer tier-oriented segment of members in a hotel LP. Conversely, the study

found a positive post-reward effect for the price-sensitive segment attracted to free hotel stays.

Table 1 provides an overview of these mixed research findings and highlights the need for additional empirical evidence in continuous reward settings where members do not have to increase their effort in pre-reward periods and, consequently, they may not feel a particular sense of accomplishment after redeeming.

2.3. Moderating effects of mailings, length of membership and socio-demographics

It is beneficial for LP providers to leverage the information that they have and target members with personalized mailings (Blattberg et al., 2008; Lewis, 2004). However, the current literature lacks a systematic examination of the impact of personalized marketing efforts on reward redemption behavior (Blattberg et al., 2008). Yet, it is important to control for the impact of mailings on members' purchase and redemption behavior in order to accurately delineate the influence of various other drivers (like goal attainment and points pressure). A complicating factor is that the possible target selection by the LP makes the mailings an endogenous decision. Such endogeneity needs to be taken into account when mailings are included as a driver of purchase behavior.

Beside mailings, various individual characteristics may influence the interplay between redemption and purchase. Members respond differently to LPs depending on their usage or spending levels (Kim et al., 2001; Liu, 2007), their experience with the LP (e.g., length of LP membership) (Bolton, Kannan, & Bramlett, 2000), or socio-demographic characteristics (Leenheer et al., 2007; Lemon & von Wangenheim, 2009; Magi, 2003).

The impact of socio-demographic differences in LPs is still ambiguous (Dorotic et al., 2012). In particular, little knowledge exists on the moderating impacts of socio-demographics and the length of LP membership on pre- and post-reward effects. Differences in individual characteristics may influence the size of the reward redemption effects: higher income members have greater purchasing power and may therefore be more flexible with their purchasing levels and respond more strongly to reward incentives. Additionally, long-term members have more experience with the LP, which may lead to higher responsiveness to the LP (Bolton et al., 2000).
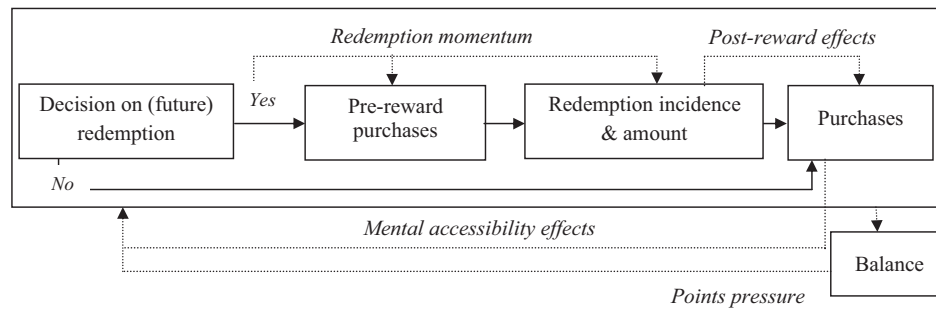
Fig. 1. Conceptual model of the interplay between redemption and purchases when customers choose what and when to redeem in a continuous, linear LP.

3. Conceptualization of the interplay between redemption and purchase

To understand the interplay between redemption and purchase behavior in a continuous and linear rewarding context, it is important to enrich the existing explanations in order to account for diverse motivations and bidirectional relationships. Rewarding may affect purchase behavior, while purchases (i.e., point collection) may in turn affect redemption. In this context, the sequence of decision-making concerning redemptions may help to explain the reward redemption effects, as illustrated in Fig. 1.

This figure outlines the sequence of decision-making that guides our research. The solid arrows indicate the decisions that members make: from the decision to redeem to the purchases after redemption. The dashed arrows, pointing at the box surrounding the process, indicate the influence of the related concept on all aspects of the process (e.g., the overall influence of an accumulated point balance).

If members have a choice to redeem all or just a fraction of their accumulated balance of LP points without being pressured or incurring sunk cost, then a potential increase in purchase behavior in the pre-redemption period is driven by an internal state rather than the points pressure. We posit that the decision to redeem a reward may by itself act as a driver of pre-redemption effects. We coined the term *redemption momentum* to refer to the redemption decision's impact on purchase behavior. The redemption momentum is active from the point in time that a redemption is planned until it occurs. The decision to redeem a reward may precede the actual redemption and induce excitement for and salience of the benefits of LP membership. This in turn may increase motivation and enhance purchase behavior before the actual redemption takes place. Applied to the LP setting, the situational benefit salience (cf. Petty & Cacioppo, 1979; Ratneshwar, Warlop, Mick, & Seeger, 1997) may refer to a temporary increase in the salience of redeeming points for a reward, which may originate from the anticipation of a specific usage situation related to the redemption (e.g., a decision to redeem points for a visit to an amusement park that reinforces the subsequent motivation for utilizing the program). Dhar, Huber, and Khan (2007) found support for a similar shopping momentum effect where the propensity of subsequent purchases is enhanced merely by an initial decision to purchase. Once the redemption decision has been made by a member, the actual redemption will typically follow within a short period of time. At the redemption event, the redemption momentum may still exist because customers can make a decision to redeem a reward in the same week as when s/he makes a purchase, or even during the purchase trip itself. After the redemption event, the post-reward effects may enhance behavior, like elaborated earlier.

Previous discussion outlines the impact of redemption decision on purchase behavior. However, purchase behavior may also affect the likelihood of redemption. Since points are directly related to purchases in an LP, obtaining points and bolstering one's balance increase awareness of the LP, i.e., increases the accessibility of the LP in memory (cf. Higgins, 1989). This in turn increases the likelihood of redeeming one's collected points. At each point-saving event, the LP becomes more mentally represented (accessible) since the customer is reminded of the LP. If the LP is accessible in the members' minds, a positive redemption decision becomes more likely. In case a member does not make a purchase in a particular week, and therefore does not obtain LP points, the mental accessibility of the LP decreases. In summary, purchase behavior increases the probability that a redemption decision will be made, which may in turn lead to the redemption momentum effects on subsequent purchases, as illustrated in Fig. 1.

4. Data description

4.1. Loyalty program description

The data for our study are derived from a nationwide coalition LP in The Netherlands. Program members can collect points by purchasing at more than ten LP partners, including both online and offline retailers, as well as service providers. Participating vendors function in the grocery retail, gas retail, insurance, and travel agency industries, among other sectors. The number of points awarded reflects spending amounts, and one LP point equates on average to a euro spent. Given that we are primarily interested in insights at the LP level (i.e., interplay between redemption and purchase behavior within the LP rather than for individual vendors), we aggregate points saved and redeemed across LP vendors.[1]

The LP provider runs periodic promotions in order to allow members to collect additional LP points or to encourage them to redeem the promoted awards. The promotions are personalized and mailed to members, highlighting accumulated points and promotional offers.

Members can redeem points for a variety of awards, ranging from kitchen utensils to travel and holidays. Therefore, the available redemption options are very heterogeneous and range from very small amounts to large awards like holiday packages. At any time, LP members can decide to redeem any amount from their accumulated balance of points to obtain rewards. Collected points do not expire.

4.2. Data and descriptive statistics

We analyze longitudinal weekly data on members' collection of loyalty points and redemptions over the course of three and a half years (184 weeks). The weekly purchase behavior reflects the number of points collected, aggregated across LP vendors per member. The LP membership card provides information on socio-demographic characteristics (age and household income) and the date that each member joined the LP. The final sample contains information on the behavior of 3094 LP members over 184 weeks. Selected members have to show at least 30 purchases and at least one redemption within the observation period. The first 10 weeks are used to initialize some dynamic

---

[1] Points are not vendor-specific and so redemption does not depend on members saving points from a particular vendor. Also, the coalition LP does not include competitors among vendors from the same industry (it rather has complementary vendors), so point saving at one vendor does not attenuate purchases at other vendors.
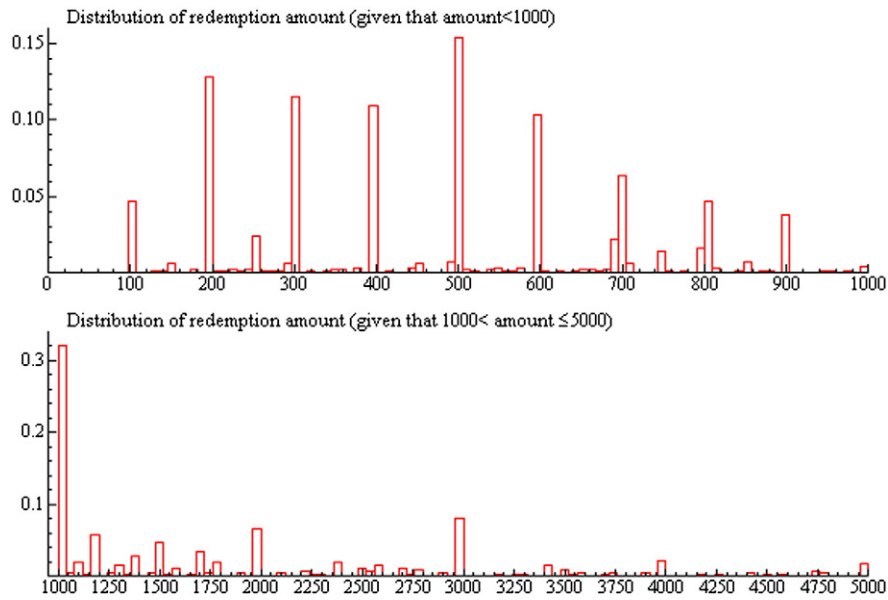
Fig. 2. Distribution of redemption amounts.

variables. To initialize the post-reward variable, we also make use of redemption data prior to the start of our estimation sample. In fact, we have information on redemptions up to 560 weeks before the start of our sample. Such prior data is not available for purchases.

On average, LP members made 0.72 purchases with the LP card per week and redeemed rewards once every 10 months (42 weeks). On average, members received 0.59 mailings per week (ranging from 0 to 2 per week across members). An average member has participated in the LP for more than 11 years, is 49 years of age, and earns a disposable annual income close to the average for The Netherlands (€17,000; Statistical Yearbook of the Netherlands, 2009).

Remarkably, there is a large variation in the number of points redeemed at a particular redemption. Although the majority of rewards obtained are worth less than 5000 points, the right-hand tail of the distribution reaches up to 60,000 points. Fig. 2 depicts the frequency distributions of the redemption amounts conditional on the amount being less than 1000; and conditional on the amount being between 1000

and 5000. The figure shows large variability in the selected (internal) redemption thresholds among LP members and yet it also indicates that certain amounts are much more common than others.

The interplay between the redemption amounts and the available points for redemption (balance) at the redemption occasion is critical to understanding the (theoretical) drivers of pre- and post-rewarding effects. In Fig. 3 we compare the empirical distribution of the redemption amount to the distribution of the number of points available at each moment in time across all members. Note that the horizontal axis has a log-scale. The distribution of the number of available points is clearly to the right of the distribution of the redemption amounts. Further investigation shows that, on average, a member spends 26% of his/her balance of points upon redemption. In only 3% of the cases is more than 90% of the accumulated balance spent. Therefore, in almost all redemption occasions, redeemers utilize much fewer points than they have at their disposal. This indicates that possible purchase acceleration in the pre-reward periods cannot occur purely due to the
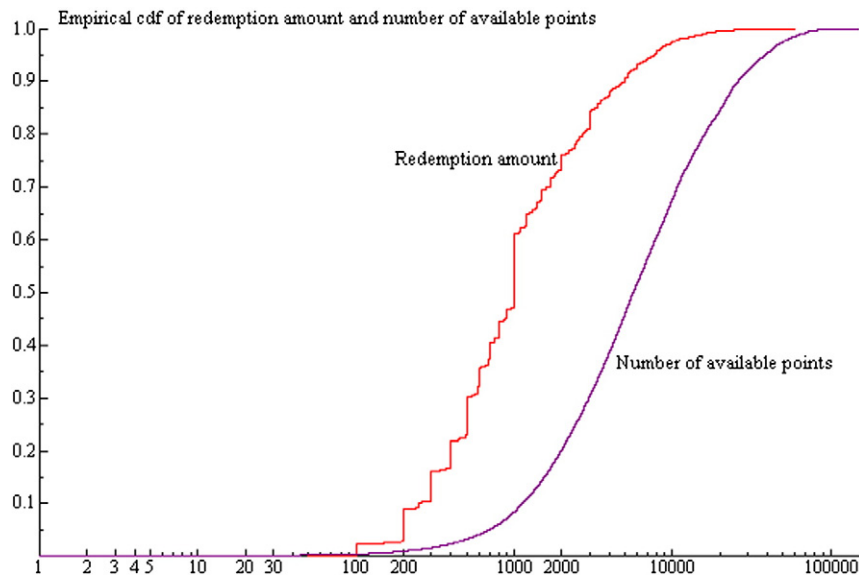

Fig. 3. Empirical cumulative distribution function of redemption amount and available points (log scale).
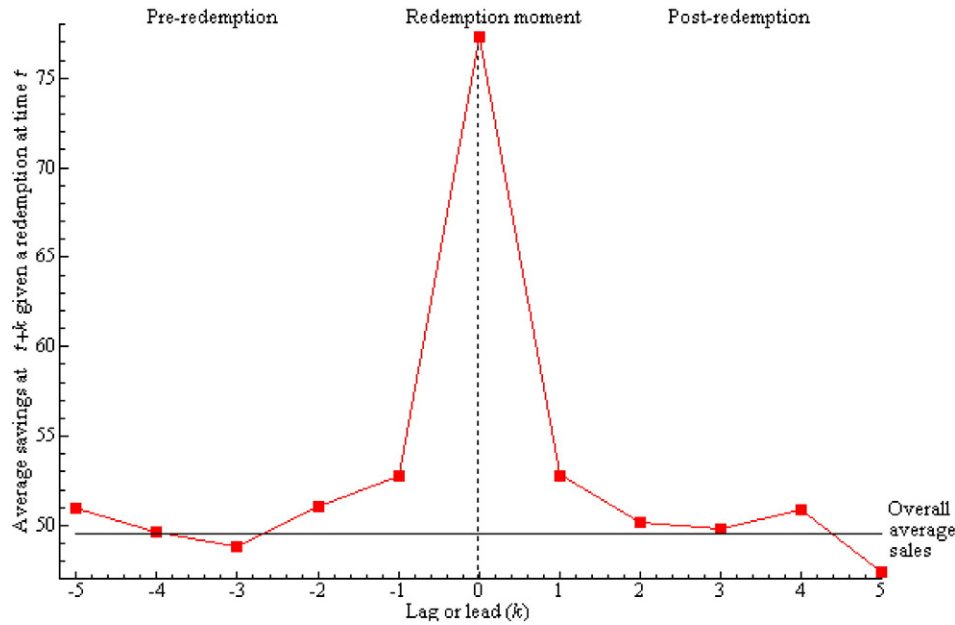
Fig. 4. Average points-saving behavior in periods close to redemption.

lack of points needed for the redemption. Theoretically, if the points pressure effect is driven by the urge to accumulate a "sufficient" amount of points to redeem the reward, our data suggest that in 97% of observed redemption cases, the theoretical arguments of "points pressure" and "sunk costs" are not applicable or at least insufficient explanations.

Before specifying the model, we provide some model-free evidence of the presence of reward effects. Fig. 4 shows the average point-saving behavior in the periods close to a redemption. The graph clearly shows an increase in average purchase behavior as redemption approaches and that behavior after the redemption stays at higher levels than average for one to two weeks. In this way this figure clearly shows the existence of pre- and post-reward effects in our LP.

5. Model

5.1. Model specification

In this section, we model the members' redemption and purchase decisions. We denote the number of saved LP points[2] in purchases by individual i in week t as S_{it}; the number of redeemed points in week t by the same individual i is denoted by R_{it}. Both actions are related to the number of loyalty points that a member i has in the beginning of week t (the balance of points), which is denoted as B_{it}. Given the redemptions, purchases and the number of points at the beginning of the week, we can calculate the number of points at the end of the week. If a member returns a purchase to the store, the balance will be corrected accordingly. We denote this correction by C_{it}. We do not model these returns, but we incorporate them in the calculation of the number of points. The updating equation for the number of points becomes:

$$B_{i,t+1} = B_{it} + S_{it} - R_{it} - C_{it}. \tag{1}$$

To address the possible bidirectional dependence between purchases and redemption, we explicitly model the sequence of decision-making (as outlined in Fig. 1). The moment in time when a positive

redemption decision is made may not coincide with the actual moment in time when the redemption incidence occurs. However, as researchers, we only observe the actual occurrence of the redemption; the timing of the *decision* is unobserved. The redemption decision may be made at any moment in week t. Once a member has planned a redemption, we assume that (s)he will not consider planning another redemption until the redemption actually happens. Next, the member decides whether to make a purchase at a participating store and use the loyalty card. In case the member decides to make a purchase, (s)he next makes a decision on the purchase amount. If the member decided to redeem in week t, (s)he finally decides on the redemption amount at the redemption incidence. As an illustration, a redemption (incidence and amount) that occurs in week t may be the result of a redemption decision at time t-2 (2 weeks before the redemption incidence). The purchases that occurred between those two events (in weeks t, t-1 and t-2) will all be affected by the redemption decision from week t-2. This impact on purchases before the redemption incidence contributes to the pre-reward effect. In fact, it is the shape of the pre-reward effect that identifies the redemption timing decision (see also Fig. 4). Note that by making the timing of the redemption decision endogenous, our assumptions of the order of decisions become less restrictive than they may seem at first. Although the redemption incidence and amount are placed last in the sequence of decisions, the redemption decision may have actually happened before the purchase decisions. However, we do not impose this. A member could also decide to redeem at the point-of-sale. In this case, there would only be a potential impact of redemptions in week t on the purchases in the same week.

5.2. Operationalization and modeling of main dependent variables

We introduce four main dependent variables: purchase behavior is analyzed through purchase incidence and purchase amount, while redemption behavior is analyzed through redemption decision and the redemption fraction (amount redeemed from the total balance).

We model purchase behavior with a hurdle or two-part model (Cragg, 1971; see Cameron & Trivedi, 2005 for a textbook treatment). In this model, the decision to purchase is modeled separately from the purchase amount. In other words, we model the log points–savings amount conditional on the points–savings incidence. The log

---

[2] Given that LP point-saving is directly related to purchase behavior, we refer to points savings as purchases.

transformation on the purchase amounts ascertains that purchase amounts remain positive.

The redemption decision is modeled using a probit model. Then, conditional on redemption incidence, we model the logit transformation for the redeemed fraction of the available number of points. This transformation ensures that the redemption amount is bounded by zero and the number of available points. Note that the number of points that can be spent in week t equals the initial number of points plus the saved points in that week. The redemption fraction is given by

$$f_{it} = \frac{R_{it}}{B_{it} + S_{it} + 1}, \tag{2}$$

where we add 1 to the number of points available to ensure that the logit transformation of f_{it} exists even if all available points are redeemed, that is, $R_{it} = B_{it} + S_{it}$.

In the section below, we first discuss our modeling approach for the redemption; afterward, we specify the purchase equations. As explained earlier, members may make a redemption decision ahead of the actual redemption incidence. The model for the timing of redemptions by member i at time t consists of two parts. First, we use a probit model to describe whether a new redemption is planned at a particular point in time. This probit model is described in terms of a latent variable RD_{it}^*, which symbolizes *redemption decision*. Next, in case a redemption is planned (i.e., a member has made a decision to redeem in a future), we model the time until the redemption incidence; this time is denoted by k_{it}. The two variables RD_{it}^* and k_{it} together completely describe the redemption incidence. To summarize the member's position in the redemption process, we introduce the *redemption timing* variable (RT_{it}). RT_{it} can take on the following finite set of values $RT_{it} \in \{-1, 0, 1, \ldots m\}$. If $RT_{it} = -1$, no redemption is planned for the near future. If $RT_{it} \neq -1$, the variable gives the number of purchase opportunities until the next redemption event (counting from the beginning of the week). Hence, if $RT_{it} = 1$ or $RT_{it} = 0$, the redemption occurs in week t itself. In the former case, the decision to redeem was made before the purchases were made in this week; in the latter case, the decision was made after the purchase. Note that in the case where $RT_{it} = 0$, the timing of the redemption decision does not induce a pre-reward effect, because the decision to redeem occurs after the purchase. Finally, when $RT_{it} > 1$, a redemption event will occur in the near future, e.g., if $RT_{it} = 2$, the redemption happens in the next week. In this way RT_{it} summarizes the decisions that member i has made, the likes of which may impact current and future behavior. Of course, the variable RT can only be partly observed. For example, if no redemption occurs at time t for individual i, we know that RT_{it} does not equal 0 or 1 and that $RT_{i,t-1}$ does not equal 2. However, the exact timing of each redemption decision remains unobserved. Therefore, RT_{it} should be seen as a latent variable.

The dynamic process for RT_{it} can be formally represented by

$$RT_{it} = \begin{cases} RT_{i,t-1} - 1 & \text{if there is a previously planned redemption } \left(RT_{i,t-1} > 1\right) \\ k_{it} & \text{if a future redemption is planned now } \left(RT_{i,t-1} \leq 1 \text{ and } RD_{it}^* > 0\right) \\ -1 & \text{if no future redemption is planned } \left(RT_{i,t-1} \leq 1 \text{ and } RD_{it}^* \leq 0\right). \end{cases} \tag{3}$$

The first line in Eq. (3) corresponds to the case where a redemption was already planned at (or before) $t-1$, so the time until the redemption incidence needs to be updated by reducing it by 1. The second and third lines correspond to the case where a new redemption could be planned (i.e., a redemption incidence occurred in the previous week or no redemption was planned before; both cases correspond to the condition $RT_{i,t-1} \leq 1$). This decision is governed by the latent variable RD_{it}^* (redemption decision). A new redemption will be planned if $RD_{it}^* > 0$, whereas no new redemption will be planned if $RD_{it}^* \leq 0$. In case of a positive redemption decision, the variable k_{it} gives the number of purchase occasions until the redemption and it is modeled as a draw

from the set of numbers $0,1,\ldots,m$, with probabilities $\pi_0, \pi_1, \ldots, \pi_m$. The number m will be relatively small; based on the model-free evidence of the pre-reward effect, we expect m to equal 2 or 3 at most.

As said, the redemption decision is modeled by a probit model. The latent redemption decision (RD_{it}^*) variable therefore follows

$$RD_{it}^* = \mu_{i1} + \gamma_{i1}t + Z_{it}^{R'}\beta_{i1} + W_t'\delta_1 + \xi_{it}, \text{ with } \xi_{it} \sim N(0,1), \tag{4}$$

where

$$Z_{it}^R = \begin{pmatrix} \log B_{it} \\ PntPre_{it} \\ PostRed_{it} \\ Access_{it} \\ Mailings_{it} \end{pmatrix}. \tag{4a}$$

In this vector of explanatory variables, B_{it} gives the balance at the start of week t, while $PntPre_{it}$, $PostRed_{it}$, $Access_{it}$ and $Mailings_{it}$ respectively give the points pressure (for eligible members), post-reward effect following a redemption incidence, accessibility of the LP due to purchases, and mailing decay variables. The exact operationalization of these variables will be discussed later. Finally, the variable t denotes a time trend and W_t captures seasonal dummies. For each member the time trend is defined relative to the moment at which the member subscribed to the LP. This variable therefore captures the length of the membership in the LP.

The logit transformed redemption fraction is modeled as

$$\log(f_{it}/(1-f_{it})) = \mu_{i2} + \gamma_{i2}t + Z_{it}^{R'}\beta_{i2} + W_t'\delta_2 + \nu_{it},$$

for all t where $RT_{it} = 0$ or 1, $\tag{5}$

with $\nu_{it} \sim N(0, \sigma_{R,i}^2)$.

To model purchases, we denote model purchase (points-saving) incidence by the binary variable SI_{it}. This variable is also modeled using a probit model, that is,

$$SI_{it} = \begin{cases} 0 & \text{if } SI_{it}^* \leq 0 \\ 1 & \text{if } SI_{it}^* > 0, \end{cases} \tag{6}$$

with

$$SI_{it}^* = \mu_{i3} + \gamma_{i3}t + Z_{it}^{S'}\beta_{i3} + W_t'\delta_3 + \varepsilon_{it}, \text{ with } \varepsilon_{it} \sim N(0,1), \tag{7}$$

where

$$Z_{it}^S = \begin{pmatrix} I(RT_{it} \geq 1) \\ \log B_{it} \\ PntPre_{it} \\ PostRed_{it} \\ Access_{it} \\ Mailings_{it} \end{pmatrix}. \tag{7a}$$

The first row of this vector gives the pre-reward effect due to redemption momentum (as an indicator related to the previously specified RT_{it}); the other rows correspond to the variables used in Eq. (4a). The indicator in the first row equals 1 if a redemption was planned before the focal purchase decision, which would allow for redemption momentum to occur. The corresponding parameter measures the impact of having made the decision to redeem on the purchase incidence.

Conditional on purchase incidence ($SI_{it} = 1$), the member's purchase (points–savings) amount follows

$$\log S_{it} = \mu_{i4} + \gamma_{i4}t + Z_{it}^{S'}\beta_{i4} + W_t'\delta_4 + \eta_{it}, \text{ for all } t \text{ where } SI_{it} = 1, \tag{8}$$

with $\eta_{it} \sim N(0, \sigma_{S,i}^2)$.

The complete set of heterogeneous parameters is related to member-specific explanatory variables (V_i) such as the individual's age, income and membership duration at the start of the dataset. Denote $\theta_i = (\mu_{i1}, \mu_{i2}, \mu_{i3}, \mu_{i4}, \gamma_{i1}, \gamma_{i2}, \gamma_{i3}, \gamma_{i4})'$ and $\beta_i = (\beta_{i1}', \beta_{i2}', \beta_{i3}', \beta_{i4}')'$. The vector θ_i contains all member-specific intercepts and member-specific trends; for this vector we specify a model including random effects, that is,

$$\theta_i = \Gamma_1 V_i + \omega_i, \tag{9}$$

where $\omega_i \sim N(0,\Omega)$. For parsimony, we do not include random effects for the parameters in β_i and we set $\beta_i = \Gamma_2 V_i$. In other words, we include interaction effects between the variables in Z_{it} and those in V_i. Therefore, the heterogeneity in β_i is only related to observed characteristics. For the ease of interpretation, we have standardized all moderating variables in V_i to have mean 0 and variance 1.

In the purchase and redemption equations above, we have introduced four error terms. The two error terms in the purchase (or redemption) equations are assumed to be independent. In principle, a correlation between the two errors can be specified; such a correlation is often included in sample selection models. In these cases, there is usually a separate process that determines whether an observation is sampled—for example, if someone participated in a job training program, then including the correlation would allow one to draw conclusions regarding the potential impact of the training program on those who decided to forgo the training. Unlike that setting, however, behavior within the LP program is not susceptible to sample selection and represents a corner-solution model (Wooldridge, 2011). Corner-solution (two-part) models separately describe the incidence and the amount conditional on incidence. The error terms in both equations are usually assumed to be independent (see the discussion in Wooldridge, 2011, p. 691). In theory, the correlation is identified; it would quantify the impact that unobserved factors may jointly have on the incidence and quantity decision, but in practice such a correlation is usually very difficult to estimate without imposing exclusion restrictions. However, models in practice often yield similar insights with or without a correlation—see for example, Madden (2008) and Konus, Neslin, and Verhoef (2014).

Another possible correlation is the one between the redemption decisions and the purchase decisions. This correlation would capture the impact that unobserved events may have on redemption and purchase decisions simultaneously. However, there are already three processes in the model that link redemption decisions to purchase decisions: (i) All decisions are tied together through the balance variable: one cannot redeem points that were not saved; (ii) The decision to redeem may precede the actual redemption moment and this has an impact on purchase behavior (redemption momentum); (iii) We allow for correlated, individual-specific parameters. The latter link captures individual specific patterns—for example, that members who purchase a lot may also redeem often. Furthermore, for each individual there tends to be only a few redemptions. This circumstance, together with the rich dependence between redemption and purchases that is already in the model, leaves little scope for estimating additional correlations.

In tandem with members' intertwined purchase and redemption decisions, LPs usually send mailings to a selected group of members to encourage redemption and purchase. In other words, these mailings may be endogenous. Without correction, this may lead to biased estimates (Franses, 2005). For example, if the mailings are sent to those who are likely to purchase, we would overestimate the impact of the mailings on purchase. From discussions with the LP manager, we know that only the frequency of the mailings is endogenous; its timing is not set based on individual behavior. This observation allows us to easily correct for the endogeneity, namely by including the average number of mailings received in the vector V_i (see Mundlak, 1978; Risselada, Verhoef, & Bijmolt, 2014 for an application in marketing). By doing so we identify the true impact of mailings on the redemption and purchase

decisions. Note that the parameters in Γ_1 and Γ_2 related to the number of mailings should *not* be seen as measuring the causal impact of mailings. In most cases, these parameters will mainly provide information on how the mailing strategy is set.

5.3. Operationalization of main explanatory factors

In this subsection, we discuss how we operationalize some of our main explanatory factors. We acknowledge that pre-reward effects may occur through the points pressure effect for those members who have an insufficient balance for their preferred redemption amount. The points pressure effect is the result of members' internal redemption thresholds, which are based on the members' preferences for the available awards. If the points pressure is active, then the member is close to a threshold, and thus (s)he is inclined to wait and save points until the threshold is reached. However, these preferences, and by extension the thresholds, are not observed.

In our LP, there is a reward available for almost every number of points; nonetheless, some common redemption thresholds can be observed across all redemptions. We therefore operationalize the internal thresholds using the most common amounts of points spent across the entire population (see Fig. 2). In our specification, we used all redemption amounts that occur more than 200 times in our sample.

We next specify the points pressure effect as a function of the relative distance between the current balance and the next redemption threshold, that is,

$$PntPre_{it} = \begin{cases} 0 \text{ if } B_{it} > \tau_K \\ \left(\dfrac{B_{it}-\tau_{k-1}}{\tau_k-\tau_{k-1}}\right)^{\alpha} \text{if } \tau_{k-1} \leq B_{it} < \tau_k, \text{ for } k = 1, \dots, K, \end{cases} \tag{10}$$

where $\tau_k, k = 1,\dots, K$ denote the internal thresholds. Given that we aim to explore the shape and duration of the points pressure effect, we specify the shape of the effect using parameter $\alpha > 0$. If $\alpha > 1$, the points pressure effect starts relatively close to the redemption threshold. If $\alpha < 1$, the points pressure effect starts relatively early.

The post-reward effect following from the redemption incidence may influence members' purchases. The post-reward effect potentially lasts for a number of weeks after the redemption incidence. In our model, we capture this effect using an exponentially weighted average of lagged redemption, that is,

$$PostRed_{it} = RI_{i,t-1} + \lambda_1 PostRed_{i,t-1}, \tag{11}$$

where RI_{it} denotes a redemption incidence indicator, and $0 \leq \lambda_1 \leq 1$ gives the decay rate of the post-reward effect. We use data before the start of our estimation sample to initialize this post-reward variable.

In line with our conceptual model, the mental accessibility of the LP due to prior purchases is operationalized as an exponential decay of purchase incidence (i.e., stock of purchases), that is,

$$Access_{it} = SI_{i,t-1} + \lambda_2 Access_{i,t-1}, \tag{12}$$

where, as before, the parameter $0 \leq \lambda_2 \leq 1$ controls the decay rate. The notion that periods with increased purchases may enhance the accessibility of the LP and thereby produce a spillover effect on behavior is in line with the literature on RFM models, direct mailings, and decay effects in both advertising recall and purchase history in household scanner data (Gönül, Kim, & Shi, 2000; Leone, 1995).

Finally, we include the dynamic impact of mailings sent to the members by an exponentially weighted average of current and past mailings, that is,

$$Mailings_{it} = M_{it} + \lambda_3 Mailings_{i,t-1}, \tag{13}$$

where $M_{it} = 1$ if member i received a mailing in week t. Like before, we use pre-sample information to initialize this variable.

5.4. Parameter estimation

We opt for Bayesian techniques for parameter estimation, as our model is highly nonlinear and contains many member-specific latent variables. More specifically, we use Markov chain Monte Carlo [MCMC] sampling, where we combine Gibbs sampling and Metropolis Hastings [MH] sampling. We sample the latent variables RT_{it}, RD^*_{it}, k_{it}, and SI^*_{it} alongside the other model parameters. The estimated parameters include the decay rates λ_1, λ_2, λ_3 as well as the probabilities π_0, π_1, ..., π_m that determine the time between a redemption decision and the actual redemption incidence. In the technical appendix we present the details of our sampler.[3]

We generated 60,000 draws from the Markov Chain and removed the first 20,000 draws as a burn-in period. Of the remaining draws, we retained every 5th draw to reduce autocorrelation. As discussed before, we set the thresholds τ_k equal to all unique redemption amounts that occur more than 200 times in our sample. Using this rule we set the thresholds to 100, 200, 300, ..., 1000, 1200, 1500, 2000, and 3000.[4] Finally, we set $m = 2$. This limits the pre-redemption effects to a maximum of 2 weeks before the redemption incidence. This choice is mainly motivated by Fig. 4. However, we have also considered a model with $m = 3$ and found no substantive difference with the presented results.[5]

6. Results

We first consider a model with only the main effects (including correction for endogeneity of mailings) and then consider the full model that accounts for all interactions. The estimation results for both models are presented in Table 2. The main effects are very robust, and the overall effects stay the same even after controlling for moderating variables.

In the discussion below, we differentiate between the effects on purchase (LP points-saving) incidence, purchase amount, redemption decision/timing and redemption amount/fraction. This provides a fruitful environment for discussing the diverse mechanisms underlying the relationship between redemption and purchase behavior. Whenever we discuss a particular parameter estimate from Table 2, we present the posterior mean and refer to it as γ; if necessary we add a subscript that refers to a particular model component. Note that there are separate coefficients for all four decisions. These parameters are all part of the matrices Γ_1 or Γ_2; see Eq. (9).

6.1. Timing of the redemption decision

The starting point in analyzing the interplay between the decision to redeem points and purchase behavior is understanding whether a redemption decision precedes purchase or vice versa (i.e., the redemption occurs as a consequence of increased purchases in some period). We find that in an overwhelming majority of redemptions (around 70%), the redemption decision is made before the purchase decision. In other words, approximately 31% of redemption decisions are made at the point of redemption: members decide to redeem ad hoc and do so immediately. In the model, this percentage is represented by π_0, which indicates the proportion of members for whom the purchase

decision (in the same week when redemption occurs) is not affected by the redemption decision. Sixty-four percent of redemptions are planned ahead in the same week: customers go to a store, make a purchase and then redeem their points (π_1). At this point, redemption momentum exists because the decision to redeem still affects the purchase. Around 6% of redemptions are planned a full week ahead and subsequently affect purchase behavior until the redemption event (π_2). We emphasize that this LP is used on a weekly basis (groceries, etc.), which adds face validity to these estimates.

6.2. Pre-reward effects

We find support for the existence of pre-reward effects even when members are not "pressured" with point expiration. Positive pre-reward effects are driven both by the points pressure effect for members with insufficient balances and the redemption momentum that goes beyond the points pressure. In terms of effect size, the redemption momentum is the most important pre-reward effect (based on the evidence presented in Table 2 and in Section 7).

6.2.1. Points pressure effects on purchase and redemption behavior

For approximately 3% of members who may have experienced points pressure before the redemption, there is an increase in the likelihood of purchase ($\gamma_{PntPre} = 0.053$). The points pressure effect starts early after passing a previous threshold (the posterior mean for log $\alpha = -2.492$, which corresponds to $\alpha = .083$). However, points pressure primarily affects purchase incidence, not the purchase amount (γ_{PntPre} for purchase amount is not significant).

As members approach the next available internal threshold, they become less likely to redeem ($\gamma_{PntPre} = -.086$). This negative effect is expected since members likely postpone redemption until they pass the threshold. This also reinforces the notion that members are driven by an internally set threshold behavior and redeem rewards after reaching this internal threshold. Accordingly, when approaching a redemption threshold, the redeemed amount tends to be a smaller fraction of the total balance ($\gamma_{PntPre} = -.131$). In other words, if members do decide to redeem before the threshold, they redeem a smaller part of their balance.

6.2.2. Redemption momentum effects

As mentioned before, the effects of redemption go beyond the points pressure effect; the mere decision to redeem a reward affects members' subsequent purchase behavior (creating the redemption momentum). When the decision to redeem a reward occurs before the actual redemption (for 69.5% of members), members increase their frequency of purchase ($\gamma_{RedMom} = 1.763$) as well as their purchase amounts ($\gamma_{RedMom} = .325$) in periods between the redemption decision and the redemption event. As expected, the redemption follows shortly after members make the decision to redeem. Hence, the pre-reward effect due to redemption momentum stretches to a maximum of one week before the redemption (evident from π_0, π_1, and π_2 estimates discussed in Section 6.1).

6.3. Post-reward effects

In the post-reward periods, members tend to purchase more often ($\gamma_{PostRed} = .033$) and they increase their purchase amounts per purchase ($\gamma_{PostRed} = .031$). We thus provide empirical support for positive post-reward effects in the continuous reward setting. The estimated redemption decay parameter in post-reward periods is $\lambda_1 = .734$. The impact of redemption therefore lasts relatively long after the redemption. The post-reward effect is maximal in the week after the redemption; the effect reduced to 73.4% 2 weeks later, to 53.9% 3 weeks later (0.734^2), and so on.

On average, post-reward effects have a positive impact on the subsequent likelihood of redeeming, since the impact of post-reward

---

[3] We checked the performance (and implementation) using a simulation experiment. We generated data using known parameters and tested whether the estimation procedure is able to retrieve these parameters. The MCMC sampler proves to perform well. Details and results of this experiment are provided in the Web supplement.

[4] As a robustness check, we have considered models with higher and lower numbers of redemption thresholds. The obtained results are very robust. Therefore, we choose to present the option that includes the majority of commonly selected thresholds and avoids three thresholds chosen relatively frequently, but less than 200 times (1400, 1700, and 2400 points).

[5] More specifically, in the model with $m = 3$, the estimated value for π_3 was close to 0. This effectively reduces the model to a model with $m = 2$. [6] More details could be found in the Websupplement accompanying this document.

Table 2
Model estimation results.

			Main effects model								Full model							
			Purchase incidence		Purchase amount[a]		Redemption decision		Redemption fraction[b]		Purchase incidence		Purchase amount[a]		Redemption decision		Redemption fraction[b]	
			Est	Sig[c]	Est	Sig	Est	Sig	Est	Sig	Est	Sig[c]	Est	Sig	Est	Sig	Est	Sig
Pre-reward effect	Baseline	*Constant*	1.037	***	−5.784	***	−2.134	***	−1.565	***	1.039	***	−5.779	***	−2.125	***	−1.541	***
		Average income	0.004		0.076	***	0.008		−0.061	*	−0.008		0.067	***	0.003		−0.049	
		Age	0.052	***	−0.040	***	−0.037	***	−0.062	*	0.051	***	−0.049	***	−0.064	***	−0.087	**
		Membership yrs	0.066	***	−0.044	***	0.021	*	0.058	*	0.025		−0.051	***	0.017		0.044	
		Avg. no. mailings	0.059	***	0.106	***	−0.076	***	0.156	***	0.061	***	0.104	***	−0.078	***	0.157	***
		Variance	0.418	***	0.539	***	0.193	***	0.786	***	0.439	***	0.538	***	0.188	***	0.797	***
	Trend	*Constant*	−0.068	***	−0.251	***	0.018		−0.064	*	−0.071	***	−0.252	***	0.017		−0.069	*
		Average income	0.005		0.000		−0.002		0.010		0.006		0.000		−0.004		0.014	
		Age	−0.033	***	−0.018	**	0.018		−0.038		−0.032	***	−0.016	**	0.016		−0.039	
		Membership yrs	−0.036	***	0.035	***	−0.016		−0.032		−0.021	**	0.034	***	−0.012		−0.031	
		Avg. no. mailings	−0.034	***	−0.022	***	0.053	***	0.010		−0.034	***	−0.021	***	0.052	***	0.006	
		Variance	0.113	***	0.143	***	0.094	***	0.371	***	0.116	***	0.143	***	0.091	***	0.373	***
	Points pressure	*Constant*	0.053	***	0.002		−0.086	***	−0.131	***	0.040	***	0.001		−0.082	***	−0.091	*
		Average income									0.019		0.006		0.006		0.012	
		Age									0.000		−0.005		−0.014		0.113	**
		Membership yrs									−0.022	**	0.004		−0.037	**	−0.107	**
		Avg. no. mailings	−0.012		0.008		−0.118	***	0.032		0.002		0.008		−0.115	***	0.059	
	Redemption momentum	*Constant*	1.763	***	0.325	***					1.821	***	0.329	***				
		Average income									0.028		−0.021	**				
		Age									0.171	*	0.026	**				
		Membership yrs									−0.093		−0.016					
		Avg. no. mailings	−0.548	***	−0.020	**					−0.558	***	−0.016					
	Post-reward effect	*Constant*	0.033	***	0.031	***	0.227	***	−0.023		0.027	**	0.032	***	0.206	***	−0.020	
		Average income									0.007		0.000		0.035	***	−0.011	
		Age									−0.014		0.002		0.038	***	−0.017	
		Membership yrs									−0.009		−0.016	***	0.008		0.012	
		Avg. no. mailings	−0.006		0.002		0.006		0.008		−0.004		0.003		0.008		0.015	
	Accessibility due to purchase	*Constant*	0.282	***	0.064	***	0.039	***	−0.057	***	0.280	***	0.065	***	0.039	***	−0.053	***
		Average income									−0.005	**	−0.007	***	0.003		−0.016	
		Age									0.002		−0.011	***	−0.002		0.022	**
		Membership yrs									−0.020	***	0.004	**	−0.004		−0.005	
		Avg. no. mailings	−0.031	***	0.000		−0.015	***	−0.002		−0.029	***	−0.002		−0.015	***	0.000	
	Log balance	*Constant*	0.040	***	0.015	***	0.125	***	−0.875	***	0.028	***	0.015	***	0.128	***	−0.858	***
		Average income									0.002		−0.011	***	−0.009		−0.004	
		Age									0.005		−0.008	**	−0.020	***	−0.021	
		Membership yrs									−0.028	***	−0.006		−0.017	**	−0.062	***
		Avg. no. mailings	−0.016	***	0.006		−0.006		0.078	***	−0.006		0.006		−0.005		0.085	***
	Mailing decay	*Constant*	0.021	***	0.003	**	0.039	***	0.027	***	0.022	***	0.003	**	0.040	***	0.029	***
		Average income									0.003	*	0.001		0.000		−0.007	
		Age									0.002		−0.001		0.006	**	−0.004	
		Membership yrs									−0.001		0.002	*	−0.002		0.000	

	Poster. mean	Sig.	Poster. mean	Sig.
π_0 (red. decision just after purchase)	30.5%	***	30.8%	***
π_1 (red. decision before the purchase)	63.8%	***	63.7%	***
π_2 (red. decision a week before purchase)	5.7%	***	5.5%	***
Points pressure $\log(\alpha)$	−2.492	***	−2.2448	***
Post redemption decay	0.734	***	0.767	***
Accessibility decay	0.847	***	0.844	***
Mailing decay	0.758	***	0.743	***

Purchase = LP points saving.
[a] Defined as log savings amount.
[b] Defined as logit transformed redemption fraction of the total amount of accumulated points (balance).
[c] ***, **, *: 99%, 95%, and 90% highest posterior density regions that do not contain zero, respectively.

effects on the subsequent decision to redeem is positive and significant ($\gamma_{PostRed} = .227$). However, a non-significant, negative post-reward effect is found on redemption fraction ($\gamma_{PostRed} = -.023$).

6.4. Purchase behavior reinforces redemption

Our conceptual model proposes that increased purchases in a certain period may encourage members to make the decision to redeem, since purchasing increases the mental accessibility of the LP due to prior purchases. As the members' average "stock-of-purchases" increases, it reinforces purchase frequency and spending amounts ($\gamma_{Access} = .282$ and .064, respectively). This increase also boosts the likelihood of the redemption decision ($\gamma_{Access} = .039$), but if members decide to redeem, it reinforces the redemption of smaller fractions (rather than a redemption of all/majority of collected points) ($\gamma_{Access} = -.057$). This finding suggests that LP members become more cognizant of the ability to redeem their collected points as a result of purchases enhancing the LP's mental saliency; however, relatively larger redemptions would be planned ahead. In addition, the estimated accessibility decay parameter between purchases is .847, which indicates that the decay in accessibility between two purchase incidents is slow. In other words, members slowly forget about the LP if they do not use it.

Table 3
Correlation matrix of individual-level saving and redemption parameters (full model).

	SI$_b$	SI$_t$	logS$_b$	logS$_t$	RD$_b$	RD$_t$	logRF$_b$	logRF$_t$
SI$_{base}$	1	−0.684	0.243	−0.023	0.231	−0.194	−0.033	0.012
SI$_{trend}$		1	−0.093	0.242	−0.14	0.154	0.046	−0.024
logS$_{base}$			1	−0.582	0.076	0.008	0.089	−0.051
logS$_{trend}$				1	0.027	−0.005	−0.055	0.064
RD$_{base}$					1	−0.708	−0.466	0.254
RD$_{trend}$						1	0.216	−0.216
logRF$_{base}$							1	−0.537
logRF$_{trend}$								1

6.5. Trends and moderating effects

The simultaneous estimation of the four dependent variables allows us to assess the associations between the individual-specific effects arising from the four purchase and redemption responses. The correlations between the eight individual-specific effects (four intercepts plus four trends) are presented in Table 3. Combining those results with the results of the moderating effects presented in Table 2 reveals interesting trends for LP managers. We discuss these insights below.

6.5.1. Decreasing responsiveness to the LP

Findings in Table 3 reveal important concerns for LP managers due to the strong negative correlation between the baseline effects and the trend for purchase and redemption behavior. LP members with a high purchase propensity (frequent buyers) tend to decrease their purchase incidence over time (ρ(SI$_{base}$, SI$_{trend}$) = −.684). The same holds for purchase amount (ρ(logS$_{base}$, logS$_{trend}$) = −.582). Similar negative correlations are observed in the redemption incidence and redemption amount parameters. The propensity to redeem is likely to decrease over time for those who, at first, decide to redeem relatively often (ρ(RD$_{base}$, RD$_{trend}$) = −.708). But even more, the slope of the trend in the redemption fraction is negatively correlated with the base redemption fraction (ρ(logR$_{base}$, logR$_{trend}$) = −.537). This implies that, with time, high-baseline members become less likely to redeem, and even if they decide to redeem, their redemption amount also decreases over time. In other words, there is a mean-reversion process. Members who are initially very active become less active over time (and vice versa).

An analysis of the moderating effects in Table 2 (full model) further supports the finding of negative trends in purchase frequency and amount over time for all members (γ's are −.071 and −.252, respectively). The decline in purchase responsiveness to the LP over time is particularly pronounced among older members and long-term loyal members. Spending patterns worsen for those groups even more than for an average member: older members show stronger declining trends both in purchase frequency and amount (coefficients −.032 and −.016, respectively); meanwhile, long-term members particularly decrease their likelihood of purchase more so than their purchase amounts (coefficients −.021 and .034, respectively).

6.5.2. Moderating impact on pre- and post-reward effects

Overall, we find strong heterogeneity in the baseline purchase and redemption behavior of LP members (given the relatively large variances in the baseline estimates for all dependent variables reported in Table 2). Importantly, long-term members seem to be less responsive to LP mechanisms. Points pressure, accumulated balance and prior purchases have less impact on the purchase incidence and redemption decision of long-term members, since the positive main effects of these variables are negatively moderated by the number of years as an LP member ($\gamma_{PntPre^*MemberYrs}$ = −.022 (purchase incidence) and −.037 (redemption decision); $\gamma_{Access^*MemberYrs}$ = −.020 (purchase incidence) but positive .004 (purchase amount); $\gamma_{Balance^*MemberYrs}$ = −.028

(purchase incidence) and −.017 (redemption decision)). Similarly, in the post-reward periods, rewarded behavior has less positive effects on purchase amounts for long-term members relative to others ($\gamma_{PostRed^*MemberYrs}$ = −.016).

In addition, long-term members show a more rational redemption behavior once they decide to redeem. Long-term members are even less likely than others to redeem just before reaching the preferred threshold ($\gamma_{PntPre^*MemberYrs}$ = −.037) and even if they do, their redemption amounts tend to be a smaller fraction of their total accumulated balance ($\gamma_{PntPre^*MemberYrs}$ = −.107). Also, the amount accumulated in the balance does not increase the likelihood of redemption and its amount ($\gamma_{Balance^*MemberYrs}$ = −.017 and −.062). These results may be explained by long-term members' experience in the LP.

We also observe a positive moderating effect of age. The redemption momentum is stronger for older members, as the redemption momentum increases both their purchase frequency (γ_{RedMom^*Age}) and spending amounts (γ_{RedMom^*Age} = .026). In contrast, the redemption momentum is weaker for higher-income members ($\gamma_{RedMom^*Income}$ = −.021). Both age and income reinforce the impact of post-reward effect on the likelihood of a new redemption ($\gamma_{PostRed^*Age}$ = .038 and $\gamma_{PostRed^*Income}$ = .035).

6.5.3. Mailing effects

The impact of mailings appears in two distinct manners in the model. First, there is the direct impact of mailings through the mailing decay variable. Second, there is the moderating impact of the average number of mailings that a member received. The impact of the latter variable on the baselines is likely attributable to the LP's target selection. Our results show that those members who purchase frequently and in higher amounts tend to receive more mailings (estimated coefficients are .061 and .104 in the full model, respectively). But conversely, more frequent redeemers receive fewer mailings on average (coefficient equals −.078).

The estimated mailing decay parameter of Eq. (12) equals λ_2 = 0.743. The impact of a mailing is strongest in the week when the mailing is received, while in the second week the carryover effects reduce to 74.3%, and 3 weeks later they reduce to 55.2%, and so on. This weekly decay parameter is in line with previously reported decay parameters on advertising effects (Clarke, 1976; Leone, 1995) and direct mailing effects (van Diepen, Donkers, & Franses, 2009a).

Overall, mailings have a direct positive impact on purchase incidence (γ_{Mail} = .022) and amount (γ_{Mail} = .003). The effect is marginally larger for long-term members ($\gamma_{Mail^*MemberYrs}$ = .002). Furthermore, mailings seem to encourage redemption for older members (γ_{Mail^*Age} = .006). The impact of mailings on the purchase likelihood is marginally enhanced for high-income members ($\gamma_{Mail^*Income}$ = .003), but it is negatively moderated by the total number of mailings received ($\beta_{M^*MailsReceived}$ = −.012). Therefore, the effectiveness of mailings in encouraging purchase and redemption declines as the number of mailings increase ($\gamma_{Mail^*MailsReceived}$ = −.012 and −.020, respectively).

7. Effect size simulations

To further analyze the impact of redemption on purchase behavior, we conducted a series of simulations in which the behavior of an average member is repeatedly generated. The effects of different model components are analyzed by switching off one component at a time (i.e., setting its parameter to zero). Only for the impact of accessibility due to purchases (stock-of-purchases) do we set the accessibility to its average value over time. Given the high frequency of purchases, the accessibility variable is always much larger than zero.

For each scenario, we analyzed the average purchase behavior around the moment of redemption. In addition, we calculated the average purchases around the moment of redemption for those redemptions where the balance is below our highest points pressure threshold. In this way we fully explore the differences between the
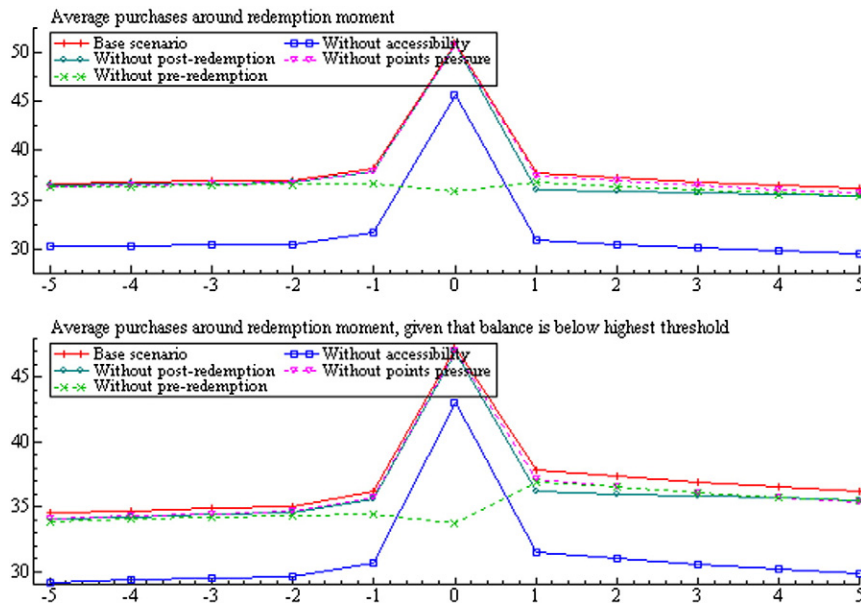
Fig. 5. Results of effect-size simulations.

points pressure and redemption momentum mechanisms in the pre-reward effects. The findings from these simulations are illustrated in Fig. 5 [6].

Looking at the difference between the base scenario (our full model) and the effects without points pressure, we can see that the points pressure presents a rather small and limited contribution to the overall rewarding effects. The same conclusion holds when we analyze the average behavior of members whose balance is below the highest points pressure thresholds. Overall, although there does seem to be a significant points pressure effect, we find its effect size to be relatively small. One explanation is that the large majority of members have a large balance of points at the time of redemption (more than sufficient for their redemption). However, even when that is not the case, the magnitude of the effect is still relatively small.

The redemption momentum clearly has the largest impact. The largest part of the peak in purchases at and before the redemption moment can be attributed to this pre-redemption effect beyond points pressure. We therefore posit that the mere decision to redeem triggers a substantial increase in purchase behavior among LP members.

The post-reward effect is significant in the model and the simulations show its substantial impact in the periods after the redemption. Our simulations indicate that post-reward effects limit the potential dip after obtaining a reward, particularly in situations where customers have a lower balance than the highest threshold.

Finally, we recommend caution when interpreting the effects in Fig. 5 that relate to the results labeled 'without the accessibility of prior purchases.' Note that the effects derived from switching off a model component are complex; all current decisions are connected to all future decisions through the accumulated balance. This also explains why the overall purchase levels are substantially lower when accessibility of prior purchases is switched off (i.e., accessibility is set to an average level). Moreover, the effects go beyond just purchase behavior: the model components also affect the redemption decision itself. For example, by switching off points pressure, we also observe a slight increase in redemption frequency (and slightly larger redemptions) at low balances. This in turn may reduce the average balance that members have, leading to less frequent purchases. So in order to understand (and explain) all effects, one needs to take the entire model into account.

8. Discussion

This study aims to better understand the LP members' reward-redemption behavior and its impact on purchase behavior—in particular, the behavior directly preceding and following a redemption. The study examined a typical LP with a continuous and linear rewarding structure (i.e., one point per euro spent), which is common among retailers of frequently purchased items. Importantly, in LPs without point expiry deadlines, members endogenously choose when and how much to redeem from a broad spectrum of potential reward options. Little is known regarding whether redemption effects occur when no expiration or binding policies exist. Because obtaining a reward in such LPs requires only little additional effort from members, some authors have postulated that pre-reward effects would not occur (Blattberg et al., 2008). Using an extensive data set of 3094 members involved in such an LP, we simultaneously modeled purchase incidence, purchase amount, redemption decision and redemption amount (as a fraction of available balance). This allowed us to empirically investigate such pre-reward effects. We summarize and discuss our key findings below.

For the majority of members (approximately 70%), the decision to redeem occurs before the actual redemption and affects their subsequent purchase decisions. Therefore, the interplay between redemption and purchases occurs in this order: (i) customer makes a decision to redeem which (ii) increases the salience of the LP and its benefits and (iii) encourages (pre-reward) purchase behavior. Once the decision to redeem is made, (iv) the redemption occurs within a short period of time (1 week). After the redemption, (v) the (post-reward) purchase behavior is enhanced by the rewarded behavior effects.

Reward redemption leads to important pre-reward increases in LP members' purchase even when they do not face point expiry or binding thresholds. The drivers of such increases go beyond points pressure. Even when the majority of LP members (97%) redeem just a (small) fraction of their overall accumulated balance at a redemption incidence, we still find strong evidence for pre-reward effects. Hence, our findings counter the notion that members' purchase behavior prior to reward redemption is motivated solely by the points pressure mechanism (in other words, that members only increase purchases in order to accumulate a sufficient number of points for their preferred reward). Our findings strongly emphasize the power of redeeming a reward in LPs: The decision to redeem motivates members and reinforces their subsequent

---

[6] More details could be found in the Websupplement accompanying this document.

behavior. Importantly, we theoretically introduce the novel concept of *redemption momentum* as an additional explanation for the existence of the pre-reward effects beyond points pressure. Hence, pre-reward effects are a general phenomenon driven by multiple underlying processes (e.g., goal attainment, increased LP engagement, salience), which occur for LPs with diverse designs. That said, alternative reward mechanisms (like redemption momentum) have substantially larger influence than external thresholds (points pressure mechanism) in a context of continuous, linear LP rewarding.

The effects of redemption also enhance behavior in post-reward periods. Hence, we provide empirical support for the reinforcing effects of the rewarded behavior mechanism. Members who just redeemed a reward demonstrate a higher purchase incidence and higher purchase amounts. This finding supports the notion that redeeming rewards may create positive attitudes and feelings that drive members to purchase more frequently and obtain higher amounts of LP points even in the absence of external pressures from the firm (Blattberg et al., 2008; Palmatier et al., 2009). It also empirically supports and extends the findings of Kopalle et al. (2012) on the existence of the post-reward effect in a continuous LP, but for a broader range of LP members (we find that the effects on purchase behavior mostly hold across various customer groups). On the other hand, our findings counter the notion of the post-reward resetting mechanism reported by Drèze and Nunes (2011), at least for the retail setting with no LP tier structure. In this respect, our findings help to clarify equivocal empirical evidence on post-reward effects in continuous LPs.

We also found support for the reinforcing impact of previous accumulated purchases on the redemption likelihood. Prior purchases not only positively affect subsequent purchases, but they also increase the redemption probability, which speaks in favor of the increased salience. Although accumulated purchases in a certain period enhance the likelihood of redemption, they do not affect the redemption fraction.

Both the pre- and post-reward effects on purchase incidence and purchase amount substantially differ between members. An important moderator is membership length. In general, the effects on purchase and redemption behavior are less pronounced among long-term LP members. This might be due to the learning effects in tandem with the more strategic redemption behavior among such members (Lal & Bell, 2003; Liu, 2007). These members have extensive experience with the LP and they may be less prone to change their purchase behavior when they redeem an award. We also observe some interesting moderating effects of age and income, which have not been shown before.

Mailings have an overall positive impact on the purchase and redemption behaviors (both on incidence and amount of purchase and redeeming). However, the impact of mailings on purchase incidence, redemption likelihood and redeemed fraction declines with the total number of mailings received. This may indicate a worrying trend of LP members becoming increasingly unresponsive to the LP's personalized mailings, or it may be the result of the LP's targeting policy (e.g., van Diepen, Donkers, & Franses, 2009b). Those members who purchase often as well as those who purchase a lot tend to receive more mailings. By contrast, those members who are more likely to redeem tend to receive relatively fewer mailings than others. Given this finding, there is possibly some untapped potential in terms of tailoring promotional strategies to increase redemption incidence.

9. Managerial implications

Firms increasingly try to remove hurdles in their LPs to improve members' experiences. As a result, LPs that preserve balances for a long time, or simply forgo the expiration of points and miles entirely, are increasingly common among retailers (e.g., Tesco's Clubcard, Nectar, Airmiles), car rental agencies (e.g., Hertz Gold Rentals), hotels (e.g., Intercontinental's Priority Club Rewards), airlines (e.g., Delta Airlines' SkyMiles, JetBlue Airlines' TrueBlue) and financial institutions (e.g., American Express Membership Rewards, Wells Fargo Rewards).

Answering the question of whether firms should encourage redemption without imposing point expiry and binding thresholds is relevant for three main reasons. First, the lack of redemption limits the LP's power to build and sustain loyalty due to missed opportunities to strengthen relationships and engage members (Levey, 2011). Second, the lack of redemption may lead to a potential decrease in LP involvement and diminish perceived program value over time, exacerbated by the trend of decreasing active participation in enrolled LPs (Gordon & Hlavinka, 2011). Third, the difference between issued and redeemed points has a profound impact on profit, since unredeemed points create an accounting liability for the firm (i.e., debt to members) (Levey, 2011; The Economist, 2005).

Our findings suggest that companies should actively try to encourage redemption in order to sensitize consumer responsiveness and increase the salience of the LP. This is particularly relevant given our findings on the overall trends of declining purchase and redemption activity. Members' accumulation of points is far above the highest common reward values, and upon redemption, members on average redeem just 26% of their available balance. Yet even in this setting, reward redemption plays an important role in increasing the purchase behavior in periods before and after the redemption.

There are two main reasons why LP managers should actively influence redemption incidence and redemption amount and thereby increase members' engagement in the LP. Firstly, lagged purchase incidence positively affects redemption incidence, as does the accumulated balance of points. This means that encouraging purchasing also increases the probability of members redeeming a reward. Secondly, redemption incidence and fraction can be stimulated with mailings. Notably, customers who recently redeemed an award are more likely to subsequently redeem another reward, suggesting that stimulating reward redemption can be a rather powerful way to increase purchase incidence directly in the short term, as well as in the long run through increased redemption incidence. But we caution here that increasing the number of mailings also has a negative implication, as it may reduce the effectiveness of each subsequent mailing due to factors such as irritation (van Diepen et al., 2009b).

To encourage the redemption frequency of LP members, companies should consider offering a wide range of potential rewards from which members can choose (see Fig. 2). On the one hand, encouraging the redemption of larger amounts decreases liability for the LP provider, but on the other hand, managers may feel they are putting the company's financial solvency at risk. However, as long as members do not decide to redeem the rewards all at the same time, firms should not experience strong problems in this respect.

Furthermore, our findings provide valuable insights on policies for managing the relationships with long-term, loyal LP members relative to more recent customers. Interestingly, long-term members are relatively more frequent purchasers (they have higher points-saving incidence); however, they have comparably lower purchase amounts (see Table 2). Over time, though, these loyal members tend to decrease their purchase incidence even more than other members. Retailers should use this insight to design promotional strategies targeted at long-term versus more novel LP members. Long-term members are relatively (albeit marginally) more apt to increase the amount spent in response to promotional strategies. This group is therefore an important target segment for policies intending to encourage redemption. Because both pre- and post-reward effects seem to be harder to evoke among long-term members, managers are advised to carefully tailor their personalized marketing strategy to encourage redemption effects.

10. Limitations and further research

This study has mainly focused on the continuous types of reward structures. Our focal LP is analogous to many LPs with a continuous rewarding structure and no point expiration. Nevertheless, this study has analyzed the effects of rewarding in only one LP in one country,

which limits its generalizability. Though the analyzed LP's structure is typical of coalition LPs in other counties, some conclusions may not automatically transfer. Moreover, being interested in effects within the LP as a whole, we aggregated LP data collected across various vendors that participate in the LP. Further research could analyze the differences in reward behavior effects across individual vendors in the context of partnership LPs, as well as for sole proprietary LPs.

Our empirical analysis of reward redemption effects is limited to LP members who had redeemed at least once in the observation period. This choice may have created a selection effect (relative to non-redeemers). However, this selection was necessary to analyze reward redemption effects. In addition, we examined point collection behavior rather than the exact amounts spent. These measures might not correspond perfectly if a member does not use his or her LP card for every purchase.

Since we could not obtain the information on the cost structure in the observed LP, we could not fully analyze the profit potential of rewarding effects. It would be beneficial to evaluate the profit implications of the rewarding effects analyzed in this study.

Our study has provided evidence that the drivers of pre-reward effects are complex and may go beyond the rational expectations of the points pressure effect. To this end, internal reward thresholds, and especially redemption momentum, may play an important role. More in-depth theoretical evidence is required on the mechanisms that drive the effects of rewarding in continuous rewarding structures without expiry deadlines. Such an investigation would require setting up a series of experimental studies. In particular, diverse psychological drivers may exist under redemption momentum; the size of this effect and its importance in engaging LP members warrant in-depth analysis of the underlying psychological mechanisms. One aspect involves the notion that members may want to maintain their accumulated balance after redemption, which may induce them to speed up purchases in pre-reward periods.

In general, we know relatively little regarding the emotional drivers of LP behavior. Future research needs to explore the notion that deciding to redeem a reward may induce excitement about and salience of the benefits of LP membership. Arousal (excitement) and valence of feelings may both be signals for action in the LPs (cf. feelings as information theory). Unfortunately, we do not have attitudinal data that would allow us to further explore these issues. Furthermore, even without explicit expiration dates, there can be a pressure to accelerate purchasing (and point accumulation) if customers believe (1) that the company will go bankrupt or (2) there will be a devaluation in points, both of which have occurred in the airline industry (see The Economist, 2005). While this is not the case in the analyzed LP, the issue of how customers perceive the value of their reward currency is an important research question.

In our analysis we had to make the assumption that points pressure thresholds are common across all members. Most likely there are differences across members. However, these differences are very difficult, if not impossible, to identify. Redemptions are relatively rare events and it is also rare that the balance of a randomly selected member is close to a particular threshold. This is due to the fact that most redemptions correspond to only a fraction of the balance. Nevertheless, we think it important for future research to explore the topic of heterogeneous points-pressure thresholds in LPs (cf. Stourm et al., 2013). Moreover, since we find that redemption momentum dominates points pressure in continuous and linear LPs, it is also important to consider the existence of such alternative mechanisms in other LP designs.

Furthermore, in our analysis we assessed the number and timing of mailings that LP members received, without looking deeply at the contents of the said mailings (e.g., Feld, Frenzen, Krafft, Peters, & Verhoef, 2013). Further research may better account for the contents of mailings. Finally, in this study we included multiple relevant moderators of the reward effects on our studied dependent variables. Future research might include other moderators, specifically some soft moderators such as attitudes toward the program and the participating retailers.

Acknowledgments

Technical appendix

In this appendix we discuss all the steps of our MCMC sampler in detail. We first introduce some common notation. The four main equations (redemption decision, redemption fraction, purchase (points saving) incidence, and purchases (points saving) amount) are summarized in vector notation; that is, we group all observations of a single member. We write these equations such that we can simplify the derivations below. For redemption decision we write

$$y_i^1 = M_i^1 \begin{pmatrix} \mu_{i1} \\ \gamma_{i1} \end{pmatrix} + Z_i^1 \beta_{i1} + W_i^1 \delta_1 + \zeta_i^1, \tag{A1}$$

where y_i^1 denotes a $T_{i1}x1$ vector with elements RD^*_{it}, M_i^1 equals a $T_{i1}x2$ matrix consisting of a column of ones and a column with a trend. Z_i^1 collects all relevant row vectors Z_{it}^R', and W_i^1 is a matrix obtained by stacking all relevant row vectors W_t'. In Eq. (A1) we collect all weeks at which a redemption decision is made (positive or negative), that is, all weeks t for which $RT^1_{i,t-1} \le 1$. Finally ζ_i^1 is a vector of normal distributed error terms with variance $\sigma_{i1}^2 = 1$.

For the redemption fraction we write

$$y_i^2 = M_i^2 \begin{pmatrix} \mu_{i2} \\ \gamma_{i2} \end{pmatrix} + Z_i^2 \beta_{i2} + W_i^2 \delta_2 + \zeta_i^2, \tag{A2}$$

where y_i^2 is a $T_{2i}x1$ vector of $\log(f_{it}/(1-f_{it}))$, only for those observations where $RI_{it} = 1$. The matrices M_i^2, Z_i^2, and W_i^2 are defined analogous to M_i^1, Z_i^1, and W_i^1. ζ_i^2 is a vector of random errors each with variance $\sigma_{i2}^2 = \sigma_{R,i}^2$.

For the points savings incidence we define

$$y_i^3 = M_i^3 \begin{pmatrix} \mu_{i3} \\ \gamma_{i3} \end{pmatrix} + Z_i^3 \beta_{i3} + W_i^3 \delta_3 + \zeta_i^3, \tag{A3}$$

where y_i^3 is a $T_{i3}x1$ vector containing the elements SI^*_{it}. The elements of the error term have variance $\sigma_{i3}^2 = 1$. Note that $T_{i3} = T$, for all i.

Finally for the point savings amount we have

$$y_i^4 = M_i^4 \begin{pmatrix} \mu_{i4} \\ \gamma_{i4} \end{pmatrix} + Z_i^4 \beta_{i4} + W_i^4 \delta_4 + \zeta_i^4, \tag{A4}$$

with y_i^4 a $T_{4i}x1$ vector of $\log(S_{it})$, only for those observations where $SI_{it} = 1$. The obvious definitions apply to M_i^4, Z_i^4, and W_i^4. The elements of ζ_i^4 have variance $\sigma_{i4}^2 = \sigma_{S,i}^2$.

Sample λ_1, λ_2 and λ_3

To sample λ_1, λ_2 and λ_3 we employ a random walk Metropolis Hastings [RW-MH] sampler. The candidate values are obtained as $\lambda_k^{cand} \sim N(\lambda_k^{current}, s_k^2)$, for $k = 1,2,3$, where the s_k^2 are set such that we obtain an acceptance rate between 15% and 40%. As the candidate density is symmetric, the acceptance probability depends only on the likelihood of the data. In principle we could take this likelihood conditional on all other parameters, including the effect-size parameters of the mailings, accessibility, and the post-reward effect.

However, these effect-size parameters are expected to be quite dependent on λ_k. Therefore, in this step, we integrate out the effect-size parameters of mailings and the post-reward effect to obtain better mixing.

We first split Γ_2 in four parts (Γ_2^1, ..., Γ_2^4), one for each equation. Next, each Γ_2^k is split in two parts: the part including the mailing, accessibility, and post-reward effects ($\widetilde{\Gamma}_2^k$), and the remainder (Γ_2^{k*}). Our approach can be seen as sampling from the distribution of $(\lambda_1, \lambda_2, \lambda_3, \widetilde{\Gamma}_2^k \ (k = 1, ..., 4))$ given all other parameters (including $\Gamma_2^{k*}, k = 1, ..., 4$), by first sampling from $(\lambda_1, \lambda_2, \lambda_3)$ given the other parameters and next sampling from $\widetilde{\Gamma}_2^k$ given $\lambda_1, \lambda_2, \lambda_3$ and the other parameters. This first step is discussed below; the other step is discussed later in this appendix.

The acceptance probability in the RW-MH sampler depends on

$$l_\lambda(\lambda_1, \lambda_2 \lambda_3) = \prod_{k=1}^{4} \int_{\widetilde{\gamma}_k} \left(\prod_{i=1}^{N} \pi\left(y_i^k | \lambda_1, \lambda_2 \lambda_3, \widetilde{\gamma}_k, \text{other parameters} \right) \right) \pi(\widetilde{\gamma}_k) d\widetilde{\gamma}_k,$$
(A5)

where $\widetilde{\gamma}_k = vec\left(\widetilde{\Gamma}_2^k\right)$, $\pi(\widetilde{\gamma}_k) \propto 1$, and

$$\pi\left(y_i^k | \lambda_1, \lambda_2 \lambda_3, \widetilde{\gamma}_k, \text{other parameters} \right)$$
$$= (2\pi)^{\frac{T_{ik}}{2}} (\sigma_{ik})^{-T_{ik}} \exp\left(-\frac{1}{2\sigma_{ik}^2} \left(v_i^k - H_{ik}' \widetilde{\gamma}_k \right)' \left(v_i^k - H_{ik}' \widetilde{\gamma}_k \right) \right),$$
(A6)

where $v_i^k = y_i^k - M_i^k \begin{pmatrix} \mu_{ik} \\ \gamma_{ik} \end{pmatrix} - Z_i^{k*} \beta_{ik}^* - W_i^k \delta_k$, and $H_{ik} = \left(V_i' \otimes \widetilde{Z}_i^k \right)$. $\widetilde{Z}_i^k$ and Z_{ik}^{k*} are defined such that they separate the post-reward, accessibility, and mailing variables from the other variables, respectively. Note that $\widetilde{Z}_i^k$ is a function of λ_k. The product over all i of the density in Eq. (A5) is proportional to $\exp\left(-\frac{1}{2} \left(\ddot{v}^k - \ddot{H}_k' \widetilde{\gamma}_k \right)' \left(\ddot{v}^k - \ddot{H}_k' \widetilde{\gamma}_k \right) \right)$, where $\ddot{v}^k$ is obtained by stacking the vectors v_i^k / σ_{ik} and $\ddot{H}_k$ is obtained by stacking the matrices $\frac{1}{\sigma_{ik}} H_{ik}$. Next we observe that

$$\exp\left(-\frac{1}{2} \left(\ddot{v}^k - \ddot{H}_k' \widetilde{\gamma}_k \right)' \left(\ddot{v}^k - \ddot{H}_k' \widetilde{\gamma}_k \right) \right)$$
$$= \exp\left(-\frac{1}{2} \left[\left(\widetilde{\gamma}_k - \hat{\gamma}_k \right)' \left(\ddot{H}_k' \ddot{H}_k \right) \left(\widetilde{\gamma}_k - \hat{\gamma}_k \right) + \ddot{v}^{k'} \ddot{v}^k - \ddot{v}^{k'} \ddot{H}_k \left(\ddot{H}_k' \ddot{H}_k \right)^{-1} \ddot{H}_k' \ddot{v}^k \right] \right),$$
(A7)

with $\widehat{\widetilde{\gamma}_k} = \left(\ddot{H}_k' \ddot{H}_k \right)^{-1} \ddot{H}_k' \ddot{v}^k$. The integral in Eq. (A5) is therefore proportional to

$$\exp\left(\frac{1}{2} \ddot{v}^k \ddot{H}_k \left(\ddot{H}_k' \ddot{H}_k \right)^{-1} \ddot{H}_k' \ddot{v}^k \right) \left| \ddot{H}_k' \ddot{H}_k \right|^{-\frac{1}{2}} \int_{\widetilde{\gamma}_k} \left| \ddot{H}_k' \ddot{H}_k \right|^{\frac{1}{2}}$$
$$\exp\left(-\frac{1}{2} \left(\widetilde{\gamma}_k - \widehat{\widetilde{\gamma}_k} \right)' \left(\ddot{H}_k' \ddot{H}_k \right) \left(\widetilde{\gamma}_k - \widehat{\widetilde{\gamma}_k} \right) \right) d\widetilde{\gamma}_k.$$
(A8)

The integral above is the kernel of a multivariate normal and therefore the integral is proportional to 1. Therefore we get

$$l_\lambda(\lambda_1, \lambda_2, \lambda_3) \propto \prod_{k=1}^{4} \exp\left(\frac{1}{2} \ddot{v}^k \ddot{H}_k \left(\ddot{H}_k' \ddot{H}_k \right)^{-1} \ddot{H}_k' \ddot{v}^k \right) \left| \ddot{H}_k' \ddot{H}_k \right|^{-\frac{1}{2}}.$$
(A9)

Finally, the acceptance rate becomes

$$\min\left\{ 1, \frac{l_\lambda\left(\lambda_1^{cand}, \lambda_2^{cand} \right)}{l_\lambda\left(\lambda_1^{current}, \lambda_2^{current} \right)} \right\}.$$
(A10)

Sample α

To sample α we also use a RW-MH sampler. The procedure is similar to that presented above. However, now we split Γ_2^k into the pre-reward effect size ($\widetilde{\Gamma}_2^{-k}$) and the remainder ($\Gamma_2^{k*}$). The derivation of the acceptance probability is equivalent to the derivation above.

Sample $\theta_i = (\mu_{i1}, \mu_{i2}, \mu_{i3}, \mu_{i4}, \gamma_{i1}, \gamma_{i2}, \gamma_{i3}, \gamma_{i4})'$

We sample the elements of this vector in four steps, one for each equation. We sample μ_{ik} and γ_{ik} by combining

$$y_i^k - Z_i^k \beta_{ik} - W_i^k \delta_k = M_i^k \begin{pmatrix} \mu_{ik} \\ \gamma_{ik} \end{pmatrix} + \zeta_i^k,$$
(A11)

with the hierarchical distribution for μ_{ik} and γ_{ik} conditional on the other parameters, which follows from $\theta_i \sim N(\Gamma_1 V_i, \Omega)$. Denote the conditional mean for $(\mu_{ik}, \gamma_{ik})'$ by m_i^k and the conditional variance by V_i^k. We now draw μ_{ik} and γ_{ik} from a multivariate normal with mean

$$\left(\frac{1}{\sigma_{ik}^2} M_i^{k'} M_i^k + \left(V_i^k \right)^{-1} \right)^{-1} \left(\frac{1}{\sigma_{ik}^2} M_i^{k'} \left(y_i^k - Z_i^k \beta_{ik} - W_i^k \delta_k \right) + \left(V_i^k \right)^{-1} m_i^k \right),$$
(A12)

and variance

$$\left(\frac{1}{\sigma_{ik}^2} M_i^{k'} M_i^k + \left(V_i^k \right)^{-1} \right)^{-1}.$$
(A13)

Sample $\sigma_{R,i}^2$ and $\sigma_{S,i}^2$

Conditional on the other parameters, $\sigma_{R,i}^2$ has an inverted χ^2-distribution with degrees of freedom equal to $\nu + T_{i2}$, where ν gives the prior degrees of freedom (set to 5). The scale parameter equals $\zeta_i^{2'} \zeta_i^2 + \nu s$, where s controls the scale under the prior (set to 1). The sampling of $\sigma_{S,i}^2$ follows equivalent steps, with the same prior settings.

Sample RT_{it} and k_{it}

For every redemption occasion, we sample the moment at which the redemption decision was made. This moment defines k_{it} and RT_{it}. This moment is sampled without conditioning on RD_{it}^*. In other words, we sample from the joint distribution of RD_{it}^*, RT_{it} and k_{it} by first sampling from the marginal distribution of the latter two variables and next from the conditional for the first variable (see the step below).

To sample the moment of the redemption decision, we calculate the conditional probabilities for all possible number of purchase occasions between the moment of redemption and the redemption decision. This number is denoted by $k^* = 0,1,...,m$. Each value of k^* corresponds to a particular sequence of RD^* and k. In the rare case where there are two redemptions in m weeks, the upper bound of k^* equals the number of weeks between the redemptions, that is, 1 if the redemptions are in two consecutive weeks. To reduce notation, below we assume the upper bound equals m. Consider a redemption happening at time t^*, the conditional probability for a

particular value of k^* is proportional to

$$
\begin{cases}
\pi_0 Pr[RD_{it^*}^*>0] \prod_{k=1}^{m-1} Pr[RD_{it^*-k}^* \leq 0] \prod_{k=0}^{m-1} \left[\phi\left(\zeta_{it^*-k}^3\right) \frac{1}{\sigma_{S,i}} \phi\left(\frac{\zeta_{it^*-k}^4}{\sigma_{S,i}}\right) \right] & \text{if } k^*=0 \\[2ex]
\pi_1 Pr[RD_{it^*}^*>0] \prod_{k=1}^{m-1} Pr[RD_{it^*-k}^* \leq 0] \prod_{k=0}^{m-1} \left[\phi\left(\zeta_{it^*-k}^3\right) \frac{1}{\sigma_{S,i}} \phi\left(\frac{\zeta_{it^*-k}^4}{\sigma_{S,i}}\right) \right] & \text{if } k^*=1 \\[2ex]
\pi_{k^*} Pr[RD_{it^*-(k^*-1)}^*>0] \prod_{k=k^*}^{m-1} Pr[RD_{it^*-k}^* \leq 0] \prod_{k=0}^{m-1} \left[\phi\left(\zeta_{it^*-k}^3\right) \frac{1}{\sigma_{S,i}} \phi\left(\frac{\zeta_{it^*-k}^4}{\sigma_{S,i}}\right) \right] & \text{if } k^*=2,...,m,
\end{cases}
\tag{A14}
$$

where the final product gives the likelihood contribution of the points savings decisions at and before the moment of redemption. The terms ζ_{it}^3 and ζ_{it}^4 are defined in Eqs. (A3) and (A4) and implicitly depend on k^* through the dependence on RT_{it}.

Sample SI^*_{it} and RD^*_{it}

Given the other parameters and RT_{it}, the latent variables SI_{it}^* and RD_{it}^* have a truncated normal distribution. The latent variable SI^*_{it} (RD^*_{it}) is negative if individual i does not make a purchase (positive redemption decision) at time t. Otherwise, it is positive. Note that a redemption decision can only be made at time t if $RT_{it-1} \leq 1$. In case $RT_{it-1} > 1$, RD_{it}^* is not sampled. RD^*_{it} is sampled from the appropriate truncated normal with mean $\mu_{i1} + \gamma_{i1}t + Z_{it}^{R\,\prime}\beta_{i1} + W_t'\delta_1$ and variance 1. The mean for SI^*_{it} equals $\mu_{i3} + \gamma_{i3}t + Z_{it}^{S\,\prime}\beta_{i3} + W_t'\delta_3$.

Sample $\pi_0,... \pi_m$

To sample $\pi_0,... \pi_m$ we first count the number of times the "time gap" between redemption decision and redemption occasion equals j; we denote this count by c_j. The prior distribution for the vector π is set to a Dirichlet $(1,1,...1)$ distribution. This distribution is quite uninformative. The conditional distribution of the vector π now becomes a Dirichlet distribution with parameters $1 + c_0$, $1 + c_1$,... $1 + c_m$.

Sample Γ_1 and Ω

Given all θ_i vectors, the sampling of Γ_1 and Ω follows the standard results for the multivariate regression model (see Rossi, Allenby, & McCulloch, 2005). In order to improve performance we have an inverted Wishart prior on the variance. We set the degrees of freedom to 10 and the location parameter such that the expected value of the distribution equals 0.5 times a unit matrix.

Sample δ

Given all latent variables and the other parameters, δ_k has a multivariate normal distribution with mean

$$
\left(\sum_{i=1}^N \frac{1}{\sigma_{ik}^2} W_i^{k\,\prime} W_i^k \right)^{-1} \sum_{i=1}^N \frac{1}{\sigma_{ik}^2} W_i^{k\,\prime} \left(y_i^k - M_i^k \begin{pmatrix} \mu_{ik} \\ \gamma_{ik} \end{pmatrix} - Z_i^k \beta_{ik} \right),
\tag{A15}
$$

and variance

$$
\left(\sum_{i=1}^N \frac{1}{\sigma_{ik}^2} W_i^{k\,\prime} W_i^k \right)^{-1}.
\tag{A16}
$$

Sample Γ_2

We split the matrix Γ_2 in four parts, the part related to equation k is denoted by Γ_2^k. We now use the fact that $Z_i^k \beta_{ik} = Z_i^k \Gamma_2^k V_i = (V_i' \otimes Z_i^k) vec(\Gamma_2^k)$. This allows us to write

$$
y_i^k - M_i^k \begin{pmatrix} \mu_{ik} \\ \gamma_{ik} \end{pmatrix} - W_i^k \delta_k = \left(V_i' \otimes Z_i^k \right) vec(\Gamma_2^k) + \zeta_i^k.
\tag{A17}
$$

Collecting the equations across all members we obtain a multivariate normal distribution for $vec(\Gamma_2^k)$ with mean

$$
\left(\sum_{i=1}^N \frac{1}{\sigma_{ik}^2} \left(V_i V_i' \otimes Z_i^{k\,\prime} Z_i^k \right) \right)^{-1} \left(\sum_{i=1}^N \frac{1}{\sigma_{ik}^2} \left(V_i' \otimes Z_i^k \right)' \left(y_i^k - M_i^k \begin{pmatrix} \mu_{ik} \\ \gamma_{ik} \end{pmatrix} - W_i^k \delta_k \right) \right),
\tag{A18}
$$

and variance

$$
\left(\sum_{i=1}^N \frac{1}{\sigma_{ik}^2} \left(V_i V_i' \otimes Z_i^{k\,\prime} Z_i^k \right) \right)^{-1}.
\tag{A19}
$$

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ijresmar.2014.06.001.

References

Arkes, H. R., Joyner, C. A., Pezzo, M. V., Gradwohl Nash, J., Siegel-Jacobs, K., & Stone, E. (1994). The psychology of windfall gains. *Organizational Behavior and Human Decision Processes, 59*(3), 331–347.

Berry, J. (2013). *Bulking up: The 2013 COLLOQUY loyalty census: Growth and trends in U.S. loyalty program activity.* Colloquy (Retrieved December 12, 2013 from www.colloquy.com).

Blattberg, R. C., Kim, B.D., & Neslin, S. A. (2008). *Database marketing: Analyzing and managing customers.* New York: Springer.

Bolton, R. N., Kannan, P. K., & Bramlett, M.D. (2000). Implications of loyalty program membership and service experiences for customer retention and value. *Journal of the Academy of Marketing Science, 28*(1), 95–108.

Cameron, A.C., & Trivedi, P. K. (2005). *Microeconometrics: Methods and applications.* New York: Cambridge University Press.

Clarke, D.G. (1976). Econometric measurement of the duration of advertising effect on sales. *Journal of Marketing Research, 13*(4), 345–357.

Cragg, J. C. (1971). Some statistical models for limited dependent variables with applications to the demand for durable goods. *Econometrica, 39*(5), 829–844.

CRMtrends (2012). Loyalty marketing best practices: Breakage. Retrieved September 21, 2012 from http://www.crmtrends.com/loyalty.html

Dhar, R., Huber, J., & Khan, U. (2007). The shopping momentum effect. *Journal of Marketing Research, 44*(3), 370–378.

Dorotic, M., Bijmolt, T. H. A., & Verhoef, P. C. (2012). Loyalty programmes: Current knowledge and research directions. *International Journal of Management Reviews, 14*(3), 217–237.

Dowling, G. R., & Uncles, M. (1997). Do customer loyalty programs really work? *Sloan Management Review, 38*(4), 71–82.

Drèze, X., & Hoch, S. J. (1998). Exploiting the installed base using cross-merchandising and category destination programs. *International Journal of Research in Marketing, 15*(5), 459–471.

Drèze, X., & Nunes, J. C. (2009). Feeling superior: The impact of loyalty program structure on consumers' perceptions of status. *Journal of Consumer Research, 35*(6), 890–905.

Drèze, X., & Nunes, J. C. (2011). Recurring goals and learning: The impact of successful reward attainment on purchase behavior. *Journal of Marketing Research, 48*(2), 268–281.

Feld, S., Frenzen, H., Krafft, M., Peters, K., & Verhoef, P. C. (2013). The effects of mailing design characteristics on direct mail campaign performance. *International Journal of Research in Marketing, 30*(2), 143–159.

Franses, P. H. (2005). On the use of econometric models for policy simulation in marketing. *Journal of Marketing Research, 42*(1), 4–14.

Gönül, F. F., Kim, B.D., & Shi, M. (2000). Mailing smarter to catalog customers. *Journal of Interactive Marketing, 14*(2), 2–16.

Gordon, N., & Hlavinka, K. (2011). *Buried treasure: The 2011 forecast of U.S. consumer loyalty program points value.* Colloquy (Retrieved May 5, 2012 from www.colloquy.com).

Gwinner, K. P., Gremler, D.D., & Bitner, M. J. (1998). Relational benefits in services industries: The customer's perspective. *Journal of the Academy of Marketing Science, 26*(2), 101–114.

Haisley, E., & Loewenstein, G. (2011). It's not what you get but when you get it: The effect of gift sequence on deposit balances and customer sentiment in a commercial bank. *Journal of Marketing Research*, 48(1), 103–115.

Hartmann, W. R., & Viard, V. B. (2008). Do frequency reward programs create switching costs? A dynamic structural analysis of demand in a reward program. *Quantitative Marketing and Economics*, 6(2), 109–137.

Henderson, C. M., Beck, J. T., & Palmatier, R. W. (2011). Review of the theoretical underpinnings of loyalty programs. *Journal of Consumer Psychology*, 21(3), 256–276.

Higgins, E. T. (1989). Knowledge accessibility and activation: Subjectivity and suffering from unconscious sources. In J. S. Uleman, & J. A. Bargh (Eds.), *Unintended thought* (pp. 75–123). New York: Guilford Press.

Kim, B.D., Shi, M., & Srinivasan, K. (2001). Reward programs and tacit collusion. *Marketing Science*, 20(2), 99–120.

Kivetz, R., & Simonson, I. (2002). Earning the right to indulge: Effort as a determinant of customer preferences toward frequency program rewards. *Journal of Marketing Research*, 39(2), 155–170.

Kivetz, R., Urminsky, O., & Zheng, Y. (2006). The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress and customer retention. *Journal of Marketing Research*, 43(1), 39–58.

Konus, U., Neslin, S. A., & Verhoef, P. C. (2014). The effect of search channel elimination on purchase incidence, order size and channel choice. *International Journal of Research in Marketing*, 41(1), 49–69.

Kopalle, P. K., Sun, Y., Neslin, S. A., Sun, B., & Swaminathan, V. (2012). The joint sales impact of frequency reward and customer tier components of loyalty programs. *Marketing Science*, 31(2), 216–235.

Kumar, V., & Shah, D. (2004). Building and sustaining profitable customer loyalty for the 21st century. *Journal of Retailing*, 80(4), 317–329.

Lal, R., & Bell, D. E. (2003). The impact of frequent shopper programs in grocery retailing. *Quantitative Marketing and Economics*, 1(2), 179–202.

Land, T. (2013, June 13). *Reward expiration and redemption.* Colloquy (Retrieved October 20, 2013 from www.colloquy.com).

Leenheer, J., van Heerde, H. J., Bijmolt, T. H. A., & Smidts, A. (2007). Do loyalty programs really enhance behavioral loyalty? An empirical analysis accounting for self-selecting members. *International Journal of Research in Marketing*, 24(1), 31–47.

Lemon, K. N., & von Wangenheim, F. (2009). The reinforcing effects of loyalty program partnerships and core service usage a longitudinal analysis. *Journal of Service Research*, 11(4), 357–370.

Leone, R. P. (1995). Generalizing what is known about temporal aggregation and advertising carryover. *Marketing Science*, 14(3), 141–149.

Levey, R. H. (2011). *Marketers overlook "buried treasure" in unredeemed loyalty points.* Chief Marketer (Retrieved January 5, 2013 from http://www.chiefmarketer.com/database-marketing/loyalty-crm/marketers-overlook-buried-treasure-in-unredeemed-loyalty-points-20042011).

Lewis, M. (2004). The influence of loyalty programs and short-term promotions on customer retention. *Journal of Marketing Research*, 41(3), 281–292.

Liu, Y. (2007). The long-term impact of loyalty programs on consumer purchase behavior and loyalty. *Journal of Marketing*, 71(4), 19–35.

Madden, D. (2008). Sample selection versus two-part models revisited: The case of female smoking and drinking. *Journal of Health Economics*, 27(2), 300–307.

Magi, A. W. (2003). Share of wallet in retailing: the effects of customer satisfaction, loyalty cards and shopper characteristics. *Journal of Retailing*, 79(2), 97–106.

McCartney, S. (2012, May 16). Best airlines for redeeming miles. *Wall Street Journal* (Retrieved June 1, 2013 from http://online.wsj.com/news/articles/SB10001424052702303879604577408271116224192).

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica*, 46(1), 69–85.

Nunes, J. C., & Drèze, X. (2006). Your loyalty program is betraying you. *Harvard Business Review*, 84(4), 124–131.

Palmatier, R., Jarvis, C. B., Bechkoff, J. R., & Kardes, F. R. (2009). The role of customer gratitude in relationship marketing. *Journal of Marketing*, 73(5), 1–18.

Petty, R. E., & Cacioppo, J. T. (1979). Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology*, 37(10), 1915–1926.

Ratneshwar, S., Warlop, L., Mick, D.G., & Seeger, G. (1997). Benefit salience and consumer's selective attention to product features. *International Journal of Research in Marketing*, 14(3), 245–259.

Risselada, H., Verhoef, P. C., & Bijmolt, T. H. A. (2014). Dynamic effects of social influence and direct marketing on the adoption of high-technology products. *Journal of Marketing*, 78(2), 52–68.

Roehm, M. L., Pullins, E. B., & Roehm, H. A., Jr. (2002). Designing loyalty-building programs for packaged goods brands. *Journal of Marketing Research*, 39(2), 202–213.

Rossi, P. E., Allenby, G. M., & McCulloch, R. E. (2005). *Bayesian statistics and marketing.* New Jersey: John Wiley & Sons.

Rothschild, M. L., & Gaidis, W. C. (1981). Behavioral learning theory: Its relevance to marketing and promotions. *Journal of Marketing*, 45(2), 70–78.

Shugan, S. M. (2005). Loyalty programs? Are they shams? *Marketing Science*, 24(2), 185–193.

Smith, A., & Sparks, L. (2009a). Reward redemption behavior in retail loyalty schemes. *British Journal of Management*, 20(2), 204–218.

Smith, A., & Sparks, L. (2009b). "It's nice to get a wee treat if you've had a bad week": Consumer motivations in retail loyalty scheme points redemption. *Journal of Business Research*, 62(5), 542–547.

Statistical Yearbook of the Netherlands (2009). Income and spending. Retrieved June 2010 from http://www.cbs.nl/NR/rdonlyres/421A3A8C-956D-451D-89B6-D2113587F940/0/2009a3pub.pdf

Stauss, B., Schmidt, M., & Schoeler, A. (2005). Customer frustration in loyalty programs. *International Journal of Service Industry Management*, 16(3), 229–252.

Stourm, V., Bradlow, E. T., & Fader, P.S. (2013, November 18). *Stockpiling in linear loyalty programs (Working Paper)* (Retrieved from the SSRN website: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2142482).

Taylor, G. A., & Neslin, S. A. (2005). The current and future sales impact of a retail frequency reward program. *Journal of Retailing*, 81(4), 293–305.

The Economist (2005, December 20). Funny money. Retrieved January 6, 2013 from http://www.economist.com/node/5323615

van Diepen, M., Donkers, B., & Franses, P. H. (2009a). Dynamic and competitive effects of direct mailings: A charitable giving application. *Journal of Marketing Research*, 46(1), 120–133.

van Diepen, M., Donkers, B., & Franses, P. H. (2009b). Does irritation induced by charitable direct mailings reduce donations? *International Journal of Research in Marketing*, 26(3), 180–188.

Verhoef, P. C. (2003). Understanding the effect of customer relationship management efforts on customer retention and customer share development. *Journal of Marketing*, 67(4), 30–45.

Wiebenga, J. H., & Fennis, B.M. (2014). The road traveled, the road ahead, or simply on the road? When progress framing affects motivation in goal pursuit. *Journal of Consumer Psychology*, 24(1), 49–62.

Wooldridge, J. M. (2011). *Econometric analysis of cross-section and panel data* (2nd ed.). Cambridge, MA: MIT Press.

Zhang, J., & Breugelmans, E. (2012). The impact of an item-based loyalty program on consumer purchase behavior. *Journal of Marketing Research*, 49(1), 50–65.

Zhang, Z. J., Krishna, A., & Dhar, S. K. (2000). The optimal choice of promotional vehicles: Front-loaded or rear-loaded incentives? *Management Science*, 64(3), 348–362.

Full Length Article

Variable selection in international diffusion models

CrossMark

Sarah Gelper [a,*], Stefan Stremersch [b,c]

[a] Eindhoven University of Technology, Innovation, Technology Entrepreneurship & Marketing Group, Den Dolech 2, PO Box 513, 5600 MB Eindhoven, The Netherlands
[b] Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands
[c] IESE Business School Universidad de Navarra, Barcelona, Spain

ARTICLE INFO

ABSTRACT

Prior research comes to different conclusions as to what country characteristics drive diffusion patterns. One prime difficulty that may partially explain this divergence between studies is the sparseness of the data, in terms of the periodicity as well as the number of products and countries, in combination with the large number of potentially influential country characteristics. In face of such sparse data, scholars have used nested models, bivariate models and factor models to explore the role of country covariates. This paper uses Bayesian Lasso and Bayesian Elastic Net variable selection procedures as powerful approaches to identify the most important drivers of differences in Bass diffusion parameters across countries. We find that socio-economic and demographic country covariates (most pronouncedly so, economic wealth and education) have the strongest effect on all diffusion metrics we study. Our findings are a call for marketing scientists to devote greater attention to country covariate selection in international diffusion models, as well as to variable selection in marketing models at large.

© 2014 Published by Elsevier B.V.

1. Introduction

Since the 80s (Heeler & Hustad, 1980), international diffusion of new products has strongly established itself as a research stream within the international marketing literature. International diffusion[1] studies predominantly seek to explain variation in new product growth patterns across countries using country characteristics, such as economics, culture or demographics (for recent contributions, see Chandrasekaran & Tellis, 2008; Talukdar, Sudhir, & Ainslie, 2002; Stremersch & Lemmens, 2009; Stremersch & Tellis, 2004; Tellis, Stremersch, & Yin, 2003; Van den Bulte & Stremersch, 2004; van Everdingen, Fok, & Stremersch, 2009).

An important difference among these studies – beyond the difference in the products or countries included – is the set of country-level covariates included in the model. Model specification in terms of covariates in international diffusion models is particularly challenging. There is no consensus in the literature about which country characteristics should or should not be included in an international diffusion model. Marketing scholars justify their choice for a certain set of explanatory variables by theoretical reasoning. Especially in international diffusion, the theory is very rich and thus the number of variables that one could consider including is very large. At the same time, the data is

often sparse, in terms of periodicity, and number of countries and products. Standard statistical estimation techniques often have difficulties to fit such large models on such sparse data. Therefore, scholars may drop one or more of the available variables through subjective choice and iterative testing of smaller models, at the risk of omission.

Scholars who do not restrict their model ex ante, often face ill-conditioning of the design matrix – or harmful multicollinearity – as a significant problem (see Chandrasekaran & Tellis, 2008; Tellis et al., 2003). An ill-conditioned design matrix may pre-empt inference from the full model, by which people resort again to dimensionality reduction techniques, such as estimating nested models (Stremersch & Tellis, 2004), bivariate models (Chandrasekaran & Tellis, 2008), composite models (Gatignon, Eliashberg, & Robertson, 1989) or factor models (Helsen, Jedidi, & Desarbo, 1993; Tellis et al., 2003). Nested models and bivariate models, however, also face the risk of omitted variable bias. Composite and factor models are difficult to interpret and are unable to disentangle the effects of distinct country covariates.

This paper uses Bayesian Lasso (Hans, 2009; Park & Casella, 2008) and Bayesian Elastic Net (Hans, 2011; Li & Lin, 2010) to explore which country characteristics matter most in international diffusion. These procedures can cope with sparse data (i.e., many variables and few data points) by specifying an appropriate informative prior, which leads to a specific form of Bayesian regularization (Fahrmeir, Kneib, & Konrath, 2010). By construction of the Lasso and Elastic Net priors, some of the estimated regression coefficients will be exactly zero, identifying a subset of most important variables. The procedure simultaneously executes shrinkage and variable selection, while alternative

* Corresponding author at: Rotterdam School of Management, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands. Tel.: +31 10 4082732.
    E-mail addresses: sgelper@rsm.nl (S. Gelper), stremersch@ese.eur.nl (S. Stremersch).
[1] We use the term international diffusion as a synonym to international new product growth.

shrinkage methods (e.g. Ridge regression) do not include variable selection and alternative variable selection methods (e.g. Bayesian model averaging) do not include shrinkage. The advantage of the Lasso and Elastic Net procedures over shrinkage methods without variable selection is that it leads to more stable estimation results and to the identification of a relatively small subset of variables that exhibit the strongest effects (Tibshirani, 1996). The advantage over variable selection methods without shrinkage is that the latter methods still lack power in a sparse data setting because the shrinkage is crucial for dealing with correlated covariates, as we show in a simulation study.

We estimate a Bayesian version of the Bass diffusion model (Bass, 1969) which was introduced by Lenk and Rao (1990) and subsequently extended by Talukdar et al. (2002). Bayesian analysis is particularly well suited for international diffusion models because of the multilevel structure of the data. The model decomposes the product- and country-variance, which is important, given that the sample of countries is typically not the same for all products and the product variance is typically larger than the country variance. Also, regularization to deal with sparse data comes natural in a Bayesian setting via the use of an informative prior. Scholars in both marketing (Lenk & Orme, 2009) and statistics (Fahrmeir et al., 2010) show an increasing attention for the usefulness of Bayesian regularization by informative priors.

We have data on the penetration levels of 6 high technology products (CD players, internet, ISDN, mobile phones, personal computers, and video cameras) in a total of 55 countries around the world. These data are also used in van Everdingen et al. (2009) and were graciously made available to us by Yvonne van Everdingen. We complement these data with an extensive set of country characteristics that encompasses the country characteristics used in previous studies on new product adoption, ranging from socio-economic over cultural to demographic and geographic characteristics.

The results indicate that even though many country characteristics have been related to new product growth in the past, in our particular set of countries and products, the following small set of variables explains most of the between-country variation. A first predominant variable is economic wealth. It has a strong positive effect on all three parameters of the Bass diffusion model. A second important variable is education which positively affects both the market potential (m) and the innovation coefficient (p). Beyond economic wealth and education, income inequality has a negative effect on the market potential (m), economic openness affects the innovation coefficient (p), while mobility affects the imitation coefficient (q) in the Bass diffusion model. Future application of variable selection techniques on other samples of international diffusion data, may yield a promising path towards generalizable findings.

2. Prior literature on international diffusion

Table 1 inventories the international diffusion literature using variations of the Bass diffusion model. For every study, we list which country characteristics are studied, whether a dimensionality reduction method is used, and which country characteristics the authors found to influence diffusion. A more general overview of diffusion and new product growth models can be found in Peres, Muller, and Mahajan (2010).

Gatignon et al. (1989) construct three country-level constructs (cosmopolitanism, mobility and sex roles), using 9 variables and find that the three constructs significantly relate to the parameters of the Bass diffusion model. This finding was confirmed in Kumar, Ganesh, and Echambadi (1998). Takada and Jain (1991) use two dummies to account for cultural and communication differences in four Pacific Rim countries and find them to affect the adoption rate. Helsen et al. (1993) cluster countries based on six factors extracted from a total of 23 country characteristics and conclude that life style and health status are related to the parameters of the Bass diffusion model. Dekimpe, Parker, and

Table 1
Overview of international diffusion literature using country characteristics in the Bass diffusion model.

Reference	Included country characteristics	Dimensionality reduction method	Important country characteristics
Gatignon et al. (1989)	Quantity of foreign mail sent and received, international telegrams received, foreign travel, foreign visitors received, number of telephones in use, percentage of population owning at least one car, number of cars per inhabitant, per capita mileage driven, women in labor force.	3 composites: cosmopolitanism, mobility and sex roles	Cosmopolitanism, mobility, sex roles
Takada and Jain (1991)	Culture dummy (high vs low context), communication dummy (homophilous vs heterophilous).	No reduction	Culture dummy, communication dummy
Helsen et al. (1993)	Number of air passengers/km, air cargo, number of newspapers, population, cars per capita, motor gasoline consumption, electricity production, life expectancy, physicians per capita, political stability, imports, exports, GDP per capita, phones per capita, electricity consumption per capita, foreign visitors per capita, tourist expenditures per capita, tourist receipts per capita, consumer price index, newspaper circulation, hospital beds, education expenditures/government budget, graduate education in population per capita.	6 factors: mobility, health status, trade, life style, cosmopolitanism, miscellaneous	Life style, health status
Kumar et al. (1998)	Quantity of foreign mail sent and received, international telegrams received, foreign travel, foreign visitors received, number of telephones in use, percentage of population owning at least one car, number of cars per inhabitant, per capita mileage driven, women in labor force.	3 composites: cosmopolitanism, mobility and sex roles	Cosmopolitanism, mobility, sex roles
Dekimpe et al. (1998)	Population growth, number of population centers, GNP per capita, crude death rate, communism, number of ethnic groups.	No reduction	Population growth, no. population centers, crude death rate, no. ethnic groups
Talukdar et al. (2002)	Income per capita, dependents–working ratio, Gini index, urbanization, international trade, TV penetration, newspapers per capita, illiteracy rate, number of ethnic groups, women in labor force, minutes of international telephone calls.	No reduction	Income per capita, urbanization, international trade, illiteracy
Van den Bulte and Stremersch (2004)	Individualism, uncertainty avoidance, power distance, masculinity, GDP per capita, Gini index.	No reduction	Individualism, uncertainty avoidance, power distance, masculinity, Gini index
Albuquerque et al. (2007)	Population size, GDP per capita, sustainability, literacy, urbanization.	No reduction	Population size

Note: Composites are constructed based on a fixed set of pre-selected country characteristics per construct; factors are obtained by principle component analysis on the complete set of country characteristics; "No reduction" means that all country characteristics are included in the model without transformation.

Sarvary (1998) find a significant effect on the diffusion process of four out of six covariates under consideration, mainly related to demographics. Talukdar et al. (2002) specify a hierarchical Bayesian Bass model, in which per capita income, urbanization and international trade affect a new product's market potential, a country's illiteracy rate affects the innovation coefficient, and no country covariate affects the imitation coefficient. Van den Bulte and Stremersch's (2004) meta-analysis shows that the q/p ratio reported in prior applications of the Bass diffusion model varies with national culture, income inequality and the presence of competing standards. Albuquerque, Bronnenberg, and Corbett (2007) study cross-country spillovers in the adoption of ISO certifications and find that only population size has an influence on market potential.

While we focus on the Bass diffusion model, there are a number of notable studies on international new product growth beyond applications of the Bass model. Dekimpe, Parker, and Sarvary (2000b) study the time between a product's first worldwide introduction and a country's adoption time and identify economic wealth (GNP per capita) and number of ethnic groups to be the main drivers. Chandrasekaran and Tellis (2008), Stremersch & Tellis (2004) and van Everdingen et al. (2009) study cross-country variation in time-to-takeoff and international spill-overs in takeoff. These studies include a large set of country-level predictors, such as economic wealth, income inequality and culture, but find mixed effects as to the influence they have on time-to-takeoff. Stremersch and Tellis (2004) study the growth phase of the product life cycle, after takeoff and identify economic wealth (GDP per capita) as the main growth driver. Stremersch and Lemmens (2009) and Putsis, Balasubramanian, Kaplan, and Sen (1997) develop flexible models to study international new product growth. Stremersch and Lemmens (2009) find that, in the context of pharmaceuticals, regulatory regimes are an important determinant of cross-country variation in new product sales growth. Lemmens, Croux, and Stremersch (2012) propose a method to dynamically segment countries based on the observed penetration pattern of new products. They exploit such dynamic segments to predict the national penetration patterns of new products prior to launch. Putsis et al. (1997) fit a flexible mixing model with cross-country influence and find significant effects of GDP per capita and number of televisions in use on differences in international diffusion patterns.

If scholars have used dimensionality reduction methods in this literature, they are mainly of two kinds, often executed in parallel. A first kind is to estimate a series of shorter models that are nested in the full model (see for instance Tellis et al., 2003). The estimation of such nested models comes at the risk of omitted variable bias or pretest error bias in the remaining regression coefficients. Such bias can even result in estimated parameters that switch signs as a consequence of omission. For instance, Chandrasekaran and Tellis (2008) report a significantly negative influence of uncertainty avoidance on time-to-takeoff when it is the only variable in the model, while the same coefficient is significantly positive in a model that also includes other country characteristics.

A second dimensionality reduction method is factor analyzing the explanatory variables and only retaining a set of factors that explain a large part of the variance (for instance Chandrasekaran & Tellis, 2008; Helsen et al., 1993; Tellis et al., 2003). The most important factors capture most of the variation in the complete set of variables and represent underlying unobserved constructs. In practice, however, it may be hard to give a meaningful interpretation to these unobserved constructs and this interpretation may not be universally accepted among scholars. Another drawback of factor analysis is that the commonly used estimation procedures (i.e. principal components or maximum likelihood) do not take into account the response variable in the model. This is a limitation, because in a regression context one wants to use different information in the explanatory variables depending on the response. Partial least squares or sliced inverse regression (Li, 1991; Naik, Hagerty, & Tsai, 2000) do take into account the response variable in the construction of the factors, but the interpretation of the resulting factor model

becomes even more difficult. It is hard to argue that the factors represent an underlying construct if they by definition are different depending on the response variable in the model.

3. Method

In this section, we first review three penalized likelihood methods, Ridge regression, the Lasso and the Elastic Net. The latter two have a variable selection property which allows exploring which variables matter most. Next, we draw the analogy with Bayesian regularization through the choice of appropriate priors on the regression coefficients. We then describe the Bass diffusion model and illustrate the properties of the three regularization methods, as compared to the standard regression using diffuse normal priors, in the Bass diffusion model using a simulation study.

3.1. Penalized likelihood and Bayesian regularization

Consider the multiple linear regression model

$$y = Xb + e \tag{1}$$

where y is the response vector and X is the $(N \times k)$ matrix containing k regressors. Assume the response to be mean-centered and the regressors to be standardized such that no intercept is included. Furthermore, let $b = (b_1, ..., b_k)'$ denote the vector of regression coefficients. Assuming that the error term e follows a $N(0, \sigma_e^2)$ distribution, the penalized likelihood estimator maximizes the likelihood under a constraint on the coefficients. The constraints we consider here are designed to shrink the estimated parameters towards zero. In particular, the three penalized estimators we consider are all of the form

$$\hat{b} = \arg\min_b \sum_{i=1}^{N} (y_i - X_i b)^2 \text{ subject to } (1-\alpha)\sum_{j=1}^{k} |b_j| + \alpha \sum_{j=1}^{k} b_j^2 < t \tag{2}$$

for some positive value of t. For $\alpha = 1$, the estimator defined by Eq. (2) is the Ridge estimator, which puts a constraint on the sum of the squared coefficients. For $\alpha = 0$, the constraint is on the sum of the absolute values of the coefficients, which yields the Lasso estimator. Any value of α such that $0 < \alpha < 1$, results in the Elastic Net estimator. The Elastic Net constraint on the coefficients is a combination of the Ridge and the Lasso constraints.

To illustrate the difference between Ridge, Lasso and Elastic Net, the shrinkage obtained by each method is illustrated in Figs. 1, 2 and 3 for the case with only two regressors. The gray area in the figures specifies the region within which the coefficients on the axis are subject to the constraint in Eq. (2) for a certain value of t. A larger value of t would correspond to a less stringent constraint on the parameters, which would be represented by a larger gray constraint region. The value of t is typically chosen by cross-validation. The ellipses represent equi-mean-squared-error lines. The inner x-mark represents the maximum likelihood solution, which is the solution to problem (2) without a constraint on the coefficients. The inner ellipses are closer to the maximum likelihood solution, and thus have lower mean squared error values. For each shrinkage type (Ridge, Lasso or Elastic Net), we present a case with uncorrelated and correlated regressors in subfigures (a) and (b) respectively.

The solution to minimization problem (2) is given by the tangent point between the gray constraint region and the ellipsoid. Fig. 1 illustrates why Ridge regression does not result in variable selection. Because of the circular shape of the constraint region, the Ridge solution will only rarely result in zero coefficient estimates. A problem with Ridge regression is the sensitivity of the outcome to changes in the constraint region, especially when the regressors are correlated (Fig. 1b). If the amount of shrinkage is strong enough, the Ridge coefficients can change signs as compared to the least squares solution, as is the case in Fig. 1b.
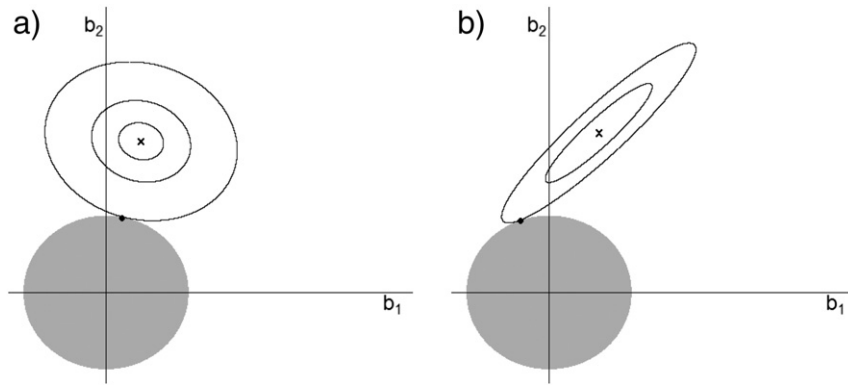
Fig. 1. Ridge solution under uncorrelated regressors (a) and correlated regressors (b). The gray region indicates the region satisfying the regression constraint, the ellipses represent equi-mean-squared-error lines, the least squares solution (x) and the regularized estimates (dot).

The variable selection property of the Lasso is illustrated in Fig. 2. Because of the squared shape of the gray constraint region, the Lasso solution can result in zero coefficients, ensuring variable selection. The tangency point between the gray constraint region and the ellipsoid is on the b_2 axis, resulting in a parameter estimate for b_1 which is exactly equal to zero, both in the uncorrelated case (Fig. 2a) and in the correlated case (Fig. 2b). The Lasso solution is in general more stable than the Ridge solution.

The Elastic Net constraint region presented in Fig. 3 for $\alpha = 0.5$ is an intermediate to the Ridge circular constraint region and the Lasso squared constraint region. The main difference with Ridge is that, similar to the Lasso, the corners of the Elastic Net constraint region facilitate variable selection. The difference with the Lasso is that due to the rounding of the constraint region in between the axes, the Elastic Net tends to select strongly correlated variables jointly in or out the model, which is often referred to as the *grouping effect* (Zou & Hastie, 2005). For instance, in Fig. 3b the correlated variables are selected together by the Elastic Net, while only one variable is selected by the Lasso in Fig. 2b.

The solution to Eq. (2) has a Bayesian interpretation as well. The link between regularization methods and hierarchical Bayes is well documented (e.g. Evgeniou, Pontil, & Toubia, 2007; Fahrmeir et al., 2010). In particular, the solution is equivalent to the posterior mode of the regression coefficients under a specific prior. Bayesian Ridge specifies a normal prior given by

$$b\big|\sigma_e^2,\lambda_r^2 \;\sim\; N\!\left(\mathbf{0},\frac{\sigma_e^2}{\lambda_r^2}I_k\right), \tag{3}$$

where the prior mean is zero for all regression parameters and the shrinkage parameter λ_r^2 controls the precision of the prior. A more precise posterior is obtained for larger values of the shrinkage parameter. Taking the prior mean equal to zero in combination with a tight prior is a conservative choice. If after combination with the data the posterior of a parameter is located away from zero, we safely conclude that the corresponding regressor is important in the model. A prior specification for the shrinkage parameter is defined as

$$\lambda_r \;\sim\; \text{Gamma}(r_r, s_r) \tag{4}$$

and for the error variance

$$\sigma_e^{-2} \;\sim\; \text{Gamma}(u_r, v_r)\;. \tag{5}$$

Posterior evaluation is obtained via the Gibbs sampler.

The disadvantage of Ridge regression is that it does not achieve variable selection. Moreover, the amount of effective shrinkage is hard to control. It not only depends on the shrinkage parameter but also on the amount of correlation in the data. The more correlation, the less stable Ridge regression becomes, which makes it a poor method for data with harmful multicollinearity like ours. This instability is shown by Tibshirani (1996) in a penalized likelihood setting, but also holds in the Bayesian setting as we illustrate in Appendix A.

Following the work of Hans (2009) and Park and Casella (2008), the Lasso point estimator for regression model (1), is defined as the mode of the posterior density of the regression parameters when imposing an independent Laplace prior with mean zero on the regression coefficients

$$b_j\big|\sigma_e^2,\lambda_l \;\sim\; \text{Laplace}\!\left(0,\frac{\sigma_e}{\lambda_l}\right) = \frac{\lambda_l}{2\sigma_e}\exp\!\left(\frac{-\lambda_l}{\sigma_e}\big|b_j\big|\right). \tag{6}$$
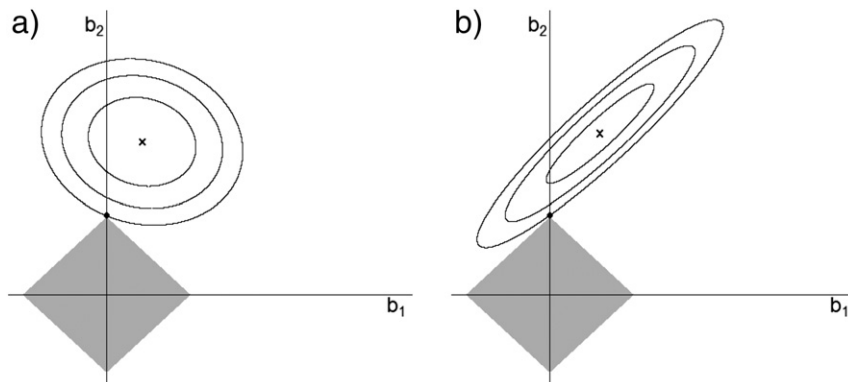

Fig. 2. Lasso solution under uncorrelated regressors (a) and correlated regressors (b). The gray region indicates the region satisfying the regression constraint, the ellipses represent equi-mean-squared-error lines, the least squares solution (x) and the regularized estimates (dot).

Fig. 3. Elastic Net solution for $\alpha = 0.5$ under uncorrelated regressors (a) and correlated regressors (b). The gray region indicates the region satisfying the regression constraint, the ellipses represent equi-mean-squared-error lines, the least squares solution (x) and the regularized estimates (dot).
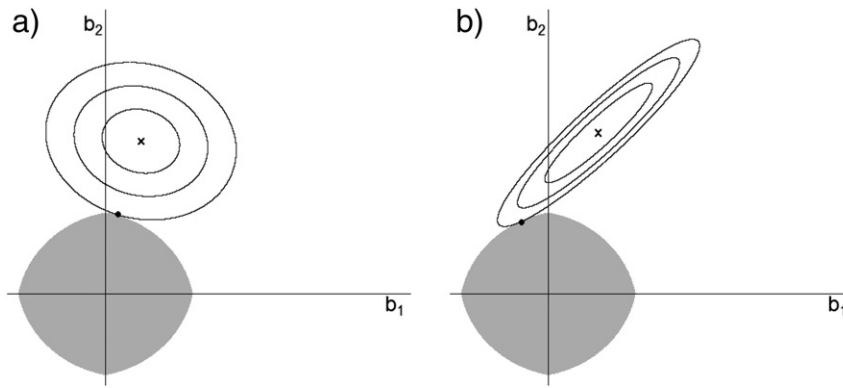
As for the Ridge, a more precise posterior is obtained for larger values of the shrinkage parameter at the cost of more shrinkage. Similar to the term $|b_j|$ in the constraint in Eq. (2), the term $|b_j|$ in the prior in Eq. (6) facilitates variable selection. The key to variable selection using this procedure is that, depending on the value of the shrinkage parameter, the posterior mode of some regression coefficients can become exactly zero. Even though there is posterior mass located away from zero, whether the posterior mode of a regression coefficient is zero or not has important consequences for model interpretation. By construction, the mode will always be included in the highest posterior density region. Therefore, a regression parameter with zero posterior mode will never be "significant". Posterior evaluation is achieved via the Gibbs sampler described in Hans (2009). The latter requires a rejection sampling step to draw from the conditional distribution of the scale parameter, which we implemented using the R package ars by Perez Rodriguez (2009).

The Laplace prior puts more prior mass close to zero and in the tails as compared to a normal prior, as illustrated in Fig. 4, reflecting the idea that there are many small effects and a number of important effects. Other variable selection procedures build on the belief that some of the true regression coefficients are exactly zero, which is hard to defend (O'Hara & Sillanpaa, 2009). Especially in the international diffusion model, it is likely that all country characteristics influence the diffusion process, but some variables to a much lesser extent than others. In this context, variable selection should be considered as a tool to help the researcher distinguish between the small and the large important effects

rather than identifying zero-effects. The Elastic Net prior on the regression coefficients is a compromise between the Gaussian prior of Ridge regression and the Laplace prior of the Lasso (Li & Lin, 2010)

$$p\left(b_j | \sigma_e^2, \lambda_{1n}, \lambda_{2n}\right) \propto \exp\left(-\frac{1}{2\sigma_e^2}\left(\lambda_{1n}\left|b_j\right| + \lambda_{2n}b_j^2\right)\right). \tag{7}$$

A comparison between the priors is given in Fig. 4. The elastic net prior is an intermediate between the Normal and the Laplace prior. The spike at zero facilitates variable selection. The Bayesian Elastic Net has been used in marketing research before by Rutz, Trusov, and Bucklin (2011) in the context of paid search advertising.

3.2. Bayesian representation of the international Bass diffusion model

We use the Bayesian regularization methods as described in the previous section to identify which country characteristics best explain differences in diffusion patterns. To specify a Bayesian version of the Bass diffusion, denote by $S_{ij}(t)$ the penetration level of product j in country i at period t after commercialization. The diffusion process of product j in country i is given by

$$\Delta S_{ij}(t) = \left(p_{ij} + q_{ij}\frac{S_{ij}(t-1)}{m_{ij}}\right)\left(m_{ij} - S_{ij}(t-1)\right) + \varepsilon_{ij}(t), \tag{8}$$

where $\Delta S_{ij}(t) = S_{ij}(t) - S_{ij}(t-1)$, and $\varepsilon_{ij} \sim N(0, \sigma_j^2)$. The first parameter m_{ij} captures the market potential, p_{ij} is the coefficient of innovation, and q_{ij} is the coefficient of imitation for product j in country i. We include an additive error term in Eq. (8) following Albuquerque et al. (2007) to ensure that penetration levels are allowed to show small decreases over time, as is observed in our data.

To know which country characteristics influence the diffusion process, the diffusion parameters m_{ij}, p_{ij} and q_{ij} are first decomposed into a country- and product-specific component after controlling for the product-country specific introduction lag denoted by L_{ij}. Denote the vector of Bass model parameters for product j in country i by $\theta_{ij} = (m_{ij}, p_{ij}, q_{ij})$, then the variance decomposition is given by

$$\text{logit}\left(\theta_{ij}\right) = \alpha_i + \beta_j + \gamma L_{ij} + \xi_{ij} \quad \text{with} \quad \xi_{ij} \sim N\left(\mathbf{0}, \Sigma_\xi\right), \tag{9}$$

where we allow a full covariance matrix Σ_ξ. Since the values of $\theta_{ij} = (m_{ij}, p_{ij}, q_{ij})$ are between zero and one, we use a logit transformation to obtain values on the whole real line, which is similar to the approach in Lenk and Rao (1990). The first component of the γ vector is fixed at zero because the introduction lag only affects the growth rate towards the market potential (determined by p_{ij} and q_{ij}), and not the market potential m_{ij} itself.
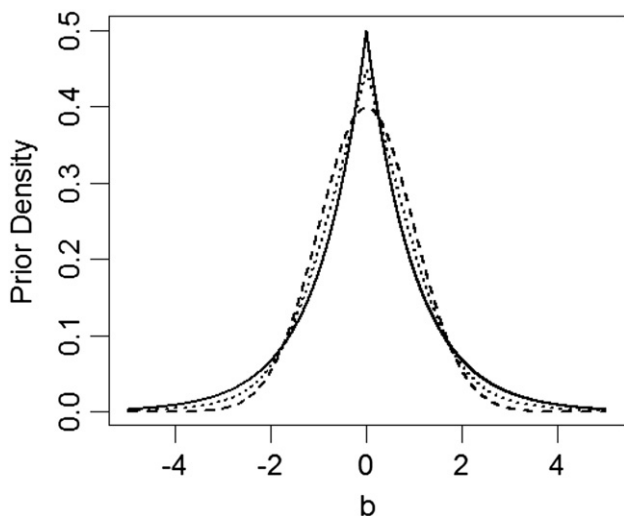

Fig. 4. Normal prior used for Ridge (dashed line), Laplace prior used for Lasso (full line) and mixed prior used for Elastic Net (dotted line).

Since our interest is in the country-specific parameters in vector α_i, it is further regressed on the country characteristics. These are represented in the matrix X of dimension $(C \times k)$, with C the number of countries and k the number of country characteristics. The third level of the Bass diffusion model then is of the form

$$\alpha_i = X_i\delta + \eta_i \qquad \text{with} \qquad \eta_i \sim N\left(\mathbf{0}, \Sigma_\eta\right), \tag{10}$$

where X_i is the row vector of length k with country characteristics for country i. The regression parameter matrix δ is of dimension $(k \times 3)$ and captures the effect of the country characteristics on the diffusion process. The matrix δ is our primary object of interest – it captures the influence of the country characteristics on the diffusion pattern – and is estimated using Bayesian regularization as described in Section 3.1.

The product-specific effects are captured in the parameter vector β_j which is modeled as a random effect with mean zero (for identification)

$$\beta_j \sim N\left(\mathbf{0}, \Sigma_\beta\right). \tag{11}$$

We assume Σ_η and Σ_β to be diagonal. All prior specifications are given in Appendix B1. The posterior and estimation details of the first level are given in Appendix B2.

Posterior evaluation of the parameters is achieved through MCMC draws. In the Lasso and Elastic Net case, apart from the posterior MCMC draws we are interested in the posterior mode of the regression coefficients in δ because the mode marks selection. The mode is obtained by maximum a posteriori (MAP) estimation. MAP estimation in the Bayesian Lasso setting is common, see e.g. Figueiredo (2003) and Genkin, Lewis, and Madigan (2007). The MAP estimator is obtained using Rao-Blackwellization as in Hans (2009) and Hans (2011). For each draw in the MCMC chain, we store the conditional distribution of δ on a fine grid. This conditional distribution is orthant normal for both Lasso and Elastic Net and sometimes has a zero-mode due to the shape of the prior. We then average the stored conditionals over the MCMC draws for each grid point to obtain an estimate of the marginal posterior from which we can easily obtain the mode as the Lasso or Elastic Net point estimate.

3.3. Simulation study

We run a simulation study to assess the performance of the Bayesian regularization methods described in Section 3.1 for estimating the country-level regression model parameters in the Bass diffusion model of Section 3.2. To assess in which conditions the Bayesian Lasso and Elastic Net perform better than Ridge or regression using diffuse normal priors, we run a 2×2 simulation design. As country covariates are typically highly correlated, the first dimension we vary is the amount of multicollinearity. We compare the accuracy of the regularization procedures across two settings, one in which covariates are correlated and one in which covariates are uncorrelated. The second dimension we take into consideration is the sparseness of the true model, i.e. whether some of the country covariates have an actual *zero* effect on the diffusion process. Due to their variable selection properties, these sparse models are favored by the Lasso and the Elastic Net. But since we do not know whether there truly are zero effects, we study the methods' performance in a situation where all country covariates

have an effect but some have a stronger effect than others. This leads to four simulation settings where we have either correlated covariates or uncorrelated covariates, and either true model sparseness or not.

The specifics of the simulation setting are as follows. We simulate data according to the multi-product multi-country Bass diffusion model specified by Eqs. (8) to (11). The dimensions of the model are the same as in our data, i.e. we simulate 6 products, 55 countries and 17 country covariates ($k = 17$). We generate the country covariates X from a normal distribution with mean zero. In the correlated settings, the correlation between x_i and x_j equals $\rho^{|i-j|}$ with $\rho = 0.5$, following the simulation setup of Tibshirani (1996). In the uncorrelated settings, we set $\rho = 0$. In the sparse settings, we again follow Tibshirani (1996) and set $\delta_{\cdot j} = (3,1.5,0,0,2,0,\ldots,0)'$ for $j \in \{1,2,3\}$ corresponding to the diffusion metrics m, p and q respectively. In the non-sparse settings, we set $\delta_{\cdot j} = (3, 1.5, 1, 3/4, 3/5, \ldots, 3/17)'$ such that each covariate influences the diffusion process but the last covariates are gradually less important than the first. To make the simulation specification complete, we set $\sigma_j^2 = 0.01$, $L_{ij} = 0$ for all i and j, $\gamma = \mathbf{0}$, $\Sigma_\xi = \Sigma_\eta = \Sigma_\beta = I_3$ and we generate $N_s = 200$ data sets in each simulation setting.

As our main interest is in the performance to retrieve the parameters of the country-level regression models in δ, we compare the mean squared error

$$\text{MSE} = \frac{1}{3kN_s} \sum_{j=1}^{3} \sum_{i_s=1}^{N_s} \left(\hat{\delta}_{\cdot j} - \delta_{\cdot j}\right)'\left(\hat{\delta}_{\cdot j} - \delta_{\cdot j}\right), \tag{12}$$

where $\hat{\delta}$ is the vector of point estimates of the country-covariate effects. For the Lasso and Elastic Net, we use the posterior mode as described above. For Ridge regression and regression using diffuse normal priors, we use the posterior median as a point estimator. All MSE values are computed based on strandardized variables.

For the sparse simulation settings, we also assess how well the Lasso and Elastic Net perform in terms of identifying those variables that have a non-zero coefficient. We compute the true positive rate (TPR) as the proportion of non-zero coefficients that are estimated to be non-zero, i.e. are correctly selected into the model. We also compare the true negative rate (TNR) as the proportion of zero coefficients that are estimated to be zero, i.e. correctly estimated as having a zero-effect:

$$\begin{aligned} \text{TPR} &= \frac{\#\left\{(ij) : \hat{\delta}_{ij} \neq 0 \text{ and } \delta_{ij} \neq 0\right\}}{\#\left\{\delta_{ij} \neq 0\right\}}, \\ \text{TNR} &= \frac{\#\left\{(ij) : \hat{\delta}_{ij} = 0 \text{ and } \delta_{ij} = 0\right\}}{\#\left\{\delta_{ij} = 0\right\}}. \end{aligned} \tag{13}$$

The mean squared error values are presented in Table 2. Overall, the Lasso achieves the best MSE values in all simulation settings. The benefit of the Lasso over the other methods, however, differs across the settings. The advantage of the Lasso is most pronounced when the covariates are correlated and the true model is sparse (Setting 4) and least pronounced when there is no multicollinearity and all predictors have an influence on the diffusion process (Setting 1). When the covariates are uncorrelated and the true model is sparse, both the Lasso and Elastic Net – which favor models with zero-coefficients – perform better than estimation based on diffuse normal priors and Ridge (Setting 2).

Table 2
Estimation accuracy (MSE * 1000).

			Lasso	Elastic Net	Diffuse normal priors	Ridge
Setting 1	Uncorrelated covariates	Non-sparse model	8.96	9.10	10.96	9.10
Setting 2		Sparse model	6.12	8.04	11.81	11.52
Setting 3	Correlated covariates	Non-sparse model	9.28	10.98	15.47	11.89
Setting 4		Sparse model	7.00	9.13	16.88	12.18

Table 3
Variable selection accuracy.

			True positive rate (TPR)		True negative rate (TNR)	
			Lasso	Elastic Net	Lasso	Elastic Net
Setting 2	Uncorrelated covariates	Sparse model	0.93	0.99	0.81	0.34
Setting 4	Correlates covariates	Sparse model	0.90	0.98	0.60	0.23

Table 4
Products and countries in our data set.

Products	Mobile phone, CD player, video camera, PC, Internet, ISDN
Countries	Argentina, Australia, Austria, Belgium, Brazil, Bulgaria, Canada, Chile, China, Colombia, Croatia, Czech Republic, Denmark, Ecuador, Estonia, Finland, France, Germany, Greece, Hong Kong, Hungary, India, Indonesia, Ireland, Israel, Italy, Japan, Malaysia, Mexico, Morocco, Netherlands, New Zealand, Norway, Pakistan, Peru, Philippines, Poland, Portugal, Romania, Russia, Singapore, Slovakia, Slovenia, South Africa, South Korea, Spain, Sweden, Switzerland, Taiwan, Thailand, Turkey, UK, USA, Venezuela, Vietnam

When the covariates are correlated and all of them have an effect, the Lasso and Elastic Net clearly outperform estimation based on diffuse normal priors and Ridge (Setting 3). In sum, whether the true model is sparse or not, methods like the Lasso and Elastic Net should be considered as methods that lead to superior outcomes when multicollinearity is present in the data.

The variable selection accuracy of the Lasso and the Elastic Net are reported in Table 3. Elastic Net has a better true positive rate than the Lasso, at the cost of a lower true negative rate. This holds true in a setting where the covariates are uncorrelated as well as when they are correlated. The correlated setting (Setting 4) is especially of interest because the Elastic Net was introduced as a method that performs better when the covariates are correlated. The *grouping effect* states that the Elastic Net tends to select groups of correlated variables jointly. In our sparse settings, the first two variables both have an effect and are correlated. The fifth variable also has an effect and is correlated with variables that have a zero-effect. In this setting, the Lasso has a true positive rate of 90% while the Elastic Net achieves 98%. However, the Elastic Net tends to select too many variables into the model that are correlated with those variables that have an effect. As a result, the true negative rate of the Elastic Net is only 23%, while that of the Lasso is 60%, which

illustrates the difference between both methods in terms of variable selection when the covariates are correlated.

4. Data

We use penetration data of six consumer durables in 55 countries listed in Table 4, gathered from publicly available sources, such as Euromonitor and the International Telecommunications Union. The country characteristics were gathered from publicly available sources such as the Statistical Yearbook of the United Nations, CIA World Factbook, World Development Indicators, U.S. Census Bureau, Euromonitor online, and Hofstede (2001). Country characteristics with multiple data points over the observation period were averaged.

We rely on the new product adoption and diffusion literature to specify our model in terms of country covariate inclusion. Table 5 gives an overview of the covariates we include, where the inclusion criterion is whether the variable has been used in previous diffusion literature. The country characteristics cover socio-economic, cultural, communication and demographic dimensions. The last column of Table 5 indicates to which growth metric (market potential, coefficient of innovation or coefficient of imitation) prior studies related each country covariate. To showcase the ability of variable selection methods to deal with long models, we link all available country characteristics to each diffusion metric. This procedure will allow us to explore whether or not there are important relationships that have not been identified or theorized on before.

To assess the degree of multicollinearity in our dataset, we compute the condition index of X as in Belsley, Kuh, and Welsh (1980). To obtain the condition index, we scale the variables in the X-matrix to have unit variance. According to Belsley et al. (1980), condition indices above 30 indicate moderate to strong multicollinearity. In our case, we obtain a condition index of 79.63, which is well beyond the threshold.

Table 5
List of included country characteristics.

Dimension	Variable	Operationalization	Related to diffusion metrics in previous literature
Socio-economic	Economic wealth	GDP per capita	m, p, q
	Inequality	GINI index on the household level based on net income	m, p, q
	Poverty	Under 5 year mortality rate	p, q
	Economic openness	(import + export)/GDP	m
	Education	Number of third-degree (university) students as a percentage of total population.	p
	Activity rate of women	% of women employed in nonagricultural sector	p, q
	Economic participation	Working to dependents ratio	m, p
Cultural	Individualism	Hofstede IND	p, q
	Uncertainty avoidance	Hofstede UAI	p, q
	Masculinity	Hofstede MAS	p, q
	Power distance	Hofstede PDI[a]	p, q
Communication	Media intensity	Number of newspapers per 1000 inhabitants	p
	Mobility	Number of cars per 1000 inhabitants	q
	Tourism	Number of incoming tourists per 1000 inhabitants	p, q
Demographic	Population growth	Annual population growth rate	p, q
	Population concentration	Population per square km	p, q
	Urbanization	% op people living in urban areas	m

[a] We do not include the fifth cultural dimension, Hofstede later added to his cultural framework, long-term orientation, because it is not available for a large number of countries in our dataset.

5. Results

5.1. Variable selection: Bayesian Lasso and Elastic Net

Table 6 presents the selected variables obtained by the Lasso and the Elastic Net and the posterior mode for a sequence of 10,000 draws after 2000 burn-in draws. The *prop*-values are the proportion of draws on the other side of zero than the mode. Because a variable is unselected from the model when the posterior mode is equal to zero, a *prop*-value cannot be calculated in such cases.

For all diffusion metrics, the predominant variable is economic wealth. Both the Lasso and the Elastic Net find that economic wealth has a positive effect on all three diffusion metrics. Talukdar et al. (2002) also found a strong effect of economic wealth on market potential but did not allow for an effect on the innovation and imitation coefficients, while according to our results this effect is strong. A second important variable is education, which influences both the market potential (m) and the innovation coefficient (p).

Apart from economic wealth and education, we find a distinct set of additional country covariates that affect the three diffusion metrics. We find a negative effect of income inequality on market potential. That is, all other things being equal, product adoption reaches a lower ceiling in such countries. We also find a positive effect of tourism on the market potential. Cosmopolitanism, which includes tourism, was one of the core constructs in the early studies of Gatignon et al. (1989) and Helsen et al. (1993). For the innovation coefficient, we find a positive effect of education and economic openness. We find a positive effect of mobility on the imitation coefficient, supporting the hypothesis that if people are more mobile they get in contact with more people and thus have a higher probability of influencing each other. All the remaining variables were not selected. Thus, after controlling for the included variables, they do not provide any additional information about the diffusion process, in our sample of products and countries. The latter subsentence is important and applies to all our findings reported in the present paper; to our experience, findings on international diffusion of new products are sensitive not only to the variable selection technique employed, but also to the sample composition in terms of which products and countries are covered as well as the extent to which such sample is balanced (i.e., the same products are covered across the same set of countries).

Table 5 summarizes which variables have been used as a driver of which metric in the previous literature. Including all variables as determinants of all diffusion metrics allowed us to extract three new findings on international diffusion. The first is the effect of education on market potential. All else equal, in a more educated population a higher proportion of the population will adopt new technologies. The second is the effect of tourism on market potential. The more touristic a country is, the more the population will get into contact with new technologies and thus the more people will eventually adopt. The third new effect is that of economic openness on the innovation coefficient.

The variable selection results obtained by the Lasso and the Elastic Net are very similar. Even though the Elastic Net is more sophisticated – as it chooses the intermediate between Ridge and Lasso in a data-adaptive way – this extra level of sophistication does not lead to substantially different insights in our setting. Fig. 5 compares the marginal densities of the effects of economic wealth and population growth on the market potential as estimated by the Lasso and the Elastic Net. Both methods identify economic wealth as an important variable. The Elastic Net posterior shrinks a bit more to zero, but there is no difference in substantive interpretation. As an illustration of an unselected variable, the right panel of Fig. 5 plots the posterior densities of the effect of population growth on market potential. Both posterior modes are zero, while the Lasso posterior is a bit more spiked. Similar comparisons between Lasso and Elastic Net posteriors are reported in Hans (2011) who finds small differences in the Lasso and Elastic Net posteriors using prostate cancer data (Stamey et al., 1989).

5.2. Diffuse normal priors and ridge

In Table 7, we report the results after estimating the Bass diffusion model using diffuse normal priors on all regression coefficients in Eq. (10) and Ridge regression. Diffuse normal priors are the most standard choice in Bayesian regression and are used in the Bass diffusion model by Talukdar et al. (2002), while Ridge regression is an alternative shrinkage method without the variable selection property as described above in Section 3. In the case of diffuse normal priors, none of the estimated effects is significant and so no conclusions can be drawn with respect to which variables influence which metric. When we do Bayesian regularization using Ridge, we only identify a positive effect of economic wealth on market potential. These poor conclusions with respect to

Table 6
Variable selection and significance for the Bayesian Lasso and Elastic Net procedure for each diffusion metric.

	Lasso						Elastic Net					
	Market potential		Innovation coefficient		Imitation coefficient		Market potential		Innovation coefficient		Imitation coefficient	
	Posterior mode	*Prop*-val	Posterior mode	*Prop*-val	Posterior mode	*Prop*-val	Posterior mode	*Prop*-val	Posterior mode	*Prop*-val	Posterior mode	*Prop*-val
Economic wealth	**.49**	**.02**	**.39**	**.02**	**.21**	**.03**	**.38**	**.02**	**.29**	**.05**	**.02**	**.04**
Inequality	−**.05**	**.04**	0		0		−**.04**	**.03**	0		0	
Poverty	0		0		0		0		0		.01	.42
Economic openness	.11	.16	**.10**	**.03**	0		.03	.32	.20	.23	0	
Education	**.20**	**.02**	**.15**	**.04**	0		**.01**	**.04**	.06	**.04**	.03	.11
Activity rate of women	0		0		0		0		0		0	
Economic participation	.05	.33	.04	.43	0		−.08	.20	0		0	
Individualism	0		0		0		0		0		−.25	.15
Uncertainty avoidance	0		0		0		−.03	.34	0		0	
Masculinity	0		0		.10	.45	0		.05	.35	.11	.24
Power distance	0		0		0		0		0		0	
Media intensity	0		0		0		0		0		0	
Mobility	.11	.22	0		**.46**	**.04**	.21	.09	0		**.30**	**.03**
Tourism	**.07**	**.03**	.10	.39	0		0		−.09	.31	0	
Population growth	0		0		0		0		.15	.25	0	
Population concentration	0		0		.05	.41	0		0		0	
Urbanization	0		.03	.25	0		.17	.07	0		.01	.40

Note: The posterior mode is the point estimate of the Lasso or Elastic Net. The *prop*-value is the proportion of draws on the other side of zero than the mode, indicating significance. The *prop*-value cannot be computed if the mode is zero, resulting in blank entries. Parameter estimates and *prop*-values are indicated in bold when the *prop*-value is less than .05.
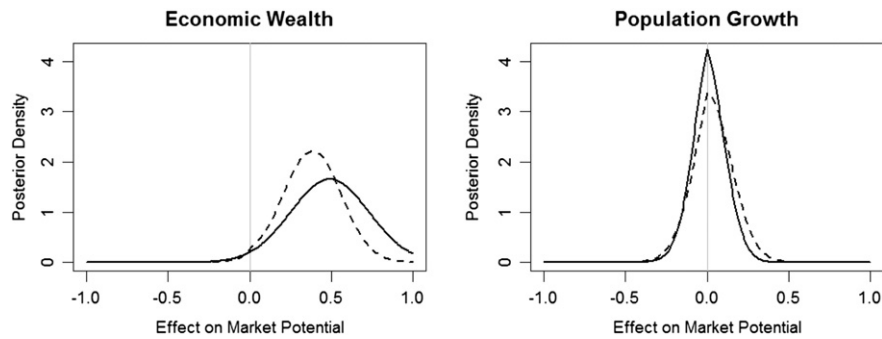
Fig. 5. Density estimates of the marginal posterior distributions of the effect of economic wealth (left) and population growth (right) on the market potential estimated using Lasso (full line) and Elastic Net (dashed line).

which country characteristic influences which diffusion metric are the result of the sparseness of the data. As we illustrated in the simulation section above, diffuse normal priors and ridge regression are poorly suited for a multicollinear setting like ours.

6. Discussion

Using the Bayesian Lasso and Elastic Net estimation procedures, we have shown that international variation in new product growth in our sample of products and countries is predominantly driven by economic wealth and education. In addition, economic inequality limits a new product's market potential. The innovation coefficient is also higher the higher the level of economic openness in a country. The imitation coefficient is higher, the higher the mobility of a country's citizens.

The application of Bayesian Lasso and Elastic Net on a larger sample of new products beyond high technology products (our present sample), such as laundry and appliances (e.g. Kumar & Krishnan, 2002), fast moving consumer goods, services, pharmaceuticals and entertainment products, may bring strong generalizable insights (on main effects or contingencies) to the international diffusion literature. The set of countries and products used in international diffusion studies will always have large effects on the findings given large product-country

interactions (Talukdar et al., 2002). An update to the meta-analytic approach, such as in Van den Bulte and Stremersch (2004) could therefore prove to be a valuable contribution to the international diffusion literature.

Such applications could also easily further enlarge the set of country covariates to variables that so far received little attention, such as distribution infrastructure, competition, or regulation (see Stremersch & Lemmens, 2009, for an exception), to yield newly discovered strong determinants of international diffusion patterns. The methodology we propose is ideally suited to handle even larger covariate sets. One particular fruitful challenge lies in the study of interaction effects among country covariates. While, the Bayesian Lasso and Elastic Net cannot guarantee the inclusion of a main effect conditional on the inclusion of an interaction, Bien, Taylor, and Tibshirani (2013) propose a non-Bayesian variant of the Lasso which does exactly that. There is room for a methodological contribution to extend such an approach to the Bayesian world.

In addition to the above applications, the present paper shows several additional limitations the reader should be aware of. It is well known that model averaging approaches substantially improve the prediction accuracy as opposed to fitting one single model (Eklund & Karlsson, 2007; Raftery, Madigan, & Hoeting, 1997; Wright, 2008). A model

Table 7
Parameter estimates and significance for procedures using diffuse normal priors and Ridge for each diffusion metric.

| | Diffuse normal priors | | | | | | Ridge | | | | | |
| | Market potential | | Innovation coefficient | | Imitation coefficient | | Market potential | | Innovation coefficient | | Imitation coefficient | |
	Posterior median	*Prop*-val	Posterior median	*Prop*-val	Posterior median	*Prop*-val	Posterior median	*Prop*-val	Posterior median	*Prop*-val	Posterior median	*Prop*-val
Economic wealth	.42	.15	−.10	.42	−.22	.32	**.30**	**.01**	.04	.24	.10	.40
Inequality	−.07	.40	−.01	.49	−.03	.46	−.06	.26	.12	.12	−.10	.30
Poverty	−.08	.32	−.02	.46	.02	.45	−.06	.21	.01	.36	−.07	.42
Economic openness	.07	.39	.09	.39	.01	.48	.05	.32	.00	.39	−.10	.35
Education	.02	.46	−.05	.44	−.03	.46	.04	.33	−.01	.27	−.02	.32
Activity rate of women	−.02	.46	.02	.48	.00	.50	.01	.44	−.02	.26	−.10	.21
Economic participation	−.11	.33	.01	.48	−.05	.43	−.08	.19	.02	.26	.06	.37
Individualism	−.10	.37	−.26	.25	−.33	.19	.00	.50	−.41	.39	.38	.39
Uncertainty avoidance	−.06	.41	.11	.38	.06	.44	−.04	.32	.01	.36	.25	.35
Masculinity	−.01	.48	.09	.34	.14	.25	−.03	.35	.01	.46	−1.32	.44
Power distance	−.12	.31	−.09	.39	−.03	.47	−.11	.13	.02	.26	−.06	.22
Media intensity	.10	.37	.04	.45	.07	.42	.15	.08	−.03	.25	.10	.23
Mobility	.24	.25	.10	.41	.07	.43	.19	.05	−.05	.34	.09	.30
Tourism	−.03	.45	−.14	.31	−.06	.41	.01	.47	−.06	.43	−.19	.47
Population growth	.04	.45	.19	.30	.16	.32	.03	.38	.29	.40	−.08	.38
Population concentration	.00	.49	−.02	.48	−.12	.38	.06	.28	.00	.39	.34	.31
Urbanization	.20	.21	.05	.44	.08	.40	.16	.04	.00	.28	−.26	.18

Note: The posterior median is the point estimate of the estimation using diffuse normal priors and Ridge. The *prop*-value is the proportion of draws on the other side of zero than the median, indicating significance. Parameter estimates and *prop*-values are indicated in bold when the *prop*-value is less than .05.

averaging approach to the Bayesian Lasso and Elastic Net as in Hans (2010) and Hans (2011) respectively could be used for predicting the diffusion metrics of a product in a certain country before launch (van Everdingen, Aghina, & Fok, 2005).

Second, the Bayesian version of the international Bass diffusion model could be formulated more flexibly. Previous research suggested making the error term of the Bass diffusion model autocorrelated and heteroskedastic (Fok & Franses, 2007). The Bayesian Lasso and Elastic Net procedures we use in this paper can be easily implemented in such alternative diffusion models.

Third, the variable selection techniques introduced in this paper can be extended to other models that capture international new product growth patterns, such as duration models for time-to-adoption, time-to-takeoff (e.g. Tellis et al., 2003; van Everdingen et al., 2009) or duration of the growth stage (Stremersch & Tellis, 2004), as well as to semiparametric sales models (e.g. splines), etc.

Despite these limitations, this paper contributes to marketing scholars' knowledge on dealing with sparse data, and offers a solution that is relatively easy to implement. It is clear that the on-going modeling practice, as documented here for international diffusion, can be improved substantially by implementing regularization methods, such as the Lasso and Elastic Net used in this paper. Such improvement would not only benefit the reliability of scholarly evidence, but would also allow to simultaneously explore a larger set of covariates and derive new empirical evidence on factors that remained uninvestigated so far. Within the research area of international diffusion, duration models (Dekimpe, Parker, & Sarvary, 2000a; Dekimpe et al., 2000b) such as the time-to-takeoff model (Chandrasekaran & Tellis, 2008; Stremersch & Tellis, 2004; Tellis et al., 2003; van Everdingen et al., 2009) could lead to additional insights into which variables explain international differences in the timing pattern of diffusion. Areas that come to mind outside the research area of new product diffusion (Naik et al., 2008) include churn modeling, in which datasets with more than 100 explanatory variables are not an exception (Lemmens & Croux, 2006; Naik, Wedel, & Kamakura, 2010), or the vector autoregressive (VAR) modeling tradition of marketing effectiveness (Dekimpe & Hanssens, 1999; Srinivasan, Pauwels, & Nijs, 2010), in which the number of regression parameters dramatically explodes as the number of variables increases. We hope that the benefit of the procedure proposed in the present paper does not remain contained to international diffusion models, but rather diffuses to other research areas in marketing science as well.

Appendix A. The Lasso versus the Ridge prior

Tibshirani (1996) shows that Ridge regression is sensitive to the amount of correlation between regressors, while the Lasso is not. We show that the same conclusions hold in a Bayesian setting. We generate 1000 data points according to the following model

$$y = \beta_1 x_1 + \beta_2 x_2 \ ,$$

without error term. The regression parameters are fixed at $\beta_1 = 6$ and $\beta_2 = 3$. The regressors x_1 and x_2 are obtained from a standard normal distribution with correlation ρ. Since no error term is included in the data generating process, we use the original Lasso prior introduced in Tibshirani (1996), which is unconditional on the scale parameter (Hans, 2009).

The parameter estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ for different shrinkage parameters and different correlations ($\rho = 0, 0.23, 0.45, 0.68, 0.9$) are plotted in Fig. A.1. It clearly illustrates the sparseness of the Lasso. If much shrinkage is applied, the estimate of β_2 is zero. In contrast to the Lasso, the Ridge solution strongly depends on ρ. For high correlations ($\rho = 0.9$), the Ridge procedure sometimes even overestimates the true parameter instead of shrinking it to zero. The variation in the Lasso estimates for
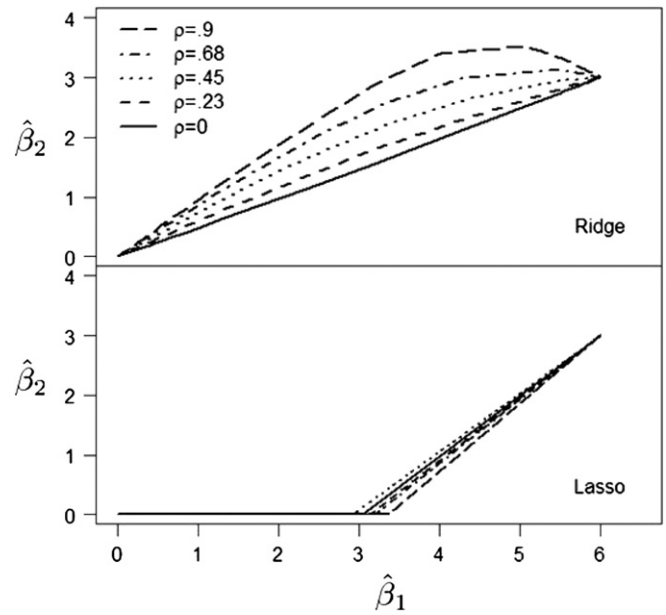

Fig. A.1. Ridge (upper panel) and Lasso (lower panel) regression estimates for different levels of shrinkage and for different values of the correlation coefficient ρ.

different values of ρ is not systematic and only due to the variation in the random generation of the regressors.

Appendix B. Model specification and estimation

Appendix B1. Prior specifications

Diffuse normal priors

$\sigma_j^{-2} \sim \text{Gamma}(1, 10)$
$\gamma|\Sigma_\xi \sim N(\mathbf{0}, 100\Sigma_\xi)$
$\Sigma_\xi \sim \text{Wishart}(5, 0.1I_3)$
$\delta_j^c|\sigma_{\eta,c}^2 \sim N(0, 100\sigma_{\eta,c}), \ c = 1, 2, 3 \text{ and } j = 1, ..., k$
$\sigma_{\eta,c}^{-2} \sim \text{Gamma}(1, 10)$
$\sigma_{\beta,c}^{-2} \sim \text{Gamma}(1, 10)$

Ridge priors

$\sigma_j^{-2} \sim \text{Gamma}(1, 10)$
$\gamma|\Sigma_\xi \sim N(\mathbf{0}, 100\Sigma_\xi)$
$\Sigma_\xi \sim \text{Wishart}(5, 0.1I_3)$
$\delta_j^c|\sigma_{\eta,c}^2 \sim N(0, \sigma_{\eta,c}/\lambda_{r,c}^2), \ c = 1, 2, 3 \text{ and } j = 1, ..., k$
$\sigma_{\eta,c}^{-2} \sim \text{Gamma}(1, 10)$
$\lambda_{r,c}^2 \sim \text{Gamma}(1, 1)$
$\sigma_{\beta,c}^{-2} \sim \text{Gamma}(1, 10)$

Lasso priors

$\sigma_j^{-2} \sim \text{Gamma}(1, 10)$
$\gamma|\Sigma_\xi \sim N(\mathbf{0}, 100\Sigma_\xi)$
$\Sigma_\xi \sim \text{Wishart}(5, 0.1I_3)$
$\delta_j^c|\sigma_{\eta,c}^2 \sim \text{Laplace}(0, \sigma_{\eta,c}/\lambda_{l,c}), \ c = 1, 2, 3 \text{ and } j = 1, ..., k$
$\sigma_{\eta,c}^{-2} \sim \text{Gamma}(1, 10)$
$\lambda_{l,c} \sim \text{Gamma}(1, 1)$
$\sigma_{\beta,c}^{-2} \sim \text{Gamma}(1, 10)$

Elastic Net priors

$\sigma_j^{-2} \sim \text{Gamma}(1, 10)$
$\gamma|\Sigma_\xi \sim N(\mathbf{0}, 100\Sigma_\xi)$
$\Sigma_\xi \sim \text{Wishart}(5, 0.1I_3)$
$\delta_j^c|\sigma_{\eta,c}^2 \propto exp((-1/(2\sigma_{\eta,c}))(\lambda_{1n,c}|\delta_j^c| + \lambda_{2n,c}(\delta_j^c)^2)), \ c = 1, 2, 3 \text{ and } j = 1, ..., k$
$\sigma_{\eta,c}^{-2} \sim \text{Gamma}(1, 10)$
$\lambda_{1n,c}^2 \sim \text{Gamma}(1, 1)$
$\lambda_{2n,c}^2 \sim \text{Gamma}(1, 1)$
$\sigma_{\beta,c}^{-2} \sim \text{Gamma}(1, 10)$

Appendix B2. MCMC draws in the first level of the Bass diffusion model

The parameters are estimated by drawing from their conditional posterior. In the first level of the Bass diffusion model, $\theta_{ij} = (m_{ij}, p_{ij}, q_{ij})'$ is obtained by a Metropolis–Hastings step. The posterior of θ conditional on σ^2 can be written as

$$p\left(\theta | \Delta S(t), \sigma^2\right) \propto \ell(\theta)\, p(\theta)\,,$$

where we drop subscripts to avoid notational clutter. The first component on the right hand side is the likelihood function and the second component the prior. The likelihood is given by

$$\ell(\theta) \propto \prod_{t=2}^{T} \exp\left\{ -\frac{1}{2\sigma^2} \left[\left(p + q\frac{S(t-1)}{m}(m - S(t-1)) \right) \right]^2 \right\}.$$

The prior follows a logistic normal distribution given by

$$p(\theta) = \frac{\prod_{i=1}^{3} \theta_i (1-\theta_i)}{|2\pi\Sigma|^{\frac{1}{2}}} exp\left\{ -\frac{1}{2} [\text{logit}(\theta) - \mu]' \Sigma^{-1} [\text{logit}(\theta) - \mu] \right\}.$$

More details on the logistic-normal distribution can be found in Atchison and Shen (1980). The parameter vector μ and matrix Σ are obtained from the second-level estimation. To obtain a candidate draw from $p(\theta|\Delta S(t), \sigma^2)$, we use a normal random-walk candidate generating function with variance such that the acceptance level is approximately 0.3. Denote the current value of θ by θ^c, then the candidate θ^* is accepted with probability $\min(1, p(\theta^*|\Delta S(t), \sigma^2)/p(\theta^c|\Delta S(t), \sigma^2))$.

Next, we draw σ_j from its conditional posterior distribution

$$\sigma_j^{-2}\Big|\Delta S(t), \theta \sim \text{Gamma}\left(1 + n_j, \frac{10 + n_j s_j^2}{1 + n_j} \right),$$

where n_j is the total number of observations for product j and s_j^2 is given by

$$s_j^2 = \sum_i \sum_t \left[\left(\Delta S_{ij}(t)\right) - \left(p_{ij} + q_{ij}\frac{S_{ij}(t-1)}{m_{ij}} \right) \left(m_{ij} - S_{ij}(t-1) \right)^2 \right]$$

References

Atchison, J., & Shen, S. M. (1980). Logistic-normal distributions: Some properties and uses. *Biometrika*, 67(2), 261–272.

Albuquerque, P., Bronnenberg, B. J., & Corbett, C. J. (2007). A spatiotemporal analysis of the global diffusion of ISO9000 and ISO14000 certification. *Management Science*, 53(3), 451–468.

Bass, F. (1969). A new product growth model for consumer durables. *Management Science*, 15(5), 215–227.

Belsley, D., Kuh, E., & Welsh, R. E. (1980). *Regression diagnostics.* New York: John Wiley & Sons, Inc.

Bien, J., Taylor, J., & Tibshirani, R. (2013). A Lasso for hierarchical interactions. *Annals of Statistics*, 41(3), 1111–1141.

Chandrasekaran, D., & Tellis, G. J. (2008). Global takeoff of new products: Culture, wealth, or vanishing differences? *Marketing Science*, 27(5), 844–860.

Dekimpe, M. G., & Hanssens, D.M. (1999). Sustained spending and persistent response: A new look at long-run marketing profitability. *Journal of Marketing Research*, 36(4), 397–412.

Dekimpe, M. G., Parker, P.M., & Sarvary, M. (1998). Staged estimation of international diffusion models: An application to global cellular telephone adoption. *Technological Forecasting and Social Change*, 57(1–2), 105–132.

Dekimpe, M. G., Parker, P.M., & Sarvary, M. (2000a). Global diffusion of technological innovations: A coupled-hazard approach. *Journal of Marketing Research*, 37(1), 47–59.

Dekimpe, M. G., Parker, P.M., & Sarvary, M. (2000b). Globalization: Modeling technology adoption timing across countries. *Technological Forecasting and Social Change*, 63(1), 25–42.

Eklund, J., & Karlsson, S. (2007). Forecast combination and model averaging using predictive measures. *Econometric Reviews*, 26(2–4), 329–363.

Evgeniou, T., Pontil, M., & Toubia, O. (2007). A convex optimization approach to modeling consumer heterogeneity in conjoint estimation. *Marketing Science*, 26(6), 805–818.

Fahrmeir, L., Kneib, T., & Konrath, S. (2010). Bayesian regularisation in structured additive regression: A unifying perspective on shrinkage, smoothing and predictor selection. *Statistics and Computing*, 20(2), 203–219.

Figueiredo, M.A. T. (2003). Adaptive sparseness for supervised learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(9), 1150–1159.

Fok, D., & Franses, P. H. (2007). Modeling the diffusion of scientific publications. *Journal of Econometrics*, 139(2), 376–390.

Gatignon, H., Eliashberg, J., & Robertson, T. S. (1989). Modeling multinational diffusion patterns: An efficient methodology. *Marketing Science*, 8(3), 231–247.

Genkin, A., Lewis, D.D., & Madigan, D. (2007). Large-scale Bayesian logistic regression for text categorization. *Technometrics*, 49(3), 291–304.

Hans, C. (2009). Bayesian lasso regression. *Biometrika*, 96(4), 835–845.

Hans, C. (2010). Model uncertainty and variable selection in Bayesian lasso regression. *Statistics and Computing*, 20(2), 221–229.

Hans, C. (2011). Elastic Net regression modeling with the orthant normal prior. *Journal of the American Statistical Association*, 106(496), 1383–1393.

Heeler, R. M., & Hustad, T. P. (1980). Problems in predicting new product growth for consumer durables. *Management Science*, 26(10), 1007–1020.

Helsen, K., Jedidi, K., & Desarbo, W. S. (1993). A new approach to country segmentation utilizing multinational diffusion patterns. *Journal of Marketing*, 57(4), 60–71.

Hofstede, G. (2001). *Culture's consequences comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). London: Sage Publications.

Kumar, V., Ganesh, J., & Echambadi, R. (1998). Cross-national diffusion research: What do we know and how certain are we? *Journal of Product Innovation Management*, 15(3), 255–268.

Kumar, V., & Krishnan, T. V. (2002). Multinational diffusion models: An alternative framework. *Marketing Science*, 21(3), 318–330.

Lemmens, A., & Croux, C. (2006). Bagging and boosting classification trees to predict churn. *Journal of Marketing Research*, 43(2), 276–286.

Lemmens, A., Croux, C., & Stremersch, S. (2012). Dynamics in the international market segmentation of new product growth. *International Journal of Research in Marketing*, 29(1), 81–92.

Lenk, P. J., & Orme, B. (2009). The value of informative priors in Bayesian inference with sparse data. *Journal of Marketing Research*, 46(6), 832–845.

Lenk, P. J., & Rao, A. G. (1990). New models from old: Forecasting product adoption by hierarchical Bayes procedures. *Marketing Science*, 9(1), 42–53.

Li, K. C. (1991). Sliced inverse regression for dimension reduction. *Journal of the American Statistical Association*, 86(414), 316–327.

Li, Q., & Lin, N. (2010). The Bayesian Elastic Net. *Bayesian Analysis*, 5(1), 151–170.

Naik, P. A., Hagerty, M. R., & Tsai, C. L. (2000). A new dimension reduction approach for data-rich marketing environments: Sliced inverse regression. *Journal of Marketing Research*, 37(1), 88–101.

Naik, P. A., Wedel, M., Bacon, L, Bodapati, A., Bradlow, E., Kamakura, W., et al. (2008). Challenges and opportunities in high-dimensional choice data analyses. *Marketing Letters*, 19(3–4), 201–213.

Naik, P. A., Wedel, M., & Kamakura, W. (2010). Multi-index binary response analysis of large data sets. *Journal of Business and Economics Statistics*, 28(1), 67–81.

O'Hara, R. B., & Sillanpaa, M. J. (2009). A review of Bayesian variable selection models: What, how and which? *Bayesian Analysis*, 4(1), 85–118.

Park, T., & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, 103(482), 681–686.

Peres, R., Muller, E., & Mahajan, V. (2010). Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing*, 27(2), 91–106.

Perez Rodriguez, P. (2009). *ars: Adaptive rejection sampling original C++ code from Arnost Komarek based on ars.f written by P. Wild and W. R. Gilks. R package version 0.4.* (http://CRAN.R-project.org/package=ars).

Putsis, W. P., Balasubramanian, S., Kaplan, E. H., & Sen, S. K. (1997). Mixing behavior in cross-country diffusion. *Marketing Science*, 16(4), 354–369.

Raftery, A. E., Madigan, D., & Hoeting, J. A. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, 92(437), 179–191.

Rutz, O. J., Trusov, M., & Bucklin, R. E. (2011). Modeling indirect effects of paid search advertising: Which keywords lead to more future visits? *Marketing Science*, 30(4), 646–665.

Srinivasan, S., Pauwels, K., & Nijs, V. (2010). Demand-based pricing versus past-price dependence: A cost–benefit analysis. *Journal of Marketing*, 72(2), 15–27.

Stamey, T. A., Kabalin, J. N., McNeal, J. E., Johnstone, I. M., Freiha, F., Redwine, E. A., et al. (1989). Prostate specific antigen in the diagnosis and treatment of adenocarcinoma of the prostate: II. Radical prostatectomy treated patients. *Journal of Urology*, 141(5), 1076–1083.

Stremersch, S., & Lemmens, A. (2009). Sales growth of new pharmaceuticals across the globe: The role of regulatory regimes. *Marketing Science*, 28(4), 690–708.

Stremersch, S., & Tellis, G. J. (2004). Understanding and managing international growth of new products. *International Journal of Research in Marketing*, 21(4), 421–438.

Takada, H., & Jain, D. (1991). Cross-national analysis of diffusion of consumer durable goods in Pacific Rim countries. *Journal of Marketing*, 55(2), 48–54.

Talukdar, D., Sudhir, K., & Ainslie, A. (2002). Investigating new product diffusion across products and countries. *Marketing Science*, 21(1), 97–114.

Tellis, G. J., Stremersch, S., & Yin, E. (2003). The international takeoff of new products: The role of economics, culture, and country innovativeness. *Marketing Science*, 22(2), 188–208.

Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Series B*, 58(1), 267–288.

Van den Bulte, C., & Stremersch, S. (2004). Social contagion and income heterogeneity in new product diffusion: A meta-analytic test. *Marketing Science*, 23(4), 530–544.

van Everdingen, Y. M., Aghina, W. B., & Fok, D. (2005). Forecasting cross-population innovation diffusion: A Bayesian approach. *International Journal of Research in Marketing,* *22*(3), 293–308.

van Everdingen, Y. M., Fok, D., & Stremersch, S. (2009). Modeling global spillover of new product takeoff. *Journal of Marketing Research, 46*(5), 637–652.

Wright, J. H. (2008). Bayesian model averaging and exchange rate forecasts. *Journal of Econometrics, 146*(2), 329–341.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B, 67*(2), 301–320.

Full Length Article

A comparison of different pay-per-bid auction formats

CrossMark

Ju-Young Kim [a,*], Tobias Brünner [b,1], Bernd Skiera [a,2], Martin Natter [a,3]

[a] Goethe-University Frankfurt, Faculty of Business and Economics, Department of Marketing, Grueneburgplatz 1, 60323 Frankfurt am Main, Germany
[b] Goethe-University Frankfurt, Faculty of Business and Economics, Department of Management and Microeconomics, Grueneburgplatz 1, 60323 Frankfurt am Main, Germany

A R T I C L E I N F O

A B S T R A C T

Pay-per-bid auctions are a popular new type of Internet auction that is unique because a fee is charged for each bid that is placed. This paper uses a theoretical model and three large empirical data sets with 44,614 ascending and 1,460 descending pay-per-bid auctions to compare the economic effects of different pay-per-bid auction formats, such as different price increments and ascending versus descending auctions. The theoretical model suggests revenue equivalence between different price increments and descending and ascending auctions. The empirical results, however, refute the theoretical predictions: ascending auctions with smaller price increments yield, on average, higher revenues per auction than ascending auctions with higher price increments, but their revenues vary much more strongly. On average, ascending auctions yield higher revenues per auction than descending auctions, but results differ strongly across product categories. Additionally, revenues per ascending auction also vary much more strongly.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Pay-per-bid auctions offered by retailers, such as Quibids, Bidcactus and MadBid, are exciting, fast-paced business-to-consumer online auctions that are attracting significant interest from consumers, popular press and start-up companies. Unlike other well-known auction sites, such as eBay, pay-per-bid auctions charge a fee for each bid that is placed, regardless of whether one wins the auction. Additionally, a bid placed increases the price by a certain increment that is chosen by the auctioneer.

At first glance, fee-based bidding does not sound attractive because the bidder encounters the risk of having to pay bidding fees without winning the auction. However, the compelling part of this model is that the bidders who win an auction can potentially save more than 99% off the current retail price (CRP) of the product. For example, on MadBid.com, a new MINI One car was sold for €8.47 rather than its retail price of €15,000. Similarly, a new Kymco scooter, which regularly sells for €1,240, was sold for €0.40.

Popular magazines, newspapers and online blogs are replete with heated discussions regarding this emerging type of auction.[4] Although some commentators are enthusiastic about the attractive deals offered by pay-per-bid auctions and how enjoyable they are, others strongly warn consumers against participating in them. Such commentators point to potentially huge losses for bidders as a result of the high

---

* Corresponding author. Tel.: +49 69 798 34634; fax: +49 69 798 35001.
  E-mail addresses: jukim@wiwi.uni-frankfurt.de (J.-Y. Kim),
bruenner@econ.uni-frankfurt.de (T. Brünner), skiera@skiera.de (B. Skiera),
natter@wiwi.uni-frankfurt.de (M. Natter).
  [1] Tel.: +49 69 798 34831; fax: +49 69 798 35021.
  [2] Tel.: +49 69 798 34649; fax: +49 69 798 35001.
  [3] Tel.: +49 69 798 34650; fax: +49 69 798 35001.

[4] For magazines, see for example, Last, Jonathan V. (February 23, 2009) "Take a chance on an auction; Swoopo and the rise of Entertainment Shopping" in The Weekly Standard; (December, 2011) "Can you buy products dirt cheap on QuiBids? Behind the hype" in Consumer Reports. For newspapers, see for example, King, Marc (January 28, 2012) "How penny auction websites can leave you with a hole in your pocket" in The Guardian; Zimmerman, Ann (August 17, 2011) "Penny auctions draw bidders with bargains, suspense" in The Wall Street Journal; Choi, Candice (June, 24, 2011) "Penny auction sites could cost a chunk of change" in The Washington Times; ConsumerReports.org (December 2011) "With penny auctions, you can spend a bundle but still leave empty-handed"; McCarthy, John (June, 2, 2011) "Penny auctions promise savings, overlook downsides" in USA Today; Migoya, David (January 25, 2010) "BARGAIN BIDDERS On penny auction sites, buyers can get gift cards for a fraction of face value — or pick up a new addiction" in The Denver Post; and Thaler, Richard H. (November 15, 2009) "Paying a Price For the Thrill of the Hunt" in The New York Times. For blogs, see for example, http://www.foxbusiness.com/personal-finance/2012/07/10/online-auction-virtual-bargain-or-real-rip-off/, http://www.scambook.com/blog/2012/07/penny-auction-fraud-alert-how-you-can-lose-by-winning/, http://www.usatoday.com/tech/columnist/kimkomando/2011-05-13-komando-penny-auctions_n.htm, http://www.thesimpledollar.com/2010/11/20/a-deeper-look-at-quibids-and-why-i-dont-think-its-worth-it/, and http://www.theregister.co.uk/2009/01/02/swoopo_startrup/.

bidding costs, which can easily be in the range of several hundred dollars per auction. However, all commentators have based their conclusions on a fairly limited number of observations, some of which are quite anecdotal.

Furthermore, auctioneers lack knowledge regarding how different auction formats influence their profitability. They frequently adjust their auction formats, and many auctioneers, such as the pioneer of this type of auction, Swoopo, have become bankrupt. Thus far, only a few researchers have analyzed pay-per-bid auctions by developing theoretical models (Augenblick, 2012; Gallice, 2011; Platt, Price, & Tappen, 2013) and by testing these models with actual sales data. Others have empirically compared the effect of the buy-now price feature on bidders' behavior in ascending penny auctions (Reiner, Natter, & Skiera, 2014). Analysis of the economic effects of different pay-per-bid auction formats that differ in the sizes and signs of price increments has thus far been neglected. We are the first researchers to close this gap.

We aim to theoretically and empirically assess the economic effects of different pay-per-bid auction formats. In particular, we compare different price increments (penny vs. ten-cent auctions) of ascending auctions as well as of ascending and descending pay-per-bid auctions. Therefore, we adapt and extend previous theoretical models, formulate predictions regarding the influence of auction formats on auctioneer revenues and empirically analyze them using three unique and large empirical data sets. Our data include the results of 44,614 ascending pay-per-bid auctions and 1,460 descending pay-per-bid auctions along with 1,142,738 bids.

The remainder of this manuscript is structured as follows. In the next section, we compare the most prominent pay-per-bid auctioneers and outline previous literature on pay-per-bid auctions. In Section 3, we describe our theoretical models and formulate predictions for the economic effects of different pay-per-bid auction formats. We investigate the economic influence of different formats of ascending pay-per-bid auctions in Section 4 and those for descending pay-per-bid auctions in Section 5. In Section 6, we compare the results of ascending and descending auctions. Section 7 summarizes our findings, discusses implications and points to topics for future research.

2. Pay-per-bid auctions

Pay-per-bid auctions are characterized by the association of bidding with tangible costs. Using traffic data (May 5 to August 5, 2013) from Alexa.com, Table 1 outlines some of the largest pay-per-bid auctioneers (with a reach of more than 0.001% of all global Internet users) and the characteristics of their auctions. All ascending auctioneers begin with a price of zero, but they differ by how much they change the price for each bid. Quibids offer various price increments that range from €0.01 to €0.15, whereas others increase the price by only €0.01. The start price of descending auctions is equal to the CRP. Bidding fees are substantial in all auction formats, ranging from €0.50 to €1.50.

2.1. Description of pay-per-bid auctions

Fig. 1 is a graphic illustration of ascending and descending pay-per-bid auctions. An ascending auction opens with a starting price that is usually €0.00. Each bid increases the price, and the bidder must pay for each bid. For example, in a typical auction at Quibids, each bid costs approximately €0.40 and increases the price by €0.01. Additionally, placing a bid delays the end of the auction by a countdown time (often 20 s). The auction ends when the countdown time has elapsed without an additional bid. The last bidder wins the auction and has the option to purchase the product from the auctioneer for the price of the final bid.

In contrast, in a descending auction, such as those offered by vipauktion, each placed bid costs €1.00 to €2.00 and decreases the current price by €0.40. After placing the bid, the bidder receives information regarding the current price. Hence, in a descending auction, the bid is not simply a bid in the narrow sense, as it is not a bid on a specific price. However, every placed bid reveals additional information regarding the current price. When the bidder accepts the current price, the product is purchased, and the auction ends. Otherwise, the auction continues, and the bidder can wait and place an additional bid to reveal information regarding an updated (lower) price.

2.2. Previous literature

Research on online auctions has recently been increasing in popularity (Barrot, Albers, Skiera, & Schäfers, 2010; Dholakia, Basuroy, & Soltysinski, 2002; Haruvy & Popkowski Leszczyc, 2009; Jap & Naik, 2008; Pinker, Seidmann, & Vakrat, 2003). Ever since the broad acceptance of the Internet online auctions such as eBay have become more popular. As a consequence, a variety of auction formats have emerged, such as name-your-own-price auctions (Amaldoss & Jain, 2008; Hinz & Spann, 2008; Spann, Skiera, & Schäfers, 2004) and pay-per-bid auctions.

Knowledge of ascending pay-per-bid auctions in particular is currently growing. Augenblick (2012), Hinnosaar (2010) and Platt et al. (2013) were the first researchers to provide theoretical models of ascending pay-per-bid auctions. Independently of one another, they show that any subgame perfect equilibrium of an ascending pay-per-bid auction that receives more than one bid must be in mixed strategies. A mixed strategy in this context means that bidders randomly choose between bidding and not-bidding in each round of the auction.

According to their theoretical models, Augenblick (2012) and Platt et al. (2013) find deviations with actual revenues being well above expected revenues. Therefore, Platt et al. (2013) extend their model to allow for risk preferences (risk-loving/risk-averse vs. risk-neutral), which leads to expected revenues that better match actual revenues. Byers, Mitzenmacher, and Zervas (2010) build on the theoretical model developed by Platt et al. (2013) and Augenblick (2012) and analyze information asymmetries across bidders. Their model shows that

Table 1
Comparison of the most popular pay-per-bid auctions.

Provider	Quibids	Dealdash	MadBid	Beezid	Bidcactus	ClicxaBids	Vipauktion
Auction format	Ascending	Ascending	Ascending	Ascending	Ascending	Descending	Descending
Starting price	$0.00	$0.00	£0.00	$0.00	$0.00	CRP	CRP
Bidding fee	$0.60	$0.60	£0.25–£1.20	$0.55–$0.90	$0.75	Varies	€1.00–€2.00
Price increment	$0.01–$0.20	$0.01	£0.01	$0.01	$0.01	Varies	€0.40
Operating countries	US/Europe/Canada/ Australia	US only	UK/Spain/Germany/ Italy/Ireland	US, ships worldwide	US, also ships to Canada	US/Ireland	Germany
Market share May–Aug 2013	75.69%	7.60%	5.06%	3.46%	2.95%	n.a.	n.a.
(% reach)	0.050	0.005	0.003	0.002	0.002	n.a.	n.a.

Market share based on reach between May and Aug 2013; % reach = percentage of all Internet users visiting this site; CRP: current retail price, with all prices in local currencies; n.a. = not available.
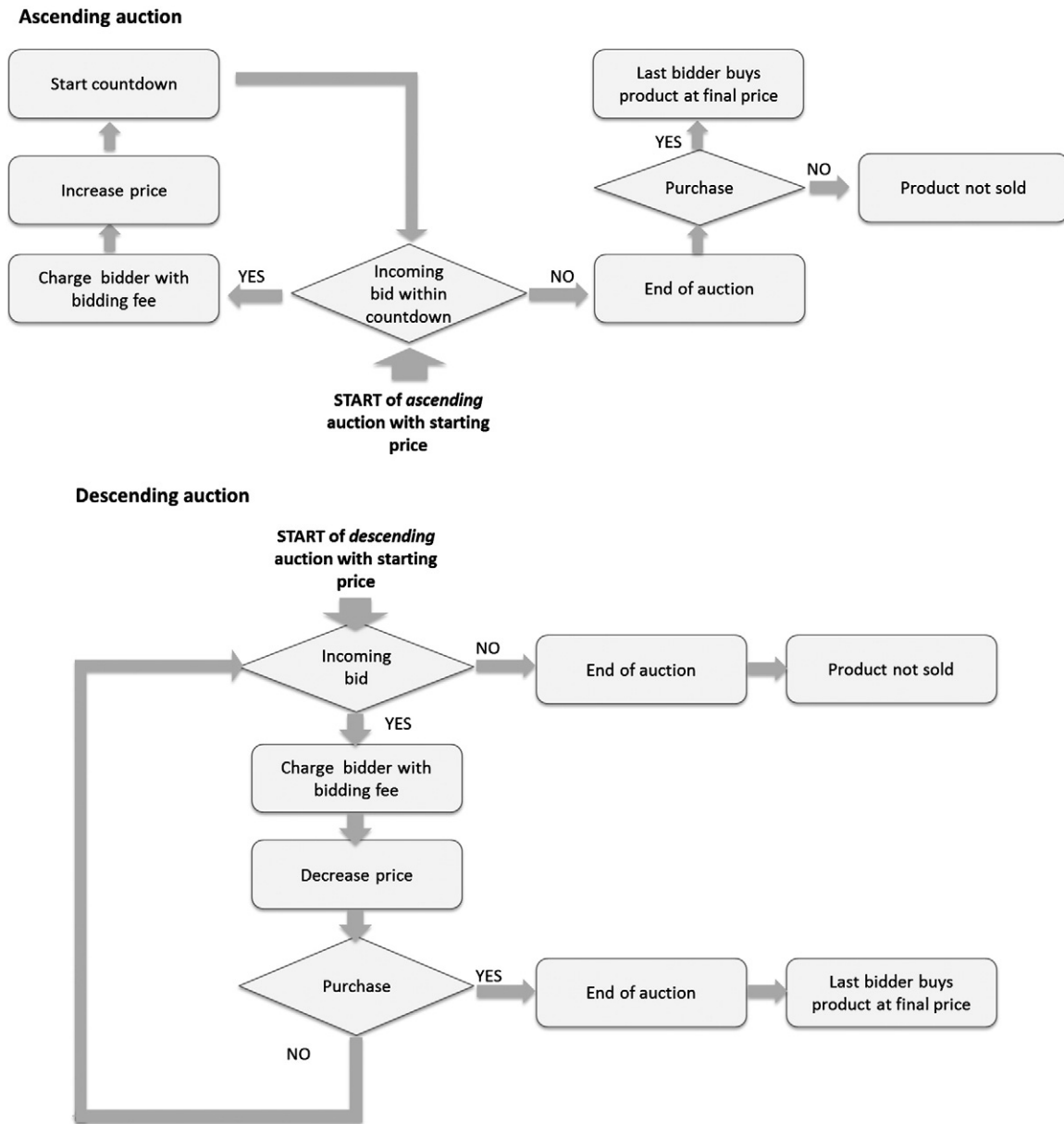
Fig. 1. Graphic illustration of ascending and descending pay-per-bid auctions.

when bidders underestimate the true number of bidders, the duration of the auction and thus the auctioneer's revenue increases. In an experimental study, Caldara (2012) shows that neither risk-loving bidders nor incorrect beliefs regarding the parameters of the auction (such as the number of bidders) are necessary for observing revenues that exceed the product's value and thus the expected revenues. Moreover, he finds that revenues move closer to the revenues of the theoretical model as bidders gain experience. The important role of experience is supported by Wang and Xu (2013), who analyze bid-level data from a large ascending pay-per-bid auction website and show that losing bidders stop participating in these auctions while others learn, continue to bid and make profits.

By contrast, there is little research on descending pay-per-bid auctions, as Gallice (2011) is the only researcher who derives equilibrium bidding behavior in descending pay-per-bid auctions. He shows that in equilibrium, only two situations can arise: either the product is purchased at the starting price, or no bid occurs. The reason is that if at

least one bidder is willing to buy at the starting price, then this bidder will buy immediately; if no bidder is willing to buy at the starting price, then no one will ever observe the price, and the product will not be sold. However, contrary to his prediction, Gallice (2011) finds (similar to what ascending pay-per-bid researchers have found) that actual revenues were well above the expected revenues. He explains the deviations with bounded rationality of the bidders.

Thus far, there is no study that compares the averages and variances of actual and expected revenues among pay-per-bid auctions with varying sizes (penny vs. ten-cent) or signs (ascending vs. descending) of price increments. In the following, we will present the theoretical models on which our empirical analysis is based.

3. Economic analysis of pay-per-bid auctions

For both auction formats, ascending and descending pay-per-bid auctions, we use the same assumptions as Platt et al. (2013) and

Augenblick (2012). These assumptions lead to a theoretical model that is similar to the model that Platt et al. (2013) and Augenblick (2012) developed for the ascending auction and to a special case of the theoretical model that Gallice (2011) suggested for the descending auction.

3.1. Model assumptions

3.1.1. Risk neutral bidders

We assume throughout the study that there are n risk-neutral bidders. However, the assumption of risk neutrality is not innocuous. Platt et al. (2013) show that expected revenues of ascending pay-per-bid auctions are decreasing in the degree of risk aversion. Expected revenue is lower (higher) when bidders are risk averse (loving) than when they are risk-neutral. Platt et al. (2013) find evidence for modest degrees of risk-loving preferences. However, the reported range of estimated degrees of risk aversion is wide (varying between -0.0017 (for $1,000) and -0.03 (for the 50 free bids)), and the degree of risk aversion appears to depend on the product that is auctioned off. Moreover, in contrast to ascending pay-per-bid auctions (Reiner, Brünner, Natter, & Skiera, 2014; Platt et al., 2013) there are no estimates of the risk attitudes of bidders in descending pay-per-bid auctions. Therefore, we assume risk-neutral bidders in the theoretical models of all auction formats.

3.1.2. Common valuations of products

Each bidder values the product to be auctioned off at the commonly known willingness-to-pay (WTP) of v. For art, antique furniture and other collectors' items that are typically associated with auctions, the assumption that all bidders value a product equally is certainly improbable. However, because the products that are offered at pay-per-bid auctions are brand new products that are readily available at alternative shopping websites or high street retailers, we believe that the assumption is reasonable. Augenblick (2012) shows that a theoretical model in which bidders have independent private valuations converges to the full information common valuation case as the differences in valuation decrease.

3.1.3. Current retail price (CRP) = willingness-to-pay (WTP)

Pay-per-bid auction providers display a product's CRP for the entire time of the auction. This price is often higher than the prices that other shopping websites post for the same product. For example, Augenblick (2012) reports that the average price of Amazon is only 79% of the CRP. However, bidders' WTP is in turn estimated to be 15% to 65% higher than the Amazon price (Platt et al., 2013). These numbers lead to a range for the WTP that is between 90.85% ($1.15 \times 79\%$) and 130.35% ($1.65 \times 79\%$) of the CRP. Because this value is posted prominently on the auction website, it may also serve as an anchor. Therefore, we use the CRP as a proxy for the WTP, which should work well for cash or vouchers because the CRP is simply equal to the amount of cash or the monetary value of the voucher.

3.2. Economic analysis of ascending pay-per-bid auctions

In the following, we present the baseline model of the ascending pay-per-bid auction developed by Platt et al. (2013) and Augenblick (2012). There are n bidders. The ascending auction starts at a price of zero. Each bid increases the price by d and costs b. After each bid, all bidders (except the current highest bidder) decide whether to place a bid or not. If several bidders decide to bid, then one of them is randomly selected, pays the bidding fee b and becomes the new highest bidder. The current price is raised by d. Thus, after q bids, the auction price is qd. If none of the bidders places a new bid, then the auction ends, and the current highest bidder buys the product for the current auction price.

Platt et al. (2013) and Augenblick (2012) show that there is a symmetric subgame perfect equilibrium in mixed strategies; after $q - 1$

bids, every bidder who is not the current highest bidder places a bid with probability β_q:

$$\beta_q = \begin{cases} 1-(1-\mu_0)^{\frac{1}{n}} & \text{if } q = 1 \\ 1-\left(\frac{b}{v-d(q-1)}\right)^{\frac{1}{n-1}} & \text{if } 1<q\leq\frac{v-b}{d} \\ 0 & \text{if } q>\frac{v-b}{d}. \end{cases} \quad (1)$$

To obtain an intuition for the equilibrium strategy in Eq. (1), let us first argue why the behavior of bidders is stochastic or, more technically, why there is no symmetric equilibrium in pure strategies. Suppose that the number of bids is low, such that the WTP for a product exceeds the current price plus the bidding cost: $v > qd + b$. If a bidder knew that all other bidders would not bid, then she would bid and make a bargain. As a result, a situation in which no bidder places a bid cannot be an equilibrium. Similarly, when all bidders always place a bid early in the auction, it will become advantageous for a bidder to wait and allow the other bidders to pay the bidding fees until there is a positive probability that the auction will end. Thus, a situation in which all bidders always bid is also not an equilibrium. Therefore, the only symmetric equilibrium is in mixed strategies. The probability of making a bid β_q is determined such that bidders are indifferent between placing a bid and not placing a bid.

The parameter μ_0 in the equilibrium strategy in Eq. (1) is the probability that at least one bidder will place a bid in the first period. This parameter is not uniquely determined in equilibrium. However, if no one places a bid in the first round, which occurs with probability $1 - \mu_0$, then the auction ends, and the auctioneer can immediately set up a new, identical auction. If, again, no bidder places a bid, then the auctioneer can restart the auction repeatedly until there is at least one bid in the first round. In our data set, we only have auctions that attracted at least one bid. Therefore, we set $\mu_0 = 1$ to concentrate on auctions that have at least one active bidder.

Platt et al. (2013) show that in an ascending auction with WTP v, these bidding strategies lead to the following expected value and variance of the revenue of one auction R_a:

$$E(R_a|v) = v \quad (2a)$$

$$Var(R_a|v) = \frac{b}{b + 2d}(v-d)^2. \quad (2b)$$

The result for the expected revenue can be obtained by the following logic. The total surplus of auctioneer and bidders together is then the WTP, v. However, in equilibrium, bidders are indifferent between placing a bid and not bidding. Consequently, bidders are also indifferent between placing no bid at all and actively participating in the auction. Because placing no bid at all yields a utility of zero, participating in the auction must also have an expected utility of zero. If the bidders' expected utilities are zero, the entire surplus accrues to the auctioneer and thus, the auctioneer's expected revenue is v. The expression for the variance of revenues in Eq. (2b) is the variance of a continuous approximation of the distribution of revenues implied by the bidding strategies in Eq. (1).

The expectation and variance of revenue presented in Eqs. (2a) and (2b) were derived for a single auction in which the bidders' WTP is v. In our empirical analysis, we calculate expected revenues by averaging revenues of auctions that were conducted during the period from August 2007 to March 2009. Similarly, the variance is estimated by the sample variance of revenues from auctions in that period. Because the WTP changed over that period, the relevant benchmarks are not the

conditional moments in Eqs. (2a) and (2b) but the unconditional expectation and variance given by the following[5]:

$$E(R_a) = E(v) \tag{3a}$$

$$Var(R_a) = \frac{b}{b+2d} E(v-d)^2 + Var(v), \tag{3b}$$

where $E(v)$ and $Var(v)$ are the expectation and the variance of the WTP, respectively.

The data sets analyzed below include ascending auctions with per-bid price increments of $d = €0.01$ and $d = €0.10$. Therefore, we seek to determine the effect of a change in d on revenues. From Eqs. (3a) and (3b), we obtain the following predictions:

Prediction 1: An increase in the price increment d reduces the variance of auctioneer revenues in ascending pay-per-bid auctions.
Prediction 2: An increase in the price increment d leaves the expected revenue in ascending pay-per-bid auctions unaffected.

3.3. Economic analysis of descending pay-per-bid auctions

For the descending auction, we adapt the theoretical model by Gallice (2011), such that the assumptions are the same as in the theoretical model by Augenblick (2012) and Platt et al. (2013). We show below that the main results of Gallice (2011) are not affected by this modification.

The descending pay-per-bid auction begins at a price, s, which is usually equal to the current retail price (CRP). The starting price is publicly observable. Each bid decreases the current auction price by e and costs b. The bidding costs are greater than the price decrement, $b > e$.[6] In contrast to the ascending auction, the current auction price is not publicly observable. Only the current bidder can view this price. After observing the current auction price, the bidder can decide whether to buy the product or not. If she decides to buy, then she pays the current auction price, and the auction ends. If she decides not to buy, then the auction continues. The other bidders do not know whether and how often someone has observed the price. Otherwise, a bidder could count the number of times that the price has been observed and would know the current price.

Gallice (2011) studies a descending pay-per-bid auction, assuming that bidders have independently distributed private valuations for the product that is for sale. He shows that when the highest valuation of a bidder is below the starting price, no bidder will place a bid. We build on his theoretical model, but in contrast to Gallice (2011), we assume that bidders have commonly known valuations for the products auctioned and, hence, a common WTP, v, that is known. If the starting price is excessively high, i.e., $v < s - e + b$, then no bidder will see the price, and the product will not be sold in equilibrium. If the starting price is sufficiently low, i.e., $v > s - e + b$, then all bidders will want to see the price, and the first bidder who does so will buy the product. The highest starting price that attracts participation by rational bidders is $\bar{s} = v + e - b$. If a bidder places a bid, then she pays b and observes the price $\bar{s} - e = v - b$. Thus, the maximum revenue per auction that the auctioneer can extract from rational bidders is $R_d = v$.[7]

Note that for a given common WTP v, the auctioneer's revenue in descending auctions is deterministic. However, as for the ascending auction, the relevant benchmarks for our empirical results are the unconditional expectation and variance:

$$E(R_d) = E(v) \tag{4a}$$

$$Var(R_d) = Var(v). \tag{4b}$$

Note that the starting price in the descending pay-per-bid auctions in our data set is the CRP. We assume that the WTP is equal to the CRP. Because of the bidding cost, the CRP is slightly above the maximum starting price $\bar{s}$.[8] This characteristic implies that in equilibrium, we should expect no bids in the descending auctions. This result is clearly at odds with the empirical results of real descending pay-per-bid auctions in which we observe active participation.

The fact that bidders observe the price although the starting price is above $\bar{s}$ can be explained by considering their curiosity. If a bidder is curious about the hidden value of the price, then she might receive some additional utility from lifting the veil and observing the price. If this additional utility is greater than the difference of bidding costs and price decrement, $b - e$, then a curious bidder will place a bid. Even when all bidders are rational (i.e., not curious) but believe that some bidders are curious, it is rational to participate because curious bidders might have placed a bid already and it would thus be profitable to place the next bid.[9] However, because the difference between the true starting price and the maximum starting price $\bar{s}$ is minuscule compared to the average CRP of the products sold in descending pay-per-bid auctions (€153.48), we assume that the auctioneer sets the maximum starting price at $\bar{s}$.

It is easy to observe from Eqs. (4a) and (4b) that the bidding costs, b, and the price decrement, e, do not affect the expected revenue and the variance of revenue. Unfortunately, neither the bidding costs nor the price decrement was altered during the time period that our sample encompasses. Therefore, there are no descending auction counterparts to Predictions 1 and 2 for ascending auctions.

3.4. Theoretical comparison of ascending and descending pay-per-bid auctions

Having derived the revenues for both auction formats, we can now compare revenues per auction and their variance for ascending and descending pay-per-bid auctions. Examining Eqs. (3a)–(4b), we obtain the following predictions:

Prediction 3: The variance of an auctioneer's revenue per auction is higher in ascending pay-per-bid auctions than in descending auctions.
Prediction 4: Ascending pay-per-bid auctions generate the same expected revenues as descending auctions.

4. Empirical study of ascending auctions

Based on our economic analyses of ascending and descending auctions, we now aim to empirically test our predictions (1–4) by comparing the expected revenues (serving as a benchmark) derived from the theoretical model with actual revenues and by explaining the resulting differences.

In the following, we will first analyze ascending and descending auctions separately (Sections 4 and 5). In Section 6, we will study the differences between ascending and descending pay-per-bid auctions.

4.1. Data

We collected data from a European ascending pay-per-bid auctioneer that provided us with two unique data sets. In contrast to a platform such as eBay that includes three parties (auctioneer, buyer, and seller),

---

[5] Here, we use the result that $Var(Y) = E(Var(Y|X)) + Var(E(Y|X))$.
[6] If b were smaller than e, then a bidder could observe the price v/e times and buy for a price of 0. The total bidding costs of this strategy would be bv/e, which is smaller than the WTP v, given that $b < e$.
[7] The formal presentation of the equilibrium of the descending pay-per-bid auction and the proof are given in Appendix A.

[8] For the descending auctions in our data set, bidding costs are $b = 0.49$, and the price decrement is $e = 0.40$. Thus, the maximum starting price is $\bar{s} = CRP - 0.09$, which is slightly below the actually chosen starting price, which is equal to the CRP.
[9] The websites of descending pay-per-bid auctioneers display the results of past auctions. Bidders can observe that past auctions attracted bids and that the product was finally sold at a price below the current retail price. Thus, a rational bidder can conclude that there must be some curious bidders.

the auctioneer is also the seller. Additionally, the auctioneer sells only brand new products that are currently available at retailers (i.e., common collectibles). Used products and older-generation products are not auctioned.

The first data set (A1) contains the results of all ten-cent (N = 42,042) and penny (N = 1,112) ascending auctions from December 2007 to November 2008 (43,154 auctions in total). For each completed auction, we received information about what product was auctioned, the final price, the current retail price (CRP), the end time and the number of bids placed by the winners of the auction. All auctions start at a price of €0.00, and each bid costs €0.50.

The second data set (A2) contains the results of 1460 ascending auctions completed in March 2009, plus additional information regarding the bidding histories of these auctions, including 949,750 bids and their respective bidders. Data set A2 also includes information regarding the participating bidders beyond these bids, such as their overall bidding experience with the auctioneer (e.g., the number of auctions won, the number of auctions participated in) and demographic information.

The two data sets differ in their numbers of auctions and in the details that are provided for each auction. Data set A1 contains considerably more auctions, whereas data set A2 provides more details regarding all bids (e.g., bidding time, nickname of bidder) and information regarding bidders beyond their behavior in these auctions (their first bidding date, the number of auctions with bids placed, the number of auctions won, total bids placed, gender and age).

All data sets include information on the auction end time, the nickname of the winner, the number of bids made by the winner, the total number of bids made (by the winner and the losers), the product sold and its CRP. According to the auctioneer, the CRP represents an average price value for the product on the basis of other online retailers. Table 2 provides an overview of the product categories of the auctioned products from both data sets.

4.2. Comparison of actual and expected revenues per auction

First, we compare actual revenues and expected revenues. For all revenues, we calculate the average revenues across all categories standardized by their CRP, i.e. we divided the revenues per auction by the corresponding CRP. We use Eq. (2a) to calculate the expected revenues and use a t-test to compare them with actual revenues. To perform the t-test, we also apply the variances of standardized expected revenues from Eq. (2b). Table 3 depicts the standardized means of the actual and expected revenues per auction as well as the percentage difference across all categories for both the ten-cent auctions and the penny auctions. As noted above, we assume that the WTP is equal to the CRP. This assumption is certainly valid for products such as vouchers or cash, as these products are of a defined value that is common to everyone (i.e., every bidder values a €50 voucher or €50 in cash equally). However, this assumption may also hold for the remaining products: all auctioned products are brand new, are in their original packaging and are currently sold at competitive retailers. Additionally, the CRP represents an average price value for each product from common online retailers.

Table 3
Comparison of means of actual and expected standardized revenues per auction from ascending auctions.

Product category	Ten-cent auction				Penny auction			
	N	Actual revenue	Expected revenue	Δ – %	N	Actual revenue	Expected revenue	Δ – %
Cash	11	2.07	1.00 ***	107%				
voucher	43	0.98	1.00 n.s.		185	4.21	1.00 ***	321%
Video game console	10,887	1.75	1.00 ***	75%				
Fast-moving electronic appliances	1257	1.34	1.00 ***	34%	11	0.73	1.00 n.s.	
Software	9190	1.26	1.00 ***	26%				
Computer hardware	2579	0.94	1.00 ***	–6%	510	1.77	1.00 ***	77%
DVD	922	0.94	1.00 n.s.	–6%				
GPS	623	0.91	1.00 **	–9%	525	1.58	1.00 ***	58%
Toys	1389	0.90	1.00 ***	–10%				
Home appliances	2266	0.90	1.00 ***	–10%	13	0.76	1.00 n.s.	
Perfume	1408	0.87	1.00 ***	–13%				
Small electronic goods	2096	0.86	1.00 ***	–14%	66	1.31	1.00 **	31%
TV + audio-visual	639	0.83	1.00 ***	–17%	158	1.21	1.00 *	21%
Computer accessories	4517	0.75	1.00 ***	–26%				
Housewares	407	0.74	1.00 ***	–26%				
Others	1491	0.72	1.00 ***	–28%				
Jewelry	3276	0.21	1.00 ***	–79%				
Total	43,057	1.12	1.00	12%	1.557	1.90	1.00 ***	90%

Δ − %: percentage differences between the means of actual and expected standardized revenues; positive differences are illustrated in green cells and negative differences in red. All revenues are standardized. Standardized revenue is defined as revenue/CRP; CRP: current retail price.
*** = $p < 0.01$, two-tailed; ** = $p < 0.05$, two-tailed; n.s. = not significant.

Table 2
Description of the data sets.

Product category	Typical products in category	Ascending auction (A1)		Ascending auction (A2)	
		# ten-cent auctions	# penny auctions	# ten-cent auctions	# penny auctions
Video game console	Nintendo DS, Nintendo Wii, PSP, PS3, Xbox 360	10,465	1	422	0
Software	Programs, PC games, video games	8,949	1	244	0
Computer accessories	USB, Computer bags, keyboards	4,332	1	185	0
Jewelry	Watches, bracelets	3,231	1	48	0
Computer hardware	Desktop, notebook, printer, monitors	2,585	440	0	71
Home appliances	Coffee machine, washer, dental care, shaver	2,230	13	38	0
Small electronic goods	Mobiles, Telephones, digital frame, modem	2,102	28	0	38
Perfume	Roma, D&G, Hugo, Boss, Calvin Klein	1,407	0		
Toys	Lego, Fisher-Price	1,389	0		
Fast-moving electronic appliances	Mp3, digital camera	1,191	11	66	0
GPS	Falk, Navignon, TomTom	623	480	0	45
DVD	Blockbuster, TV series	922	0		
TV + audio-visual	Samsung, LG, Philips	627	133	7	30
Housewares	Cutlery, pots	407	0		
Cash	Cash and 5 kg gold	11	0		
Vouchers	iTunes 25 €, 150 free bids, 300 free bids	43	0	0	266
Others	Bags, key rings	1,528	3		
Total		42,042	1,112	1,010	450

First, observing only the categories with obvious common values (cash and voucher), we find significant differences between the actual and expected revenues for the cash category. Actual revenues are more than double with cash, meaning that the auctioneer sold e.g. €100 for €207. The deviation between actual and expected revenues for the voucher category is not significant for ten-cent auctions. However, we do find significant and surprising differences for this category in penny auctions: selling vouchers generated revenues that were sold for four times above their expected value.

For ten-cent auctions, the results in Table 3 illustrate that the auctioneer additionally generated higher revenues per auction than expected in the video game console, fast-moving electronic appliances and software categories. In the remaining categories, all revenues per auction are significantly lower. One explanation may be that hedonic products, such as game consoles, mp3 players (e.g., iPods), and video games, induce more emotions and consequently more bids than utilitarian products such as home appliances or GPS devices.

However, in penny auctions, it is salient that the significant deviations from the expected revenues are always in favor of the auctioneer. We do not find significant differences in the fast-moving electronic appliances and electronic appliance categories. This may be due to the low number of observations in these categories.

4.3. Explanations for differences between actual and expected revenues per auction

To investigate these differences in greater detail, we conduct a regression analysis with the difference in standardized revenues per auction ((actual revenue − expected revenue) / CRP) as the dependent variable. Furthermore, we add a binary variable (penny auction) to indicate whether an auction was a penny auction (value = 1) or a ten-cent auction (value = 0) and account for the number of bidders in each auction and the type of category, whether it is perceived as hedonic, utilitarian, or both hedonic and utilitarian. Additionally, we investigate the role of average product value in each category (measured by the average CRP in a category). Hence, the concept of hedonic and functional products, which was previously tested as a reliable construct by Strahilevitz and Myers (1998), as well as the value of the product (CRP) are used to explore potential differences in product categories (see Appendix B).

Table 4 displays the results of the linear regression analysis of the ascending auction. We use data set A2 only because A1 did not include information regarding the number of bidders participating in each auction.

The analysis includes 1,003 ten-cent and 334 penny auctions (123 auctions were excluded because of missing values) and explains 31.4% of the variance in the dependent variable. In contrast to theoretical models, which assume that the number of bidders has no influence, Table 4 shows that a higher number of competing bidders leads to a higher difference between actual and expected standardized revenues. As this number does not impact the expected standardized revenues, it means that a higher number of bidders yield higher actual revenues per auction, which benefits the ascending pay-per-bid auctioneer. This finding can be explained by relaxing the assumption that all bidders know the exact number of participants. Byers et al. (2010) show that the expected revenue exceeds its equilibrium level v if bidders underestimate the true number of participants. Conversely, if bidders overestimate the number of participants, the auctioneer's revenue will decrease. A large (small) number of bidders in our regression might pick up situations in which bidders underestimate (overestimate) the true number of participants, therefore leading to higher (lower) actual revenue than expected.

We also find that those categories perceived as either only hedonic or both hedonic and utilitarian (here, the reference category) additionally drive the auctioneer's revenues. Thus, hedonic categories may cause emotional arousal, which results in less rational bidding behavior and increased bidding efforts (Hirschman & Holbrook, 1982). Finally, more valuable products (with higher CRPs), such as jewelry, negatively affect the relative difference between actual and expected standardized revenues.

To understand this surprising result, we examine whether bidders increase the number of bids in accordance with a higher product value (as measured by the CRP). The analysis shows that the number of bids is highly correlated ($r = 0.939$) with the (log) product value ($p < 0.01$). However, bidders only slightly increase their number of bids in accordance with a higher CRP. More specifically, we find that the discount off the final price relative to the CRP is positively correlated with the product value ($r = 0.52$). Thus, more valuable products are sold at greater (percentage) discounts. Finally, small price increments (i.e., penny auctions) positively affect auctioneers' revenues. This finding is not surprising, as the results in Table 3 already suggest systematic differences between ten-cent and penny auctions with respect to revenues. However, according to Prediction 2, revenues should be unaffected regardless of varying changes in price. In the following section, we analyze Predictions 1 and 2 in greater detail.

4.4. Comparison of actual revenues in the context of different changes in price

Our economic analysis suggests that an increase in the price increment per bid d reduces the variance in auctioneers' revenues (Prediction 1). Thus, in our data set, penny auction revenues should exhibit wider variation than those of ten-cent auctions. Fig. 2 shows the distribution of revenues using the example of a GPS product (TomTom Go 930T, CRP = €549, data set A1) that was auctioned off in both penny and ten-cent auctions.

Fig. 2 indicates that the variance differs between the two price increments. The volatility of achieved revenues is much higher in penny auctions than in ten-cent auctions.

A two-group variance comparison test supports this indication (see Table 5). We compare the variances of penny and ten-cent auctions across three product categories of data set A1 (GPS, Computer Hardware and TV/Audio-visual), in which we have at least 100 penny auctions and ten-cent auctions. To attain comparability across the products, we again use standardized revenues. We find that the variance of penny auctions is always significantly greater ($p < 0.01$) than that of ten-cent auctions, thus supporting Prediction 1 from the theoretical model.

To empirically compare the revenues per auction of penny and ten-cent auctions (Prediction 2), we use a two-independent-sample t-test, which additionally accounts for the unequal variances between the two price increments. In contrast to our predictions, we find that the revenues of penny auctions are always significantly higher ($p < 0.01$) than the revenues of ten-cent auctions. Thus, penny auctions may lead to rather unsteady revenues compared with ten-cent auctions but appear to be more profitable for the auctioneer.
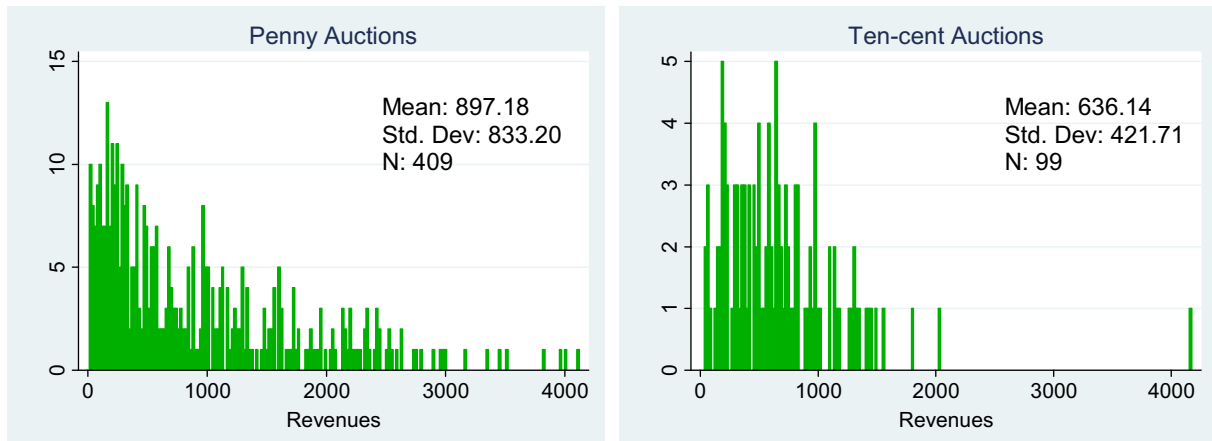
Table 4

Drivers of differences between actual and expected standardized revenues per auction in ascending auctions.

Variable	Parameter
Penny auction	0.779[***]
Number of bidders per auction	0.013[***]
LN current retail price	−1.278[***]
Hedonic category	0.422[**]
Utilitarian category	−0.205[*]
Hedonic & utilitarian category[a]	
Constant	6.044[***]

adj. $R^2 = 0.314$, N = 1,337.

Standardized revenue is defined as revenue/CRP; CRP: current retail price.

[a] Reference category.

[***] $p < 0.01$, two-tailed.

[**] $p < 0.05$, two-tailed.

[*] $p < 0.10$, two-tailed.

Fig. 2. Distribution of revenues per auction (in euro) for TomTom Go 930T at penny and ten-cent auctions.

5. Empirical study of descending auctions

5.1. Data

We also received data (D1) from a descending pay-per-bid auctioneer, including all completed auctions (1460) from August 2007 to October 2008. Similar to A2, this data set contains both the results of all auctions and the corresponding bidding histories (N = 192,988). Here, auctions began at the products' CRPs and each bid, which cost €0.49, decreased the price by €0.40. Information regarding the bidding fees, the change in price and the starting price, the CRP, was publicly available; however, the current price was not publicly available. After placing a bid (and paying the bidding fee), the current price of the product was shown to the bidder. Viewing further updates of the current price required placing additional bids. The bidders also had no idea about the number of competitors and the starting time of the auction. Thus, they could not infer current prices.

Similar to the ascending auctions, only one seller auctions brand new products in original packaging. Table 6 gives an overview of the auctioned products.

5.2. Comparison of actual and expected revenues per auction

We again standardize all revenues and then compare actual revenues with expected revenues that were derived from Eq. (4a) (see Table 7). We again assume that the WTP is a given common value that is equal to the CRP.

The actual revenues indicate that the variance is significantly different from zero. This result can be explained by fluctuations in the WTP or CRP over time (see Eq. (4b)). However, the variance is also different from zero in the voucher category, in which we expect the same common value over time. Thus, alternatively, bidders who seek to achieve a large discount may not decide to buy the product when the WTP is lower than the price in the auction. Rather, such bidders wait until the price decreases further, hoping that other bidders will think similarly.

Information regarding final prices, which is provided on the website, may support this behavior: knowing that other bidders do not directly buy when the price is below their own WTP may convince bidders to wait as well or to place multiple bids.

Table 7 also shows that actual revenues are significantly higher than expected revenues in all categories; on average, the descending pay-per-bid auctioneer generated higher revenues with the auctions compared with selling the products at their CRP across all categories.

Similar to the results of the ascending auction, the deviation of actual to expected revenues is highest for the vouchers category. However, in contrast to the results of the ascending auction, categories such as video game console or fast-moving electronic appliances do not stand out.

5.3. Explanations for differences between actual and expected revenues per auction

To investigate the differences between actual and expected revenues per auction in descending auctions, we again conduct a regression analysis with the difference in standardized revenues ((actual revenue − expected revenue) / CRP) as the dependent variable. Furthermore, we account for the number of bidders in each auction, the type of category and the average product value. Table 8 displays the results of the linear regression analysis of the descending auctions.

Our estimation includes 1,460 auctions and explains 18.6% of the variance of the dependent variable. The results reveal that the number of bidders in each auction significantly affects revenue differences. Actual revenues increase with the number of bidders because the number of bidders do not impact expected revenues. Thus, the number of bidders affects the revenues per auction in both ascending and descending auctions.

Furthermore, purely hedonic product categories negatively affect revenue per auction (compared with the reference category, hedonic and utilitarian). A possible explanation for this finding is that in descending auctions, bidders may encounter an increasing trade-off between ownership and additional savings. If the bidder has a strong

Table 5
Comparison of average standardized revenues per auction and variances of penny and ten-cent auctions.

	Penny auction			Ten-cent auction			Δ − %	
	N	Mean[a]	Std. dev.	N	Mean[a]	Std. dev.	Mean[a]	Std. dev
GPS	480	1.57	1.48	623	0.91[***]	0.77[***]	72%	92%
Computer hardware	440	1.83	1.89	2,585	0.94[***]	0.80[***]	95%	137%
TV + audio	133	1.16	1.32	627	0.83[***]	0.75[***]	40%	75%
Total	1,112	1.58	1.64	42,042	1.12[***]	1.02[***]	41%	61%

Δ − %: percentage differences between the means and standard deviations of penny and ten-cent auctions.
[a] Standardized revenues, defined as revenue/CRP; CRP: current retail price.
*** $p < 0.01$, two-tailed.

Table 6
Description of data set (D1) with descending auctions.

Product category	Products per category	Number of auctions
Video game console	Nintendo DS, Nintendo Wii, PSP, PS3, Xbox 360	126
Software	Programs, PC games, video games	69
Computer accessories	USB, computer bags, keyboards	208
Jewelry	Watches, bracelets	18
Computer hardware	Desktop, notebook, printer, monitors	101
Home appliances	Coffee machine, washer, dental care, shaver	105
Small electronic goods	Mobile, telephones, digital frame, radio	142
Perfume	Hugo Boss, Lagerfeld	14
Toys	Lego, board games	64
Fast-moving electronic appliances	Mp3, digital camera	196
GPS	Falk, Navignon, TomTom	26
DVD	Blockbuster, TV series	81
TV + audio-visual	Samsung, LG, Philips	40
Housewares	Fondue pots	16
Vouchers	Free bids, 100€ voucher	161
Others	Bags, magazine subscription	93
Total		1,460

desire to own the product, then she will buy earlier, thus leading the actual revenues to converge with the expected revenues. If, however, the bidder is interested in obtaining a large discount, then she will wait until there are sufficient bids. By contrast, in ascending auctions, there exists a longer period in which the bidders can seek to attain both ownership and a large discount. Bidders who seek to buy hedonic goods most likely fear that another bidder could buy the product first. Finally, more valuable products negatively affect auctioneer revenues.

Table 7
Comparison of the means of actual and expected standardized revenues from descending auctions.

Product category	N	Actual revenue	Std. dev.	Expected revenue	Δ − %
Vouchers	161	1.34	0.39	1.00 ***	34%
Perfume	14	1.24	0.37	1.00 **	24%
Computer Accessories	208	1.18	0.19	1.00 ***	18%
Toys	64	1.18	0.19	1.00 ***	18%
Others	93	1.18	0.17	1.00 ***	18%
DVD	81	1.18	0.09	1.00 ***	18%
Small electronic goods	142	1.17	0.15	1.00 ***	17%
Software	69	1.16	0.09	1.00 ***	16%
Housewares	16	1.13	0.08	1.00 ***	13%
Home appliances	105	1.13	0.12	1.00 ***	13%
Video game Console	126	1.12	0.10	1.00 ***	12%
Fast-moving Electronic Appliances	196	1.12	0.09	1.00 ***	12%
Jewelry	18	1.09	0.10	1.00 ***	9%
TV + audio-visual	40	1.08	0.11	1.00 ***	8%
GPS	26	1.08	0.08	1.00 ***	8%
Computer Hardware	101	1.07	0.09	1.00 ***	7%
Total	1460	1.17	0.19	1.00 ***	17%

Δ − %: percentage differences between the means of actual and expected standardized revenues; positive differences are illustrated in green cells.
Standardized revenue is defined as revenue/CRP; CRP: current retail price.
*** = $p < 0.01$, two-tailed; ** = $p < 0.05$, two-tailed.

Table 8
Drivers of differences between actual and expected standardized revenues per auction in ascending auctions.

Variable	Parameter
Number of bidders per auction	0.002***
LN current retail price	−0.085***
Hedonic category	−0.022**
Utilitarian category	0.006 n.s.
Hedonic & utilitarian category[a]	
Constant	0.483***

adj. $R^2 = 0.186$, N = 1,460.
Standardized revenue is defined as revenue/CRP; CRP: current retail price.
n.s. = not significant.
[a] Reference category.
*** $p < 0.01$, two-tailed.
** $p < 0.05$, two-tailed.

6. Comparison of revenues of descending auctions and ascending auctions

Having analyzed ascending and descending pay-per-bid auctions separately, we now turn to the comparison between the two auction formats. When comparing two auction formats in the field, one ideally wants to compare them under the same conditions (e.g., by keeping the set of bidders and products constant). This comparability is easier to achieve in laboratory experiments, but that increase in internal validity could come at the expense of external validity. Our field data provide advantages with respect to external validity but also provide some challenges, as the comparison between auction formats may be confounded by differences in the sets of bidders and products.

We aim to limit the effect of those unobservable differences by selecting two auction websites that operate in the same geographic region, namely, Germany. We also restrict our comparisons to auctions from the same period of time, namely, December 2007 to October 2008. Moreover, we focus on identical products (vouchers, iPods, video game consoles, and USB sticks) that were sold sufficiently often on both auction websites (more than 40 times).

As a result, both auction websites address potential bidders who live in the same geographic region and are interested in buying the same products during the same period of time. The bidders then self-select themselves into one (or maybe both) of these auction formats. Thus, although we control for many effects, self-selection may still affect our results. Our comparison measures differences between the two auction formats under the condition that bidders can freely choose between the two auction formats. This difference is still relevant information for auctioneers, as they must also consider the self-selection decisions of bidders.

Our economic analysis suggests that the revenues of ascending auctions and descending auctions are equal (Prediction 4) and that the variance of descending auctions is lower than the variance of ascending auctions (Prediction 3).

Table 9 displays the standardized revenues and variances across four identical products that were sold on both websites from December 2007 to October 2008, taken from data sets A1 und D1. All ascending auctions are ten-cent auctions. Consistent with our expectations, variances are significantly higher in ascending auctions than in descending auctions. This result holds across all observed categories. In the voucher category, the standard deviation is nearly forty times higher in ascending auctions.

Fig. 3 illustrates these strong deviations using the example of iPods. For the ascending auction, the scale of the x-axis begins at zero because there are numerous auctions in which the achieved revenues are smaller than the CRP. In contrast, the scale begins at 1 for descending auctions, where revenues are at least as high as the CRP. The scale ends at 5 for ascending auctions; hence, in some auctions, revenues are five times as high as the CRP. The maximum standardized revenue for descending auctions is reached with revenues that are 1.3 times higher than the CRP.

J.-Y. Kim et al. / Intern. J. of Research in Marketing 31 (2014) 368–379

377

Table 9
Comparison of mean standardized revenues and variances of ascending and descending auctions.

	Ascending auction			Descending auction			Δ − % in means	Δ − % in variance
	N	Mean[a]	Std. dev.	N	Mean[a]	Std. dev.		
Voucher	43	0.98	0.69	111	1.22	0.02	−20%[***]	3977%[***]
Video game console	9,376	1.73	1.06	66	1.09	0.09	59%[***]	1117%[***]
iPod	607	1.46	0.95	130	1.11	0.09	31%[***]	942%[***]
USB stick	1,984	0.73	0.65	69	1.13	0.09	−35%[***]	645%[***]

Δ − %: percentage differences between the means of ascending and descending auctions.

[a] Standardized revenue, defined as revenue/CRP; CRP: current retail price.

[***] $p < 0.01$, two-tailed.

Thus, ascending pay-per-bid auctions are associated with higher risks but result in a much wider range of standardized revenues. The results in Table 9 also indicate that the revenues of this specific category (iPods) are even significantly higher in ascending auctions than in descending auctions. This finding contradicts Prediction 4, which posited revenue equivalence. Table 9 shows that the reverse is true for vouchers and USB sticks. Here, standardized revenues are 20% to 35% lower in ascending auctions. However, standardized revenues for video game console and iPods, defined as revenues relative to the CRP, are significantly higher in ascending ten-cent auctions, ranging from 37% to 59% higher in the video game console category.

The differences in our results may again be explained by category differences. As we have previously shown, the increased revenues for hedonic products such as video game consoles and iPods in ascending auctions may be caused by irrational bidding behavior resulting from emotional arousal or overestimation of bargain. In contrast, hedonic products result in decreased revenues in descending auctions. Here, we expect that if the bidder has a strong desire to own a product, then she will buy early for fear of another bidder taking the product.

7. Summary, implications and future research

7.1. Summary of results

The objective of this paper was to theoretically and empirically analyze the economic effects of alternative formats of pay-per-bid auctions, in particular different auction formats (ascending versus descending auction) and different price increments (one-cent versus ten-cent auctions). For this purpose, we adapted and extended existing theoretical models on pay-per-bid auctions, formulated predictions regarding auctioneers' revenues and tested them empirically with three large, unique data sets.

Analyzing *ascending pay-per-bid auctions*, we found that an increase in the price increment for each bid reduces the variance of the auctioneer's revenue, confirming our prediction: a higher change in price increment reduces the risk that is associated with selling the product. We found that the use of ten-cent auctions yields revenues per auction that are less volatile and consequently less risky than the use of penny auctions.

However, penny auctions led to higher revenues per auction compared with ten-cent auctions. In contrast to Prediction 2, our analysis provides evidence that an increase in the price increment affects the expected revenue.

We further explained the observed differences between the actual revenues and the expected revenues that were derived from the theoretical model, and we found that factors such as the number of bidders and the type of the product category (whether hedonic or utilitarian) are drivers of these differences.

Our empirical data set of *descending pay-per-bid auctions* did not include different price increments; thus, we could not determine their economic effects. However, the data showed differences between actual revenues and expected revenues. Again, the number of bidders and the characteristic of the category (hedonic or utilitarian) were found to affect these differences.

Finally, we compared *ascending and descending pay-per-bid auctions*. Confirming Prediction 3, we found that the variance of the revenue per auction is higher in ascending auctions than in descending auctions. However, in contrast to Prediction 4, which postulated revenue equivalence between ascending and descending pay-per-bid auctions, we found significant differences in revenues per auction.

The theoretical model helps to place our empirical results into perspective. Average revenues per auction above the CRP are not supported
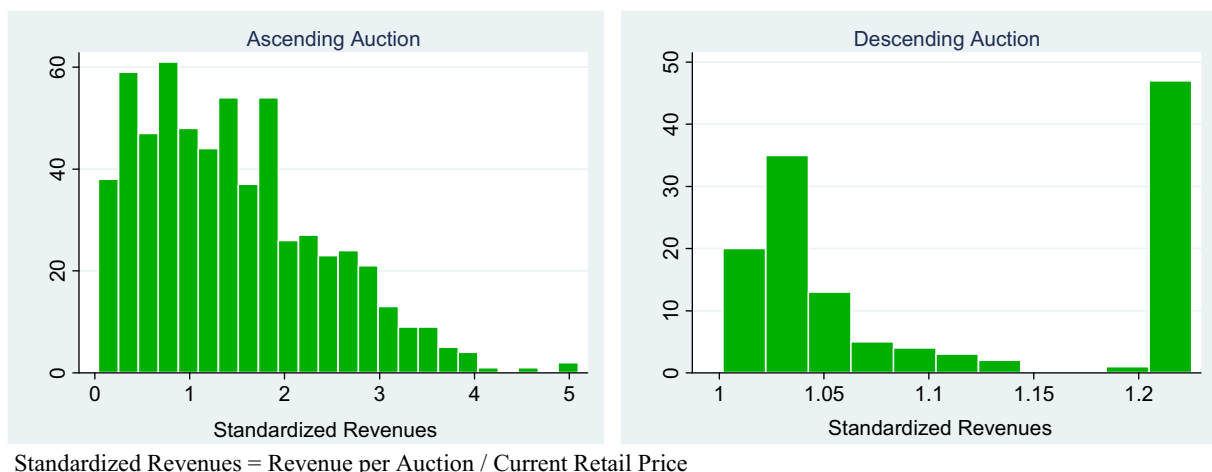

Standardized Revenues = Revenue per Auction / Current Retail Price

Fig. 3. Distribution of the standardized revenues of ascending and descending auctions for iPods.

by the predictions of the theoretical model. Consequently, average revenues are the result of a consumer behavior that is not consistent with the theoretical model, such as non-equilibrium play, the overvaluation of products, risk-loving preferences or other forms of behavior that are inconsistent with our assumptions. However, all these phenomena may be transitory. The longer these new auction formats are available, the more experience users obtain. Ultimately, individuals may learn to play the equilibrium. Risk-seeking shoppers may move on to newer entertainment shopping venues. Irrational individuals could learn to act more rational in pay-per-bid auctions or avoid them altogether. Therefore, our empirical findings should be interpreted with caution because average revenues above CRP may not be sustainable over a long period.

7.2. Implications

Our findings provide a number of implications for marketers and researchers. First, auctioneers can use the findings from our study to consider different auction formats that may improve their revenues. For ascending pay-per-bid auctioneers, our results indicate that penny auctions yield higher revenues per auction but are also associated with higher risks. Thus, an ascending auctioneer must weigh the individual advantages and disadvantages of such a method. Our results further imply that an ascending pay-per-bid auctioneer may benefit from a higher number of bidders and may benefit from selling products in hedonic categories rather than utilitarian categories in the short term. Thus, such an auctioneer could make a greater effort to enhance traffic (e.g., through advertising) and to auction off more hedonic products than utilitarian products.

On the contrary, our results suggest that a descending pay-per-bid auctioneer should sell utilitarian products rather than hedonic products in their auctions. Auctions with hedonic products end earlier, leading to lower revenues per auction.

Finally, we would recommend that auctioneers who cannot decide between ascending and descending formats should choose the descending format when they are risk-averse and to choose an ascending format when they are less risk-averse because the latter method offers potential for a much wider range of revenues.

7.3. Future research

A crucial question that is beyond the aim of this paper would involve determining how long the differences between actual and expected revenues occur. Although our data sets cover a period of up to one year, this period may be too short to fully capture bidders' learning. Such learning would reduce the differences between actual and expected revenues per auction. Future research may thus aim to analyze the development of the observed differences over time.

Additionally, our comparison of the revenue per auction between ascending and descending pay-per-bid auctions could not fully separate the effect of the auction format from the self-selection effect of bidders. Although the difference that includes both effects is relevant information for auction providers, future research may be capable of better distinguishing these effects.

Acknowledgments

Appendix A. Equilibrium of the descending pay-per-bid auction

In this appendix we prove that the following strategies constitute a subgame perfect equilibrium: (i) The auctioneer chooses the starting price $s = v + e - b$. (ii) All bidders want to place a bid if $s \leq v + e - b$. If the starting price is above $v + e - b$, no bidder wants to place a bid. (iii) Once a bidder observes the price she buys the product and the auction ends.

Proof of part (iii). Suppose all players except bidder j follow the equilibrium strategies given above. Assume first that bidder j is the first bidder who observes the price. She observes the price $v - b$. The starting price is $v + e - b$ and bidder j's bid reduces this price by e. Since she values the product at v, buying the product gives her utility of b. If she does not buy the product, then the next bidder who observes the price will buy it and bidder j will receive nothing. Thus, there is no profitable deviation for bidder j. Assume now that bidder j is not the first to observe the price. Then, the utility from buying the product is greater than b, whereas the utility from not buying is still 0. Again, there is no profitable deviation for bidder j.

Proof of part (ii). Suppose that all players except bidder j follow the equilibrium strategies given above. If bidder j places a bid, observes the price and buys the product, then she receives a utility of $v - (s - e) - b$. The first part of this expression is the value of the product, the second is the price bidder j pays and the last part is the bidding fee. When bidder j acquires the opportunity to observe the price, she must be the first bidder to observe the price, as part (iii) tells us that otherwise, the auction would have ended already. Not bidding yields a utility of zero. Thus, bidder j wants to place a bid, if $v - (s - e) - b \geq 0$ which can be rewritten as $s \leq v + e - b$.

Proof of part (i). Suppose that all bidders follow the equilibrium strategies given above. The auctioneer's revenue is $s - e + b$, if $s \leq v + e - b$ and 0, otherwise. This revenue is maximized for $s = v + e - b$ and the maximum revenue is equal to v.

Appendix B. Description of scales and categorization of products

To categorize products as hedonic or utilitarian, we surveyed a sample of 86 people and asked our respondents to classify the products as hedonic or utilitarian. For this purpose, we introduced the respondents to the concept of hedonic and utilitarian products and asked them to classify the products according to the method of Strahilevitz and Myers (1998) as utilitarian (practical), hedonic (frivolous), both utilitarian and hedonic, or neither utilitarian nor hedonic. The classification of products was then based on the modal classification of the respondents.

Categorization of hedonic/utilitarian products

Product	Current retail price in €	Categorization of product as	
		Hedonic	Utilitarian
Nintendo Wii	172	1	0
Apple iPod Touch	142	1	1
Nintendo WiiFit	71	1	0
Nintendo DS Lite	107	1	0
TomTom GO 740	399	0	1
Nikon D90 Camera	992	1	1
Kaspersky Internet Security	35	0	1
Phillips Full HD TV	1,269	1	0
Braun Oral-B Triumph	121	0	1
Panasonic KX	74	0	1
Acer Aspire	1,000	1	1
Samsung SGH-i900	601	1	1
Rothenschild Kryptonite	236	0	1
Kingston Data Traveler 32GB	59	0	1
Voucher 50 bids	25	0	1

1: yes; 0: no.

Appendix C. Supplementary data

Estimation code for this article can be found online at http://www.runmycode.org. Interested scholars may contact either the corresponding author or IJRM's editorial office in order to request the dataset.

References

Amaldoss, W., & Jain, S. (2008). Joint bidding in the name-your-own-price channel: A strategic analysis. *Management Science*, *54*(10), 1685–1699.

Augenblick, N. (2012). *Consumer and producer behavior in the market for penny auctions: A theoretical and empirical analysis.* Working paper (Available at: http://faculty.haas.berkeley.edu/ned/Penny_Auction.pdf).

Barrot, C., Albers, S., Skiera, B., & Schäfers, B. (2010). Vickrey vs. eBay: Why second-price sealed-bid auctions lead to more realistic price–demand functions. *International Journal of Electronic Commerce*, *14*(4), 7–38.

Byers, J. W., Mitzenmacher, M., & Zervas, G. (2010). *Information asymmetries in pay-per-bid-auctions.* Paper presented at the ACM Conference on Electronic Commerce. Massachusetts: Harvard University.

Caldara, M. (2012). *Bidding behavior in pay-to-bid auctions: An experimental study.* Working paper (Available at: http://www.chapman.edu/research-and-institutions/economic-science-institute/_files/BrownBag/michaelcaldera2012.pdf).

Dholakia, U. M., Basuroy, S., & Soltysinski, K. (2002). Auction or agent (or both)? A study of moderators of the herding bias in digital auctions. *International Journal of Research in Marketing*, *19*(2), 115–130.

Gallice, A. (2011). *Price reveal auctions on the internet.* Working paper Available at: http://www.andreagallice.eu/PRAs%202011.pdf.

Haruvy, E., & Popkowski Leszczyc, P. T. L. (2009). Bidder motives in cause-related auctions. *International Journal of Research in Marketing*, *26*(4), 324–331.

Hinnosaar, T. (2010). *Penny auctions are unpredictable.* Working paper (Available at: http://toomas.hinnosaar.net/pennyauctions.pdf).

Hinz, O., & Spann, M. (2008). The impact of information diffusion on bidding behavior in secret reserve price auctions. *Information Systems Research*, *19*(3), 351–368.

Hirschman, E. C., & Holbrook, M. B. (1982). Hedonic consumption: Emerging concepts, methods and propositions. *Journal of Marketing*, *46*(3), 92–101.

Jap, S. D., & Naik, P. A. (2008). BidAnalyzer: A method for estimation and selection of dynamic bidding models. *Marketing Science*, *27*(6), 949–960.

Pinker, E. J., Seidmann, A., & Vakrat, Y. (2003). Managing online auctions: Current business and research issues. *Management Science*, *49*(11), 1457–1484.

Platt, B. C., Price, J., & Tappen, H. (2013). The role of risk preferences in pay-to-bid auctions. *Management Science*, *59*(9), 2117–2134.

Reiner, J., Brünner, T., Natter, M., & Skiera, B. (2014). *Using cumulative prospect theory to explain why bidders participate in pay-per-bid auctions.* Working paper.

Reiner, J., Natter, M., & Skiera, B. (2014). Buy-now features in pay-per-bid auctions. *Journal of Management Information Systems* (forthcoming).

Spann, M., Skiera, B., & Schäfers, B. (2004). Measuring individual frictional costs and willingness-to-pay via name-your-own-price mechanisms. *Journal of Interactive Marketing*, *18*(4), 22–36.

Strahilevitz, M., & Myers, J. G. (1998). Donations to charity as purchase incentives: How well they work may depend on what you are trying to sell. *Journal of Consumer Research*, *24*, 434–446 (March).

Wang, Z., & Xu, M. (2013). *Selling a dollar for more than a dollar? Evidence from online penny auctions.* RFF discussion paper (Available at: http://www.rff.org/RFF/Documents/RFF-DP-13-15.pdf).

ELSEVIER

Full Length Article

Predicting consumer behavior with two emotion appraisal dimensions: Emotion valence and agency in gift giving

CrossMark

Ilona E. de Hooge *

Department of Marketing Management, Rotterdam School of Management, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands

ARTICLE INFO

ABSTRACT

Decades of emotion research have demonstrated the unique influences of many specific emotions on consumer behaviors. These countless numbers of emotion effects can make it difficult to understand the role of emotions in consumer behavior. The current research introduces a parsimonious framework that can predict the effects of emotions on the consumer behavior of gift giving with just two appraisal dimensions: valence and agency. A series of studies examining gift giving reveals that positive emotions exert positive effects on gift giving, independent of their agency. In contrast, agency does predict the effects of negative emotions on gift giving. Negative self-caused emotions increase gift giving, whereas negative other-caused emotions decrease gift giving. These findings seem to hold for inactive and active emotions, and for uncertain and certain emotions. Together, these findings make a unique theoretical and empirical contribution to the understanding of emotions in gift giving. Moreover, it provides a pragmatic framework for both academics and practitioners.

1. Introduction

Over the last couple of decades, numerous studies have shown how specific emotions can influence consumer behaviors. For example, we currently know that feelings of anger may motivate consumers to complain about a company (Nyer, 1997), that feelings of pride can encourage people to buy public display products (Griskevicius, Shiota, & Nowlis, 2010), and that dissatisfied customers experiencing regret may switch to a different service provider (Zeelenberg & Pieters, 2004). With at least twenty emotions that play a role in marketing settings (Richins, 1997), it can be difficult for marketing academics and practitioners to identify and understand the influences of all these emotions on consumer behaviors. It is uncertain whether it is necessary to distinguish between similar emotions such as regret and disappointment, or shame and guilt when examining the influences of emotions on consumer behaviors. Some emotion scholars suggest that we should make this distinction (e.g., De Hooge, Zeelenberg, & Breugelmans, 2007; Griskevicius, Shiota and Nowlis, 2010; Zeelenberg & Pieters, 2004). The present research proposes that it is possible to predict the influences of

emotions for at least some consumer behaviors on the basis of a limited number of emotion appraisal dimensions.

According to most scholars, emotions can be defined with a restricted number of cognitive or emotion appraisal dimensions. These include, for example, whether the emotion is positive or negative (valence), whether the outcomes are certain or uncertain (certainty), and whether the person feels powerful or powerless (power) (Bagozzi, Gopinath, & Nyer, 1999; Frijda, 1986; Frijda, Kuipers, & Ter Schure, 1989; Roseman, Wiest, & Swartz, 1994; Smith & Lazarus, 1993). Emotion valence (the extent to which an emotion is positive or negative) and emotion agency (the extent to which an emotion is caused by oneself or caused by another person) are the two central appraisal themes that form the basis of the current research. Numerous emotion researchers have proposed valence and agency as appraisal dimensions (Bagozzi, Gopinath and Nyer, 1999; Frijda, 1986; Ortony, Clore, & Collins, 1988; Ruth, Brunel, & Otnes, 2002; Smith & Ellsworth, 1985). Because these appraisal dimensions separately have been found to influence prosocial behaviors and altruism (e.g., Butt & Choi, 2006; Chaudhuri, 1997; Fredrickson, 2001; Kelley & Hoffman, 1997; Moll et al., 2007), the present research suggests that the interaction between valence and agency can predict, to a certain extent, the influence of specific emotions on a consumer behavior that generates millions of dollars every year: gift giving. Many scholars have examined the emotions that influence gift giving (e.g., Belk, 1976; Komter & Vollebergh, 1997; Ruth, 1996; Schwartz, 1967), but there is hardly any empirical work that studies the effects of multiple different emotions on gift

* Tel.: +31 10 408 14 87; fax: +31 10 408 90 11.
  E-mail address: IHooge@RSM.nl.

giving. The current research also includes the appraisals of emotion activity (the extent to which an emotion makes the person feel active or inactive) and emotion certainty (the extent to which the outcomes of the emotion-generating situation are certain or uncertain), to test whether the effects of valence and agency generalize to active and inactive as well as certain and uncertain emotions.

The present research consists of studies which examine the effects of four different emotions on gift giving. These effects concern integral emotion influences, such that givers experience emotions towards receivers. The findings reveal that, at least for most of the tested emotions, the effects on gift giving decisions can be predicted on the basis of just two appraisal dimensions: valence and agency. More specifically, it appears that positive valence exerts positive influences on gift giving, independent of the agency. In contrast, the effects of emotions with a negative valence depend on the agency of the emotion. Negative emotions caused by the giver generate positive effects on gift giving, but negative emotions caused by the receiver generate negative effects on gift giving. This framework also holds when including emotion activity (Study 5) or emotion certainty (Study 6) as a third appraisal dimension. Thus, emotion valence and agency seem to be able to predict most emotion influences on gift giving in a parsimonious way, making it easier for marketing practitioners and academics to understand and to apply emotion influences in consumer settings.

2. Gift giving and emotions

Gift giving has been studied by scholars from anthropology, psychology, marketing, economics, sociology, and philosophy (for overview see Banks, 1979; Belk, 1982). According to most disciplines, gifts can be understood as goods or services that are voluntarily provided from one person to another person or to a group (Belk, 1979). Usually, this provision takes place in ritual-like situations, such as birthdays, weddings, or Christmas settings, and the gifts may involve physical gifts, immaterial gifts (such as time or services), or cash gifts (Belk, 1976). The gift giving process takes place in three general stages: a gestation stage in which the giver searches for and buys gifts, a prestation stage in which gifts are exchanged, and a reformulation stage in which gifts are consumed or rejected, and in which the relationship between the giver and the recipient may change (Sherry, 1983).

Numerous factors can influence gift giving (Belk, 1976; Sherry, 1983), and emotions are considered to be one of those factors. Most theories and empirical research on emotions in gift giving focus on emotion effects *during* the gift giving process (Algoe, Haidt, & Gable, 2008; Belk, 1996; Ruffle, 1999; Ruth, Brunel, & Otnes, 2004; Ruth, Otnes, & Brunel, 1999; Sherry, McGrath, & Levy, 1993; Wooten, 2000). Instead, the current research focuses on how emotions generated *before* the gift giving process can influence the purchase of gifts during the gestation stage. These emotions concern the giver's emotions in relation to the receiver and thus reflect integral emotion effects. Gift giving theories mention that love, joy, penitence, sadness, and gratitude might stimulate gift giving to express emotional states (Cheal, 1988; Fischer & Arnold, 1990; Ruth, 1996; Sherry, 1983), or that gifts communicate feelings of love, affection, care, esteem, and friendship (Belk & Coon, 1993; Goodwin, Smith, & Spiggle, 1990; Komter & Vollebergh, 1997; Otnes, Ruth, & Milbourne, 1994; Wolfinbarger & Yale, 1993). A giver may express feelings of joy and pride with gift giving after a recipient has achieved something (Ruth, 1996; Smith & Ellsworth, 1985), and may try to lessen feelings of guilt by purchasing gifts (Wolfinbarger, 1990). A number of empirical studies support the notion that emotions might influence gift giving in the gestation stage. The positive emotions of pride and confidence have been found to stimulate self-gift giving (Mick & Faure, 1998), and feelings of (agapic) love have been shown to exert an effect on the money, time, and effort that is spent on finding a gift (Belk & Coon, 1993; Goodwin, Smith and Spiggle, 1990). Thus, many different emotions seem to influence gift giving during the gestation stage, although little empirical research has examined the effects of

multiple different emotions on gift giving. If different emotions indeed exert different effects on gift giving, how can these emotion influences be summarized in a parsimonious framework?

In general, emotions arise in response to evaluative judgments and interpretations of events that are relevant for consumers' well-being (Bagozzi, Gopinath and Nyer, 1999; Nyer, 1997). Put differently, emotions reflect a goal that is potentially threatened (in the case of negative emotions) or served (in the case of positive emotions) (Zeelenberg, Nelissen, Breugelmans, & Pieters, 2008). Different combinations of these evaluative judgments, often called cognitive appraisals or emotion appraisals, yield different emotional responses (Frijda, 1986; Ortony, Clore and Collins, 1988; Roseman, Antoniou, & Jose, 1996; Smith & Lazarus, 1993). Most scholars mention the existence of five emotion appraisal dimensions: valence or pleasantness (the extent to which an emotion is positive or negative), activity or arousal (the degree to which one feels active or inactive), certainty (the degree to which the outcome of the event is certain or uncertain), power or control (the degree to which one feels powerful or powerless), and agency (the degree to which the emotion is caused by oneself or by other people) (Bagozzi, Gopinath and Nyer, 1999; Frijda, 1986; Smith & Ellsworth, 1985). Of these appraisal dimensions, I argue that the two appraisal dimensions of valence and agency can be used to predict emotion effects on gift giving.

The appraisal dimension of valence makes a distinction between positive emotions and negative emotions (Bagozzi, Gopinath and Nyer, 1999; Frijda, 1986). Research has demonstrated that a distinction between emotions on the basis of their positivity or negativity is essential in understanding emotion influences on many different kinds of behavior that are related to altruism and prosociality (e.g., Chaudhuri, 1997; Fredrickson, 2001; Gino & Schweitzer, 2008; Reisenzein & Hofmann, 1990, 1993; Ruth, Brunel and Otnes, 2002). For example, compared to negative emotions, positive emotions have been found to encourage cooperation in groups and in social dilemma games (Haselhuhm & Mellers, 2005; Hertel, Neuhof, Theuer, & Kerr, 2000). They have also been shown to increase prosocial actions, to increase helping, to increase altruistic behaviors towards colleagues, and to reduce harmful actions towards others (Batson, 1998; Isen & Levin, 1972; Kelley & Hoffman, 1997). Because gift giving is considered to be a form of prosocial or altruistic behavior (Fischer, Gainer, & Arnold, 1996; Homans, 1961; Otnes & Beltramini, 1996), these findings indicate that the appraisal dimension of valence should be taken into account when predicting emotion effects on gift giving.

Since gift giving is an inherently social process that involves at least one other person (the receiver), I claim that the social aspects of emotions should also be taken into account when explaining the role of emotions in gift giving. The appraisal dimension of agency (also named causality or responsibility) is the social aspect of emotions, and it distinguishes emotions that are caused by the self (self-caused emotions) from those caused by other people (other-caused emotions) (Bagozzi, Gopinath and Nyer, 1999; Frijda, 1986; Ortony, Clore and Collins, 1988; Roseman, Wiest and Swartz, 1994). Translated to the giver's perspective in a gift giving context, agency distinguishes self-caused emotions that result from actions of givers and other-caused emotions that result from actions of receivers. Previous research has found that agency is an important appraisal dimension for emotions in a social context (Lazarus, 1991; Reisenzein & Hofmann, 1993; Smith & Ellsworth, 1985; Weiner, 1986). For example, agency has been found to predict collaborative and competitive motives and subsequent compromise behaviors in negotiations (Butt & Choi, 2006; Butt, Choi, & Jaeger, 2005). It has also been demonstrated to influence social behaviors such as helping (or avoiding) others, attaching to others, correcting other people's mistakes (Moll et al., 2007), complaining and protest behaviors towards companies (Grappi, Romani, & Bagozzi, 2013; Soscia, 2007), and concerns with the welfare of comparable others (Choshen-

Hillel & Yaniv, 2011). These findings suggest that the appraisal dimension of agency should also be taken into account when predicting emotion effects on the social act of gift giving.

There is some empirical support for the notion that the influence of emotions on gift giving can be predicted on the basis of the appraisal dimensions of valence and agency. Ruth, Brunel and Otnes (2002) used a gift giving setting to examine whether consumption emotions could be summarized into a limited number of appraisal dimensions. In two studies, participants were asked to think of a situation in which they felt a certain emotion during or after having received a gift. Both studies found that most variance of ten different consumption emotions could be explained with two appraisal dimensions: valence and agency. Although the findings concern emotions experienced after gift receipt and not the influence of emotions on the purchase of gifts, they do reveal that valence and agency might play a role in gift giving.

3. Hypotheses

In sum, I argue that the emotion effects of most emotions on gift giving can be predicted on the basis of just two appraisal dimensions: valence and agency. I propose that self-caused and other-caused emotions exert distinct influences on gift giving, which are contingent on the valence of the emotion. More specifically, I put forth the following two propositions.

Firstly, I hypothesize that positive valence (e.g., pride, joy, satisfaction, gratitude, love) increases gift giving, independent of the agency of the emotion. Positive emotions can broaden consumers' momentary thoughts to focus on a wider (than typical) range of thoughts and actions (Derryberry & Tucker, 1994; Fredrickson, 1998, 2001; Fredrickson & Branigan, 2005). This range of thoughts and actions mostly focuses upon activities that are evolutionary adaptive and that build enduring personal resources (Fredrickson & Cohn, 2008). One such evolutionary adaptive activity that builds enduring personal resources is developing and maintaining social relationships. Therefore, *all* positive emotions are presumed to motivate social approach behaviors and actions that maintain one's social relationships (Carver & Scheier, 1990; Fredrickson, 1998, 2001; Frijda, 1986). Indeed, "the function common to all positive emotions has been conceptualized as facilitating approach behavior" (Fredrickson & Cohn, 2008, p. 778). For example, positive affect has been found to motivate helping, independent of the cause of these positive feelings (Isen & Levin, 1972). In a gift giving context, this suggests that positive emotions would increase gift giving to maintain relationships with receivers, independent of whether the positive feelings are caused by the giver (e.g., pride, satisfaction) or caused by the receiver (e.g., gratitude, love). Indeed, some positive other-caused emotions such as gratitude have been found to motivate cooperation and prosocial behavior (DeSteno, Bartlett, Baumann, Williams, & Dickens, 2010; McCullough, Kimeldorf, & Cohen, 2008). For positive self-caused emotions there is also some research suggesting that such emotions (e.g., pride) might stimulate prosocial behavior (Harth, Kessler, & Leach, 2008; Wubben, De Cremer, & Van Dijk, 2012).

Secondly, I hypothesize that negative valence (e.g., shame, guilt, anger, fear, disgust) can increase or decrease gift giving, depending on the agency of the emotion. My argument is based on the idea that negative emotions narrow consumers' momentary thoughts to specific, immediate actions that address consumers' needs (Derryberry & Tucker, 1994; Fredrickson, 1998, 2001; Fredrickson & Branigan, 2005). Negative emotions are believed to be developed for specific, survival–critical situations (Fredrickson & Cohn, 2008; Frijda, 1986). When there are threatening situations that demand action, negative emotions are thought to shrink perceptions and thoughts to actions that are necessary to deal with the threat (Fredrickson & Branigan, 2005; Zeelenberg, Nelissen, Breugelmans and Pieters, 2008). For instance,

every specific negative emotion has been connected with specific action tendencies, whereas these tendencies are mostly underspecified for specific positive emotions (Fredrickson & Cohn, 2008). Moreover, recent research has demonstrated that negative emotions, despite their similar valence, can have different effects on prosocial behavior and decision making (De Hooge, Zeelenberg and Breugelmans, 2007; Lerner & Keltner, 2000; Raghunathan & Pham, 1999). This suggests that negative emotions in a gift giving context will probably *not* stimulate gift giving automatically. Instead, negative emotions will motivate givers to analyze the gift giving situation in relation to their own needs. The agency of the emotion thereby provides information.

In the case of negative self-caused emotions such as shame or guilt, givers have done something wrong themselves. As a consequence, the emotion signals that gift giving might be undertaken to improve the relationship with the receiver. This prediction might seem surprising at first given that previous literature on emotions showed that self-conscious emotions such as regret and guilt motivate a self-focus (Tracy, Robins, & Tangney, 2007), which may suggest that such emotions would decrease gift giving. Yet, some recent studies have challenged this view by reporting that guilt and shame can motivate prosocial behavior (De Hooge, Breugelmans, & Zeelenberg, 2008; De Hooge, Zeelenberg and Breugelmans, 2007; Ketelaar & Au, 2003), and that negative self-caused emotions can motivate compromise behaviors in negotiations (Butt, Choi and Jaeger, 2005). These scholars argue that negative self-caused emotions make people feel less self-confident and subsequently stimulate behaviors that are positively regarded by others and by society in order to avoid more wrongdoing. This suggests that such negative self-caused emotions might stimulate gift giving.

On the contrary, negative other-caused emotions such as anger, fear, or contempt indicate that receivers have done something wrong. As a consequence, the emotion signals that the receiver should undertake action to mend the relationship with the giver, or, alternatively, that the giver should decrease gift giving in order to weaken the relationship with the receiver. This prediction is in line with the findings that negative other-caused emotions can motivate dominating behaviors in negotiations (Butt, Choi and Jaeger, 2005) and conflict-creating behaviors in relationships (Sanford & Rowatt, 2004). In summary, I hypothesize that negative self-caused emotions *increase* gift giving, whereas negative other-caused emotions *decrease* gift giving.

4. Examining emotion dimensions in gift giving

To study the proposition that the influence of emotions on gift giving depends on the appraisal dimensions of valence and agency, I conducted six studies in which different emotions were induced and different measures for gift giving were used. In all studies, I explored the effects of valence and agency on gift giving by first introducing an emotion induction task and then measuring the effects on gift giving. There are different ways to measure increases and decreases in gift giving. Following conventional gift giving research, I used the amount of money that the giver spends on a gift, the effort that the giver puts into finding a gift (e.g., Flynn & Adams, 2009; Goodwin, Smith and Spiggle, 1990; Katz, 1976), how personal the gift is, and how big it is (Goodwin, Smith and Spiggle, 1990; Ward & Broniarczyk, 2011) as gift giving measures. To generalize the findings further, Study 2 explored emotion influences on the decision to buy a gift, and Studies 4 to 6 also measured the time that the giver intends to spend on searching for a gift.

In all studies, standard emotion induction measures from emotion research were used (De Hooge, Breugelmans and Zeelenberg, 2008; Frijda, Kuipers and Ter Schure, 1989; Roseman, Wiest and Swartz, 1994). In Studies 1 to 3, participants described a personal situation in which they experienced a certain emotion (the autobiographical recall procedure). To examine the effects beyond specific emotions, Study 4 directly manipulated the valence and agency dimensions. In addition,

this study investigated whether emotion effects on gift giving occurred because givers wanted to maintain (in the case of positive emotions), improve (in the case of negative self-caused emotions), or weaken (in the case of negative other-caused emotions) the relationships with receivers. To avoid confusion, these motivations will collectively be labeled relationship management from now on. Together, these studies provide support for the idea that emotion effects on gift giving can be predicted on the basis of the dimensions of valence and agency.

Studies 5 and 6 continued by examining whether other appraisal dimensions might also play a role. Study 5 included the appraisal dimension of activity whereas Study 6 included the appraisal dimension of certainty. In both studies all gift giving measures and reasons for gift giving were assessed. In addition, because one might argue that the strength of the relationship between givers and receivers might play a role in the findings, Studies 5 and 6 included relationship strength as a covariate in the analyses. The results reveal that the valence–agency framework can also predict effects of inactive and uncertain emotions on gift giving. Together, the six studies complement each other in multiple ways, and convergence in the results obtained in these different settings contribute to the generalizability of the findings.

5. Study 1: inducing pride, gratitude, guilt, and anger

5.1. Method

5.1.1. Participants and design

Two hundred seventy-one international students from a Western European university (147 males, $M_{age} = 20.41$, $SD_{age} = 2.15$) participated in partial fulfillment of a course requirement. They were randomly assigned to the control condition or to one of the conditions of a 2 (valence: positive vs. negative) × 2 (agency: self-caused vs. other-caused) between subjects design with gift giving and money spent on the gift as dependent variables.

5.1.2. Procedure and variables

To induce emotions, participants first completed an autobiographical recall procedure. In this manipulation procedure, participants are usually asked to recall a personal incident in which they experienced a certain emotion (De Hooge, Zeelenberg and Breugelmans, 2007; Ketelaar & Au, 2003; Roseman, Wiest and Swartz, 1994). In this study, participants reported a personal experience in which they felt very proud (positive self-caused condition), gratified (positive other-caused condition), guilty (negative self-caused condition), or angry (negative other-caused condition). In the control condition, participants described a regular weekday. Participants spent approximately 10 min on the emotion induction task. Next, participants were asked to think of someone towards whom they experienced the emotion (in the other-caused conditions) or someone who was present in the described event (in the self-caused conditions). If there was nobody present, participants were asked to think of someone that they had told about the event afterwards. In all cases, participants typed the name of this person.

To measure gift giving, participants then imagined that a week after the described event it was the named person's birthday. As the dependent measures, participants indicated how much they would spend on the birthday gift (amount in euros), how much effort they would put into finding a gift (0 = no effort, 10 = a lot of effort), how personal the gift would be (involvement, 0 = not personal at all, 10 = very personal), and how big the gift would be (size, 0 = smaller than normal, 10 = bigger than normal). A factor analysis on these gift items showed a clear one factor solution (see Appendix A for the items and factor loadings). The factor gift giving (Eigenvalue = 3.04) explained 76% of the variance, but only formed a reliable scale ($\alpha = .92$) when the money spent on the gift was left out ($\alpha = .36$ when included). Therefore, in this and further studies the money spent on the gift was analyzed separately. Finally, as an emotion manipulation check, participants reread their situation description and indicated how much pride, gratitude, guilt, and anger they felt (0 = not at all, 10 = very strongly).

Table 1
Emotion manipulation check means (and standard deviations) and tests in Studies 1 to 4.

Study	Condition			
	Positive		Negative	
	Self-caused	Other-caused	Self-caused	Other-caused
Study 1				
Target emotion	Pride	Gratitude	Guilt	Anger
M (SD)	8.90 (1.20)	8.40 (1.94)	8.00 (1.76)	8.74 (1.68)
Compared to other conditions	$ts(266) > 6.04^{**}$	$ts(266) > 4.01^{**}$	$ts(266) > 12.62^{**}$	$ts(266) > 10.60^{**}$
Compared to other emotions within the same condition	$ts(57) > 6.33^{**}$	$ts(52) > 5.50^{**}$	$ts(48) > 10.32^{**}$	$ts(53) > 8.63^{**}$
Study 2				
Target emotion	Pride	Gratitude	Guilt	Anger
M (SD)	8.86 (1.58)	8.39 (1.52)	8.38 (1.67)	8.46 (1.63)
Compared to other conditions	$ts(174) > 6.38^{**}$	$ts(174) > 6.35^{**}$	$ts(174) > 6.47^{**}$	$ts(174) > 7.86^{**}$
Compared to other emotions within the same condition	$ts(36) > 4.49^{**}$	$ts(35) > 6.43^{**}$	$ts(33) > 7.36^{**}$	$ts(34) > 9.96^{**}$
Study 3				
Target emotion	Satisfaction	Love	Shame	Fear
M (SD)	8.53 (1.88)	8.51 (1.53)	8.00 (1.23)	8.09 (2.13)
Compared to other conditions	$ts(211) > 2.34^{*}$	$ts(211) > 4.84^{**}$	$ts(211) > 7.89^{**}$	$ts(211) > 6.36^{**}$
Compared to other emotions within the same condition	$ts(44) > 5.49^{**}$	$ts(42) > 3.21^{**}$	$ts(40) > 8.18^{**}$	$ts(44) > 6.75^{**}$
Study 4				
Target emotion	Pride, satisfaction	Gratitude, love	Guilt, shame	Anger, fear
M (SD)	8.11 (1.60), 8.56 (1.57)	7.82 (2.42), 7.57 (2.56)	7.74 (1.81), 7.33 (1.84)	7.24 (2.61), 2.90 (2.86)
Compared to other conditions	$ts(133) > 2.83^{**}$	$ts(133) > 3.00^{**}$	$ts(133) > 4.77^{**}$	$ts(133) > 2.06^{*}$
Compared to other emotions within the same condition	$ts(26) > 3.63^{**}$	$ts(27) > 3.47^{**}$	$ts(26) > 2.09^{*}$	$ts(28) > 2.09^{*}$

Note. For every condition, the target emotion was compared to the same emotion in the other conditions ("compared to other conditions") and compared to other emotions within the same condition ("compared to other emotions within the same condition").

  * $ps < .05$.
 ** $ps < .01$.

5.2. Results and discussion

5.2.1. Emotion manipulation check

Results for the emotion manipulation checks of Studies 1 to 4 can be found in Table 1. The emotion manipulation for all those studies was successful. Participants in the positive self-caused condition reported more pride than participants in all other conditions, and more pride than other emotions. Similar effects were found for gratitude in the positive other-caused condition, for guilt in the negative self-caused condition, and for anger in the negative other-caused condition.

5.2.2. Gift giving

According to the predictions, the effects of negative emotions, but not of positive emotions, are dependent on the agency of the emotion. Positive emotions (pride and gratitude) and negative self-caused emotions (guilt) would increase gift giving, whereas negative other-caused emotions (anger) would decrease gift giving. Results for Studies 1, 3, and 4 can be found in Table 2. Please note that for Studies 1 to 4 the degrees of freedom for the ANOVAs and the contrast analyses differ. Because the control condition did not have a value on valence or agency (it is neutral on both valence and agency), this condition was only included in the contrast analyses. A 2 (valence: positive vs. negative) $\times$ 2 (agency: self-caused vs. other-caused) ANOVA with gift giving as dependent variable showed two main effects (Fs > 32.30, ps < .01, η^2s > .13) and a two-way interaction ($F(1, 214) = 38.27$, $p < .01$, $\eta^2 = .15$). Participants in the positive self-caused condition ($t(266) = 2.44$, $p = .01$), in the positive other-caused condition ($t(266) = 2.76$, $p < .01$), and in the negative self-caused condition (albeit marginally, $t(266) = 1.87$, $p = .06$) all bought bigger gifts in terms of effort, involvement, and size compared to participants in the control condition. There were no differences across these three conditions (ts < 1). In contrast, participants in the negative other-caused condition bought a smaller gift compared to all other conditions (ts > 7.11, ps < .01).

5.2.3. Money spent on the gift

A 2 (valence: positive vs. negative) $\times$ 2 (agency: self-caused vs. other-caused) ANOVA with money as dependent variable showed two main effects (Fs > 6.79, ps < .01, η^2s > .03) and a two-way interaction

($F(1, 214) = 7.37$, $p < .01$, $\eta^2 = .03$). Participants in the positive self-caused condition ($t(266) = 3.25$, $p < .01$), the positive other-caused condition ($t(266) = 3.26$, $p < .01$), and the negative self-caused condition ($t(266) = 1.94$, $p = .05$) all spent more on the gift compared to control participants. There were no differences across these three conditions (ts < 1.24, ps > .22). In contrast, participants in the negative other-caused condition spent less on the gift compared to all other conditions (ts > 2.04, ps < .04).

5.2.4. Discussion

Study 1 provides first support for the idea that the valence–agency framework can predict emotion effects on gift giving. Whereas positive emotions (pride and gratitude) appeared to increase gift giving, the effects of negative emotions depended on the agency of the emotion. The negative self-caused emotion guilt increased gift giving, but the negative other-caused emotion anger decreased gift giving. However, one might question whether measures such as the type of gift consumers would buy and the amount they would be willing to spend reflect all possible gift giving behaviors. For example, emotions might exert different effects on the decision to buy a gift. Therefore, Study 2 replicates the results of Study 1 with other gift giving measures.

6. Study 2: other gift giving measures

6.1. Method

6.1.1. Participants and design

One hundred seventy-nine international students from a Western European university (70 males, $M_{age} = 21.49$, $SD_{age} = 1.98$) participated in partial fulfillment of a course requirement. They were randomly assigned to the control condition or to one of the conditions of a 2 (valence: positive vs. negative) $\times$ 2 (agency: self-caused vs. other-caused) between subjects design with gift giving decision and type of gift as dependent variables.

6.1.2. Procedure and variables

To induce emotions, participants first completed the autobiographical recall procedure of Study 1. Participants then typed the name of the person who was present in the described event or who was told about

Table 2
Gift giving and gift giving reason means (and standard deviations) as a function of condition in Studies 1, 3, and 4.

		Condition			
Study	Control	Positive		Negative	
		Self-caused	Other-caused	Self-caused	Other-caused
Dependent variable	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Study 1					
Total gift giving	6.08 (1.65)[a]	7.06 (1.82)[b]	7.22 (1.80)[b]	6.87 (1.79)[ab]	3.17 (3.31)[c]
Money	24.54 (20.09)[a]	40.07 (27.89)[b]	40.47 (29.92)[b]	34.20 (28.47)[b]	14.61 (20.37)[c]
Study 3					
Total gift giving	6.09 (2.13)[a]	7.10 (1.39)[b]	7.27 (1.75)[b]	6.97 (1.57)[ab]	3.86 (3.36)[c]
Money	22.60 (21.30)[a]	33.07 (27.89)[b]	53.23 (33.33)[c]	32.29 (24.71)[ab]	13.89 (16.23)[ad]
Study 4					
Total gift giving	5.74 (1.96)[a]	7.38 (1.25)[b]	7.73 (1.25)[b]	6.77 (2.08)[b]	2.68 (2.67)[c]
Money	22.22 (21.63)[a]	43.11 (31.16)[b]	37.89 (29.36)[b]	33.63 (27.43)[ab]	7.76 (8.83)[c]
Time	44.41 (22.67)[a]	62.67 (27.76)[b]	72.54 (24.71)[b]	64.48 (27.89)[b]	23.76 (31.44)[c]
Relationship management	6.24 (1.78)[a]	7.40 (1.44)[b]	6.80 (1.40)[a]	7.10 (1.55)[ab]	3.03 (2.58)[c]
Express feelings	4.74 (2.77)[a]	7.67 (1.37)[b]	6.68 (2.67)[b]	5.33 (3.20)[a]	4.97 (3.18)[a]
Perceived cost	4.28 (1.98)[ab]	4.95 (2.07)[b]	3.55 (2.72)[a]	3.32 (2.01)[a]	3.83 (2.25)[a]

Note. Total gift giving was the mean of effort, involvement, and size scores (ranging from 0 to 10). Money was measured in euros, time in minutes. There are no significant differences between means with the same superscript, with all ts < 1.24, ps > .22. Means with different superscripts differ significantly with all ts > 1.94, ps < .05, and means with letters ab and ad differ marginally significantly from means with letter a with all ts > 1.60, ps < .11.

the event afterwards. To measure gift giving, participants imagined that a week after the described event it was the named person's birthday. As the dependent measures, participants indicated whether they would buy a gift for this person (yes vs. no) and what type of gift they would buy (a present, a gift card, money, or nothing). It was thereby assumed that buying a present would entail greater effort than buying a gift card or giving money (Ruth, Otnes and Brunel, 1999). Finally, participants reread their situation and answered the emotion manipulation check of Study 1.

6.2. Results and discussion

6.2.1. Gift giving decision

According to the hypotheses, the effects of negative emotions, but not of positive emotions, are dependent on the agency of the emotion. Positive emotions (pride and gratitude) and negative self-caused emotions (guilt) are expected to increase gift giving, whereas negative other-caused emotions (anger) are expected to decrease gift giving. Indeed, a chi-square test showed that agency had no influence on gift giving in the positive conditions. Both self-caused (95%) and other-caused emotions (94%) motivated participants to give a gift ($\chi^2 < 1$). In contrast, agency did have an effect on gift giving in the negative conditions. Most participants in the negative self-caused condition intended to give a gift (85%), whereas only 57% of participants in the negative other-caused condition intended to do so (χ^2 (1, $N = 69$) = 6.64, $p = .01$, $\varphi = .31$). Of the control participants, 92% intended to give a gift.

6.2.2. Type of gift

A chi-square test showed that emotions also influenced the type of gift that participants intended to buy. Agency had no effect on the type of gift in the positive conditions. Both self-caused (95%) and other-caused emotions (89%) motivated participants to give a present (χ^2 (2, $N = 73$) = 2.12, $p = .35$). In contrast, agency did influence the type of gift in the negative conditions. Negative self-caused emotions mostly motivated participants to give a present (77%), whereas only 40% of participants in the negative other-caused condition intended to do so (χ^2 (3, $N = 69$) = 9.87, $p = .02$, $\varphi = .38$). 14% of participants in the negative other-caused condition intended to give a gift card, and 43% intended to give nothing. Of the participants in the control condition, 81% intended to give a present.

6.2.3. Discussion

Study 2 replicates the results of Study 1 with different dependent measures. It seems that the valence–agency framework can predict emotion influences on gift giving aspects such as whether a gift is bought and what type of gift is bought. Positive emotions (pride and gratitude) appeared to stimulate gift giving, whereas the effects of negative emotions depended on the agency. The negative self-caused emotion guilt stimulated gift giving, but the negative other-caused emotion anger did not. Even though Studies 1 and 2 find similar results on different dependent measures, there may be some doubts about the generalizability of the findings. Both studies examined the effects of the emotions of pride, gratitude, guilt, and anger. Study 3 excludes the possibility that the emotion effects on gift giving might be based on the specific emotions used in Studies 1 and 2. In this next study, four other emotions, namely satisfaction, love, shame, and fear were induced.

7. Study 3: inducing four other emotions

7.1. Method

7.1.1. Participants and design

Two hundred twenty international students from a Western European university (107 males, $M_{age} = 21.67$, $SD_{age} = 2.45$) participated in partial fulfillment of a course requirement. There were four

participants who did not answer the autobiographical recall induction, resulting in two hundred sixteen participants (104 males, $M_{age} = 21.67$, $SD_{age} = 2.46$). They were randomly assigned to the control condition or to one of the conditions of a 2 (valence: positive vs. negative) × 2 (agency: self-caused vs. other-caused) between subjects design with gift giving and money spent on the gift as dependent variables.

7.1.2. Procedure and variables

Participants completed the same procedure as in Study 1. This time, however, they were asked to report a personal experience in which they felt very satisfied (positive self-caused condition), loved (positive other-caused condition), ashamed (negative self-caused condition), or afraid (negative other-caused condition). Participants indicated what gift they would buy for the named person (a person towards whom the emotion was experienced, who was present in the described event, or who was told about the event afterwards) by answering the gift giving items of Study 1. Participants ended with the emotion manipulation check, on which they indicated how much satisfaction, love, shame, and fear they felt (0 = not at all, 10 = very strongly).

7.2. Results and discussion

7.2.1. Gift giving

Studies 1 and 2 suggested that the positive emotions satisfaction and love and the negative self-caused emotion shame increased gift giving, whereas the negative other-caused emotion fear decreased gift giving. The ANOVA on gift giving showed two main effects ($Fs > 19.79$, $p < .01$, $\eta^2s > .10$) and a two-way interaction ($F(1, 173) = 24.56$, $p < .01$, $\eta^2 = .13$). Both positive emotions and the negative self-caused emotion (albeit marginally for the negative self-caused emotion, $ts > 1.85$, $p < .06$) increased gift giving compared to the control condition. These three conditions did not differ ($ts < 1$). In contrast, negative other-caused emotions decreased gift giving compared to all other conditions ($ts > 4.79$, $ps < .01$).

7.2.2. Money spent on the gift

The ANOVA on money spent on the gift showed only a main effect of valence ($F(1, 173) = 25.41$, $p < .01$, $\eta^2 = .13$; agency $F < 1$) and a two-way interaction ($F(1, 173) = 23.49$, $p < .01$, $\eta^2 = .12$). Participants in the positive self-caused condition (albeit marginally, $t(211) = 1.93$, $p = .06$), in the positive other-caused condition ($t(211) = 5.57$, $p < .01$), and in the negative self-caused condition (albeit marginally, $t(211) = 1.74$, $p = .08$) all spent more on the gift than the control condition. Unexpectedly, participants in the positive other-caused condition also spent more compared to the positive self-caused and the negative self-caused conditions ($ts > 3.73$, $ps < .01$). Participants in the negative other-caused condition instead spent *less* compared to all other emotion conditions ($ts > 3.36$, $ps < .01$), and marginally less compared to the control condition ($t(211) = 1.60$, $p = .11$).

7.2.3. Discussion

Although there are some variations in the findings of Study 3 compared to Studies 1 and 2, the findings indicate that the valence–agency framework can mostly predict the effects of specific emotions such as satisfaction, love, shame, and fear on gift giving. These four emotions showed that both positive emotions and negative self-caused emotions can increase gift giving, and that negative other-caused emotions can decrease gift giving. The framework could not predict the finding that the positive other-caused emotion love stimulated givers to spend more money on gifts compared to the other tested emotions.

According to the hypotheses, emotions tell the giver something about the relationship with the receiver. Both positive emotions and negative self-caused emotions increase gift giving, because they signal that the relationship with the receiver should be maintained (positive emotions) or improved (negative self-caused emotions). In contrast,

negative other-caused emotions decrease gift giving because they signal that the relationship with the receiver should be weakened. Alternatively, emotions could influence gift giving for different reasons. For example, the giver may want to express feelings and use gift giving as a way to do so. Emotions could also influence perceptions of how easy or difficult it is to buy a gift, such that positive emotions and negative self-caused emotions would decrease the perceived costs of gift giving, and negative other-caused emotions would increase the perceived costs of gift giving. To examine why emotions affect gift giving, Study 4 measured the reasons underlying gift giving. The next study directly manipulated valence and agency instead of specific emotions to test whether the effects can indeed be explained by the valence and agency of emotions and not by the specific emotions.

8. Study 4: inducing valence and agency

8.1. Method

8.1.1. Participants and design

After excluding four participants who did not answer the autobiographical recall procedure, one hundred thirty-eight participants (62 males, $M_{age} = 20.96$, $SD_{age} = 2.37$) participated in partial fulfillment of a course requirement. They were randomly assigned to the control condition or one of the conditions of the 2 (valence: positive vs. negative) × 2 (agency: self-caused vs. other-caused) between subjects design with gift giving, money spent on the gift, and time spent on the gift as the dependent variables.

8.1.2. Procedure and variables

Participants first recalled a situation in which they felt very positive due to their own behavior (positive self-caused condition), positive due to the behavior of other people (positive other-caused condition), negative due to their own behavior (negative self-caused condition), or negative due to the behavior of other people (negative other-caused condition). In the control condition, participants recalled a normal weekday. Next participants indicated how much pride, satisfaction (both positive self-caused), gratitude, love (both positive other-caused), guilt, shame (both negative self-caused), fear, and anger (both negative other-caused) they felt in the described situation (0 = not at all, 10 = very strongly).

Participants followed the same procedure as in the other studies and typed in the name of a person towards whom they experienced the feeling (in the other-caused conditions) or someone who was present in the described event (in the self-caused conditions). If there was nobody present, participants thought of someone that they had told about the event afterwards. In all cases, participants typed the name of this person. They then answered the gift giving measures of Study 1, but also indicated how much time they would spend on searching for a gift (in minutes) (item five of Appendix A). Next, participants responded to twelve items about the reasons why they decided to give a gift. These items were specifically developed to measure relationship management, expression of feelings, and perceived costs of gift giving. A factor analysis with Oblimin rotation was conducted to allow the factors to be correlated. This analysis showed a clear three-factor solution (see Appendix B for the items and factor loadings of the Pattern matrix). The first factor, relationship management (Eigenvalue = 4.79), consisted of seven items, explained 40% of the variance, and formed a reliable scale ($\alpha = .91$). The second factor, express feelings (Eigenvalue = 2.30), explained 19% of the variance ($\alpha = .93$). Finally, the third factor, perceived cost (Eigenvalue = 1.65), explained 14% of the variance ($\alpha = .75$). Two negatively formulated items did not load on any factor, and were therefore left out of the analyses. For each item, participants were asked to indicate the extent to which this motivated their choice for the gift (0 = not at all, 10 = very strongly).

8.2. Results and discussion

8.2.1. Gift giving

Both positive valence conditions would increase gift giving, independent of whether the positive feelings were self-caused or other-caused. For the negative valence conditions, self-caused feelings would increase gift giving and other-caused feelings would decrease gift giving. Two main effects ($Fs < 26.25$, $ps < .01$, $\eta^2 s > .20$) and a two-way interaction ($F(1, 107) = 36.77$, $p < .01$, $\eta^2 = .26$) were found. Again, both positive emotions and negative self-caused emotions increased gift giving compared to the control condition ($ts > 1.95$, $ps < .05$) and did not differ from each other ($ts < 1.18$, $ps > 24$). Participants in the positive other-caused condition bought marginally bigger gifts than participants in the negative self-caused condition ($t(133) = 1.85$, $p = .07$). Negative other-caused emotions decreased gift giving compared to all other conditions ($ts > 5.93$, $ps < .01$).

8.2.2. Money spent on the gift

For money spent on the gift two main effects ($Fs > 10.20$, $ps < .01$, $\eta^2 s > .09$) and a two-way interaction ($F(1, 107) = 4.50$, $p = .04$, $\eta^2 = .04$) were found. Both positive emotions and negative self-caused emotions increased gift giving compared to the control condition (albeit marginally for negative self-caused emotions, $ts > 1.68$, $ps < .09$) and did not differ from each other ($ts < 1.40$, $ps > .16$). Negative other-caused emotions decreased gift giving compared to all other conditions ($ts > 2.17$, $ps < .03$).

8.2.3. Time spent searching for a gift

For time spent two main effects ($Fs > 8.36$, $ps < .01$, $\eta^2 s > .07$) and a two-way interaction ($F(1, 107) = 22.48$, $p < .01$, $\eta^2 = .17$) were found. Both positive emotions and negative self-caused emotions increased time spent on the gift compared to the control condition ($ts < 2.47$, $ps < .01$) and did not differ from each other ($ts < 1.35$, $ps > .18$). Participants in the negative other-caused condition spent *less* time searching compared to all other conditions ($ts > 2.85$, $ps < .01$).

8.2.4. Reasons for gift giving

According to the hypotheses, emotions would influence gift giving because they signal that the relationship with the receiver should be maintained (in the case of positive emotions), improved (in the case of negative self-caused emotions), or weakened (in the case of negative other-caused emotions). It was not expected that emotions would influence gift giving because givers wanted to express their feelings, or because emotions influenced perceptions of the costs of gift giving. Indeed, the ANOVA on relationship management showed two main effects ($Fs > 34.47$, $ps < .01$, $\eta^2 s > .24$) and a two-way interaction ($F(1, 107) = 25.00$, $p < .01$, $\eta^2 = .19$). Participants in the positive self-caused condition ($t(133) = 2.34$, $p = .02$) and in the negative self-caused condition (albeit marginally, $t(133) = 1.73$, $p = .09$) were more interested in maintaining their relationship compared to participants in the control condition. There was no difference between the positive other-caused condition and the control condition ($t(133) = 1.13$, $p = 26$). In contrast, participants in the negative other-caused condition were less interested in maintaining their relationship compared to all other conditions ($ts > 6.61$, $ps < .01$).

The effects of emotions on gift giving could not be explained by a motivation to express feelings or by changes in perceived costs of gift giving. The ANOVA on express feelings showed only a main effect of valence ($F(1, 107) = 15.34$, $p < .01$, $\eta^2 = .13$). A similar ANOVA with perceived cost as dependent variable showed a two-way interaction ($F(1, 107) = 4.85$, $p = .03$, $\eta^2 = .04$), but the pattern of results did not reflect the pattern found for emotion effects on gift giving. Participants in the positive self-caused condition found it easier to buy gifts compared to participants in all other conditions ($ts > 1.98$, $p < .06$). There were no other differences across the conditions ($ts < 1.59$, $p > .12$).

Because the hypotheses predict mediation of only one gift giving reason (relationship management) and not of the other two reasons (express feelings and perceived cost), I analyzed the data by means of the PROCESS macro, model 4 (the parallel multiple mediator model). PROCESS uses an OLS regression-based path analytical framework for estimating direct and indirect effects, combined with bootstrap methods to make inferences about the significance of the indirect effects (see Hayes, 2013 for an extensive description). Condition was first recoded into four dummy variables. Mediation analyses for every dummy variable separately demonstrated that relationship management ($bs > 0.59$, $ts > 7.60$, $ps < .01$) and express feelings ($bs > 0.12$, $ts > 2.48$, $ps < .02$) were predictors of gift giving for all conditions, while perceived cost ($bs < 0.09$, $ps > .19$) was not. Supporting the hypotheses, bootstrap confidence intervals for the indirect effects via relationship management ($bs > 0.69$) based on 10,000 bootstrap samples were entirely above zero ($0.18 < 95\%$ CIs < 1.99), while all bootstrap confidence intervals via express feelings included zero ($-0.003 < 95\%$ CIs < 0.46). There was no evidence that emotions influenced gift giving independent of its effect on relationship management ($bs < 0.22$, $ts < 1$) (with the exception of the positive other-caused condition, $b = 1.29$, $p < .01$, and the negative other-caused condition, $b = 1.83$, $p < .01$).

For money spent on the gift, relationship management ($bs > 1.97$, $ts > 1.69$, $ps < .09$) and express feelings ($bs > 1.83$, $ts > 2.29$, $ps < .03$) were predictors, while perceived cost ($bs < 0.58$, $ps > .55$) was not. Bootstrap confidence intervals for the indirect effects via relationship management ($bs > 3.28$) were entirely above zero ($2.05 < 95\%$ CIs < 16.77), while all bootstrap confidence intervals via express feelings included zero ($-5.43 < 95\%$ CIs < 8.44). There was no evidence that emotions influenced money independent of its effect on relationship management ($bs < 7.93$, $ts < 1.31$, $ps > .19$) (with the exception of the negative other-caused condition, $b = 16.48$, $p = .02$). Finally, for time spent on the gift, only relationship management was a predictor ($bs > 4.49$, $ts > 3.54$, $ps < .01$) (express feelings $bs < 1.22$, $ps > .15$; perceived cost $bs < 1.53$, $ps > .14$). Bootstrap confidence intervals for the indirect effects ($bs > 5.61$) were entirely above zero ($1.71 <$ bootstrapped 95% CIs < 28.55), and emotions did not influence time independent of its effect on relationship management ($bs < 5.56$, $ts < 1$).

8.2.5. Discussion

Study 4 replicates the findings of the previous studies with a different emotion manipulation and also sheds some light on why emotions influence gift giving. A direct manipulation of the emotion dimensions of valence and agency again shows that the valence–agency framework can predict emotion effects on gift giving. Moreover, one of the reasons why these effects take place is that emotions

signal whether the relationship with the receiver should be maintained, improved, or weakened. Positive emotions and negative self-caused emotions increase gift giving because they signal that the relationship should be maintained (positive emotions) or improved (negative self-caused emotions). In contrast, negative other-caused emotions decrease gift giving because they signal that the relationship should be weakened. Motivation to express feelings, or changed perceptions in the cost of gift giving could not explain the emotion effects.

Although Studies 1 to 4 provide converging evidence for the valence–agency framework, they do not test the role of other appraisal dimensions such as activity (also called arousal) or certainty. According to emotion literature, emotions can make people feel active (e.g., pride or anger) or inactive (e.g., happiness or sadness) (Frijda, Kuipers and Ter Schure, 1989; Roseman, Wiest and Swartz, 1994). All previously tested emotions are considered to be active emotions. Study 5 included activity as an appraisal dimension to test whether the hypothesized effects also apply to inactive emotions. Emotions can also be based on situations that entail certain outcomes (e.g., sadness, anger, or pride), or on situations in which there is no certainty concerning outcomes yet (e.g., anxiety or hope). All previously tested emotions are certain emotions. Study 6 included certainty as an appraisal dimension to test whether the hypothesized effects also apply to uncertain outcomes.

Moreover, Study 4 demonstrated that relationship management is one reason why emotions influence gift giving. Another possible reason could be mood repair. Givers could change their gift giving in order to make themselves feel better. To test this possible reason, Studies 5 and 6 included mood repair as a reason for gift giving. Finally, it is possible that the emotion effects on gift giving could be explained by differences in the strength of relationships between givers and receivers. To exclude this possibility, relationship strength was included as a covariate in the analyses of Studies 5 and 6.

9. Study 5: inducing inactive emotions

9.1. Method

9.1.1. Participants and design

After excluding seven participants who did not answer the autobiographical recall procedure, two hundred forty-three US citizens (130 males, $M_{age} = 32.64$, $SD_{age} = 11.85$) participated in a study on Amazon Mechanical Turk in exchange for a monetary reward. They were randomly assigned to one of the conditions of the 2 (valence: positive vs. negative) × 2 (agency: self-caused vs. other-caused) × 2 (activity: inactive vs. active) between subjects design with gift giving

Table 3
Emotion manipulation check means (and standard deviations) and tests in Studies 5 and 6.

Study	Condition				
	Positive		Negative		
Study 5					
Condition	Inactive	Active	Inactive	Active	
Target emotion	Happiness	Pride	Sadness	Anger	
M (SD)	8.90 (1.20)	8.40 (1.94)	8.00 (1.76)	8.74 (1.68)	
Compared to other conditions	$ts(266) > 6.04$**	$ts(266) > 4.01$**	$ts(266) > 12.62$**	$ts(266) > 10.60$**	
Compared to other emotions within the same condition	$ts(57) > 6.33$**	$ts(52) > 5.50$**	$ts(48) > 10.32$**	$ts(53) > 8.63$**	
Study 6					
Condition	Uncertain	Certain	Uncertain	Certain	
Target emotion	Hope	Pride	Anxiety	Anger	
M (SD)	5.81 (1.12)	5.83 (1.36)	5.86 (1.31)	6.13 (1.13)	
Compared to other conditions	$ts(238) > 3.84$**	$ts(238) > 2.73$**	$ts(238) > 4.02$**	$ts(238) > 8.82$**	
Compared to other emotions within the same condition	$ts(57) > 3.26$**	$ts(64) > 5.04$**	$ts(62) > 7.43$**	$ts(55) > 6.07$**	

Note. For every condition, the target emotion was compared to the same emotion in the other conditions ("compared to other conditions") and compared to other emotions within the same condition ("compared to other emotions within the same condition").

* $ps < .05$.
** $ps < .01$.

decision, gift giving, money spent on the gift, and time spent on the gift as the dependent variables.

9.1.2. Procedure and variables

According to most emotion theories, inactive emotions do not have a clear agency. In other words, they can be self-caused or other-caused depending on the situation (Frijda, 1986; Frijda, Kuipers and Ter Schure, 1989; Nyer, 1997; Ortony, Clore and Collins, 1988; Reisenzein & Hofmann, 1990, 1993; Richins, 1997; Roseman, Wiest and Swartz, 1994). Therefore, in this study four different emotions were induced (happiness as a positive inactive emotion, pride as a positive active emotion, sadness as a negative inactive emotion, and anger as a negative active emotion). Participants in the self-caused condition recalled a situation in which each of these four emotions was felt due to something they had done to/for another person, and participants in the other-caused condition recalled a situation in which the emotion was felt due to the behavior of other people. Next, participants indicated how much pride, happiness, anger, and sadness they felt in the described situation (1 = not at all, 7 = very strongly). They also indicated to what degree they felt negative and positive (valence), whether they themselves or another person was the cause of the situation (agency), whether they felt inactive or active (activity), whether they felt uncertain or certain (certainty), and whether they felt powerless or powerful (power) (1 = not at all, 7 = very strongly).

Participants typed the name of the person in the described situation. They answered the gift giving dependent measures of Studies 1 and 2, and indicated their reasons for gift giving on the items from Study 4. This time, the reasons also included three items related to mood repair (Appendix B, items 13 to 15, $\alpha = .94$). Finally, to control for the strength of the relationship with the receiver, participants indicated how strong their relationship was with the person in question, how close they were, how satisfied they were with their relationship, and how much they liked the person in question (1 = not at all, 7 = very strong/close/satisfied/much liked) ($\alpha = .93$).

9.2. Results and discussion

9.2.1. Emotion manipulation check

The emotion manipulation in both Studies 5 and 6 worked (see Table 3 for specific emotions). The appraisal dimension items also demonstrated that the agency manipulation was successful. Participants in the self-caused conditions felt that they were the cause of the situation ($M = 5.26, SD = 1.74$) more and other people less ($M = 3.25, SD = 2.06$) than those in the other-caused conditions ($M = 2.58, SD = 1.89$ and $M = 5.50, SD = 1.90, ts > 8.84, ps < .01$). Participants in the positive conditions felt more positive ($M = 6.18, SD = 1.11$) and less negative

($M = 1.50, SD = 1.19$) than those in the negative conditions ($M = 1.56, SD = 1.04$ and $M = 5.66, SD = 1.55, ts(241) > 23.12, ps < .01$).

Unfortunately, participants in the inactive conditions did not report feeling more inactive ($M = 2.35, SD = 1.82$) or less active ($M = 3.88, SD = 2.18$) compared to those in the active conditions ($M = 2.30, SD = 1.72$ and $M = 4.01, SD = 2.06, ts < 1$). Because participants in the inactive conditions did not report feeling more inactive than active either, and because most emotion theories agree that happiness and sadness are inactive emotions and pride and anger are active emotions (Lewis & Haviland-Jones, 2000), participants might not have understood the items "I felt inactive" and "I felt active" as intended. A posttest with 101 US citizens (38 males, $M_{age} = 37.06, SD_{age} = 14.21$) on Amazon Mechanical Turk confirmed this idea. In the posttest, participants were asked to report to what extent they felt inactive, active, low on energy, energized, aroused, stimulated, not motivated to do anything, motivated to do something, and motivated to take action (1 = not at all, 7 = very strongly) when they felt happiness, pride, sadness, and anger (in random order). Results showed that experiences of the inactive emotions of happiness and sadness were reported as being lower on energy ($M = 3.76, SD = 0.86$), less energized ($M = 3.97, SD = 0.89$), less arousing ($M = 3.70, SD = 1.09$), less stimulating ($M = 3.81, SD = 0.92$), less motivating to do something ($M = 3.94, SD = 0.95$), more motivating not to do anything ($M = 3.89, SD = 1.03$), and less motivating to take action ($M = 3.87, SD = 1.07$) compared to experiences of the active emotions of pride and anger ($Ms < 2.09, SDs < 1.03$ for low on energy and motivating not to do anything, $Ms > 4.63, SDs < 1.66$ for the other items, $ts > 5.72, ps < .01$).

9.2.2. Gift giving decision

Studies 1 to 4 demonstrated that the effects of active negative emotions, but not of active positive emotions, depend on the agency. Thus, pride and self-caused anger were expected to increase gift giving, and other-caused anger was expected to decrease gift giving. More importantly, Study 5 tested whether this pattern of results would also be found for the inactive emotions happiness and sadness (see Table 4). Even though the results partially differed across gift giving dependent measures, overall the findings revealed that the same pattern applies to inactive emotions. Chi-square tests with gift giving decision as dependent variable showed that agency had no influence on gift giving for the positive inactive emotion happiness. Both self-caused (97%) and other-caused happiness (93%) motivated participants to give a gift ($\chi^2 < 1$). In contrast, agency did affect gift giving marginally for the negative inactive emotion sadness. Most participants in the self-caused sadness condition intended to give a gift (84%), whereas only 64% of participants in the other-caused sadness condition intended to do so ($\chi^2 (1, N = 67) = 3.39, p = .07, \varphi = .23$). This finding replicated the pattern found for active emotions in Study 2. Unexpectedly, the results for positive active emotions did not replicate the findings of Study 2.

Table 4
Gift giving and gift giving reason means (and standard deviations) as a function of condition in Study 5.

Dependent variable	Inactive conditions				Active conditions			
	Positive (happiness)		Negative (sadness)		Positive (pride)		Negative (anger)	
	Self-caused	Other-caused	Self-caused	Other-caused	Self-caused	Other-caused	Self-caused	Other-caused
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Total gift giving	6.76 (1.89)[a]	6.50 (2.49)[a]	6.43 (3.54)[a]	3.77 (3.64)[b]	6.00 (3.49)[a]	7.33 (1.45)[a]	5.71 (3.27)[a]	2.42 (3.24)[b]
Money	60.33 (45.37)[a]	86.96 (123.19)[a]	99.42 (130.62)[b]	23.61 (26.55)[c]	68.75 (73.49)[a]	59.14 (42.13)[a]	48.03 (41.85)[a]	15.94 (22.59)[c]
Time	96.63 (85.67)[a]	86.96 (66.60)[a]	124.19 (170.45)[a]	47.03 (76.64)[b]	79.17 (81.18)[a]	101.79 (89.84)[a]	86.42 (100.43)[a]	23.59 (34.18)[bc]
Relationship management	7.65 (1.42)[a]	7.73 (2.08)[a]	7.80 (2.74)[a]	5.35 (3.28)[a]	7.19 (2.48)[a]	8.07 (1.79)[a]	7.59 (2.73)[a]	4.00 (3.17)[c]
Express feelings	9.10 (1.95)[a]	8.11 (3.03)[a]	8.37 (3.40)[a]	6.85 (3.62)[ab]	7.44 (3.89)[ab]	9.53 (1.84)[c]	7.88 (3.27)[a]	5.47 (4.00)[d]
Perceived cost	6.23 (2.24)[a]	6.38 (3.15)[a]	6.10 (3.19)[a]	4.87 (3.15)[b]	7.01 (3.14)[a]	7.24 (2.33)[a]	5.34 (2.79)[a]	4.84 (3.71)[a]
Mood repair	4.62 (2.71)[a]	5.15 (3.52)[a]	5.91 (3.27)[a]	5.65 (3.55)[a]	4.82 (3.11)[b]	5.37 (3.22)[a]	6.88 (2.98)[c]	4.41 (3.72)[b]

Note. Total gift giving was the mean of effort, involvement, and size scores (ranging from 0 to 10). Money was measured in dollars, time in minutes. There are no significant differences between means with the same superscript, with all $ts < 1.40, ps > .16$. Means with different superscripts differ significantly with all $ts > 2.04, ps < .05$, and means with letters ab differ marginal significantly from means with letter a, all $ts > 1.67, ps < .09$.

Although the majority of participants in the self-caused pride condition reported they would buy a gift (79%), even more participants would do so in the other-caused pride condition (100%, χ^2 (1, $N = 53$) =6.67, $p = .01$, $\varphi = .36$). The results for anger did replicate the findings of Study 2. Most participants in the self-caused anger condition intended to give a gift (85%), whereas only half of the participants in the other-caused anger condition intended to do so (χ^2 (1, $N = 65$) =9.02, $p < .01$, $\varphi = .37$).

9.2.3. Gift giving

The valence–agency framework can predict the effects of inactive emotions on gift giving. A 2 (valence: positive vs. negative) × 2 (agency: self-caused vs. other-caused) × 2 (activity: inactive vs. active) ANOVA with gift giving as dependent variable and relationship strength added as a covariate variable showed only a main effect of valence ($F(1, 234) = 4.43$, $p = .04$, $\eta^2 = .02$; $Fs < 2.40$, $ps > .12$ for agency and activity), a two-way interaction of valence and agency ($F(1, 234) = 11.47$, $p < .01$, $\eta^2 = .05$), and no three-way interaction ($F < 1$). Agency had no influence on gift giving ($ts < 1.61$, $ps > .11$) for both inactive and active positive emotions. However, it did have an effect on gift giving ($ts > 3.62$, $ps < .01$) for both inactive and active negative emotions. Self-caused sadness and self-caused anger increased gift giving compared to other-caused sadness and other-caused anger.

9.2.4. Money spent on the gift

The valence–agency framework can predict the effects of inactive emotions on money spent on the gift. The ANOVA with money spent on the gift as dependent variable and relationship strength as a covariate showed only a marginal main effect of activity ($F(1, 234) = 2.77$, $p = .09$, $\eta^2 = .01$; for valence and agency $Fs < 2.57$, $ps > .12$), a two-way interaction of valence and agency ($F(1, 234) = 6.49$, $p < .01$, $\eta^2 = .03$), and a three-way interaction ($F(1, 234) = 6.49$, $p < .01$, $\eta^2 = .03$). Separate 2 (valence: positive vs. negative) × 2 (agency: self-caused vs. other-caused) ANOVAs with relationship strength as covariate for inactive and active emotions demonstrated a two-way interaction of valence and agency ($F(1, 120) = 8.31$, $p < .01$, $\eta^2 = .07$) for *inactive* emotions only. While there was no difference between self-caused and other-caused happiness on money spent on the gift ($t(235) = 1.39$, $p = .17$), there was a difference between self-caused and other-caused sadness ($t(235) = 4.24$, $p < .01$). Participants in the self-caused sadness condition intended to spend more on the gift. For active emotions, there was no difference between the self-caused and other-caused pride conditions on money spent on the gift ($t < 1$), and there was a marginal significant difference between the self-caused and other-caused anger conditions ($t(235) = 1.77$, $p = .08$). Participants in the self-caused anger condition intended to spend marginally more on the gift.

9.2.5. Time spent on the gift

The valence–agency framework can predict the effects of inactive emotions on time spent on the gift. The ANOVA with time spent on the gift as dependent variable and relationship strength as a covariate showed only a marginal main effect of agency ($F(1, 234) = 2.90$, $p = .09$, $\eta^2 = .01$; for valence and activity $Fs < 1$), a two-way interaction of valence and agency ($F(1, 234) = 5.16$, $p = .02$, $\eta^2 = .02$), and no three-way interaction ($F < 1$). Agency had no influence on gift giving ($ts < 1$) for both inactive and active positive emotions. However, it did affect gift giving ($ts > 2.65$, $ps < .01$) for both inactive and active negative emotions. Self-caused sadness and self-caused anger increased time spent searching on a gift compared to other-caused sadness and other-caused anger.

9.2.6. Reasons for gift giving

It was expected that emotions would influence gift giving because they signal that the relationship with the receiver should be maintained, improved, or weakened. A second possible reason was that givers would use gift giving to repair their mood. Again, the PROCESS macro model 4 was used to test these hypotheses. Mediation analyses for every of the seven dummy variables separately for every gift giving dependent measure demonstrated that relationship management ($bs > 0.88$, $ts > 14.98$, $ps < .01$ for gift giving, $bs > 8.93$, $ts > 4.62$, $ps < .01$ for money, and $bs > 10.14$, $ts > 5.34$, $ps < .01$ for time) was the only predictor of gift giving for all emotion conditions (mood repair $bs < 2.18$, $ps > .19$). Supporting the hypotheses, bootstrap confidence intervals for the indirect effects via relationship management ($bs > 0.34$) based on 10,000 bootstrap samples were entirely above zero ($0.05 < 95\%$ CIs < 4.01 for gift giving, $0.17 < 95\%$ CIs < 47.98 for money, and $0.53 < 95\%$ CIs < 76.56 for time), while all bootstrap confidence intervals via mood repair included zero ($-9.78 < 95\%$ CIs < 11.93). There was no evidence that emotions influenced gift giving independent of its effect on relationship management ($bs < 1.03$, $ps > .11$ for gift giving, $bs < 21.33$, $ps > .12$ for money, $bs < 19.53$, $ps > .30$ for time) (with the exception of other-caused pride for gift giving, $b = 0.80$, $p = .05$, and self-caused sadness for money, $b = 38.56$, $p < .01$, and for time, $b = 36.16$, $p = .04$).

9.2.7. Discussion

One of the remaining questions was whether the valence–agency framework would also hold for inactive emotions. Overall the findings of Study 5 seem to suggest that the valence–agency framework can also account for the effects of inactive emotions such as happiness and sadness on gift giving. Moreover, the effects of both active and inactive emotions on gift giving seem to be mediated by a motivation to maintain, improve, or weaken the relationship with the receiver, and not by a motivation to feel good (mood repair). Finally, the emotion effects on gift giving were found even when relationship strength was included as a covariate, suggesting that the findings could not be explained by differences in the relationship between the giver and the receiver. Study 6 examines whether the hypothesized framework also applies to uncertain emotions such as hope and anxiety.

10. Study 6: inducing uncertain emotions

10.1. Method

10.1.1. Participants and design

After excluding eight participants that did not answer the autobiographical recall procedure, two hundred forty-two US citizens (147 males, $M_{age} = 31.33$, $SD_{age} = 11.05$) participated in a study on Amazon Mechanical Turk in exchange for a monetary reward. They were randomly assigned to one of the conditions of the 2 (valence: positive vs. negative) × 2 (agency: self-caused vs. other-caused) × 2 (certainty: uncertain vs. certain) between subjects design with gift giving decision, gift giving, money spent on the gift, and time spent on the gift as the dependent variables.

10.1.2. Procedure and materials

This study had the same design as Study 5. This time, however, two uncertain emotions (hope and anxiety) and two certain emotions (pride and anger) were induced. To manipulate agency, for every emotion participants in the self-caused condition remembered a situation in which the emotion was felt due to something they had done to/for another person, and participants in the other-caused condition remembered a situation in which the emotion was felt due to the behavior of other people. Participants continued with the manipulation check, gift giving dependent measures, and reasons for gift giving from Study 5.

Table 5
Gift giving and gift giving reason means (and standard deviations) as a function of condition in Study 6.

Dependent variable	Uncertain conditions				Certain conditions			
	Positive (hope)		Negative (anxiety)		Positive (pride)		Negative (anger)	
	Self-caused	Other-caused	Self-caused	Other-caused	Self-caused	Other-caused	Self-caused	Other-caused
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Total gift giving	6.51 (2.45)[a]	6.96 (2.19)[a]	5.94 (3.68)[a]	4.47 (3.43)[b]	6.42 (2.49)[a]	7.08 (1.62)[a]	6.59 (2.87)[a]	3.09 (3.59)[b]
Money	62.14 (40.70)[a]	85.23 (91.09)[a]	80.16 (89.47)[a]	30.81 (38.86)[b]	57.94 (56.43)[a]	52.19 (39.33)[ab]	70.00 (73.80)[a]	29.17 (43.29)[b]
Time	80.32 (102.87)[a]	113.17 (87.32)[a]	84.06 (102.30)[a]	50.16 (47.78)[ab]	82.06 (87.09)[a]	84.69 (56.38)[a]	72.69 (67.50)[a]	27.97 (40.02)[b]
Relationship management	7.26 (1.93)[a]	7.39 (2.19)[a]	7.30 (3.05)[a]	5.64 (2.81)[b]	8.18 (2.22)[a]	7.91 (1.76)[a]	7.49 (2.82)[a]	4.40 (2.97)[b]
Express feelings	8.45 (2.29)[a]	8.57 (3.19)[a]	7.52 (3.85)[a]	7.47 (3.65)[ab]	8.97 (2.38)[a]	9.50 (1.56)[a]	8.56 (3.11)[a]	6.18 (4.23)[b]
Perceived cost	7.50 (2.07)[a]	5.56 (2.83)[b]	5.63 (3.41)[b]	4.47 (2.84)[b]	7.10 (2.44)[a]	7.16 (2.40)[a]	6.36 (2.90)[a]	4.83 (3.42)[b]
Mood repair	5.04 (2.76)[a]	4.66 (2.75)[a]	5.63 (3.57)[a]	4.02 (2.96)[ab]	5.48 (2.75)[a]	5.65 (3.12)[a]	6.88 (2.86)[c]	4.10 (3.56)[ab]

Note. Total gift giving was the mean of effort, involvement, and size scores (ranging from 0 to 10). Money was measured in dollars, time in minutes. There are no significant differences between means with the same superscript, with all $ts < 1.62$, $ps > .11$. Means with different superscripts differ significantly with all $ts > 2.02$, $ps < .05$, and means with letters ab differ marginal significantly from means with letter a, all $ts > 1.74$, $ps < .08$.

10.2. Results and discussion

10.2.1. Emotion manipulation check

Participants in the self-caused conditions felt themselves to be the cause of the situation ($M = 4.13$, $SD = 2.29$) more and other people less ($M = 3.74$, $SD = 2.18$) than the other-caused conditions ($M = 2.15$, $SD = 1.60$ and $M = 5.39$, $SD = 1.97$, $ts > 6.18$, $ps < .01$). Participants in the positive conditions felt more positive ($M = 5.72$, $SD = 1.36$) and less negative ($M = 1.73$, $SD = 1.15$) than the negative conditions ($M = 1.96$, $SD = 1.52$ and $M = 5.39$, $SD = 1.72$, $ts > 19.49$, $ps < .01$). Finally, participants in the uncertain conditions felt more uncertain ($M = 4.13$, $SD = 1.99$) and less certain ($M = 3.14$, $SD = 1.92$) than the certain conditions ($M = 2.80$, $SD = 1.99$ and $M = 3.90$, $SD = 2.06$, $ts > 2.97$, $ps < .01$).

10.2.2. Gift giving decision

The certain emotions pride and self-caused anger were expected to increase gift giving, whereas other-caused anger was expected to decrease gift giving. Study 6 examined whether this pattern of results would also be found for the uncertain emotions hope and anxiety (for results see Table 5). The valence–agency framework did not explain the effects of uncertain emotions for gift giving decision. Agency had no influence on gift giving for the positive uncertain emotion hope. Both self-caused (100%) and other-caused hope (97%) motivated participants to give a gift ($\chi^2 < 1$). Agency also did not have an influence on gift giving for the negative uncertain emotion anxiety. The majority of participants in both the self-caused anxiety condition (75%) and the other-caused anxiety condition (71%) intended to give a gift ($\chi^2 < 1$). The valence–agency framework did explain the effects of certain emotions for gift giving decision. Both self-caused (94%) and other-caused pride (100%) motivated participants to buy a gift (χ^2 (1, $N = 65$) = 2.00, $p = .16$). In addition, most participants in the self-caused anger condition intended to give a gift (89%), whereas only 60% of the participants in the other-caused anger condition intended to do so (χ^2 (1, $N = 56$) = 5.75, $p = .02$, $\varphi = .32$).

10.2.3. Gift giving

The valence–agency framework can predict the effects of uncertain emotions on gift giving. A 2 (valence: positive vs. negative) × 2 (agency: self-caused vs. other-caused) × 2 (certainty: uncertain vs. certain) ANOVA with gift giving as dependent variable showed main effects of valence and agency ($Fs > 6.79$, $ps < .01$, $\eta^2s > .03$; certainty $F < 1$), a two-way interaction of valence and agency ($F(1, 234) = 16.76$, $p < .01$, $\eta^2 = .07$), and no three-way interaction ($F(1, 234) = 2.31$, $p = .13$). Agency had no influence on gift giving ($ts < 1$) for both uncertain and certain positive emotions, whereas it did affect gift giving ($ts >$ 2.02, $ps < .05$) for both uncertain and certain negative emotions. Self-caused anxiety and self-caused anger increased gift giving compared to other-caused anxiety and other-caused anger.

10.2.4. Money spent on the gift

The valence–agency framework can predict the effects of uncertain emotions on money spent on the gift. The ANOVA on money spent on the gift showed a main effect of agency ($F(1, 234) = 5.07$, $p = .03$, $\eta^2 = .02$; for valence and certainty $Fs < 2.16$, $ps > .13$), a two-way interaction of valence and agency ($F(1, 234) = 11.09$, $p < .01$, $\eta^2 = .05$), and no three-way interaction ($F(1, 234) = 1.34$, $p = .25$). Agency had no influence on money spent on the gift ($ts < 1.41$, $ps > .16$) for both uncertain and certain positive emotions, but did affect money spent on the gift ($ts > 2.43$, $ps < .02$) for both uncertain and certain negative emotions. Self-caused anxiety and self-caused anger increased the money spent on the gift compared to other-caused anxiety and other-caused anger.

10.2.5. Time spent on the gift

Finally, it seems that the valence–agency framework can also predict the effects of uncertain emotions on time spent on the gift. The ANOVA on time spent on the gift showed a main effect of valence ($F(1, 234) = 9.90$, $p < .01$, $\eta^2 = .04$; agency and certainty $Fs < 2.29$, $ps > .13$), a two-way interaction of valence and agency ($F(1, 234) = 8.20$, $p < .01$, $\eta^2 = .03$), and no three-way interaction ($F < 1$). Agency had no influence on time ($ts < 1.62$, $ps > .11$) for both uncertain and certain positive emotions, but did affect gift giving (albeit marginally for anxiety, $t(235) = 1.74$, $p = .07$; for anger $t(235) = 2.16$, $p = .03$) for both uncertain and certain negative emotions. Self-caused anxiety and self-caused anger increased time spent on the gift compared to other-caused anxiety and other-caused anger.

10.2.6. Reasons for gift giving

Relationship management was thought to be the mediating factor for uncertain emotions. Mediation analyses for every of the seven dummy variables separately for every gift giving dependent measure demonstrated that relationship management ($bs > 0.87$, $ps < .01$) was the only predictor of gift giving for almost all conditions. With the exception of other-caused hope, bootstrap confidence intervals for the indirect effects via relationship management ($bs > 0.31$) based on 10,000 bootstrap samples were entirely above zero ($0.30 < 95\%$ CIs <27.96). Other-caused hope was the only emotion that influenced gift giving independent of its effect on relationship management ($bs > .75$, $ps < .06$ for other-caused hope, $bs < 33.73$, $ps > .14$ for the other emotions).

10.2.7. Discussion

In general, the findings seem to suggest that the valence–agency framework can also predict the effects of uncertain emotions on gift giving. It appears that uncertain positive emotions increase gift giving, independent of their agency. With the exception of the decision to buy a gift, the influence of uncertain negative emotions on gift giving depend on their agency. When excluding other-caused hope, the motivation to maintain, improve, or weaken the relationship with the receiver also seems to be the mediator for uncertain emotions.

11. General discussion

A large number of studies have taught us that consumers can experience a plethora of emotions, each with their own influences on consumer behaviors. Yet, the huge amount of detailed knowledge on consumer emotions makes it difficult to understand the role of emotions in consumer behaviors, let alone to manage such emotion influences. The current research shows that many emotion influences on for example gift giving can be predicted with just a limited number of emotion appraisal dimensions. Together, the negativity or positivity of an emotion (valence) and the cause of the emotion (agency) can explain how various specific emotions influence gift giving. Six studies with different emotions and different gift giving measures indeed revealed that positive emotions and negative self-caused emotions increase gift giving, and that negative other-caused emotions decrease gift giving. Furthermore, a meta-analysis of the effects observed across all six studies, using Winer's (1971) method of pooling *t*'s, validated the results. For positive emotions, the effect of agency was not significant ($z = 1.84$, $p = .08$, $-0.005 < 95\%$ CI < 0.16) and the effect size was small ($r = .08$). On the contrary, for negative emotions the effect of agency was significant ($z = 9.53$, $p < .01$, $0.32 < 95\%$ CI < 0.48) and the effect size was large ($r = .41$). These findings attest to the robustness of the valence–agency framework. Such a parsimonious framework can help both academics and practitioners in managing consumer emotions.

11.1. Theoretical and practical contributions

The present findings constitute an important contribution to gift giving research. Gift giving scholars agree that emotions play a central role in all stages of the gift giving process. Many theories mention different specific emotions that might evoke gift giving, and some studies have provided empirical support for the idea that emotions might influence gift giving. However, there has been very little empirical research that has studied how at least two different emotions might influence gift giving, or that has provided a theoretical framework for how different emotions can influence gift giving. The current research addressed this issue by presenting a series of empirical studies on the role of givers' emotions in gift giving, and by providing a parsimonious framework that captures most emotion effects on gift giving.

The current findings also provide some new insights into the reasons underlying emotion effects on gift giving. According to most gift giving research, emotions can affect gift giving because givers want to express their feelings. These include the expression of emotional states such as love, joy, penitence, sadness, and gratitude (Cheal, 1988; Fischer & Arnold, 1990; Ruth, 1996; Sherry, 1983), or the communication of feelings of love, affection, care, pride, esteem, and friendship to receivers (Belk & Coon, 1993; Goodwin, Smith and Spiggle, 1990; Komter & Vollebergh, 1997; Otnes, Ruth and Milbourne, 1994; Ruth, 1996; Smith & Ellsworth, 1985; Wolfinbarger & Yale, 1993). Indeed, the results of Studies 4 to 6 demonstrate that givers may feel a need to express their feelings when they experience positive emotions. Yet, givers experiencing negative emotions did not indicate a need to express their feelings, and the reported need to express feelings did not predict the found emotion effects on gift giving. Instead, a motivation to maintain, improve, or weaken relationships with receivers seemed to be able to predict the emotion pattern in gift giving. We now know that when givers experience a positive emotion or a negative self-caused emotion, they feel a need to maintain or improve their relationships with receivers, and consequently increase gift giving, and when givers feel a negative other-caused emotion, they will want to weaken their relationships with receivers, and consequently decrease gift giving. These findings seem to suggest that emotions influence consumer behaviors because consumers want to maintain or change their relationships with others.

On a more general level, I believe that the study of emotion appraisal dimensions in gift giving can also broaden the view on how emotions influence consumer behaviors. Until recently, most research concerning emotion effects on consumer behaviors focused on the effects of one or more specific emotions (Aaker, Drolet, & Griffin, 2008; Griskevicius, Shiota and Nowlis, 2010). This research may have provided a detailed picture of how specific emotions influence consumer behaviors, but it does not give a clear picture of how emotions play a role. For example, a florist trying to sell flowers as a gift by using emotion appeals will not realize that he could increase sales by gratitude-inducing sales messages. A more parsimonious model, which distinguishes emotions on the basis of a limited number of appraisal dimensions, such as the one applied in the current research, may provide a solution. In that case, the florist would realize that not only gratitude but also multiple different positive emotions or negative self-caused emotions could increase sales. Thus, the understanding of consumer behavior would benefit from research that identifies which appraisal dimensions are best in distinguishing different emotions, capturing the interplay among those emotions, and predicting behavioral outcomes.

The present findings have multiple managerial implications. In interpersonal selling, the current framework might be used by retailers to induce emotions that fit with their sales targets. Similar to the emotion induction methods used in the current studies, retailers might ask consumers if they ever encountered events caused by themselves or by other people that made them feel positive or negative. Recalling such events might influence consumers' decisions whether to buy a gift, what gift to select from a gift registry, or how much money to spend. Such tactics could also be applied in direct marketing using social media. On a more general level, the current framework could help to identify what types of emotions can be used for emotion appeals in advertisements that focus on gift giving. The findings suggest that especially advertisements aiming for positive emotions or negative self-caused emotions, independent of what the specific emotion might be, should be effective in stimulating gift giving. Moreover, the findings seem to indicate that promotional activities for gift giving should mostly be held at locations that generate positive emotions or negative self-caused emotions. For instance, amusement parks, zoos, sex shops, catholic churches, and conferences aimed at specific topics such as emotions or interpersonal relationships seem to be good locations for generating gift giving. Finally, sales promotions aimed at gifts could be more effective when using the current valence–agency framework. Slogans such as "did someone make you feel positive?" printed on coupons might help.

11.2. Limitations and future research

Four observations can be made concerning the current studies. First, the effects of negative self-caused emotions on gift giving do not seem to be as strong as the effects of other emotions on gift

giving. In all studies, positive emotions and the negative other-caused emotions exerted significantly different effects on gift giving compared to situations in which givers did not experience any emotion. Yet, in most studies the effects of negative self-caused emotions on gift giving were marginally different from situations in which givers did not experience any emotion. This might suggest that negative self-caused emotions do not always have a positive effect on gift giving. It is important to note that, although the comparisons with neutral conditions mostly revealed marginally significant results, the positive effects of the negative self-caused emotions on gift giving did not differ from the positive effects of the positive emotions. Thus, negative self-caused emotions and positive emotions have similar effects on gift giving. Similarly, the effects of negative self-caused emotions on gift giving did differ from negative other-caused emotions, suggesting that agency is relevant when predicting the effects of negative emotions on consumer behaviors such as gift giving. Future research is needed to further demonstrate the effects of negative self-caused emotions on consumer behaviors.

Second, there are some variations in the findings across studies. Most notably are the findings that the positive other-caused emotion love increased gift giving more than all other emotions in Study 3, and that other-caused anxiety did not influence the decision to buy a gift in Study 6. Interestingly, both love and anxiety have been mentioned by multiple scholars to play a significant role in gift giving (Belk & Coon, 1993; Goodwin, Smith and Spiggle, 1990; Komter & Vollebergh, 1997; Wolfinbarger & Yale, 1993; Wooten, 2000). Together with the current findings, this might indicate that love and anxiety are two emotions that exert additional influences on gift giving above and beyond their valence and agency. Moreover, these variations were not predicted on the basis of the valence–agency framework. This implies that even though the valence–agency framework is largely sufficient to predict emotion effects on gift giving and is useful because it is parsimonious, it may at times fail to capture some specific emotion effects. These findings can serve as the seeds of future research that examines the applications and boundaries of the valence–agency framework.

Third, readers may question the importance of the motivation to improve one's social relationships that is generated by negative self-caused emotions. Negative self-caused emotions such as shame, guilt, and regret may arise in situations where there are no other people present. When this was the case, participants of Studies 1 to 4 instead thought of people who were told about the event afterwards. One may wonder whether givers would experience a need to improve anything about relationships with such "told-about-event-afterwards" receivers. After all, these receivers had not been hurt in any way. Yet, one may argue that the perceived need to improve our social relationships, particularly for impression management reasons, may also occur when we give negative information about us to others. For example, the relationship management scale included aspects such as "I care about what the receiver thinks of me" and "I want to make a good impression on the receiver". Indeed, a comparison between participants that reported on a receiver being present or not present during the negative self-caused emotion-inducing event on relationship management in Study 4 did not show a significant difference ($t(25) = 0.37$, $p = .72$). Future research is needed to further study the role of relationship improvement motivations in negative self-caused emotions when receivers are not related to the emotion-inducing events.

The fourth and final observation concerns the generalizability of the findings. Gift giving is an often-occurring consumer behavior in which millions of dollars are spent each year (Belk, 1976; Cheal, 1988; Mauss, 1925; Otnes & Beltramini, 1996; Ruth, Otnes and Brunel, 1999). Some researchers consider gift giving to be a form of prosocial behavior or to describe a range of different behaviors

that could also be labeled charitable giving (Belk, 1979; Fischer, Gainer and Arnold, 1996; Sherry, 1983). For example, gift giving can be related to blood and organ donations, governmental foreign aid, church relief work, contributions to charitable causes, and community service. Following this line of reasoning, the current findings could apply to all of these behaviors. However, gift giving also has its specific elements. For example, it can be perceived as a form of social or interpersonal consumer behavior in which the consumer interacts with one other person. The findings of Studies 4 to 6 indeed indicate that the emotion effects on gift giving occur because the giver wants to maintain, improve, or weaken the relationship with the receiver. This might limit the generalizability of the current findings to other social consumer behaviors. Future research could examine whether the emotion effects demonstrated in the current research hold for other consumer behaviors such as donations to charity, environmentally friendly behaviors, and social media use.

11.3. Conclusion

Decades ago, emotion scholars demonstrated that the effects of positive emotions on consumer behaviors could be different from the effects of negative emotions on the same behaviors. Since then, we have acquired significantly more knowledge on emotions and their effects. We currently know that there are at least twenty different consumer emotions, each with their own causes and consequences. Yet, this huge amount of data can prevent us from identifying the essential aspects of emotion influences on consumer behaviors. The current findings provide a solution, and demonstrate that most emotion influences on at least some consumer behaviors can be understood by taking into account only a limited number of appraisal dimensions. After decades of detailed studies, it thus appears that it may be time to take a step back and focus on the basics of emotions, namely emotion appraisal dimensions.

Appendix A. Items and factor loadings of the gift giving measures

Item	Study 1	Study 3	Study 4
Gift giving measure			
How much money would you spend on X's birthday?	.61	–	–
How much effort would you put into finding a gift for X?	.91	.93	.95
How personal would the gift be that you would buy for X?	.91	.91	.88
How big would the gift be that you would buy for X?	.87	.75	.88
How much time would you spend searching for a gift for X?	–	–	–
Reliability (α)	.36/.92	.89	.93

Note. With the exception of money (item 1) and time (item 5), items were answered using 10-point scales with end points labeled 0 (no effort/not personal at all/very small) and 10 (much effort/very personal/very big). Money was measured in euros in Studies 1 to 4 and in dollars in Studies 5 and 6, and time was measured in minutes.

Appendix B. Items and factor loadings of the reasons for gift giving in Studies 4 to 6

Item	Relationship management	Express feelings	Perceived cost	Mood repair
Wanted to make [receiver] feel better	**.70**	.05	.06	
Wanted to have a good interaction with [receiver]	**.81**	.04	.07	
Wanted to receive a positive reaction from [receiver]	**.93**	.07	.00	
Care about what [receiver] thinks of me	**.81**	.17	.00	
Wanted to make a good impression on [receiver]	**.83**	.05	.02	
Wanted to look good	**.60**	.01	.25	
Don't care about what [receiver] thinks of me (recoded)	**.74**	.11	.10	
Wanted to show [receiver] how I feel	.01	**.95**	.01	
Wanted to express my feelings	.03	**.90**	.01	
Buying a gift for [receiver] was very easy	.25	.08	**.67**	
Buying a gift for others is easy	.03	.04	**.71**	
Buying a gift for [receiver] required little effort	.19	.01	**.77**	
Wanted to make me feel better				**.87**
Wanted to feel good again				**.95**
Wanted to improve my mood				**.92**
Reliability (α)	.91	.93	.75	.94

Note. Items were complements to the sentence "I chose to give this gift because I…" and were answered using 10-point scales with end points labeled 0 (not at all) and 10 (very strongly). In the participant's text the [receiver] was the name of the mentioned person. Items 13 to 15 were only included in Studies 5 and 6, and the factor loadings of these items concern the factor analysis run in Study 5.
Factor loadings printed bold reflect the items included in the calculation of the factor.

References

Aaker, J., Drolet, A., & Griffin, D. (2008). Recalling mixed emotions. *Journal of Consumer Research, 35*, 543–557.

Algoe, S. B., Haidt, J., & Gable, S. L. (2008). Beyond reciprocity: Gratitude and relationships in everyday life. *Emotion, 8*, 425–429.

Bagozzi, R. P., Gopinath, M., & Nyer, P. U. (1999). The role of emotions in marketing. *Journal of the Academy of Marketing Science, 27*, 184–206.

Banks, S. K. (1979). Gift-giving: A review and an interactive paradigm. In W. L. Wilkie, & A. Abor (Eds.), *Advances in consumer research, Vol. 6.* (pp. 319–324). Association for Consumer Research.

Batson, C. D. (1998). Altruism and prosocial behavior. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology, Vol. 2.* (pp. 282–316). Boston, MA: McGraw-Hill.

Belk, R. W. (1976). It's the thought that counts: A signed diagraph of gift giving. *Journal of Consumer Research, 3*, 155–162.

Belk, R. W. (1979). Gift-giving behavior. In J. N. Sheth (Ed.), *Research in marketing, Vol. 2.* (pp. 95–126). Greenwich, CT: JAI.

Belk, R. W. (1982). Effects of gift-giving involvement on gift selection strategies. In A. Mitchell, & A. Abor (Eds.), *Advances in consumer research, Vol. 9.* (pp. 408–412). Association for Consumer Research.

Belk, R. W. (1996). The perfect gift. In C. Otnes, & R. F. Beltramini (Eds.), *Gift giving: A research anthology* (pp. 59–84). Bowling Green: Bowling Green State University Popular Press.

Belk, R. W., & Coon, G. S. (1993). Gift given as agapic love: An alternative to the exchange paradigm based on dating experiences. *Journal of Consumer Research, 20*, 393–417.

Butt, A. N., & Choi, J. N. (2006). The effects of cognitive appraisal end emotion on social motive and negotiation behavior: The critical role of agency of negotiator emotion. *Human Performance, 19*, 305–325.

Butt, A. N., Choi, J. N., & Jaeger, A.M. (2005). The effects of self-emotion, counterpart emotion, and counterpart behavior on negotiator behaviour: A comparison of individual-level and dyad-level dynamics. *Journal of Organizational Behavior, 26*, 681–704.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review, 97*, 19–35.

Chaudhuri, A. (1997). Consumption emotion and perceived risk: A macro-analytic approach. *Journal of Business Research, 39*, 81–92.

Cheal, D. (1988). *The gift economy.* London: Routledge.

Choshen-Hillel, S., & Yaniv, I. (2011). Agency and the construction of social preference: Between inequality aversion and prosocial behavior. *Journal of Personality and Social Psychology, 101*, 1253–1261.

De Hooge, I. E., Breugelmans, S. M., & Zeelenberg, M. (2008). Not so ugly after all: When shame acts as a commitment device. *Journal of Personality and Social Psychology, 95*, 933–943.

De Hooge, I. E., Zeelenberg, M., & Breugelmans, S. M. (2007). Moral sentiments and cooperation: Differential influences of shame and guilt. *Cognition and Emotion, 21*, 1025–1042.

Derryberry, D., & Tucker, D.M. (1994). Motivating the focus of attention. In P.M. Niedenthal, & S. Kitayama (Eds.), *The hearts' eye: Emotional influences in perception and attention* (pp. 167–196). San Diego, CA: Academic Press.

DeSteno, D., Bartlett, M. Y., Baumann, J., Williams, L. A., & Dickens, L. (2010). Gratitude as a moral sentiment: Emotion-guided cooperation in economic exchange. *Emotion, 10*, 289–293.

Fischer, E., & Arnold, S. J. (1990). More than a labor of love: Gender roles and Christmas gift shopping. *Journal of Consumer Research, 17*, 333–345.

Fischer, E., Gainer, B., & Arnold, S. J. (1996). Gift giving and charitable donating: How (dis)similar are they? In C. Otnes, & R. F. Beltramini (Eds.), *Gift giving: A research anthology* (pp. 175–184). Bowling Green: Bowling Green State University Popular Press.

Flynn, F. J., & Adams, G. S. (2009). Money can't buy love: Asymmetric beliefs about gift price and feelings. *Journal of Experimental Social Psychology, 45*, 404–409.

Fredrickson, B.L. (1998). What good are positive emotions? *Review of General Psychology, 2*, 300–319.

Fredrickson, B.L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*, 218–226.

Fredrickson, B.L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought–action repertoires. *Cognition and Emotion, 19*, 313–332.

Fredrickson, B.L., & Cohn, M.A. (2008). Positive emotions. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 777–796) (Third Edition). New York: The Guilford Press.

Frijda, N. H. (1986). *The emotions.* New York: Cambridge University Press.

Frijda, N. H., Kuipers, P., & Ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology, 57*, 212–228.

Gino, F., & Schweitzer, M. E. (2008). blinded by anger or feeling the love: How emotions influence advice taking. *Journal of Applied Psychology, 93*(5), 1165–1173.

Goodwin, C., Smith, K. L., & Spiggle, S. (1990). Gift-giving: Consumer motivation and the gift purchase process. In M. E. Goldberg, G. Gorn, & R. W. Pollay (Eds.), *Advances in consumer research, Vol. 17.* (pp. 690–698). Provo, UT: Association for Consumer Research.

Grappi, S., Romani, S., & Bagozzi, R. P. (2013). Consumer response to corporate irresponsible behavior: Moral emotions and virtues. *Journal of Business Research, 66*, 1814–1821.

Griskevicius, V., Shiota, M. N., & Nowlis, S. M. (2010). The many shades of rose-colored glasses: An evolutionary approach to the influence of different positive emotions. *Journal of Consumer Research, 37*, 238–250.

Harth, N. S., Kessler, T., & Leach, C. W. (2008). Advantaged group's emotional reactions to intergroup inequality: The dynamics of pride, guilt, and sympathy. *Personality and Social Psychology Bulletin, 34*, 115–129.

Haselhuhm, M. P., & Mellers, B.A. (2005). Emotions and cooperation in economic games. *Cognitive Brain Research, 23*, 24–33.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis.* New York: The Guilford Press.

Hertel, G., Neuhof, J., Theuer, T., & Kerr, N. L. (2000). Mood effects on cooperation in small groups: Does positive mood simply lead to more cooperation? *Cognition and Emotion, 14*, 441–472.

Homans, G. C. (1961). *Social behavior: Its elementary forms.* New York: Harcourt, Brace, & World.

Isen, A.M., & Levin, P. F. (1972). Effect of feeling good on helping: Cookies and kindness. *Journal of Personality and Social Psychology, 21*, 384–388.

Katz, J. M. (1976). How do you love me? Let me count the ways (the phenomenology of being loved). *Sociological Inquiry, 46*, 17–22.

Kelley, S. W., & Hoffman, K. D. (1997). An investigation of positive affect, prosocial behaviors, and service quality. *Journal of Retailing, 73*, 407–427.

Ketelaar, T., & Au, W. T. (2003). The effects of guilt on the behaviour of uncooperative individuals in repeated social bargaining games: An affect-as-information interpretation of the role of emotion in social interaction. *Cognition and Emotion, 17*, 429–453.

Komter, A., & Vollebergh, W. (1997). Gift giving and the emotional significance of family and friends. *Journal of Marriage and Family, 59*, 747–757.

Lazarus, R. S. (1991). *Emotion and adaptation.* New York: Oxford University Press.

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition and Emotion, 14*, 473–493.

Lewis, M., & Haviland-Jones, J. M. (2000). *Handbook of emotions* (Second ed.). New York: The Guilford Press.

Mauss, M. (1925). *The gift: Form and functions of exchange in archaic societies.* New York: Norton.

McCullough, M. E., Kimeldorf, M. B., & Cohen, A.D. (2008). An adaptation for altruism? The social causes, social effects, and social evolution of gratitude. *Current Directions in Psychological Science, 17*(4), 281–285.

Mick, D.G., & Faure, C. (1998). Consumer self-gifts in achievement contexts: The role of outcomes, attributions, emotions, and deservingness. *International Journal of Research in Marketing, 15*, 293–307.

Moll, J., de Oliveira-Souza, R., Garrido, G. J., Bramati, I. E., Caparelli-Daquer, E. M.A., Paiva, M. L. M. F., et al. (2007). The self as a moral agent: Linking the neural bases of social agency and moral sensitivity. *Social Neuroscience, 2*, 336–352.

Nyer, P. U. (1997). A study of the relationships between cognitive appraisals and consumption emotions. *Journal of the Academy of Marketing Science, 25*, 296–304.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions.* New York: Cambridge University Press.

Otnes, C. C., & Beltramini, R. F. (1996). *Gift giving: A research anthology.* Bowling Green: Bowling Green State University Popular Press.

Otnes, C. C., Ruth, J. A., & Milbourne, C. C. (1994). The pleasure and pain of being close: Men's mixed feelings about participation in valentine's day gift exchange. In C. T. Allen, & D. R. John (Eds.), *Advances in consumer research, Vol. 21.* (pp. 159–164). Provo, UT: Association for Consumer Research.

Raghunathan, R., & Pham, M. T. (1999). All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational Behavior and Human Decision Processes, 79,* 56–77.

Reisenzein, R., & Hofmann, T. (1990). An investigation of dimensions of cognitive appraisal in emotion using the repertory grid technique. *Motivation and Emotion, 14,* 1–26.

Reisenzein, R., & Hofmann, T. (1993). Discriminating emotions from appraisal–relevant situational information base-line data for structural models of cognitive appraisals. *Cognition and Emotion, 7,* 271–293.

Richins, M. L. (1997). Measuring emotions in the consumption experience. *Journal of Consumer Research, 24,* 127–146.

Roseman, I. J., Antoniou, A. A., & Jose, P. E. (1996). Appraisal determinants of emotions: Constructing a more accurate and comprehensive theory. *Cognition and Emotion, 10,* 241–277.

Roseman, I. J., Wiest, C., & Swartz, T. S. (1994). Phenomenology, behaviors, and goals differentiate discrete emotions. *Journal of Personality and Social Psychology, 67,* 206–221.

Ruffle, B. J. (1999). Gift giving with emotions. *Journal of Economic Behavior and Organization, 39,* 399–420.

Ruth, J. A. (1996). It's the feeling that counts: Toward an understanding of emotion and its influence on gift-exchange processes. In C. Otnes, & R. F. Beltramini (Eds.), *Gift giving: A research anthology* (pp. 195–214). Bowling Green: Bowling Green State University Popular Press.

Ruth, J. A., Brunel, F. F., & Otnes, C. C. (2002). Linking thoughts to feelings: Investigating cognitive appraisals and consumption emotions in a mixed-emotions context. *Journal of the Academy of Marketing Science, 30,* 44–58.

Ruth, J. A., Brunel, F. F., & Otnes, C. C. (2004). An investigation of the power of emotions in relationship realignment: The gift recipient's perspective. *Psychology and Marketing, 21*(1), 29–52.

Ruth, J. A., Otnes, C. C., & Brunel, F. F. (1999). Gift receipt and the reformulation of interpersonal relationships. *Journal of Consumer Research, 25,* 385–401.

Sanford, K., & Rowatt, W. C. (2004). When is negative emotion positive for relationships? An investigation of married couples and roommates. *Personal Relationships, 11,* 329–354.

Schwartz, B. (1967). The social psychology of the gift. *The American Journal of Sociology, 73,* 1–11.

Sherry, J. F. J. (1983). Gift-giving in anthropological perspective. *Journal of Consumer Research, 10,* 147–168.

Sherry, J. F. J., McGrath, M.A., & Levy, S. J. (1993). The dark side of the gift. *Journal of Business Research, 28,* 225–244.

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology, 48,* 813–838.

Smith, C. A., & Lazarus, R. S. (1993). Appraisal components, core relational themes, and the emotions. *Cognition and Emotion, 7,* 233–269.

Soscia, I. (2007). Gratitude, delight, or guilt: The role of consumers' emotions in predicting postconsumption behaviors. *Psychology and Marketing, 24*(10), 871–894.

Tracy, J. L., Robins, R. W., & Tangney, J. P. (2007). *The self-conscious emotions Theory and research.* New York: The Guilford Press.

Ward, M. K., & Broniarczyk, S. M. (2011). It's not me, it's you: How gift giving creates giver identity threat as a function of social closeness. *Journal of Consumer Research, 38,* 164–181.

Weiner, B. (1986). *An attributional theory of motivation and emotion.* New York: Springer-Verlag.

Winer, B. J. (1971). *Statistical principles in experimental design.* New York: McGraw-Hill.

Wolfinbarger, M. F. (1990). Motivations and symbolism in gift-giving behavior. In M. E. Goldberg, G. Gorn, & R. W. Pollay (Eds.), *Advances in consumer research, Vol. 17.* (pp. 699–706). Provo, UT: Association for Consumer Behavior.

Wolfinbarger, M. F., & Yale, L. J. (1993). Three motivations for interpersonal gift giving: Experiential, obligated, and practical motivations. In L. McAlister, & M. L. Rothschild (Eds.), *Advances in consumer research, Vol. 20.* (pp. 520–526). Provo, UT: Association for Consumer Research.

Wooten, D. B. (2000). Qualitative steps toward an expanded model of anxiety in gift-giving. *Journal of Consumer Research, 27,* 84–95.

Wubben, M. J. J., De Cremer, D., & Van Dijk, E. (2012). Is pride a prosocial emotion? Interpersonal effects of authentic and hubristic pride. *Cognition and Emotion, 26,* 1084–1097.

Zeelenberg, M., Nelissen, R. M.A., Breugelmans, S. M., & Pieters, R. (2008). On emotion specificity in decision making: Why feeling is for doing. *Journal of Judgment and Decision Making, 3,* 18–27.

Zeelenberg, M., & Pieters, R. (2004). Beyond valence in customer dissatisfaction: A review and new findings on behavioral responses to regret and disappointment in failed services. *Journal of Business Research, 57,* 445–455.

Full Length Article

Consumer participation in the design and realization stages of production: How self-production shapes consumer evaluations and relationships to products

CrossMark

S. Sinem Atakan [a,*], Richard P. Bagozzi [b], Carolyn Yoon [b]

[a] Graduate School of Business, Ozyegin University, Istanbul, Turkey
[b] Stephen Ross School of Business, University of Michigan, Ann Arbor, MI, 48109, USA

A B S T R A C T

Psychological responses of consumers to specific stages of self-production activities are investigated in four studies. Findings reveal that consumer participation in the realization stage (physical production) enhances affective commitment to the product. However, physical production without opportunity to express choice or creativity during the production process does not change the symbolic meaning of the product (how self-expressive it is) and, therefore, does not result in identification with the product. Participation during the design stage (input-specification) enhances identification, leading to affective commitment, which in turn enhances evaluation of the self-made product. Finally, engaging consumers in both the realization and design stages of the production process does not create value for consumers over and above the main effects created by a high level of participation in either stage.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Self-production in consumption, the active engagement in the creation of end products by consumers, is increasingly common as companies develop more ways for consumers to participate in the production process. It enhances not only the potential input of innovative ideas, but also the subjective valuation of one's self-produced items. When consumers play an active role in the production of products, they come to overvalue their own creations (Norton, Mochon, & Ariely, 2012).

Production consists of design (specification of input), realization (manufacturing, throughput), and use, according to the service systems perspective (Lengnick-Hall, 1996; Van Raaij & Pruyn, 1998). The present paper focuses on consumers' participation in the design and realization stages of production. During the design stage, characteristics of the product or service (e.g., physical layout, design, quality) are decided on. During the realization stage, the actual creation and execution of the product or service take place. Consumers' participation in self-production primarily takes place at the design stage (such as while designing a t-shirt on a website or kitchen layout for a new home) or the realization stage (such as while assembling furniture using step-

by-step instructions or cooking using a dinner kit). In some instances, consumers engage in both steps of self-production, both designing and physically creating the product.

The form of the self-production activities and the type of control consumers have over the products often vary between the design and realization stages. In the design stage, consumers engage in activities that require them to create, choose, or specify the form, layout, colors, and so on. They are mostly intellectually involved in the creation process and control the representational outcome as they specify the attributes. In the realization stage, consumers physically interact with the input materials and exert physical effort to create the product. They exercise manual control and power over the product during its creation. Also, to the extent that consumers physically shape the product, they are exposed to haptic cues in the process. Given these differences in the nature of consumer participation in design versus realization stages of production, an important question is how consumers' relationships with, and evaluation of, self-made products are differentially shaped during the two stages.

Our present research distinguishes between the design and realization stages, which at a first glance, seem to overlap with Buechel and Janiszewski's (2014) distinction between customization and physical assembly activities. However, there are key differences between our work and Buechel and Janiszewski's studies. We focus on the specific effects of each type of the self-production process (design vs. realization) on person–object relationship and valuation of the completed end-product. In particular, we investigate the underlying mechanisms governing final

* Corresponding author at: Graduate School of Business, Ozyegin University, Nisantepe Mah, Orman Sok, No: 13, Cekmekoy, Istanbul, Turkey 34794. Tel.: +90 216 564 95 04; fax: +90 216 564 90 44.
E-mail address: sinem.atakan@ozyegin.edu.tr (S.S. Atakan).

product evaluation (i.e., identification and affective commitment) by fully separating and integrating the two different stages of self-production activities. By contrast, Buechel and Janiszewski focus on the valuation of input materials in kits as function of the timing of the design decision (i.e., before or during the realization process); they do not manipulate actual physical effort exerted during assembly (i.e., realization).

We propose and test whether the psychological processes that shape consumers' evaluations depend on the production stage or type of activities that consumers engage in during self-production. Some research has examined self-design (e.g., Franke, Schreier, & Kaiser, 2010; Fuchs, Prandelli, & Schreier, 2010; Moreau & Herd, 2010), while other studies have scrutinized realization (e.g., Norton et al., 2012; Troye & Supphellen, 2012). We draw upon both streams of research and explore why and how consumer participation in design and realization stages shapes what the self-produced product means to consumers.

The rest of the paper proceeds as follows. First, we review extant findings from several literature streams, including research on the person–object relationship, extended self, and self-production. Then, we draw on touch, self-design, and organizational behavior literatures to develop our theoretical arguments, and from this generate predictions to be tested. We present four studies examining how consumers relate to self-made products as a result of participating in different stages of self-production. Study 1 focuses on the realization stage; Study 2 scrutinizes the design stage; and Studies 3A and 3B investigate whether involvement in both stages creates value for consumers, over and beyond that created by participation in either stage alone. Finally, we consider how our findings add to a better understanding of consumers and conclude by discussing implications and suggested avenues for future research.

2. Conceptual framework and research hypotheses

2.1. Psychological responses to participation in the production of products

Research on person–object relationships (Belk, 1988; Pierce, Kostova, & Dirks, 2003) indicates that creating, shaping, or physically producing a product result in powerful associations between the self and the product. The product becomes part of the extended-self due to the labor, time, and values invested into it. A product that has taken one's effort, attention and time becomes integrated into the self, since it has grown or emerged from the self (Csikszentmihalyi & Rochberg-Halton, 1981). People tend to exhibit emotional reactions to products and form feelings of attachment to objects that are connected to the self (Belk, 1988; Kleine, Kleine, & Allen, 1995). In addition to affect-laden aspects of product–self relationship, symbolic meanings of the product–self relationship contribute to one's self-concept (Grubb & Grathwohl, 1967; Kleine, Kleine, & Kernan, 1993; Atakan, Bagozzi, & Yoon, 2014). People continually compare their self-identity to the image of products and accommodate and assimilate ones that have similar or desired identities to their self-concept. During the self-production process, the product is formed in the image of its creator, revealing his tastes, preferences, and identity. The product gains symbolic meaning as a result of self-production and starts to reflect one's identity to the self as well as the outside world. Hence, the creator comes to identify with the self-made product, especially so with a self-expressive one, which self-production fosters.

Previous research suggests that people exhibit emotional reactions (affective commitment), and/or cognitively compare their identities (identification), to self-made products. In fact, Bloch (1995), in a theoretical paper, suggests that consumers may exhibit cognitive and/or affective responses to product design. However, to the best of our knowledge, no empirical research to date has investigated the dimensions of person–object relationships in the context of self-production.

Several literatures, including psychology, marketing, and organizational behavior, have conceived of emotional connections between consumers and animate or inanimate objects, such as people, groups, ideas, brands, or products. For instance, attachment theory (Bowlby,

1969; Ainsworth & Bowlby, 1991; Hazan & Shaver, 1994) informs our understanding of the emotional connection between an infant and a parent, as well as the romantic relationship between partners. It predicts relationship quality and functioning and indicates that physical contact has important positive implications for close relationships. One line of research on brand attachment within marketing utilizes attachment theory to conceptualize the relationship between consumers and brands as comprising both cognitive and emotional bonds. These bonds reflect connections developed over a relatively long time span (Park, MacInnis, & Priester, 2008). In a somewhat similar approach, Aaker, Fournier, and Brasel (2004) investigate brand–consumer relationships and discuss sincere (warm and caring) as well as exciting brands (see also Batra, Ahuvia, & Bagozzi, 2012).

Another tradition looks at person–product relationships but more specifically considers how an object helps consumers define and maintain a sense of self. Ball and Tasaki (1992) propose the term, "product attachment", and define it as the extent to which an object is used to maintain a cognitive structure of self. It is similar to the conceptualization posited by Kleine et al. (1995), who regard attachment as a reflection of the extent of "me-ness" associated with the product. Both traditions imply that product attachment is identical to self-extension and consists of cognitive (e.g., identity, a sign of self-worth) as well as affective (e.g., feelings associated with the product) components. Our conceptualization of affective commitment differs from this overall encompassing conceptualization and is more in line with the tradition of social identity research in the social psychology and organizational behavior literatures.

Research in social psychology and organizational behavior (Tajfel, 1978; Turner, 1985) informs our understanding of the relationship between individuals and groups or institutions. This literature makes a distinction between affective and cognitive components of organizational or group membership (Bergami & Bagozzi, 2000; Ellemers, Kortekaas, & Ouwerkerk, 1999). The affective component refers to the sense of belongingness, emotional involvement, and attachment to the target object (Allen & Meyer, 1990; Bergami & Bagozzi, 2000). It is measured with items such as "having a strong sense of belonging," and "feeling emotionally attached" (Allen & Meyer, 1990). The greater is the felt belongingness, the greater the affective commitment. The cognitive component, on the other hand, refers to a form of identification whereby a person comes to view himself as a member of the focal entity. This happens through cognitive processes of self-categorization, where one recognizes one's similarities with others in the organization or to the organization itself. The greater is the perceived similarity in values, goals, and characteristics, the greater the identification. It is through the perception of oneness with the focal entity and the degree to which one defines oneself by the same attributes that one develops identification with the focal entity (Dutton, Dukerich, & Harquail, 1994). Identification reflects the degree of congruence between one's own self-image and the image of the focal object and is a cognitive process.

By analogy to the relationship between a person and a group or organization, we propose two distinct dimensions of social identity, affective commitment and cognitive identification, to explain how self-production processes change consumers' relationships with products. First, we adopt the term "affective commitment" as the emotional response of consumers to their self-made product. It represents the emotional bond between a person and product and the warm feelings that one has for the object. In a similar vein, Norton et al. (2012) also talk about an emotional reaction to self-made products and propose that labor (physically building a product) leads to "love", an emotional reaction that was measured by a single item, "liking", in their study. Our conceptualization of affective commitment, based mainly on social psychology and organization behavior research, differs from that of brand attachment (cognitive as well as emotional bonds formed over intense and long-term interactions with the brand over the course of the

product's life) or the extended-self (how much the product contributes to the definition and maintenance of self-identity). We focus on emotional reactions to the product occurring as a function of self-production that happens in a relatively short time during and immediately after participating in its construction or formulation. It entails affect-laden responses of consumers to self-made products in the moments of personal engagement during self-production rather than how much the product contributes to the self-identity, which typically unfolds over a long period of time and involves considerable cognitive self-reflection.

Second, we use the term "identification" to refer to consumers' cognitive perception of how similar one is to the focal object (perception of oneness with the target). It is the perceived overlap between the product's identity and a person's identity, how much the image of the product applies to the self. Previous research indicates that innate objects, such as products (Belk, 1988) and brands (Aaker, 1997), may be compared to living beings and personified to take on human traits (e.g., "cool" cars, "friendly" computers; cf. Yoon, Gutchess, Feinberg, & Polk, 2006). Identification, thus, entails the awareness that a product has similar properties to one's sense of self, wherein the product takes on attributes of the self, and the self takes on attributes of the product, making it a type of cognition and entailing a kind of psychological or social construction.

In short, we propose that although often empirically associated, affective commitment and identification are conceptually distinct and have different antecedents and consequences. Together, they constitute consumers' experienced self-identity with respect to self-made products. Designing and physically making the product are parts of the act of self-production. Our research explores how different stages of self-production distinctively shape the relationship between consumers and their self-made products.

2.2. Participation in the realization (physical production) stage

During the realization stage, consumers put physical effort into making, assembling, or modifying a product. The process involves varying levels of physical exertion from simple manual labor, such as cutting, hammering, or knitting, to more effortful physical tasks, such as carrying heavy parts, carving hard surfaces, or extensive painting. Kinesthetic movements and physical effort are the predominant forms of engagement during this stage.

The process also entails touching and physical handling of the product in a purposive manner before product use. Human beings start to respond to touch even before they are born (Krishna, 2012). Research on interpersonal touch shows that touch enhances one's general positive feelings toward the target (Fisher, Rytting, & Heslin, 1976; Patterson, Powell, & Lenihan, 1986), increases the attachment level in a relationship, and causes one to feel closer to the other person (Anisfeld, Casper, Nozyce, & Cunningham, 1990; Bell, Daly, & Gonzalez, 1987). While interpersonal touch generates affective responses, particularly feelings of closeness and attachment to human beings, research on the endowment effect and the "IKEA effect" reveal that touch generates psychological responses to even inanimate objects. Wolf, Arkes, and Muhanna (2008) find that touching objects increases people's willingness to pay for those objects. Research by Peck and Shu (2009) shows that merely touching an object results in an increase in perceived ownership of that object. Moreover, Norton et al. (2012) suggest that physical labor leads to love for the object. Overall, these results imply that the physical handling of a product during its construction process may result in an affective response to the product. Therefore, we suggest that physically handling the product during the realization stage creates an emotional bond to the product that, in turn, enhances evaluation of the final product:

H1. Participating in the production of a product during the realization stage enhances evaluation of the final product.

H2. Affective commitment to the product mediates the impact of consumer participation during the realization stage on product evaluation.

Physical engagement during the realization stage often provides little opportunity to modify the product according to one's wishes, tastes, or preferences, and to forge a tangible congruency between one's own and the product's identity. Therefore it is less likely to lead to identification. We suggest that to the extent that the production process involves minimal opportunity to modify the actual product, the realization process will contribute little to the consumer's self-identity in this early stage of one's relationship to the product before one has a chance to solidify the idea of a shared identity of self and product. In this case, production fails to signal the identity of the individual, and therefore, identification with the product is likely to be low or negligible. Greater evidence of a tangible nature is needed to solidify and create a strong sense of shared identity, such as what occurs through ongoing product use over time.

2.3. Participation in the design (input specification) stage

During the design stage, consumers design, make choices (such as color, materials, shape), or use their creativity to modify the shape or other aspects of a product. A growing body of experimental evidence indicates that being the designer of a product results in economic value for consumers, whereby the additional value does not merely accrue from functional value, i.e., better fit between the consumer's underlying preferences and product attributes (Franke et al., 2010; Moreau & Herd, 2010). There are psychological reactions to participating in the design stage, such as the "I designed it myself" effect that originates from awareness of being the creator of the design (Franke et al., 2010). In addition to the array of typically utilitarian or tangible values (e.g., customization of product attributes) that consumers derive from self-production, we propose a symbolic source of value that heretofore has not been elucidated.

Participating in the design process enables investment of mental energy and a sense of one's being (ideas, values, choices) into the product. As a consumer changes the visual appearance of a product, the product starts to reflect his or her tastes, preferences, and identity. The consumer is connected to a sense of the self (free will), body parts, personal attributes, possessions, and even one's own abstract ideas, in addition to other people and objects in close proximity to the self (McClelland, 1951; Prelinger, 1959). A self-made product is a vehicle for imbedding these aspects of the self into it. As the product becomes formed into the image of its creator, it gains symbolic meaning (Belk, 1988). Hence, we propose that to the extent that consumers can construct and change the visual representation of a product during the design stage, they start to identify with the product. Formally, we propose:

H3. Participating in the design stage of the self-production process enhances identification with the product.

To understand the relationship between the identification and affective commitment dimensions of person–object relationships, we draw on insights from the extended-self literature and organization research. Ahuvia (2005) and Kleine et al. (1995) reveal that products linked to an identity (from past, present, or possible futures) tend to be loved by their owners. In addition, Belk (1988) and Pierce et al. (2003) propose that people form feelings of attachment to products that express their self-identity. Identification with the product, a largely cognitive process, sets the stage for this. Self-image reflected in a product and the degree of self-reflection therein determines the emotional response to the product. In addition to the person-object literature, quantitative research from organizational theory (Bergami & Bagozzi, 2000; Bhattacharya & Sen, 2003) suggests that identification (the extent to which the individual sees the focal object as part of one's self-identity) leads to affective

commitment (emotional response to the object). We, therefore, expect identification to indirectly affect product evaluation through affective commitment, whereas affective commitment directly affects product evaluation:

H4. Identification affects product evaluation through affective commitment.

We propose that identification with and affective commitment to the product represent the overall bond between the consumer and product. Peck and Shu (2009) show that stronger bonds are likely to result in higher valuation of products. We hypothesize that participating in the production process during the design stage strengthens identification with and affective commitment to the product, which in turn enhances evaluation of the product.

H5. Identification with, and affective commitment to, the product mediate the effect of consumer participation during the design stage on product evaluation.

3. Study 1: participation in the realization stage

In Study 1, we test whether higher levels of engagement in production during the realization stage (where no participation in design has occurred) lead to higher evaluation of the product due to increased affective commitment to, but not identification with, the product.

3.1. Method

3.1.1. Design and procedure
The study was a one-factor between-subjects design with two experimental treatment groups and one control group. The task involved making a picture frame from cardboard. Seventy-five undergraduate students, recruited from a paid subject pool in a large Midwestern university, were randomly assigned to control, low-realization, and high-realization groups. In the control condition, participants were given a cardboard picture frame and asked to examine it. They spent about 29 s on average ($SD = 2.21$) to examine the frame. In the low-realization condition, participants only had to glue together pre-cut, ready-to-assemble pieces by following step-by-step instructions. They spent about 9.42 min on average ($SD = .38$) assembling the pieces. In the high-realization condition, participants were given step-by-step instructions to make the frame from scratch. They spent 23.68 min on average ($SD = 2.07$). Detailed step-by-step instructions allowed for no specification of inputs (i.e., no opportunities for designing the product) in either the low- or high-realization conditions.

Previous literature on the endowment effect (Strahilevitz & Loewenstein, 1998) indicates that time spent with a product may increase attractiveness and valuation of the product. Hence, in order to equate the time spent with the frame across all conditions, before evaluating the frame, participants worked on a filler task in the control and the low-realization conditions for 25 and 20 min, respectively, while having the product in front of them during the whole time, allowing the respondents to see and touch it if desired. The filler task was unrelated to the product and consisted of a "count the articles" procedure. The filler task was introduced as an irrelevant study that was about identifying how definite and indefinite articles interfere with reading. The participants read short, historical stories about various cities in the United States and counted and entered the number of definite (e.g., the) and indefinite (e.g., a, an) articles in each paragraph. This task was chosen because it was easy to perform yet required concentration to complete.

3.1.2. Measures
After making the frame, participants in the experimental conditions rated the amount of physical engagement needed to make the frame

($1 =$ none, $7 =$ a great deal). Then, all participants evaluated the frame using three 7-point bipolar evaluative items (negative/positive, bad/good, unfavorable/favorable; $\alpha = .92$) and reported on four 7-point scales ($1 =$ not at all, $7 =$ extremely) how much they identified with the product: "The frame represents who I am," "I identify with the frame," and "It reflects the type of person that I am", adapted from Reed, Aquino, and Levy (2007), while the last item was "The image of the frame fits my self-image" ($\alpha = .90$). Participants also indicated their degree of affective commitment to the product on four 7-point scales ($1 =$ not at all, $7 =$ extremely): the specific items were "like", adapted from Norton et al. (2012); "attached" and "connected", adapted from Thomson, MacInnis, and Park (2005); and "warm" ($\alpha = .87$). Identification and affective commitment measures were counterbalanced.

3.2. Results

3.2.1. Manipulation check
Participants in the high condition ($M = 3.78$, $SD = 1.48$) indicated higher levels of physical engagement needed to make the frame than did those in the low condition ($M = 2.46$, $SD = .74$), $t(53) = 4.19$, $p < .001$. The confidence intervals for high ($CI_{.95} = 3.22, 4.34$) and low ($CI_{.95} = 2.18, 2.74$) conditions did not include 1 ("none"), indicating that the physical engagement in both conditions was higher than none (the control condition). The manipulation was, therefore, successful.

#### 3.2.2. Discriminant validity for affective commitment and identification measures
Confirmatory Factor Analysis (CFA) and Structural Equation Models (SEM) revealed that the three constructs (product evaluation, identification, and affective commitment) are distinct factors. For identification and affective commitment latent variables, the four items of each were combined to produce two indicators each, using the partial disaggregation model (Bagozzi & Heatherton, 1994). The first indicator was the average of two (out of four) items, and the remaining two measures were averaged to form the second indicator. This approach yields models with fewer parameters to estimate and reasonable ratios of cases to parameters while smoothing out measurement error to a certain extent. Fig. 1a reports the results of the SEM analyses including the standardized path coefficients. Overall, the goodness-of-fit measures ($\chi^2(11) = 12.86$, $p \approx .30$, SRMR $= .030$, NNFI $= .99$, CFI $= 1.00$, RMSEA $= .04$) show an excellent fit. An analysis of the ϕ_{ij} entries (correlations between constructs, corrected for attenuation) indicated that the correlation between product evaluation and affective commitment was .66 ($SE = .09$; $CI_{.95} = .48, .84$), product evaluation and identification was .40 ($SE = .10$; $CI_{.95} = .20, .60$), and identification and affective commitment was .59 ($SE = .10$; $CI_{.95} = .39, .79$). None of the confidence intervals included the value of one, providing evidence of discriminant validity for measures of product evaluation, identification, and affective commitment.

3.2.3. Test of hypotheses
An ANOVA on product evaluation indicated a significant main effect of participation during the realization stage ($F(2, 72) = 12.21$, $p < .001$), thereby providing support for H1. Planned contrasts revealed that, participants in the high ($M = 5.56$; $t(72) = 4.24$, $p < .001$) as well as low ($M = 5.65$; $t(72) = 4.52$, $p < .001$) conditions evaluated the product more favorably than those in the control condition ($M = 4.02$); no difference was found between high and low conditions ($t(72) = .26$, $p = .79$). Even low levels of engagement during the realization stage enhanced product evaluation.
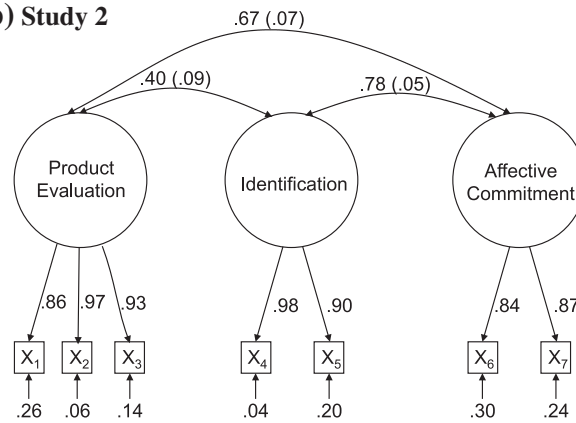
An ANOVA on affective commitment revealed a significant main effect of participation during the realization stage ($F(2, 72) = 9.98$, $p < .001$). Participants in the high ($M = 4.01$) condition indicated higher affective commitment to the product than the ones in the low ($M = 3.12$; $t(72) = 2.48$, $p < .05$) or control ($M = 2.27$; $t(72) =$
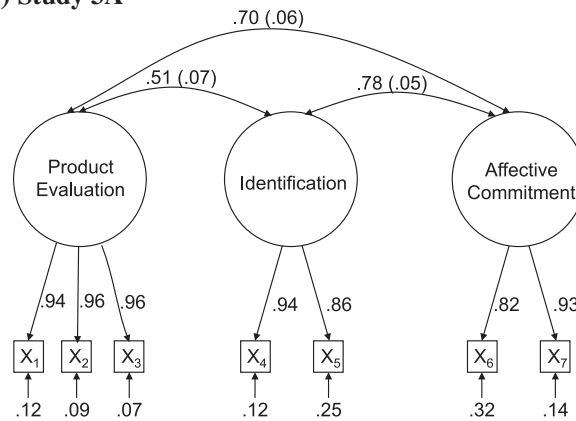
a) Study 1
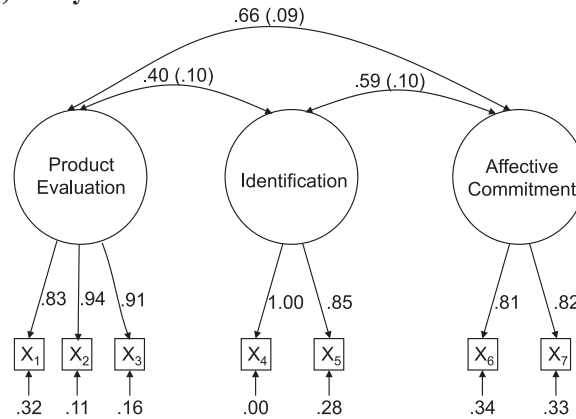
$x^2(11) = 12.86$; $\underline{p} \simeq .30$ RMSEA = .042; SRMR = .030; NNFI = .99; CFI = 1.00

b) Study 2

$x^2(11) = 33.38$; $\underline{p} \simeq .001$ RMSEA = .14; SRMR = .059; NNFI = .94; CFI = .97

c) Study 3A

$x^2(11) = 20.62$; $\underline{p} \simeq .04$ RMSEA = .089; SRMR = .035; NNFI = .98; CFI = .99

Fig. 1. Structural equation modeling results.

4.44, $p < .001$) conditions; the difference between low and control conditions was also statistically significant ($t(72) = 2.19$, $p < .05$).

Next, to test whether affective commitment mediates the positive impact of participation during the realization stage on product evaluation (H2), mediation analysis followed the bootstrapping method for multicategorical causal agents (Hayes & Preacher, 2013; Preacher & Hayes, 2008). Using dummy coding with the control group as the reference, two separate models were run: one for low-realization level and one for high-realization level. For both models, the bootstrapping confidence intervals based on 5000 samples yielded 95% confidence intervals for the relative indirect effects that exclude zero (low-realization: indirect effect = .39, $SE = .18$, $CI_{.95} = \{.05, .76\}$; high-realization: indirect

effect $= .80$, $SE = .23$, $CI_{.95} = \{.38, 1.29\}$), indicating that both low- and high-realization conditions indirectly influence product evaluation through affective commitment; we thus obtain support for H2. As expected, an ANOVA on identification revealed no effect of participation during the realization stage ($M_{high} = 2.71$, $M_{low} = 2.14$, $M_{control} = 2.11$; $F(2, 72) = 2.09$, $p = .13$).

3.3. Discussion

Findings from Study 1 provide empirical evidence that participation during the realization stage (a) affects how consumers relate to self-made products, and b) changes their product evaluations. CFA and SEM analyses indicate that identification and affective commitment are two distinct dimensions of person–object relationships. An emotional bond (affective commitment) to the product is formed as a result of physical investment of self into the product during the realization stage. This emotional bond mediates the impact of participation during the realization stage on product evaluation. However, participation in the realization stage alone does not necessarily result in identification with the product because little opportunity is provided to design the product in concert with one's self-image. Study 2 focuses on identification processes during participation in the design stage.

4. Study 2: participation in the design stage

We investigate how participation during the design stage alone shapes person–object relationships and whether it has different effects than the realization stage in shaping how consumers relate to self-made products. Study 2 tests Hypotheses 3–5.

4.1. Method

4.1.1. Design and procedure

Similar to Study 1, Study 2 is a one-factor between subjects design with two experimental treatment groups and one control group. The study involved designing an insert for a travel coffee mug. One hundred and three undergraduate students at a large Midwestern university completed the study in partial fulfillment of course requirements. Participants were told that they would participate in several unrelated studies. The first task, presented as an investigation of PowerPoint (PPT) in terms of ease of use, was designed to control the PPT skills of participants when creating a design. It involved a basic tutorial on how to insert and modify figures, text, and ClipArt in PPT. After the tutorial, participants reported how difficult it was to edit figures, to edit text, and to do the tutorial examples (1 = very easy, 7 = very difficult; $\alpha = .77$) in PPT.

Participants were then randomly assigned to control, low-design, and high-design conditions. All participants were given a travel mug with a removable blank insert and told that they would have the option to keep the mug at the end. The base of the mug could be twisted off to remove the inner insert, and one could draw or write on the insert before reinserting and twisting on the base. In the control condition, the participants were encouraged to examine the mug but could not modify or change it in any way. They spent about 2.5 min reading the directions and examining the mug. In the low- and high-design conditions, the participants designed the insert using PPT. In the low-design condition, the participants could insert only one figure/image from PPT ClipArt, change its size, and place it anywhere they wanted. They spent an average of 5.3 min on this task. In the high-design condition, they could design the insert in any way they wanted (e.g., insert figures or images from ClipArt, change colors, write on the insert) and were encouraged to be creative. The participants spent about 11 min designing the mug. In the experimental conditions, after the design stage, the insert was printed and placed back in the travel mug by the experimenter in another room, and the mug was returned back to the participant in less than a minute.

In order to equate the time spent with the product, the participants in the control and the low-design conditions worked on a filler task for 12 and 7 min, respectively, before evaluating the mug. The filler task involved a "word generation" procedure in which participants were asked to generate words starting with the letters N, D, C and words with letters E or S in the middle (Liu, 2008).

4.1.2. Measures

Participants in the experimental conditions answered the manipulation check questions regarding the amount of effort exerted on design. They indicated the degree of original thinking and creativity that went into the design and how intellectually stimulating they found the task (1 = none at all, 7 = very much; $\alpha = .89$). They then completed the dependent measures identical to those used in Study 1: product evaluation ($\alpha = .94$), identification with ($\alpha = .94$), and affective commitment to ($\alpha = .88$) the product.

4.2. Results

Reported PPT difficulty levels did not differ across the three conditions ($F < 1$). Therefore, difficulty is excluded from subsequent analyses.

4.2.1. Manipulation check

Participants in the high-design condition ($M = 4.24$, $SD = 1.54$) reported higher levels of effort that went into the design than did those in the low-design condition ($M = 2.74$, $SD = .94$), $t(64) = -4.77$, $p < .001$. The confidence intervals for high ($CI_{.95} = 3.68, 4.79$) and low ($CI_{.95} = 2.42, 3.07$) conditions did not include 1 (none), indicating that the level of effort that went into the design in both conditions was higher than none (the control condition). The manipulation was thus successful.

4.2.2. Discriminant validity for affective commitment and identification measures

As in Study 1, CFA and SEM revealed that product evaluation, identification, and affective commitment are distinct constructs. Fig. 1b reports the results of the SEM analyses including the standardized path coefficients. The model yields a good representation of the data ($\chi^2(11) = 33.38$, $p < .001$). Three out of four goodness-of-fit measures (SRMR = .059, NNFI = .94, CFI = .97, RMSEA = .14) give a satisfactory fit, which points to an acceptable model (Hu & Bentler, 1998). An analysis of the ϕ_{ij} entries indicated that the correlation between product evaluation and affective commitment was .67 ($SE = .07$; $CI_{.95} = .53$, .81), between product evaluation and identification was .40 ($SE = .09$; $CI_{.95} = .22$, .58), and between identification and affective commitment was .78 ($SE = .05$; $CI_{.95} = .68$, .88). None of the CIs included the value of one, providing evidence of discriminant validity.

4.2.3. Test of hypotheses

An ANOVA on product evaluation showed that the effect of level of participation during the design stage was significant ($F(2, 100) = 5.22$, $p < .01$). Evaluation of the product was higher in the high-design ($M = 5.85$) than in the control condition ($M = 4.95$; $t(100) = 3.19$, $p < .01$). There was no difference between low-design ($M = 5.47$) and high-design conditions ($t(100) = 1.31$, $p = .19$). The difference between the control and the low-design conditions was marginally significant ($t(100) = 1.90$, $p = .06$). Thus, higher levels of participation during the design stage enhanced evaluation of the mug.

As expected, an ANOVA on identification revealed a significant effect of level of participation during the design stage ($F(2, 100) = 29.87$, $p < .001$); H3 was supported. Identification was lower in the control ($M = 2.28$) than in the low- ($M = 3.65$; $t(100) = 4.14$, $p < .001$) or high-design ($M = 4.89$; $t(100) = 7.70$, $p < .001$) conditions. The difference between the high- and low-design conditions was also statistically significant ($t(100) = 3.56$, $p < .01$).

An ANOVA on affective commitment revealed a significant main effect ($F(2, 100) = 23.02$, $p < .001$). Contrasts indicated that affective commitment was significantly lower in the control ($M = 3.13$) than in the low- ($M = 4.54$; $t(100) = 4.90$, $p < .001$) or high-design ($M = 5.03$; $t(100) = 6.43$, $p < .001$) conditions. Although directionally consistent with what was expected, there was no difference between the high- and low-design conditions ($t(100) = 1.61$, $p = .11$). Unlike participation in the realization stage (Study 1), participation in the design stage is found to enhance both identification with, and affective commitment to, the product.

To examine whether one mediator (identification) causally affects the other mediator (affective commitment), the multiple-step, multiple mediator model (Hayes, Preacher, & Myers, 2011) was used. Consistent with H5, we found that level of participation in the design stage enhances identification, and identification influences affective commitment, which in turn augments evaluation of the product; see Fig. 2b. The independent variable (level of participation during design stage — a categorical variable with three levels) was dummy coded for the analysis. We ran two models, using one of the dummy codings as the independent variable and the second one as the covariate in each model; see Table 1a for estimates of the path coefficients and the bootstrapping results. The indirect path from level of participation to product evaluation through identification and affective commitment, in that order, was significant whether there was low-design (indirect effect = .36, $SE = .15$, $CI_{.95} = \{.14, .73\}$) or high-design (indirect effect = .69, $SE = .24$, $CI_{.95} = \{.34, 1.30\}$) participation; H5 is supported. As predicted by H4, the indirect path from level of participation to product evaluation through identification (independent of affective commitment) was not significant when level of participation was low (indirect effect = −.11, $SE = .12$, $CI_{.95} = \{−.35, .10\}$) or high (indirect effect = −.20, $SE = .21$, $CI_{.95} = \{−.69, .15\}$). The indirect path from level of participation to product evaluation through affective commitment (independent of identification) was significant when level of participation was low (indirect effect = .39, $SE = .16$, $CI_{.95} = \{.12, .75\}$) but not when

a) Study 1
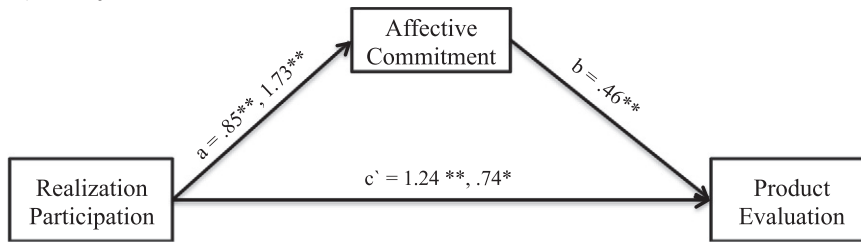
The first value represents the path estimate of low-realization participation and the second one represents the path estimate of high-realization participation.

b) Study 2
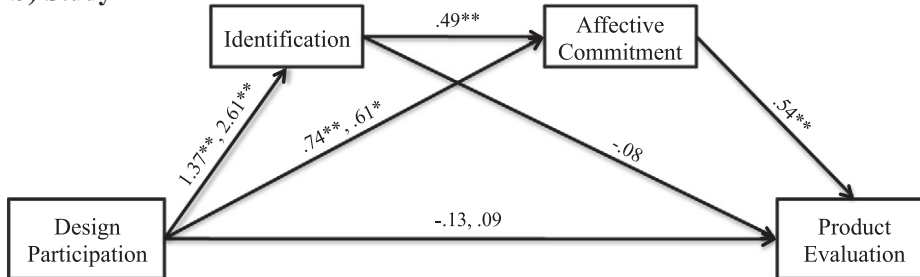
The first value represents the path estimate of low-design participation and the second one represents the path estimate of high-design participation.

c) Study 3A

Path estimates represent unstandardized regression coefficients.
*$p < .05$, **$p < .01$
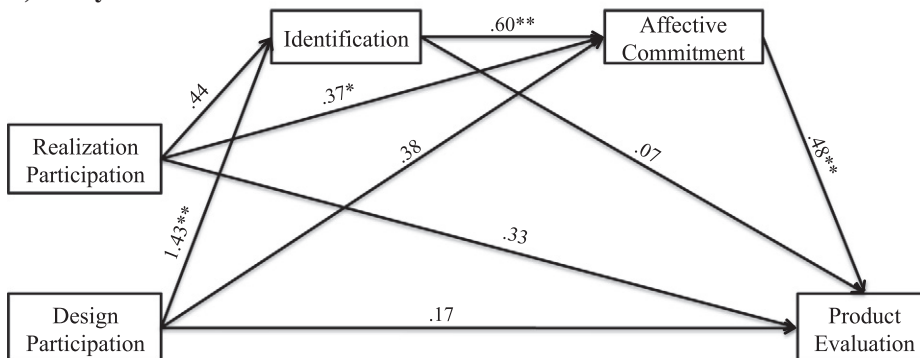
Fig. 2. Mediation models. a. Study 1: The first value represents the path estimate of low-realization participation and the second one represents the path estimate of high-realization participation. b. Study 2: The first value represents the path estimate of low-design participation and the second one represents the path estimate of high-design participation. c. Study 3A: Path estimates represent unstandardized regression coefficients. *$p < .05$, **$p < .01$.

Table 1
Path coefficients and indirect effects for the model.

a. Study 2	Path coefficients			b. Study 3A	Path coefficients		
	To product evaluation	To identification	To affective commitment		To product evaluation	To identification	To affective commitment
From low-design participation	−.13 (.27)	1.37** (.33)	.74** (.33)	From realization participation	.33 (.20)	.44 (.23)	.37* (.18)
From high-design participation	.09 (.31)	2.61** (.34)	.61* (.31)	From design participation	.17 (.23)	1.43** (.23)	.38 (.21)
From identification	−.08 (.09)		.49** (.07)	From identification	.07 (.10)		.60** (.07)
From affective commitment	.54** (.10)			From affective commitment	.48** (.10)		

	Indirect effects				Indirect effects		
	Estimate	Bootstrap 95% CI			Estimate	Bootstrap 95% CI	
Independent variable: Low-design participation				Independent variable: Realization participation			
Total	.65	.32, 1.10		Total	.33	.11, .60	
Specific: Low-design → I → PE	−.11	−.35, .10		Specific: Realization → I → PE	.03	−.06, .16	
Specific: Low-design → A → PE	.39	.12, .75		Specific: Realization → A → PE	.18	.01, .37	
Specific: Low-design → I → A → PE	.36	.14, .73		Specific: Realization → I → A → PE	.13	−.01, .32	
Independent variable: High-design participation				Independent variable: Design participation			
Total	.82	.36, 1.30		Total	.70	.38, 1.08	
Specific: High-design → I → PE	−.20	−.69, .15		Specific: Design → I → PE	.10	−.22, .42	
Specific: High-design → A → PE	.33	−.03, .76		Specific: Design → A → PE	.18	−.01, .42	
Specific: High-design → I → A → PE	.69	.34, 1.30		Specific: Design → I → A → PE	.41	.17, .75	

Presented are estimates of the path coefficients and the bootstrapping results, parentheses contain the standard errors.
I = Identification, A = Affective Commitment, PE = Product Evaluation.
 * $p \leq .05$.
 ** $p \leq .01$.

level of participation was high (indirect effect $= .33$, $SE = .20$, $CI_{.95}$ $= \{-.03, .76\}$). Two other models were run to test the following causal chain: design participation $\rightarrow$ affective commitment $\rightarrow$ identification $\rightarrow$ product evaluation. In these models, the indirect paths from level of design participation to product evaluation through affective commitment and identification were not significant whether participation involved low-design (indirect effect $= .07$, $SE = .08$, $CI_{.95} = \{-.26, .05\}$) or high-design (indirect effect $= .10$, $SE = .11$, $CI_{.95} = \{-.36, .07\}$). Thus, this outcome strengthens the findings for the hypothesized ordering from identification to affective commitment.

Next, we tested whether self-expressiveness of the product affects identification. Identification with the product should increase to the extent that the product reflects one's sense of self. Two independent raters coded the mug insert designs made by participants in the experimental conditions. The mug designs included such items as the slogans of the university that the participants attended to, their own names, sports team symbols, and pictures of drink or food items. Participants had also used various colors such as red, green, and pink to design the mugs. The raters used three seven-point scales (1 = not at all, 7 = very much) to evaluate how self-expressive the designs were ("the design is self expressive", $\alpha = .69$; "one can get a sense of the designer's personality from this", $\alpha = .72$; "it reflects the designer's self-image", $\alpha = .74$). The ratings were averaged to form a single self-expressiveness index. Identification and affective commitment were separately regressed onto the index. As expected, self-expressiveness of the design predicted the level of identification ($\beta = .62$, $t = 2.88$, $p < .01$), indicating that identification with the product increases as the self-expressiveness of the product increases. Self-expressiveness did not have any direct effects on affective commitment ($\beta = .10$, $t = .59$, $p = .55$).

4.3. Discussion

Findings from Study 2 provide further evidence that there are two dimensions through which a person relates to a self-made product: identification and affective commitment. As expected, level of participation during the design process positively influences product evaluation, and the impact of participation on product evaluation is mediated through identification and affective commitment, while the impact of identification is mediated through affective commitment. Participation during the design stage enables consumers to modify the product to reflect who they are, their self-identity, their tastes and preferences, resulting in enhanced identification with the product. Identification increases affective commitment, which in turn enhances product evaluation.

5. Studies 3A and 3B: participation in both the design and realization stages

It is possible that the divergent results for identification between Studies 1 and 2 are due to the products used in the studies. Consumers might be more likely to identify with a mug, a product used dynamically and frequently, than they would with a picture frame, a product passively displayed and viewed infrequently. We, therefore, sought to replicate our results in the next two studies with the same product for both stages of the production process.

Moreover, from a managerial standpoint, the question remains as to whether it is beneficial for a firm to invest in enabling its consumers to engage in both stages of production. In the present study, we thus test the interactive effects of participation in the design and realization stages. On the one hand, two stages could interact with each other to further enhance (additively or multiplicatively) product evaluation and strengthen the person–object relationship. On the other hand, a high level of participation in only one stage could be sufficient to enhance product evaluation to some maximum level, and any other effects resulting from participation in an

additional stage could be minimal. Studies 3A and 3B aim to address these questions.

5.1. Study 3A: method

5.1.1. Design and procedure

Study 3A was a 2 (level of participation during design stage: low vs. high) $\times$ 2 (level of participation during realization stage: low vs. high) between-subjects design. The task involved designing and making a music CD with its case. One hundred and twenty-two undergraduate students were recruited from a paid subject pool at a large Midwestern university. First, participants were administered the same PPT tutorial from Study 2, which measured their skills in PPT ($\alpha = .73$). After completing filler studies, the participants were told that the next study would investigate music preferences of students. They were asked to choose five songs from a list containing six genres with six songs under each genre that were the most popular among students according to a pretest. The instructions indicated that the chosen songs would be burned onto a CD and the CD placed in a case. Participants were told that they would have the option to keep the CD and its case. After choosing the songs, they were randomly assigned to low- or high-realization and low- or high-design conditions.

In the low-realization conditions, the songs were burned onto a CD, and its case was made for the participant by the experimenter in another room. In the high-realization conditions, a blank CD case template on PPT was provided to the participants. They made the CD case following step-by-step guidelines; first they had to type the titles of the songs and the artists, then print the template on white cardboard, and finally cut and glue the template. Following the guidelines, the participants also burned the songs onto a CD themselves. In the low-design conditions, the participants could not modify the case template except for typing up the song titles and the artists. In the high-design conditions, they could title the CD and design the case in any way they wanted using PPT.

The low-realization/low-design condition, comprising only choosing the songs, served as the control condition representing the baseline evaluation of the CD and its case. The participants spent 3.5 min on average choosing the songs. In the low-realization/high-design condition, participants spent 16.4 min on average choosing the songs and designing the case. In the high-realization/low design condition, they spent 16.6 min on average choosing the songs, burning the CD, and making the case. In the high-realization/high design condition, they spent 27 min on average choosing the songs, burning the CD, and designing and making the case. In the low-realization conditions, the final product was returned to the participant in less than three minutes. To equate the time spent with the product, participants worked on the same filler task as in Study 2 while the CD and its case were in front of them. They did so for 20, 15, and 10 min in the low-realization/low-design, low-realization/high design, and high-realization/low-design conditions, respectively.

5.1.2. Measures

All participants, except for those in the low-realization/low-design condition, indicated the level of effort (1 = none at all, 7 = very much) that went into the physical construction ("how much basic physical effort did you use", "how much simple manual labor did you use", "how much basic physical energy did you put into making the product"; $\alpha = .78$) as well as the design ("how much original thinking went into making the CD and its case", "how much creativity did you use", "how much did you think to make it"; $\alpha = .91$). Then, all participants answered the product evaluation ($\alpha = .97$), identification ($\alpha = .91$), and affective commitment ($\alpha = .88$) questions.

5.2. Study 3A: results

PPT difficulty was significantly different between low ($M = 1.16$) and high ($M = 1.32$) design conditions ($F(1, 117) = 4.65$, $p < .05$).

Analyses conducted with and without PPT difficulty yielded the same substantive results. We thus report only the findings without PPT difficulty.

5.2.1. Manipulation checks

An ANOVA on reported design effort indicated a significant main effect of design participation ($F(1, 90) = 59.71$, $p < .001$), and a non-significant effect of realization participation ($F < 1$) levels; the design manipulation was thus successful. An ANOVA on reported physical effort indicated a significant main effect of realization participation ($F(1, 90) = 9.55$, $p < .01$), and a non-significant effect of design participation ($F(1, 90) = 1.91$, $p = .17$) levels; the realization manipulation was successful.

5.2.2. Discriminant validity for affective commitment and identification measures

As in Studies 1 and 2, CFA and SEM revealed that measures of product evaluation, identification, and affective commitment are distinct (see Fig. 1c). The model yields a good representation of the data ($\chi^2(11) = 20.62$, $p \approx .04$). Three out of four goodness-of-fit measures (SRMR = .035, NNFI = .98, CFI = .99, RMSEA = .089) give a satisfactory fit, pointing to an acceptable model. An analysis of the ϕ_{ij} entries indicated that the correlation between product evaluation and affective commitment was .70 ($SE = .06$; $CI_{.95} = .58$, .82), between product evaluation and identification was .51 ($SE = .07$; $CI_{.95} = .37$, .65), and between identification and affective commitment was .78 ($SE = .05$; $CI_{.95} = .68$, .88). None of the confidence intervals included the value of one, providing evidence of discriminant validity for the measures of the three constructs.

5.2.3. Test of hypotheses

Replicating findings from Studies 1 and 2, an ANOVA on product evaluation showed that the main effects of levels of participation during realization ($F(1, 118) = 9.01$, $p < .01$) and design ($F(1, 118) = 14.81$, $p < .001$) were significant. The main effects were qualified by a marginally significant interaction ($F(1, 118) = 2.91$, $p = .09$). We did not have an a priori hypothesis regarding the interaction; however, we explored further what happens when consumers engage in both stages of production by decomposing the interaction. Simple effects tests indicated that during low levels of design participation, evaluation of the product was significantly more favorable when realization was high ($M = 5.26$) rather than low ($M = 4.18$) ($F(1, 118) = 10.43$, $p < .01$). However, during high levels of design participation, evaluation of the product did not differ between the high ($M = 5.75$) and low ($M = 5.46$) realization conditions ($F < 1$). Similarly, when realization participation was low, higher levels of design participation enhanced evaluation of the product ($F(1, 118) = 15.02$, $p < .001$). However, when realization participation was high, design participation did not enhance evaluation of the product ($F < 1$). A high level of participation in either stage of the production process was enough to enhance evaluation of the final product. Participation in an additional stage of production did not necessarily enhance the evaluation of the product.

An ANOVA on identification revealed a significant main effect for design participation ($M_{low} = 2.63$, $M_{high} = 4.06$; $F(1, 118) = 37.81$, $p < .01$), but only a marginally significant effect for realization participation ($M_{low} = 3.13$, $M_{high} = 3.56$; $F(1, 118) = 3.40$, $p = .07$). The interaction effect was not significant ($F < 1$). As hypothesized, design participation enhanced identification with the product; however, realization participation exhibited a minimal effect on identification.

As anticipated, an ANOVA on affective commitment revealed significant main effects for both realization ($F(1, 118) = 8.04$, $p < .01$) and design ($F(1, 118) = 30.21$, $p < .001$) participation. The interaction was not significant ($F < 1$). Participants reported higher affective commitment in the high ($M = 4.21$) than low ($M = 3.56$) realization condition, and in the high ($M = 4.51$) than low ($M = 3.26$) design condition.

To test the proposed mediations, bootstrapping analyses were conducted to estimate direct and indirect effects with two independent variables and two mediators; see Fig. 2c. Product evaluation was the dependent variable; realization and design participation were the independent variables. Identification and affective commitment were hypothesized mediators for the effects of design and realization participation. Two separate models were run using bootstrapping. In each of the models, design or realization participation was specified as the independent variable and the other was treated as a covariate. Covariates are treated exactly like independent variables in the estimation, with paths to all mediators and the outcome. Including the other independent variable as a covariate in the model corrects for the effect of the independent variable, and each model generates the desired indirect effect for the variable currently listed as the independent variable.

First, realization participation was the independent variable and design participation was the covariate. The results indicated that the total (indirect + direct) effect of realization participation on product evaluation (total effect = .66, $p < .01$) was nonsignificant when the mediators were included in the model (direct effect of realization participation = .33, $p = .11$). The total indirect effect of realization participation on product evaluation was significant, with a point estimate of .33 and a 95% CI of .11 to .60. Only the indirect effect through affective commitment (indirect effect = .18, $SE = .09$, $CI_{.95} = \{.01, .37\}$) was significant. The indirect paths through identification (indirect effect = .03, $SE = .05$, $CI_{.95} = \{-.06, .16\}$), and through identification and affective commitment (indirect effect = .13, $SE = .08$, $CI_{.95} = \{-.01, .32\}$) were not significant, because their confidence intervals contained zero. The impact of realization participation on product evaluation was mediated only through affective commitment. Furthermore, consistent with predictions, the path estimates from realization participation to affective commitment (path estimate = .37, $SE = .18$, $p < .05$) and from affective commitment to product evaluation (path estimate = .48, $SE = .10$, $p < .01$) were significant, whereas those from realization participation to identification (path estimate = .44, $SE = .23$, $p = .06$) were not.

In the second model, design participation was the independent variable and realization participation was the covariate. The results indicate that the total effect of realization participation on product evaluation (total effect = .87, $p < .01$) was non-significant when the mediators are included in the model (direct effect of design participation = .17, $p = .46$). The total indirect effect of design participation on product evaluation was significant, with a point estimate of .70 and a 95% CI of .38 to 1.08. The indirect effects through identification (indirect effect = .10, $SE = .16$, $CI_{.95} = \{-.22, .42\}$), and through affective commitment (indirect effect = .18, $SE = .11$, $CI_{.95} = \{-.01, .42\}$) were not significant. The indirect path to product evaluation through both identification and affective commitment, in that order, was however significant (indirect effect = .41, $SE = .15$, $CI_{.95} = \{.17, .75\}$). Consistent with predictions, the effects from design participation to identification (path estimate = 1.43, $SE = .23$, $p < .001$), from identification to affective commitment (path estimate = .60, $SE = .07$, $p < .001$), and from affective commitment to product evaluation (path estimate = .48, $SE = .10$, $p < .001$) were significant. See Table 1b for the estimates and bootstrapping results. Another model where design participation was again the independent variable and realization participation was the covariate tested the following causal chain: design participation → affective commitment → identification → product evaluation. In this model, the indirect path to product evaluation through affective commitment and identification was not significant (indirect effect = .06, $SE = .09$, $CI_{.95} = \{-.10, .24\}$). This result demonstrates that affective commitment does not influence identification and thereby bolsters the reported findings for the hypothesized sequence.

In sum, the analyses indicate that the effect of realization participation on product evaluation is mediated through affective commitment

only, whereas the impact of design participation on product evaluation is mediated through both identification and affective commitment. Furthermore, identification precedes affective commitment in the case of design participation.

Next, we investigated whether self-expressiveness of the design affects identification with the product. Two independent raters coded the CD case designs made by the participants in the high-design participation conditions. The CD cases from the low-design conditions only listed the songs, without any particular design, and therefore were not rated. The raters used the same scales from Study 2 to evaluate how self-expressive the designs were (the design is self expressive, $\alpha =$ .93; one can get a sense of the designer's personality from this, $\alpha =$ .88; it reflects the designer's self-image, $\alpha = .64$). The ratings were averaged to form a self-expressiveness index. Identification and affective commitment were regressed onto the index separately. As expected, self-expressiveness of the design predicted the level of identification ($\beta = .51, t = 6.76, p < .001$). It also affected the level of affective commitment ($\beta = .41, t = 5.25, p < .001$).

5.3. Study 3B: method

The above studies presented the manipulation check questions before the main dependent variable questions. Thus we cannot rule-out the possibility that the manipulation checks might have influenced participants' responses to product evaluation, as well as identification and affective commitment measures, by priming them to the physical and/or design effort that they have invested into the product. Thus, Study 3A was rerun (n = 144) using the process of designing and physically making a picture frame (instead of a CD with its case), with the measures of manipulation check questions asked at the end. The participants were recruited at a large private university in Turkey. All measurement scales from Study 3A were translated into Turkish using a back-translation procedure (Brislin, 1976; Cavusgil & Das, 1997).

5.4. Study 3B: results

The results replicated Study 3A findings. An ANOVA on product evaluation revealed significant main effects of realization ($F(1, 140) = 9.19, p < .01$) and design ($F(1, 140) = 17.25, p < .001$) participations, as well as a significant interaction effect ($F(1, 140) = 4.14, p < .05$). Identical to Study 3A findings, simple effects tests indicated that during low levels of design participation, evaluation of the product was significantly more favorable when realization was high ($M = 5.34$) rather than low ($M = 4.24$) ($F(1, 140) = 10.9, p < .01$). However, during high levels of design participation, evaluation of the product did not differ between the high- ($M = 5.80$) and low- ($M = 5.59$) realization conditions ($F < 1$). Similarly, when realization participation was low, higher levels of design participation enhanced evaluation of the product ($F(1, 140) = 17.09, p < .001$). However, when realization participation was high, design participation did not enhance evaluation of the product ($F(1, 140) = 2.55, p = .11$).

Also identical to Study 3A, an ANOVA on identification revealed a significant main effect for design participation ($M_{low} = 2.32, M_{high} = 4.14; F(1, 140) = 53.46, p < .001$). However, the main effect of realization participation ($F(1, 140) = 1.28, p = .26$) and the interaction effect ($F < 1$) were not significant. As in Study 3A, an ANOVA on affective commitment revealed significant main effects for both realization ($F(1, 140) = 4.89, p < .05$) and design ($F(1, 140) = 27.56, p < .001$) participation. The interaction was not significant ($F(1, 140) = 1.41, p = .24$). Participants reported higher affective commitment in the high- ($M = 3.99$) than low- ($M = 3.46$) realization condition, and in the high- ($M = 4.42$) than low- ($M = 3.07$) design condition.[1]

---

[1] The complete analyses and results for Study 3B are available upon request.

The results indicate that the order of the manipulation check questions did not affect the Study 3A findings. Moreover, replication of our findings in two different countries and languages enhances the generalizability of the results.

5.5. Discussion

Studies 3A and 3B replicate findings from Studies 1 and 2, and provide convergent evidence that participation in different stages of self-production differentially affects how consumers relate to products. We find that participation during the realization stage enhances affective commitment, but not identification; whereas participation during the design stage enhances identification with the product, which in turn results in stronger affective commitment to the product. Finally, engaging in both stages of production does not create value for consumers over and above the main effects obtained for a high level of participation in either stage alone.

These results are consistent with the suggestion by Franke et al. (2010) that marginal effects of consumer participation may diminish as the level of contribution increases. It is possible that there is a saturation point beyond which higher levels of participation may be perceived as a cost rather than a value for consumers, and that moderate levels of consumer engagement provide the highest value. We speculate that there may be even an inverse-U shape relationship between the level of effort in self-production activities and valuation of the self-made products. Future research is needed to clarify this relationship.

6. General discussion

Our research focused on elucidating the psychological responses of consumers to specific stages of self-production activities. In particular, we sought to contribute to insights about self-made products by specifying two dimensions through which consumers may relate to them: identification and affective commitment. To the best of our knowledge, this is the first empirical study that investigates the dimensions through which consumers relate to self-made products. Our results indicated that affective commitment and identification are closely related, yet distinct, concepts that predict consumers' favorable evaluation of products.

We demonstrated that consumers feel a greater emotional bond (i.e., affective commitment) with the product when they physically invest themselves in the product during the realization stage. During the design stage, consumers form a cognitive bond (i.e., identification) with the product, when they are able to manipulate the product to symbolize their self-identity. Additionally, identification with the product enhances affective commitment to the finished end-product. This latter link suggests that participation in the design of products contributes to one's identification with products, a cognitive process, and identification then enhances one's affective commitment to products, an affective process. This identification-to-affective commitment sequence has been found in research on group identity and social identity within organizations (e.g., Ellemers et al., 1999). We contribute to a better understanding of affective commitment to products by showing that identification during the design stage, as well as consumers' physical construction of a product during the realization stage, can create a sense of emotional bond (affective commitment) to the product and thereby enhance evaluation of the self-made end-product. Our findings also contribute to the literatures on self-production, co-production, and do-it-yourself products by identifying psychological and social processes underlying consumer responses and the different dimensions through which consumers may relate to a self-made product at different stages of its production.

We offer insights that go beyond what has been uncovered in prior studies that have employed a variety of operationalizations of self-production. For example, Mochon, Norton, and Ariely (2012) and Norton et al. (2012) provided step-by-step directions to participants to make origami figures. Bendapudi and Leone (2003) asked participants

to consider situations where they select and then physically build the product (e.g., bookshelf, poster frame). Franke et al. (2010) required participants to virtually design a t-shirt, scarf, or a cell phone cover. Buechel and Janiszewski (2014) had participants engage in physical assembly of a simple craft kit (Winter Holiday Elf) and manipulate the timing of the customization decision (i.e., whether design occurs before or during realization); hence they did not systematically vary the level of participation during both design and realization stages as we did in four studies.

In our research, the design and realization stages were fully separated and examined individually (Studies 1 and 2), as well as combined and examined together (Studies 3A and 3B). Buechel and Janiszewski (2014) suggest that when the design and assembly activities are segregated (vs. integrated), consumers' valuations of the input kit materials decrease. However, they report no significant effects on valuations of the finished end-product. We speculate that the null effects were due to the fact that participants used input kits that they were unlikely to perceive as being part of their self-identity. Moreover, there was no actual separation of the physical assembly and customization decisions in their study. They manipulated the timing of the customization decisions when there was already physical assembly as opposed to the engagement in only one or two different stages of self-production, as in our studies. The present research shows that, in line with previous research, even when customization (design) decisions and assembly actions (realization) are performed separately, higher levels of participation, in fact, do lead to enhanced evaluation of the finished outcome product. In addition, we provided a theoretical basis for proposing and documenting how identification and affective commitment mediate the effect of the level of design and realization participation on evaluation of self-made products.

From a managerial viewpoint, both identification and physical construction can be used to create affective commitment to products and could be the target of marketing communications and activities designed to enhance affective commitment to products. Affective commitment not only enhances evaluation of the product but also lengthens its usage duration and increases the care a consumer shows for the product and, therefore, contributes to sustainable consumption (Nieuwenhuis, 2008). From a practical perspective, the relative significance of affective commitment to, and identification with, the product may vary by context. One factor that may affect the relative significance is the type of the product that is produced or whether the product will be used privately or publicly. Publicly (vs. privately) consumed products are signals of identity to the outside world. Hence, consumers may be more likely to publicly use products that they identify with, especially if they have high needs for self-expression. Consequently, identification with the product may turn out to be more important for managers especially if their products tend to be consumed publicly. On the other hand, in some contexts, affective commitment may provide greater motivation to consumers for promoting positive word-of-mouth.

Furthermore, we highlight the importance of encouraging consumers to take part in the production process physically during the realization stage. Previous research (Moreau & Herd, 2010; Deng, Hui, & Hutchinson, 2010) focuses mostly on self-design (e.g., creativity and choice), not physical engagement. Researchers have largely neglected to study the specific role of the realization stage in the production of products. Advancements in the online environment have been providing ever-increasing opportunities for design participation at the expense of the realization stage. However, we empirically show that participation during the realization stage is distinct from participation during the design stage and can enhance product evaluations as much as design participation.

Finally, our results suggest that engaging consumers even in a limited amount of physical assembly rather than having them build from scratch, or asking them to choose a limited number of features, instead of having them design from scratch, may have similar effects in terms of enhancing evaluation of the finished end-products. Hence, from a managerial perspective, investing in relatively easy-to-implement systems

that enable consumers to participate in even limited amounts of self-production activities (either in the realization or the design stage) is likely to create customer value and prove to be useful for the firms to do. For example, while marketing ready-to-assemble furniture such as an IKEA bookcase, companies may provide stickers or special pencils that allow consumers to write on the product and enable consumers to transform products to symbols of self-identity.

7. Limitations and future research

Our research presents several other interesting questions that have considerable practical implications. For example, personality variables (i.e., creativity, liking to work with one's hands) may moderate the value created through different stages of production and are domains ripe for exploration. Consumers who enjoy working with their hands and expressing themselves through physical labor may identify with self-made products even if they only participate in the realization stage, since manual labor is part of their self-identity. Our studies involved university students who may have higher needs for cognition than many non-students and may value design more than assembly or customized construction opportunities. Hence, design participation may have been relatively more important for our population and may add more value than craftsmanship or manual effort.

In our studies, in order to equate time spent with the product, we used filler tasks in the control and low-level of participation conditions. In the high-level of participation conditions, participants evaluated the finished product right after the haptic experience. Therefore, for consumers in the control and low-level of participation conditions, the gap in time between the haptic experience and the evaluation of the product could have affected these consumers. However, in all conditions, participants were not limited in their handling of the product insofar as it was always in front of them while they worked on the filler tasks. They could look at and touch the products whenever they wanted, as well as experience the visual cue of the product to remind them of actually experiencing it a few minutes earlier. We submit that using filler tasks reflects a conservative test of our hypothesized relationships, since consumers in real-world settings would presumably spend less time with the product than what we imposed in the control and low-level of participation conditions as a result of the filler task (Strahilevitz & Loewenstein, 1998). Future research is needed to clarify whether eliminating or changing the duration of filler tasks affects consumers' evaluation of self-made products.

We measured only one consequence of the mediating mechanisms (i.e., identification and affective commitment), that is product evaluation. However, by scrutinizing product evaluation, we provide results for a central variable that determines other variables. For example, in attitude theory, evaluations are important, often the most important, antecedents to decisions, intentions, and behavior. Future research may address how affective commitment and identification affect other variables such as word-of-mouth, loyalty, length of product usage, and satisfaction with the performance of the product.

The theoretical framework that we used in our research was adopted from the social identity literature that emphasizes the distinction among cognitive identification, affective commitment, and collective self-esteem (an evaluative component) derived from the target object (i.e., group). Our research explored only two dimensions (identification and affective commitment) in this relationship. Future research might investigate whether and how self-production contributes to self-esteem that is derived from the product. Products can be used to build and restore the sense of self (Belk, 1988; Gao, Wheeler, & Shiv, 2009) as the extended-self literature indicates. As our work reveals, self-made products may be symbols of identity. To the extent that they serve as constant reminders of one's sense of self, owning or using self-made products may make consumers feel good about themselves, feel smart and confident, and help them gain respect from others, especially if the products symbolize positive aspects of identity. This implies

an interaction between the self and the product. First, the person changes the image of the product to reflect his or her identity (identifies with the product); later on, the product changes the person's sense of self-esteem as it reminds one of him- or herself and his or her actions. In fact, in the context of organizational behavior, Bergami and Bagozzi (2000) found that identification with the organization determined organization-based self-esteem. Hence, identification with, but not necessarily affective commitment to, the product is likely to affect product-based self-esteem. This too is a fruitful direction for future research.

Additionally, future research is needed to investigate other conditions under which identification and affective commitment dimensions and their functioning differ. We expect that the current versus ideal self may have unique effects. Consumers may identify with products that reflect their current identity; however, affective commitment to the product may depend on the extent to which that part of identity is perceived positively or desirable.

Finally, we found that engaging consumers in both stages of self-production did not create value over and above a high level of participation in either stage alone. In our studies, the participation level in the production process was necessarily limited due to experimental constraints. Higher levels of participation in both stages may result in additive or multiplicative effects depending on the circumstances. For instance, designing and building a home may result in a much more favorable evaluation of the final product than that resulting from building the home without participating in actual construction. Nevertheless, situations involving such extreme levels of production participation are likely to be limited, given the high cost of investing time and effort in these situations and infrequent opportunities to do so. Hence, we expect our results to hold in many everyday consumption situations.

Acknowledgments

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ijresmar.2014.05.003. Estimation code for this article can be found online at http://www.runmycode.org. Interested scholars may contact either the corresponding author or IRJM's editorial office in order to request the dataset.

References

Aaker, J. (1997). Dimensions of brand personality. *Journal of Marketing Research, 34*(3), 347–356.

Aaker, J., Fournier, S., & Brasel, S. A. (2004). When good brands do bad. *Journal of Consumer Research, 31*(1), 1–16.

Ahuvia, A.C. (2005). Beyond the extended self: Loved objects and consumers' identity narratives. *Journal of Consumer Research, 32*(June), 171–184.

Ainsworth, M.D. S., & Bowlby, J. (1991). An ethological approach to personality development. *American Psychologist, 46*, 333–341.

Allen, N. J., & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of Occupational Psychology, 63*, 1–18.

Anisfeld, E., Casper, V., Nozyce, M., & Cunningham, N. (1990). Does infant carrying promote attachment? An experimental study of the effects of increased physical contact on the development of attachment. *Child Development, 61*, 1617–1627.

Atakan, S. S., Bagozzi, R., & Yoon, C. (2014). Make it your own: How process valence and self-construal affect evaluation of self-made products. *Psychology & Marketing, 31*(6), 451–468.

Bagozzi, R. P., & Heatherton, T. F. (1994). A general approach to representing multifaceted personality constructs: Application to state self-esteem. *Structural Equation Modeling, 1*(1), 35–67.

Ball, A.D., & Tasaki, L. H. (1992). The role and measurement of attachment in consumer behavior. *Journal of Consumer Psychology, 1*(2), 155–172.

Batra, R., Ahuvia, A., & Bagozzi, R. P. (2012). Brand love. *Journal of Marketing, 76*, 1–16.

Belk, R. (1988). Possessions and the extended self. *Journal of Consumer Research, 15*(2), 139–168.

Bell, R. A., Daly, J. A., & Gonzalez, M. C. (1987). Affinity-maintenance in marriage and its relationship to women's marital satisfaction. *Journal of Marriage and Family, 49*(2), 445–454.

Bendapudi, N., & Leone, R. P. (2003). Psychological implications of customer participation in co-production. *Journal of Marketing, 67*(Jan), 14–28.

Bergami, M., & Bagozzi, R. P. (2000). Self-categorization, affective commitment and group self-esteem as distinct aspects of social identity in the organization. *British Journal of Social Psychology, 39*, 555–577.

Bhattacharya, C. B., & Sen, S. (2003). Consumer-company identification: A framework for understanding consumers' relationships with companies. *Journal of Marketing, 67*(2), 76–88.

Bloch, P. H. (1995). Seeking the ideal form: Product design and consumer response. *Journal of Marketing, 59*(3), 16–29.

Bowlby, J. (1969). *Attachment and loss: Vol. 1. Attachment.* London: Hogarth Press.

Brislin, R. W. (1976). Comparative research methodology: Cross-cultural studies. *International Journal of Psychology, 11*(3), 215–229.

Buechel, E. C., & Janiszewski, C. (2014). A lot of work or a work of art: How the structure of a customized assembly task determines the utility derived from assembly effort. *Journal of Consumer Research, 40.* http://dx.doi.org/10.1086/673846.

Cavusgil, S. T., & Das, A. (1997). Methodological issues in empirical cross-cultural research: A survey of the management literature and a framework. *Management International Review, 37*, 71–96.

Csikszentmihalyi, M., & Rochberg-Halton, E. (1981). *The meaning of things: Domestic symbols and the self.* Cambridge University Press.

Deng, X., Hui, S. K., & Hutchinson, J. W. (2010). Consumer preferences for color combinations: An empirical analysis of similarity-based color relationships. *Journal of Consumer Psychology, 20*(4), 476–484.

Dutton, J. E., Dukerich, J. M., & Harquail, C. V. (1994). Organizational images and member identification. *Administrative Science Quarterly, 39*, 239–263.

Ellemers, N., Kortekaas, P., & Ouwerkerk, J. W. (1999). Self-categorisation, commitment to the group and group self-esteem as related but distinct aspects of social identity. *European Journal of Social Psychology, 29*, 371–389.

Fisher, J.D., Rytting, M., & Heslin, R. (1976). Hands touching hands: Affective and evaluative effects of an interpersonal touch. *Sociometry, 39*, 416–421.

Franke, N., Schreier, M., & Kaiser, U. (2010). The "I designed it myself" effect in mass customization. *Management Science, 56*(1), 125–140.

Fuchs, C., Prandelli, E., & Schreier, M. (2010). The psychological effects of empowerment strategies on consumers' product demand. *Journal of Marketing, 74*, 65–79.

Gao, L., Wheeler, S.C., & Shiv, B. (2009). The "shaken self": Product choices as a means of restoring self-view confidence. *Journal of Consumer Research, 36*, 29–38.

Grubb, E. L., & Grathwohl, H. L. (1967). Consumer self-concept, symbolism, and market behavior: A theoretical approach. *Journal of Marketing, 31*, 22–27.

Hayes, A. F., & Preacher, K. J. (2013). *Statistical mediation analysis with a multicategorical independent variable.* British Journal of Mathematical and Statistical Psychology. http://dx.doi.org/10.1111/bmsp.12028.

Hayes, A. F., Preacher, K. J., & Myers, T. A. (2011). Mediation and the estimation of indirect effects in political communication research. In E. P. Bucy, & R. L. Holbert (Eds.), *Sourcebook for political communication research: Methods, measures, and analytical techniques* (pp. 434–465). New York: Routledge.

Hazan, C., & Shaver, P. R. (1994). Attachment as an organizational framework for research on close relationships. *Psychological Inquiry, 5*, 68–79.

Hu, L., & Bentler, P.M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424–453.

Kleine, S. S., Kleine, R. E., III, & Allen, C. T. (1995). How is a possession "me" or "not me"? Characterizing types and an antecedent of material possession attachment. *Journal of Consumer Research, 22*(Dec), 327–343.

Kleine, R. E., III, Kleine, S. S., & Kernan, J. B. (1993). Mundane consumption and the self: A social-identity perspective. *Journal of Consumer Psychology, 2*(3), 209–235.

Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of Consumer Psychology, 22*(3), 332–351.

Lengnick-Hall, C. A. (1996). Customer contributions to quality: A different view of the customer-oriented firm. *The Academy of Management Review, 21*(3), 791–824.

Liu, W. (2008). Focusing on desirability: The effect of decision interruption and suspension on preferences. *Journal of Consumer Research, 35*(4), 640–652.

McClelland, D. (1951). *Personality.* New York: Holt, Rinehart, and Winston.

Mochon, D., Norton, M., & Ariely, D. (2012). Bolstering and restoring feelings of competence via the IKEA effect. *International Journal of Research in Marketing, 29*, 363–369.

Moreau, C. P., & Herd, K. B. (2010). To each his own? How comparisons with others influence consumers' evaluations of their self-designed products. *Journal of Consumer Research, 36*(5), 806–819.

Nieuwenhuis, P. (2008). From banger to classic — a model for sustainable car consumption? *International Journal of Consumer Studies, 32*, 648–655.

Norton, M., Mochon, D., & Ariely, D. (2012). The IKEA effect: When labor leads to love. *Journal of Consumer Psychology, 22*, 453–460.

Park, C. W., MacInnis, D. J., & Priester, J. (2008). Brand attachment: Constructs, consequences, and causes. *Foundations and Trends in Marketing, 1*(3), 191–230.

Patterson, M. L., Powell, J. L., & Lenihan, M. G. (1986). Touch, compliance, and interpersonal affect. *Journal of Nonverbal Behavior, 10*(1), 41–50.

Peck, J., & Shu, S. B. (2009). The effect of mere touch on perceived ownership. *Journal of Consumer Research, 36*(Oct), 434–447.

Pierce, J. L., Kostova, T., & Dirks, K. T. (2003). State of psychological ownership: Integrating and extending a century of research. *Review of General Psychology, 7*(1), 84–107.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*(3), 879–891.

Prelinger, E. (1959). Extension and structure of the self. *Journal of Psychology*, *47*(1), 13–23.

Reed, A., Aquino, K., & Levy, E. (2007). Moral identity and judgments of charitable behaviors. *Journal of Marketing*, *71*, 178–193.

Strahilevitz, M.A., & Loewenstein, G. (1998). The effect of ownership history on the valuation of objects. *Journal of Consumer Research*, *25*(3), 276–289.

Tajfel, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations.* London: Academic Press.

Thomson, M., MacInnis, D. J., & Park, C. W. (2005). The ties that bind: Measuring the strength of consumers' emotional attachments to brands. *Journal of Consumer Psychology*, *15*(1), 77–91.

Troye, S. V., & Supphellen, M. (2012). Consumer participation in coproduction: "I made it myself" effects on consumers' sensory perceptions and evaluations of outcome and input product. *Journal of Marketing*, *76*(2), 33–46.

Turner, J. C. (1985). Social categorization and the self-concept: A social cognitive theory of group behavior. In E. J. Lawler (Ed.), *Advances in group processes: Theory and research*, *Vol. 2.* (pp. 77–122). Greenwich, CT: JAI Press.

Van Raaij, W. F., & Pruyn, A. Th. H. (1998). Customer control and evaluation of service validity and reliability. *Psychology and Marketing*, *15*, 811–832.

Wolf, J. R., Arkes, H. R., & Muhanna, W. A. (2008). The power of touch: An examination of the effect of duration of physical contact on the valuation of objects. *Judgment and Decision Making*, *3*(6), 476.

Yoon, C., Gutchess, A. H., Feinberg, F., & Polk, T. A. (2006). A functional magnetic resonance imaging study of neural dissociations between brand and person judgments. *Journal of Consumer Research*, *33*(1), 31–40.

Full Length Article

Impact of component supplier branding on profitability

CrossMark

Stefan Worm [a,*], Rajendra K. Srivastava [b,c,1]

[a] HEC Paris, School of Management, Department of Marketing, 1, Rue de la Libération, 78351 Jouy-en-Josas France
[b] Singapore Management University, 81 Victoria Street, Singapore 188065, Singapore
[c] Sim Kee Boon Institute for Financial Economics at Singapore Management University

ARTICLE INFO

ABSTRACT

In recent years, many business-to-business (B2B) component supplier (CS) firms have added branding to their marketing toolbox. By extending the logic of ingredient branding to B2B components, they aim to create "pull" from B2B end customers by building a strong CS brand image among their customers' customers. In contrast with the established "push" approach of building strong relationships with original equipment manufacturers (OEMs), it is unclear whether and under which conditions CS branding is a worthy strategy. On the one hand, anecdotal evidence suggests that suppliers can leverage strong CS brand image in negotiations with increasingly powerful OEMs to enhance their financial performance. On the other hand, many B2B managers believe that branding does not work in their industry context and erodes profitability. We build a data set consisting of survey measures and archival data across a broad set of industries. Our results indicate that the financial outcomes of CS branding largely depend on the characteristics of the CS and OEM industries. Unlike dyadic OEM–CS relationships, which enhance profitability invariably across industry contexts, CS branding is effective only in well-defined situations. CS branding initiatives can enhance return in CS industries with substantial levels of product differentiation and technology intensity. However, unfavorable results may arise in industry contexts in which OEM–end customer relationships or OEM brands are important.

© 2014 Published by Elsevier B.V.

1. Introduction

Many suppliers in business-to-business (B2B) or industrial markets have begun investing systematically in their brands, with the idea that branding strategies can help them stabilize or grow their profits in increasingly competitive markets (Wise & Zednickova, 2009). The most important branding option for B2B component suppliers (CSs) is CS branding, which represents an extension of the ingredient branding approach (Wiersema, 2012). CS branding can be applied to B2B components embedded in original equipment manufacturer (OEM) products that, in turn, are marketed to B2B end customers (i.e., OEMs' customers) (Ghosh & John, 2009). The strategy aims to create "pull" from B2B end customers for CS products by building a strong CS brand image.

The growing interest in CS branding is noteworthy because B2B marketers have traditionally relied on direct (i.e., "push") marketing strategies and focused on building strong relationships with OEMs. The goal of these relationship-marketing efforts is to create superior value for OEMs by providing additional benefits or reducing costs (Cannon & Homburg, 2001; Frazier, Spekman, & O'Neal, 1988; Tuli, Bharadwaj, & Kohli, 2010; Ulaga & Eggert, 2006). In contrast with these tried-and-true approaches, it is not clear whether and in which situations pull created through CS branding affects CS performance when end customers are businesses (i.e., in "B2B2B" markets). On the one hand, anecdotal evidence from popular examples such as Intel (personal computers), Bose (automobile audio systems), and DuPont (Lycra, apparel) promises superior margins when end customers are consumers (i.e., in "B2B2C" markets; Kotler & Pfoertsch, 2010). On the other hand, many B2B marketers and senior managers remain skeptical about whether this approach can be successfully transferred to their own business contexts (ISBM/BMA, 2005). Indeed, failures often occur, particularly in CS industries with limited differentiation and innovation.

Despite an emerging body of literature on B2B brands (e.g., Cretu & Brodie, 2007; Homburg, Klarmann, & Schmitt, 2011; Wuyts, Verhoef, & Prins, 2009; Zablah, Brown, & Donthu, 2010), empirical research has paid limited attention to financial performance outcomes of CS branding. This is surprising not only given the significant resources required to build and sustain brand image (Keller, 1993) but also because CS firms may be less efficient in branding from a lack of experience. To the best of our knowledge, only the study by Ghosh and John (2009) looks at CS brands, but it does not examine financial outcomes.

* Corresponding author. Tel.: +33 139 67 9601.
E-mail addresses: worm@hec.edu (S. Worm), rajs@smu.edu.sg (R.K. Srivastava).
[1] Tel.: +65 6828 0591.

Against this backdrop, this study explores the following research questions:

1. How does CS branding in B2B markets affect CS profitability?
2. Under which conditions does CS branding pay off for CSs?
3. How does the impact of CS brand image on CS profitability compare to that of value created for OEMs with traditional relationship marketing?

This empirical study combines survey measures with archival data on supplier financial performance and industry-level competition, covering a broad range of CS and OEM industries. The results suggest that CSs can leverage strong CS brands to maintain or grow their profitability. However, several critical environmental factors related to the CS and OEM industries moderate the impact of CS brand image on CS financial performance. For example, CSs in industries characterized by differentiated products and technology intensity can leverage the CS brand asset successfully. In contrast, CSs selling into OEM industries in which OEM–end customer relationships and OEM brands are highly important find it more difficult to leverage strong CS brands (i.e., CS brands compete with OEM brands). The results also show that OEM–CS relationship quality and OEMs' value perceptions of CS have a positive, non-conditional effect on CS financial performance.

This study responds to calls for additional research on the firm-level, bottom-line financial outcomes of B2B marketing strategies (ISBM, 2010; Marketing Science Institute, 2008; Wiersema, 2012). The results advance the understanding on how strong CS brands help CSs cope with increasing marketplace pressures and thus complement existing research on how consumer brand image stabilizes financial outcomes (Johansson, Dimofte, & Mazvancheryl, 2012). Our findings also contribute to the emerging contingency perspective in the marketing discipline, which examines the environmental conditions under which marketing instruments and market-based assets lead to financial performance (e.g., Reibstein, Day, & Wind, 2009). The findings help the B2B marketers in CS firms to determine whether, in their situational context, CS branding is a promising strategy to invest limited resources. Contrary to managerial intuition, CS brand building is neither a suitable instrument to cope with commoditization in a CS industry nor a tool to help deal with OEMs, which "own" end customers through OEM–end customer relationships and OEM brands. In both cases, investing in value creation in the relationship with OEMs yields better outcomes because it is effective regardless of context. Furthermore, our findings should help B2B marketers better understand and communicate the contribution of their branding actions to senior management.

In the remainder of the article, we first review the relevant literature and highlight our contributions. To answer the research questions, we then develop a theoretical framework (1) to examine the mechanisms linking the OEM–CS relationship and CS brand image with CS financial performance and (2) to identify contingency conditions under which CS branding strategies are likely to be more or less productive. Next, we report our empirical study and estimate the range of gains and losses associated with CS branding initiatives. We conclude with a discussion of the results.

2. Related literature

Table 1 summarizes selected research on performance outcomes of brands in B2B settings. Extant research finds that B2B brands have an impact on a range of performance indicators, including buyer intentions and attitudes (Cretu & Brodie, 2007; Hutton, 1997; Wuyts et al., 2009; Zablah et al., 2010), relational outcomes (Ghosh & John, 2009), and financial performance (Homburg et al., 2011). In addition, several studies find that outcomes of B2B brands are contingent on the situational context (e.g., Zablah et al., 2010). Most research relies on brand awareness or brand image as customer mindset brand metrics. Zablah et al. (2010) and Ghosh and John (2009) use measures of brand strength that are consequences of brand image and awareness.

Table 1
Overview of selected B2B branding studies.

Study	Brand metric (IV)	Controlling for B2B relationship marketing metrics		Performance metric (DV)		Examine CS branding	Contingency of brand performance	Sample	Data
		Relationship quality	Customer-perceived value	Buyer intentions & attitudes	Financial performance metric				
Homburg et al. (2011)	• Brand awareness				• Return on sales (self-reported)		Yes	Multiple B2B industries	• Survey (suppliers)
Zablah et al. (2010)	• Brand preference			• Brand importance			Yes	Multiple B2B industries	• Survey (buyer firms)
Ghosh and John (2009)	• Differentiation			• Perceived Supplier opportunism		Yes		Three B2B industry sectors	• Survey (OEMs)
Wuyts et al. (2009)	• Brand awareness	Yes		• Consideration • Buyer choice			Yes	Single B2B industry (market research providers)	• Conjoint survey (buyer firms)
Cretu and Brodie (2007)	• Brand image		Yes	• Attitudinal loyalty				Single B2B industry (shampoo for hair salons)	• Survey (buyer firms)
Aaker and Jacobson (2001)	• Brand image				• Stock return • Return on equity (both archival)			Single B2B industry (computers)	• Archival data (financials) • Survey (buyer firms)
Hutton (1997)	• Brand awareness • Brand preference			• Willingness to pay • Attitudinal loyalty			Yes	Three B2B products (office supplies)	• Survey (buyer firms)
Current study	Brand image	Yes	Yes		• Return on sales (archival)	Yes	Yes	Multiple B2B industries (manufacturing)	• Archival data (financials, industry data) • Multiple-informant survey (mindset data)

Despite this progress, understanding of how brands affect B2B firm performance is still in an early stage. We identify five aspects in the B2B brand literature that warrant attention. First, most research does not examine CS branding. Ghosh and John's (2009) study, which concludes that CS brand contracts can act as a safeguard for CS specific investments, is a noteworthy exception. However, their research focuses on opportunism as an outcome but not on firm performance. Likewise, ingredient-branding research in the consumer behavior literature has examined end-customer perceptions rather than ingredient supplier performance outcomes (e.g., Desai & Keller, 2002; Park, Jun, & Shocker, 1996).

Second, little research has examined financial performance as an outcome, with the noteworthy exceptions of Homburg et al. (2011) and Aaker and Jacobson (2001). Financial performance outcomes are important because they also reflect the considerable cost of building and sustaining brands, enabling a more accurate assessment of the return on B2B brand strategies.

Third, scant research has assessed the impact of B2B brands in the context of relationship marketing efforts. Yet evidence shows that B2B brands may not have a direct impact on supplier performance beyond that of established relationship marketing metrics. Wuyts et al. (2009), who control for relationship quality, find that brand awareness affects buyers' consideration sets but not their final choice. Similarly, Cretu and Brodie (2007) find that brand image affects customer loyalty only indirectly through customer perceptions of value and quality. Unfortunately, the B2B relationship literature has not explored the impact of brands as predictors of supplier performance, while largely focusing on relational aspects, idiosyncratic assets, and customer-perceived value (e.g., Palmatier, Dant, & Grewal, 2007; Palmatier, Dant, Grewal, & Evans, 2006; Ulaga & Eggert, 2006).

Fourth, many studies on B2B brands are limited to specific industries or products. Therefore, their findings may not transfer to other situational contexts, because converging evidence suggests that the impact of B2B brands on performance is contingent on industry-specific context factors (Homburg et al., 2011; Zablah et al., 2010). Fifth, the bulk of extant research draws on single-informant survey data. Thus, common method variance ultimately cannot be ruled out as a competing explanation for the results obtained, especially given that B2B managers may have biased perceptions of the impact of their brands (Davis, Golicic, & Marquardt, 2008).

The current study addresses these research gaps. We examine the impact of CS brand image on CS profitability while controlling for OEMs' value perceptions and relationship quality. We examine how this impact is contingent on situational contexts using a combination of archival and survey measures collected across multiple B2B industries.

3. Theoretical framework and hypotheses

Fig. 1 illustrates our study's triadic setting: CSs manufacture components that are subsequently incorporated into OEMs' products, which in turn sell to B2B end customers. As the figure shows, CS financial performance is a function of the OEM–CS dyadic relationship (through push) and CS brand image perceptions among end customers (through pull). Thus, CS marketers can invest in relationship marketing activities with OEMs (i.e., direct influence strategies) and/or invest in building a CS brand with end customers (i.e., indirect influence strategies). The former strengthen OEM–CS relationships and, in analogy with push strategies, focus on how CSs can help improve OEM products, enhance margins, and reduce costs and risks for OEMs. Conversely, by building the CS brand (i.e., developing a strong CS brand image among end customers to foster brand loyalty), CSs can also bring end customers to OEMs, in a manner similar to pull strategies in consumer markets. While CSs' relational value creation activities (push) can help improve supply-chain processes and OEM products, strong CS brands (pull) can facilitate OEMs' end-product sales, growth, and end customers' willingness to pay. B2B research has mostly focused on the path from relationship marketing to CS performance (Palmatier et al., 2006). We examine
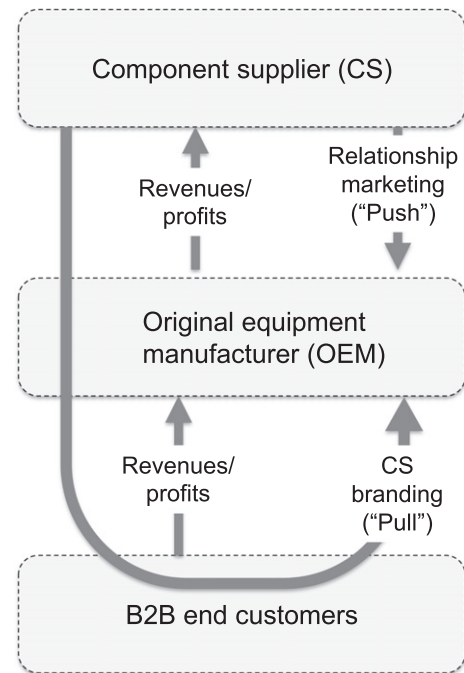

Fig. 1. Research setting.

the impact of both OEM–CS relationships and CS brand image on CS financial performance simultaneously.

We draw on the resource-based view (RBV) to examine whether CS brand image (i.e., end customers' perception of the CS brand) represents a valuable market-based asset that enables CSs to protect or grow profitability (Srivastava, Shervani, & Fahey, 1998). We develop our theoretical framework in four steps (see Fig. 2). First, we introduce our dependent variable. Second, we examine the effect of OEM–CS relationship quality on CS financial performance. Third, we analyze how CS brand image can affect CS financial performance by giving CSs a more powerful position in their negotiations with OEMs. Fourth, we examine the extent to which the direct effect of CS brand image on CS financial performance is contingent on situational factors related to CS and OEM industries.

3.1. Choice of dependent variable

OEMs have become more aggressive in their procurement of components, as strategic cost management and target costing have become common practice (Cooper & Slagmulder, 1999). They increasingly induce competition among CSs, which in turn face the challenge of either gaining key supplier status or being demoted to the role of backup supplier (Ulaga & Eggert, 2006). As a consequence, CSs face strong pressure to reduce prices on a continuous basis, resulting in ongoing erosion of CS profit margins and higher variability in profits.[2] Because firm valuations reflect investors' expectations of the level of and volatility in future cash flows (Srivastava et al., 1998), CSs' ability to shield themselves from these pressures to maintain or grow the level of profitability is of major concern. Thus, we focus our analysis on profitability growth (i.e., change in profitability) as an indicator of CS financial performance. This choice of dependent variable is in line with recent calls in the literature to demonstrate marketing's financial impact at the corporate level (Fang, Palmatier, & Grewal, 2011; Tuli & Bharadwaj, 2009; Wiersema, 2012). Because many B2B CSs are privately owned, we operationalize our measures using an accounting-based metric: return on sales (ROS, also known as "operating margin"). ROS is among the most important

---

[2] For the industries represented in our sample, return on sales on average decreased by 0.5 percentage points between 2005 and 2010; based on an average of 6 percentage points, this is a decrease of 12%.
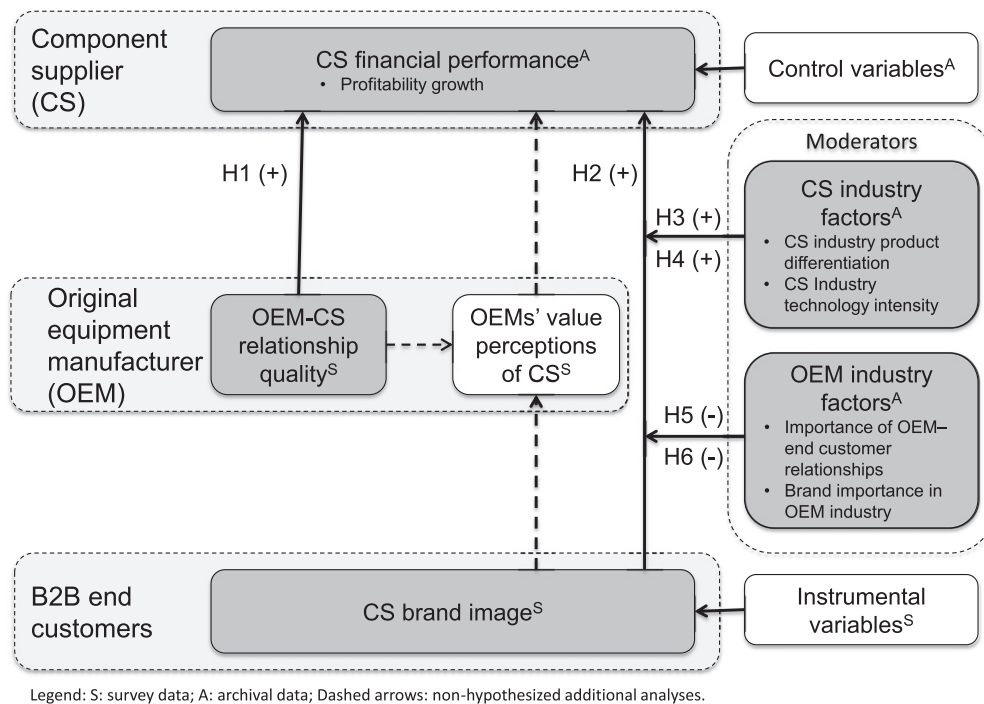
Legend: S: survey data; A: archival data; Dashed arrows: non-hypothesized additional analyses.

Fig. 2. Conceptual model.

financial ratios for marketing managers (Mintz & Currim, 2013), and previous research has frequently used it as a performance measure (e.g., Homburg, Artz, & Wieseke, 2012; Homburg et al., 2011).

3.2. Effect of OEM–CS relationships on CS profit

The B2B marketing literature has traditionally taken the perspective that CSs can compete more effectively by building strong relationships with OEMs (Palmatier, Dant, & Grewal, 2007). Following Palmatier et al. (2006), we define OEM–CS relationship quality as an overall assessment of the strength of the relationship, as reflected by OEMs' trust in CSs and commitment to the relationship. Good relationship quality positively influences relational behaviors (Morgan & Hunt, 1994), thereby facilitating collaboration between OEMs and CSs (Anderson & Narus, 1990; Doney & Cannon, 1997). In turn, enhanced collaboration promotes value creation in the relationship "beyond that which each party could achieve separately" (Palmatier et al., 2006, p. 140). For example, when relationship quality is high, OEMs are more willing to grant CSs access to private internal information, enabling CSs to identify and develop more effective customized solutions that create superior value for OEMs (Srivastava, Shervani, & Fahey, 1999; Tuli, Kohli, & Bharadwaj, 2007). The ability to create perceived value for OEMs, in turn, represents a major driver of CS financial performance in B2B markets (Anderson & Narus, 2004; Palmatier, Scheer, & Steenkamp, 2007; Payne & Frow, 2005; Ulaga & Eggert, 2006). When CSs create value for OEMs by increasing benefits or reducing costs, they reduce the attractiveness of available alternative CSs (Kelley & Thibaut, 1978). In addition, both parties share this extra value (Bagozzi, 1974), which in turn increases OEMs' willingness to pay higher prices and thus eases the pressure on CS profitability or enables profitability growth. Furthermore, OEMs show greater loyalty to CSs that provide high value, stabilizing profitability in the long run. Taken together, these arguments suggest that better OEM–CS relationship quality enhances CS profitability growth. Thus:

Hypothesis 1. The better the quality of the OEM–CS relationship, the higher is the CS's profitability growth.

3.3. Effect of CS brand image on CS profit

We define CS brand image as end customers' perceptions of the CS brand (Keller, 1993). Homburg et al. (2011) rely on brand awareness as a B2B brand metric, whereas we expect CS brand image to be a stronger measure of the CS brand asset because it is more vulnerable to customer sanctions (Rao, Qu, & Ruekert, 1999; Wernerfelt, 1988). We first establish how CS brand image affects the market performance of OEM offerings. Building on this notion, we then examine how CS brand image affects CS financial performance by giving the CS a more powerful position relative to OEMs.

3.3.1. CS brand image and market performance of OEM offerings

We first discuss the effect of CS brand image on OEM market performance to be able to understand its effect on CS performance. Consumer research has shown that the presence of a strong ingredient brand enhances end customers' product perceptions and performance because it acts as a quality signal (Desai & Keller, 2002; Park et al., 1996; Rao et al., 1999; Simonin & Ruth, 1998; Swaminathan, Reddy, & Dommer, 2012). Ghosh and John (2009) contend that this principle extends to B2B components, for which CS brands can enhance the differentiation of OEM offerings in the eyes of B2B end customers. According to prior research, we expect that CS brands facilitate the buying process for OEMs' customers (i.e., end customers) by reducing information costs for decision makers and lowering their perceived risk of purchase (Erdem & Swait, 1998; Homburg et al., 2011; Zablah et al., 2010). In summary, we expect OEM offerings to achieve better market performance among B2B end customers if they incorporate high-image CS brands (e.g., through enhanced customer satisfaction, faster adoption of innovation, more effective customer acquisition, and better retention). On the cost side, OEMs can reduce their own marketing expenses because the CS partly takes charge of marketing to end customers.

3.3.2. Effect of CS brand image on CS profitability

A strong CS brand image can also enhance the negotiation power of CSs relative to OEMs, thus preventing the erosion of CS margins or helping

CSs grow margins. According to Beier and Stern (1969), CS power is based on the OEM's dependence on the CS. Such dependence arises when a CS controls important and critical resources that OEMs need to achieve their goals and for which few alternative sources of supply exist (Beier & Stern, 1969; Buchanan, 1992). According to the relational perspective of the RBV (Dyer & Singh, 1998), CS brand image is a CS resource that the OEM accesses through the relationship.

As we discussed previously, strong CS brands can drive the market performance of OEM offerings in several ways. When the achievement of OEMs' market performance goals hinges on the availability of a strong CS brand, CS brand image becomes a critical resource that increases CSs' power over OEMs (Shervani, Frazier, & Challagalla, 2007). CSs can benefit from this elevated power as it helps balance power in the OEM–CS relationship (Antia & Frazier, 2001; Beier & Stern, 1969). CS power may manifest in a greater willingness of OEMs to comply with the CS's demands because dependent OEMs will be more tolerant of the CS's use of coercive strategies (Beier & Stern, 1969; Gundlach & Cadotte, 1994; Shervani et al., 2007). Thus, strong CS power enables CSs to negotiate more aggressively with OEMs to enforce higher prices and moderate levels of service, thereby maintaining or increasing margins and profits (Ghosh & John, 2009). In turn, OEMs will make less use of coercive strategies with more powerful CSs, which lowers transaction costs in the relationship and ensures that CSs obtain fairer, sustained, and stable returns on their investments (Frazier, Gill, & Kale, 1989). Furthermore, CSs will be more difficult to replace if they provide important resources to OEMs (Buchanan, 1992). Thus, strong CS brands enhance retention of OEMs and increase return on their customer acquisition investments. In summary, CS power enables CSs to better maintain or increase price and profitability levels over time. Thus:

Hypothesis 2. The stronger a CS's brand image, the higher is the CS's profitability growth.

Note that building and sustaining a strong CS brand image requires considerable and continuous financial commitment (Keller, 1993; Webster & Keller, 2004). In addition, the risk of inefficient utilization of brand-building budgets is high in B2B industries, where the majority of firms have less experience or lower levels of capabilities in managing brands than their B2C counterparts. Thus, the favorable direct impact of CS brand image on CS profit can only occur in industry contexts where it prevails over the costs of CS branding. The net gain from CS brand image is a function of industry-level moderators, which we introduce next.

3.4. Industry-level moderators of the effect of CS brand image on CS financial performance

We investigate the role of four industry-level context variables: CS industry product differentiation, CS industry technology intensity, OEM–end customer relationship importance in the OEM industry, and OEM brand importance in the OEM industry. Two reasons motivated the choice of moderators. First, research in strategy has long been built on the axiom that no strategy is universally superior to or independent of the environmental context, thus calling for a contingency view (Venkatraman, 1989). In line with this, research on the RBV indicates that the value that firm assets and strategies create is dependent on the external market environment in which a firm operates (Barney, 2001; Black & Boal, 1994; Eisenhardt & Martin, 2000; Priem & Butler, 2001). Consequently, industry-level moderators are well established in the literature as critical moderators of the firm strategy–firm performance relationship (Kohli & Jaworski, 1990; Porter, 1980). Second, these moderators are particularly pertinent to the study context. As mentioned previously, research suggests that OEMs' dependence on a strong CS brand is a function of (1) the importance of the resource to OEMs and (2) the availability of alternatives to OEMs (Gundlach & Cadotte, 1994). Thus, our analysis focuses on industry characteristics that affect the importance and availability of alternatives for a strong CS brand.

3.4.1. Moderating effects of CS industry characteristics

We include two CS industry characteristics as moderators: product differentiation and technology intensity. Product differentiation exists in a CS industry if various suppliers' components have meaningful differences (Reimann, Schilke, & Thomas, 2010; Smith, 1956). When product differentiation is high, end customers face higher complexity of the decision task, increased information overload, and greater perceived purchase risk. The reduction in information cost and perceived risk by CS brand image thus exerts a greater influence on end customers' purchase decisions (Kirmani & Rao, 2000; Morgan & Hunt, 1994), making the presence of a strong CS brand more important. CS brands with a strong image are also more difficult to replace when components are differentiated because end customers learn that CS brands are cues for certain unique component characteristics (Van Osselaer & Janiszewski, 2001). Product differentiation also makes it more difficult for OEMs to replace a strong CS brand because doing so adds the effort of integrating a different component to the CS brand-related switching costs. When switching costs are high, OEMs are likely more tolerant of demands of a strong CS brand (Buchanan, 1992; Gundlach & Cadotte, 1994). In summary, CS industry product differentiation increases the importance of strong CS brands while making it more difficult for OEMs to replace them with an alternative supplier. Therefore, OEMs are more dependent on strong CS brands, which enables CSs to better leverage CS brand image to protect or enhance price and profitability levels. Thus:

Hypothesis 3. The positive effect of CS brand image on a CS's profitability growth is enhanced in CS industries with strong product differentiation.

Technology intensity is the degree to which CSs in an industry emphasize R&D (Bahadir, Bharadwaj, & Srivastava, 2008). Technology-intensive components evolve more rapidly, which makes them inherently risky for end customers and more complex to evaluate (Homburg et al., 2011). End customers also want CS brands that guarantee availability of upgrades for technology-intensive components (Ghosh & John, 1999). Thus, the reduction in information cost and risk provided by a strong CS brand has greater value for end customers and exerts greater influence on OEMs' market performance (Kirmani & Rao, 2000; Zandan & Clark, 1987). Furthermore, end-customer loyalty afforded by CS brands can offer short-term protection to OEM sales in dynamic markets when competing OEMs incorporate novel, innovative components into their products (Srivastava et al., 1998). Thus, CS brand image is more important to OEMs for technology-intensive components. In addition, CSs can establish innovation benefits as a point of difference of the CS brand, creating a co-specialized asset bundle that is more difficult for rivals to imitate (Teece, 1988; Van Osselaer & Janiszewski, 2001) and thus limiting OEMs' ability to substitute a strong CS brand offering with that of an alternative CS. Finally, the benefit of using CS brands to balance dependence in the OEM–CS relationship is greater in technology-intensive settings in which contracts are usually incomplete (Ghosh & John, 2005, 2009). In summary, we expect CS industry technology intensity to increase the importance of CS brand image to OEMs, reduce the substitutability of strong CS brands, and enhance the need for CS brands as safeguards. Consequently, CSs will be able to better leverage CS brand image to enhance and sustain the level of prices and profits.

Hypothesis 4. The positive effect of CS brand image on a CS's profitability growth is enhanced in highly technology-intensive CS industries.

3.4.2. Moderating effects of OEM industry characteristics

We examine two OEM industry characteristics as moderators: OEM–end customer relationship importance and brand importance in the OEM industry. OEM–end customer relationship importance reflects the degree to which end customers enter into long-term customer

relationships with OEMs (Palmatier et al., 2006). OEM industries with high relationship importance are characterized by close relationships, while in other industries, end customers and OEMs typically maintain transactional exchanges, as procurement mostly relies on electronic platforms or tenders (Anderson & Narus, 1991).

We suggest that OEMs' end customer relationships facilitate flow of information to end customers and reduce their perceived risk (Johnston & Lewin, 1996). For example, end customers will trust that an OEM will not put the relationship at risk by using unsuitable components. Furthermore, close relationships with OEMs give end customers assurance that OEMs will respond flexibly to resolve unforeseen problems arising from any unknown component (Cannon & Homburg, 2001; Ulaga & Eggert, 2006). Both factors (i.e., information access and assurance of flexibility in response) decrease the importance of using a CS brand to reduce information cost and perceived risk. In addition, because relational contracts have greater adaptation capabilities (Noordewier, John, & Nevin, 1990), end customer relationships enable OEMs to more easily adjust or re-negotiate supply contracts to replace a strong CS brand with an alternative. Therefore, when OEM–end customer relationship importance is high, strong CS brand image will be less important for the market performance of OEMs and strong CS brands will be easier for OEMs to replace, decreasing the CS's ability to leverage CS brand image to sustain or enhance prices and profitability. Thus:

Hypothesis 5. The positive effect of CS brand image on CSs' incremental profitability is reduced when OEM–end customer relationships are important in an industry.

Brand importance in the OEM industry captures whether end customers attach importance to OEM brands when making purchase decisions (Zablah et al., 2010). Similar to CS brands, OEM brands reduce information costs and perceived risk for end customers in their purchasing decision. When end customers typically rely on OEM brands, the quality signal an OEM brand provides likely dilutes that of the CS brand (Swaminathan et al., 2012). In addition, consumer research has found that strong host (i.e., OEM) brands benefit less from ingredient (i.e., CS) brand image spillover (Simonin & Ruth, 1998). Furthermore, extant research predicts that the learning of CS brand image as a predictive cue is less effective when competing cues, such as OEM brands, are salient (Van Osselaer & Janiszewski, 2001). Thus, it is more difficult and costly for CS brands with a strong image to establish a unique brand positioning that prevents substitution. When OEM brands matter, OEMs will also protect against CSs' attempts to wrest away part of their margins (Ghosh & John, 2009), for example, by using self-brand components instead of a CS brand (Desai & Keller, 2002). In summary, we expect higher levels of OEM brand importance to reduce the importance of CS brands to OEMs and make it easier for them to replace strong CS brands with alternatives. This limits CS brands' ability to sustain or enhance prices and profitability levels.

Hypothesis 6. The positive effect of CS brand image on CSs' profitability growth is reduced when OEM brand importance is high.

4. Methodology

4.1. Data collection

We drew on three sources to build the data set for the study. First, we collected perceptual measures of CS brand image, OEM–CS relationship quality, and OEMs' value perception with a survey of executives in OEMs conducted in 2007. We randomly sampled 2200 OEMs in manufacturing industries in Germany from the database of a commercial provider. To avoid sampling bias toward smaller OEMs as a result of their larger numbers, we imposed a quota for each firm size bracket (indicated by the number of employees) to be equally represented. We

identified the heads of marketing and sales by telephone, which was successful in 989 cases. After sending an invitation and two reminders, we obtained 241 complete responses (response rate =24%). This response rate compares favorably with similar studies, which fall in the range of 15%–21% for surveys at the management level (e.g., Homburg, Droll, & Totzek, 2008; Reinartz, Krafft, & Hoyer, 2004). In the survey, we instructed informants to pick a typical component that the OEM sourced along with the brand name of a current CS for that component.

Because research repeatedly suggests that the accuracy of survey data can be improved by obtaining responses from multiple informants (Phillips, 1981; Van Bruggen, Lilien, & Kacker, 2002), we collected data on OEMs' value perception and OEM–CS relationship quality from a second manager in the same OEM who was in charge of procurement for the component.[3] We received an additional 104 usable abbreviated questionnaires from these informants (response rate = 54%). We dropped four single-informant surveys from the analysis because of low informant confidence ratings.

Next, we obtained archival financial statement information to match the CS brands provided by informants in the survey. In the European Union, profit-and-loss statements are publicly available even for many private firms. We first identified the CS firm behind each CS brand name and then gathered archival financial information for those firms from Bureau Van Dijk's ORBIS database. In case of more complex ownership structures, we used financial results at the lowest level of aggregation available. Combining the two data sets yielded 101 observations for which financial measures could be retrieved from archives for at least four years between 2006 and 2010. In the remaining cases, either the CS firm could not be unambiguously identified from archives or financial measures for the CS had not been (fully) reported. For 42 of 101 observations, we use multiple-informant survey data; the remaining 59 are based on single informants.[4]

Using each firm's four-digit core Standard Industrial Classification (SIC) code, we then obtained financial measures for the 100 largest firms within each CS and OEM industry. We chose the cutoff point of 100 because industry averages calculated this way strongly converge to overall industry averages. Ultimately, we matched each CS and OEM with the corresponding industry profiles by Datamonitor Group (www.datamonitor.com) to collect information on characteristics of each CS's and OEM's industry. By using archival sources for measures of financial performance and competition, we were able to eliminate many of the concerns with common method variance raised about survey research (e.g., Homburg, Klarmann, Reimann, & Schilke, 2012; Lindell & Whitney, 2001; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003).

We assessed non-response bias in two ways. First, a comparison of early and late responses to the survey revealed no systematic differences in the measures (Armstrong & Overton, 1977). Second, we compared the demographic information from the commercial database and found that responding firms did not differ from the firms in the initial sample in terms of industry sector ($p = 0.77$), annual revenues ($p = 0.40$), and number of employees ($p = 0.99$). Thus, we do not consider non-response bias a problem in this study.

The availability of multiple responses from two informants per firm for 104 cases enabled us to assess the reliability of ratings provided by the managers surveyed. Following James, Demaree, and Wolf (1984, 1993), we calculated an index of inter-rater agreement to examine the degree to which informants from the same firm shared similar perceptions of a CS.[5] The index values for the two scales were 0.82 and 0.86,

---

[3] To maximize the response rate, we only collected measures of CS brand image from OEM marketing and sales managers, because they should be more knowledgeable about end-customer perceptions than procurement managers.

[4] We aggregate multiple responses using confidence-based weights (Van Bruggen et al., 2002).

[5] We calculate inter-rater reliability ($r_{WG(J)}$) as = $[1 - (S_{xj}^2/\sigma_{EU}^2)]/[J[1 - (S_{xj}^2/\sigma_{EU}^2)] + (S_{xj}^2/\sigma_{EU}^2)]$, where J is the number of items reflecting the construct, $S_{xj}^2 =$ is the mean of observed variances on the J items, and $\sigma_{EU}^2 = (A^2 - 1)/12$ refers to the expected error variance, where A is the number of alternatives in the response scale.

Table 2
Sample and informant characteristics.

Sample characteristics			Informant characteristics for survey		
Industries	CS	OEM	Position[a]	Informant 1	Informant 2
Electronics	24%	30%	Marketing/sales	78%	17%
Machinery	23%	21%	Technical management	10%	17%
Chemicals	10%	6%	Purchasing	8%	64%
Measuring, analyzing, and controlling	10%	4%	General management	19%	17%
Plastics	6%	7%			
Lumber and furniture	5%	2%	Experience in current industry in years		
Metal products	5%	11%			
Stone, clay, glass, and concrete	3%	1%	<2	5%	11%
Transportation equipment	3%	4%	2–4	20%	9%
Apparel & textile products	1%	2%	5–9	37%	27%
Construction & construction equipment	0%	2%	≥10	38%	53%
Printing and publishing	0%	6%			
Other	11%	2%			
Annual revenues in millions of euros					
<10	4%	25%			
10–99	11%	43%			
100–999	31%	25%			
1000–9999	24%	5%			
10,000–99,999	28%	3%			
≥100,000	3%	0%			
Number of employees					
1–19	2%	9%			
20–49	2%	14%			
50–99	4%	16%			
100–249	2%	18%			
250–499	6%	16%			
500–999	11%	13%			
≥1000	73%	14%			

[a] Multiple replies permitted.

respectively, indicating strong inter-rater agreement. In addition, we calculated the intra-class correlations coefficient to assess reliability when perceptions are aggregated (Shrout & Fleiss, 1979). In summary, these results suggest that the measures consistently reflect shared perceptions among managers within the same firm rather than diverse individual perceptions (Bharadwaj, Bharadwaj, & Bendoly, 2007). No significant differences occurred between the two types of managers on the scales for OEMs' value perceptions and relationship quality, respectively ($M = 4.68/4.49$, $p = 0.36$; $M = 5.60/5.71$, $p = 0.65$).

4.2. Sample

Sample demographics for CSs and OEMs appear in Table 2 along with informant characteristics. The sample covers a diverse set of companies from different industry backgrounds. Firms ranged from small enterprises to multi-billion-euro companies. The strategy of obtaining a primary response from marketing and sales managers and a secondary response from purchasing was successful: 78% of primary informants had a background in marketing and sales, and most of the remaining informants were general managers who also manage customer relationships (typical of small firms). In addition, 64% of secondary informants were purchasing managers, and many of the remaining responses came from technical managers who often handle the purchasing in small firms that do not have a dedicated purchasing department. Informants had an average professional experience of 10 to 13 years (primary and secondary informants, respectively) in their current industry, and 75% had worked in their industry for at least five years.

4.3. Measures

Table 3 lists the measures of the key variables used in the study. The dependent variable in the hypotheses is profitability growth (i.e., the increase in profitability of CSs). We compute the measure of profitability growth as the difference in ROS (operating margin), between two points in time, one year before the survey, in 2006, and one year after

the survey, in 2008.[6] To account for heterogeneity among industries, we adjust the measure for between-industry variations by subtracting the corresponding average difference across the 100 largest firms with the same four-digit SIC code (Venkatraman & Ramanujam, 1986). The adjustment for average industry levels is equivalent to common survey-based measures in which managers rate their firms' performance relative to competition (e.g., Homburg & Pflesser, 2000; Jaworski & Kohli, 1993).

In developing our survey scales, we drew on existing measures of the constructs and adapted them to this study's context whenever possible. We took several steps to ensure the content validity of our measures. First, we established clear definitions of the constructs and decided on the indicator specification (formative vs. reflective). Second, we tested and refined an initial set of items in interviews with five experts from academia and practice. Third, we followed Anderson and Gerbing's (1991) item-sorting task to ensure substantive validity. The final measurement items appear in Table 3 along with their item reliabilities. We developed the measurement scale for CS brand image from a review of the literature. Building on the information economics perspective, we operationalized CS brand image as the strength of CS brands as a signal to end customers. Thus, the four items adapted from extant research capture CS brands' credibility and their ability to signal quality (Erdem & Swait, 1998; Erdem, Swait, & Valenzuela, 2006). Our measure of OEM–CS relationship quality comprises items for trust and commitment adapted from Doney and Cannon (1997) and Morgan and Hunt (1994).[7] We measured the control variable OEM's value perception of CS with four items adapted from Ulaga and Eggert (2006) and Menon, Homburg, and Beutin (2005). We calculated item scores for multiple responses available for the latter two constructs

---

[6] Because we conducted the survey in the fourth quarter of 2007, and given that many firms' 2007 fiscal years end earlier in the year, we used 2006 and 2008 data to compute profitability growth to ensure that measures were taken before and after the survey.
[7] Prior research has operationalized relationship quality as consisting of different components (Kumar, Scheer, & Steenkamp, 1995). We include the two components that are common across conceptualizations: trust and commitment.

Table 3
Summary of measures and data sources.

Variable	Measure	Data source
Dependent variable		
Industry-adjusted ROS growth	$\Delta ROSG_{ijt_1,t_2}(\%) = (ROS_{ijt_2} - ROS_{ijt_1}) - \frac{1}{K}\sum_{k=1}^{K}(ROS_{kjt_2} - ROS_{kjt_1})^{b\,a}$ CS firm i in industry j, competitors k in industry j, years t_1 and t_2	ORBIS
Independent variables		
CS brand image	Measurement items adapted from Erdem and Swait (1998) and Erdem et al. (2006): • In the eyes of our customers brand [CS] delivers what it promises. (0.81)[b] • Brand [CS] has name that our customers trust. (0.96)[b] • Our customers trust in brand [CS]'s skills. (0.92)[b] • Among our customers, brand [CS] has a reputation for high quality. (0.83)[b]	Survey
OEM–CS relationship quality	Multiple informants.[c] Measurement items adapted from Doney and Cannon (1997) and Morgan and Hunt (1994): • The relationship with supplier [CS] deserves our firm's maximum effort to maintain. (0.79)[b] • Our firm is very committed to the relationship with supplier [CS]. (0.72)[b] • We trust in supplier [B]'s integrity. (0.42)[b]	Survey
Moderator variables		
• CS industry product differentiation • Importance of OEM–end customer relationships • Brand importance in OEM industry	Obtained from archival industry profiles for the CS and OEM industries; draws on extensive primary and secondary research databases & modeling/forecasting tools. Coded on seven-point scale.	Data-monitor
CS industry technology intensity	$R\&D\,Intensity_{jt}(\%) = \frac{1}{k}\sum_{k=1}^{K}\frac{R\&D\,Expenditures_{jk}}{Revenue_{jk}}$ industry j, competitors k in industry j, year t	ORBIS
Control variables		
OEMs' value perceptions of CS	Multiple informants.[c] Measurement items adapted from Ulaga and Eggert (2006) and Menon et al. (2005): • Supplier [CS]'s components are of high value for our company. (0.52)b • The benefits we receive from [CS]'s components far outweigh the costs. (0.71)b • For the costs incurred, we find the benefits offered by [CS]'s components to be of high value. (0.48)b • Supplier [CS]'s components create more value for us when comparing all costs and benefits. (0.42)b	Survey
CS sales	CS operating revenue	ORBIS
OEM number of employees	OEM Number of employees	ORBIS
CS and OEM industry intangibles intensity	$Intangibles\,Intensity_{jt}(\%) = \frac{1}{K}\sum_{k=1}^{K}\frac{Intangible\,Assets_{jk}}{Total\,Assets_{jk}}$ industry j, competitors k in industry j, year t	ORBIS
CS and OEM industry concentration	Normalized Herfindahl Index for industry j, year t	ORBIS
OEM industry technology intensity	(See CS industry technology intensity above)	ORBIS

[a] Return on sales: $ROS_{it}(\%) = 100 * \frac{Operating\,Income_{it}}{Revenue_{it}}$.
[b] Numbers in parentheses are item reliabilities.
[c] Item scores calculated as confidence-weighted mean (Van Bruggen et al., 2002) across two informants for 42 of 101 cases.

as a confidence-based weighted mean rather than just averaging scores, following Van Bruggen et al. (2002). Because the responses of more confident informants should show less error, this strategy increases the predictive validity of the aggregate data.

We instructed informants to take the OEM's perspective when providing the measures for OEMs' value perceptions and OEM–CS relationship quality. Afterwards, we asked informants to switch to the end customers' perspective when answering the questions for CS brand image. We expected that the frequent and intensive interactions between B2B buyers and sellers would make OEM marketing and sales managers reliable informants on end customers' perceptions. For example, Homburg et al. (2008) find that B2B marketing managers provide highly accurate measures of their customers' satisfaction. High confidence of OEM marketing and sales managers in the accuracy of responses on the CS perspective confirms this expectation ($M = 5.85/7$); their confidence is not significantly lower than that of OEM procurement managers in responses regarding the OEM's perspective ($M = 5.88/7$, $p = 0.72$).

We obtained archival measures of CS industry product differentiation, importance of OEM–end customer relationships, and brand importance in OEM industry from the widely used Datamonitor[8] industry profiles. We matched CSs and OEMs with their corresponding industry. We then extracted the three industry measures from the section dedicated to competitive landscape, which reports on a standardized set of competitive forces. A doctoral student mapped the verbal ratings onto a seven-point scale.[9] We assessed reliability using a second rater who independently

coded one-third of the sample. High intra-class correlation coefficients (ICC) indicate excellent consistency of the ratings (ICC = 0.98/0.78/0.78; $p < .001$) (Cicchetti, 1994). The ISBM B2B data resources library (http://isbm.smeal.psu.edu) describes Datamonitor as a "quick and reliable way to get data on key industrial sectors." Datamonitor undertakes continuous quality control and cross-checking to ensure accuracy of the reports. The profiles draw from extensive primary and secondary research databases as well as complex modeling and forecasting tools to deliver reliable and consistent market analyses for a broad range of industries. We further tested internal consistency by extracting a measure of industry rivalry, which is inversely correlated with all three measures, providing support for their validity. Altogether, these arguments reinforce our confidence in the validity and reliability of these measures. We operationalized technology intensity in the CS industry as the average ratio of R&D expenditures to revenues across the 100 largest firms within each industry. By using archival data of industry-level competitive environments, we avoid the potential subjectivity bias of survey-based assessments by managers who may not know the external environment well enough to make precise comparisons with other industries (Homburg, Klarmann, Reimann, & Schilke, 2012).

We include firm and industry-level control variables that could offer alternative explanations for the effects observed. Given that our rationale for the impact of CS brand image is based on power of CSs over OEMs, we include an archival measure of CS revenue to control for CS firm size. We also control for OEM firm size using the number of employees because revenue data were not publicly available for 50% of the OEMs. Further, we control for industry concentration (Herfindahl index) in the CS and OEM industries because power is often a result of

[8] Datamonitor Industry Profiles have been rebranded to MarketLine Industry Profiles.
[9] We provide verbatim examples of Datamonitor measures in Appendix B.

concentration in the buyer and seller industries. OEM industry technology intensity could represent another source of power of OEMs. Furthermore, we control for intangibles intensity (average ratio of intangible assets to total assets) to rule out that the CS brand image measure merely reflects the prevalence of intangibles in the CS or OEM industry.

4.4. Scale validation

To assess measurement quality, we subjected the survey data to confirmatory factor analysis. The results for the tests of construct reliability and validity in Table 4 are indicative of good psychometric properties. More specifically, no coefficient alpha values and composite reliabilities are lower than 0.80, thus exceeding the recommended thresholds (Bagozzi & Yi, 1988). The goodness-of-fit statistics indicate good fit (comparative fit index = 0.99, root mean square error of approximation = 0.037). Furthermore, item reliabilities (see Table 3) are greater than the recommended value of 0.40 (Bagozzi & Baumgartner, 1994). Finally, factor loadings are highly significant, providing further support of convergent validity (Bagozzi, Yi, & Phillips, 1991).

We assessed discriminant validity in two ways. Table 4 shows that each construct's average variance extracted is greater than the square of latent correlations between that construct and any other construct in the study (Fornell & Larcker, 1981). In addition, we conducted chi-square difference tests for each pair of constructs in the model (Bagozzi et al., 1991). The fit of the unconstrained model is better than that of the models in which the correlation between two constructs is constrained to one ($\Delta\chi^2$ significant at $p < .001$). These results indicate discriminant validity.

5. Estimation and results

5.1. Accounting for potential endogeneity

A major concern with cross-sectional studies is the potential threat of endogeneity in biasing the results. It is possible for CS brand image to be endogenous if unobserved firm-level variables were simultaneously related to both the outcome variable and the independent variable, CS brand image. In this instance, the error term in the regression can be viewed as a sum of two terms: $\varepsilon_i = u_i + v_i$, where u_i is uncorrelated with the independent variables and v_i is correlated with the independent variables. One way to address this potential problem is to include all the unobserved firm-level variables as controls in the model. However, because it is not possible to identify all such variables, we use an instrumental variable approach (Wooldridge, 2010). Three constructs collected through the survey, capturing direct antecedents of CS brand image, serve as instruments: use of various direct communication instruments by the CS, use of joint communication by the CS, and CS brand awareness. We selected these variables because they should be correlated with CS brand image but most likely not correlated with potentially omitted variables, such as objective product quality or quality of a CS's general management (Verbeek, 2008).

We evaluate validity and strength of the instruments using a Hansen–Sargan test for over-identifying restrictions and Stock and Yogo's (2005) test for weak instruments (see also Cameron & Trivedi, 2009). Consistent with our expectations, the Hansen–Sargan test shows that the null hypothesis of valid instruments could not be rejected ($\chi^2 = 0.80$, $p = 0.67$). Furthermore, Stock and Yogo's test indicates that we can reject the null hypothesis of weak instruments because the minimum eigenvalue statistic exceeds the two-stage least squares nominal 5% Wald test critical value (Wald test critical value = 22.3, minimal eigenvalue = 28.72). Taken together, these results indicate that the instruments used were both relevant and strong, lending support to their validity and enhancing confidence in the results.

We estimate the instrumental variable models using generalized method of moments (GMM) to obtain unbiased and consistent

Table 4
Descriptive statistics, scale validity and reliability, and correlations.

	CS brand image	M	SD	α	CR	AVE	1	2	3	4	5	6	7	8	9	10	11	12	13	14
							Correlations													
							N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ
1.	Industry-adjusted ROS growth	0.59	4.81	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ														
2.	OEMs' value perceptions of CS	4.67	1.18	0.80	0.82	0.53	0.25													
3.	CS brand image	5.56	1.47	0.96	0.97	0.88	0.36	0.31												
4.	OEM–CS relationship quality	5.72	1.10	0.83	0.84	0.65	0.13	0.31	0.37											
5.	CS industry product differentiation	3.29	1.27	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	−0.03	−0.12	−0.17	−0.12										
6.	CS industry technology intensity	0.04	0.04	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	0.42	0.05	0.28	0.04	0.09									
7.	Importance of OEM–end customer relationships	4.97	1.01	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	−0.09	−0.05	0.17	−0.03	0.10	0.05								
8.	Brand importance in OEM industry	3.48	1.52	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	−0.13	−0.10	−0.15	−0.09	0.28	0.00	0.19							
9.	CS sales	2×10^7	3×10^7	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	0.27	0.08	0.14	−0.01	−0.14	0.26	0.08	−0.06						
10.	OEM number of employees	1,277	7,713	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	−0.13	0.08	0.09	0.03	0.07	−0.01	0.10	0.18	−0.05					
11.	CS industry intangibles intensity	0.05	0.05	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	0.10	0.02	0.21	0.08	0.15	0.43	0.03	0.07	−0.04	−0.01				
12.	CS industry concentration	0.14	0.22	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	0.29	0.19	0.16	0.10	−0.12	0.46	0.03	−0.02	0.40	0.03	−0.05			
13.	OEM industry intangibles intensity	0.05	0.03	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	−0.05	−0.17	0.00	−0.14	0.09	0.06	−0.02	0.00	0.12	−0.05	0.18	0.05		
14.	OEM industry concentration	0.07	0.08	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	−0.10	−0.10	−0.07	0.07	0.00	0.02	0.11	−0.05	−0.05	−0.01	0.26	−0.08	0.18	
15.	OEM industry technology intensity	0.02	0.02	N.A.ᵃ	N.A.ᵃ	N.A.ᵃ	−0.02	−0.01	0.04	0.02	0.10	0.20	0.20	0.05	−0.12	0.05	0.10	−0.02	0.27	0.14

Notes: Coefficient alpha, composite reliability, goodness-of-fit index, and average variance extracted not applicable.
All correlations greater than .19 are significant at $p < .05$.
CR = composite reliability, and AVE = average variance extracted.

ᵃ Archival measure.

parameter estimates (Arellano & Bond, 1991). An advantage of the GMM is that it does not require assumptions about the distribution of the independent variables (Hansen & West, 2002). In addition, because some CS brands appear more than once in the sample, we use clustered robust standard errors, which enable observations for the same CS brand to be correlated. We standardized all independent variables to create the interaction terms. In all cases, the variance inflation factors are substantially below the value of 10, indicating that multi-collinearity is not an issue (Neter, Kutner, Nachtsheim, & Wasserman, 1995).

5.2. Estimation results

Table 5 reports the results obtained using instrumental variable regression. We follow standard guidelines for moderated regression analysis from the literature to test the interaction hypotheses (Cohen, Cohen, West, & Aiken, 2003). With an R-square of 0.47, the model explains a considerable amount of the variance in the archival performance measure.

Furthermore, we conduct hierarchical regression analysis to examine whether the incremental variance explained by the interaction effects of CS brand image with the moderators is significant. The interaction terms of CS brand image explain a significant amount of additional variance beyond that explained by other variables in the model ($\Delta R^2 = 0.15, F = 5.96, p < .001$). In the following, we report the results of the hypotheses tests.

The model does not support Hypothesis 1; enhanced OEM–CS relationship quality does not directly affect CS profitability growth ($\beta = -0.30$, n.s.). Later in Section 5.3, we conduct a mediation test to explain the absence of a direct effect and show a significant indirect effect of OEM–CS relationship quality on CS profitability growth through OEM value perceptions. The data fail to support Hypothesis 2 ($\beta = -0.06$, n.s.). There is no significant main effect of CS brand image on CS profitability growth. As Hypothesis 3 predicts, there is a positive

interaction of CS brand image with product differentiation in the CS industry ($\beta = 0.97, p < .01$). We also find support for Hypothesis 4 ($\beta = 0.82, p < .01$); CS brand image interacts positively with technology intensity of the CS industry. The data also confirm the negative interaction of CS brand image with importance of OEM–end customer relationships ($\beta = -1.69, p < .001$), in support of Hypothesis 5. We also find support for Hypothesis 6; CS brand image has a negative interaction with brand importance in the OEM industry ($\beta = -0.80, p < .05$).

5.3. Mediation tests

To explain the lack of a direct effect of OEM–CS relationship quality on CS return on sales (ROS) growth (contrary to our expectation in Hypothesis 1), we conduct mediation analysis to examine whether OEMs' value perceptions of CS mediate the effect of OEM–CS relationship quality on financial performance. The underlying reason is that the B2B literature suggests that OEM–CS relationship quality affects CS financial performance by facilitating value creation in the OEM–CS relationship (see the rationale provided for Hypothesis 1 in Section 3.2). That is, the positive effect of OEM–CS relationship quality may be accounted for by the variable OEMs' value perceptions of CS. The mediator, OEMs' value perceptions of CS, captures the tradeoff between the benefits and costs of sourcing from a CS, as perceived by an OEM. We draw on Hayes and Preacher's (2013) MEDIATE procedure using both OEM–CS relationship quality and CS brand image as predictors. To align the mediation test with the instrumental variable estimation used for our models, we enter the predicted value for CS brand image obtained using the instrumental variables from the GMM model as an independent variable in the mediation test. The estimation results appear in Table 6. The results for the first stage show that OEM–CS relationship quality has a positive effect on OEMs' value perceptions of CS ($\beta = 0.34, p < .001$). However, the effect of CS brand image on OEMs'

Table 5
Instrumental variable regression results with CS profitability growth as dependent variable.

Predictors	Hypotheses	Instrumental variable regression results using a GMM estimator[a] (clustered robust standard errors)	
		DV: CS ROS growth[b]	
Main effects			
OEM–CS relationship quality	H1	−0.30	(0.62)
CS brand image	H2	−0.06	(0.39)
Hypothesized interactions			
CS industry product differentiation × CS brand image	H3	0.97**	(0.36)
CS industry technology intensity × CS brand image	H4	0.82**	(0.35)
Importance of OEM–end customer relationships × CS brand image	H5	−1.69***	(0.44)
Brand importance in OEM industry × CS brand image	H6	−0.80*	(0.47)
Main effects of moderators			
CS industry product differentiation		−0.02	(0.42)
CS industry technology intensity		1.39**	(0.55)
Importance of OEM–end customer relationships		−0.73*	(0.42)
Brand importance in OEM industry		0.28	(0.45)
Controls			
OEMs' value perceptions of CS		1.46***	(0.41)
CS sales		1.8×10^{-8}**	(7.5×10^{-9})
OEM number of employees		-6.1×10^{-5}***	(1.6×10^{-5})
CS industry intangibles intensity		−0.24	(0.35)
CS industry concentration		−0.011	(0.38)
OEM industry intangibles intensity		−0.45	(0.39)
OEM industry concentration		−0.42	(0.54)
OEM industry technology intensity		0.33	(0.53)
Intercept		−0.08	(0.43)
Wald χ^2		845***	
N		98	
R^2		0.47	
R^2 of alternative model excluding interactions of CS brand image (F-test statistic)		0.32	(5.96***)

***: Significant at $p < .001$; **: significant at $p < .01$; *: significant at $p < .05$ (directional one-tailed test). All parameter estimates are unstandardized estimates. Standard errors are in parentheses.
[a] Variable instrumented: CS brand image; Instruments: CS brand awareness, CS direct communication intensity, CS joint communication intensity.
[b] Variable industry-adjusted.

Table 6
Supplementary analysis: mediation by OEMs' value perceptions of CS.

Dependent variable/predictors	Mediation test using the MEDIATE procedure (Hayes & Preacher, 2013) (standard errors)[a]
	DV: CS ROS growth[b]
DV: OEMs' value perceptions of CS	
OEM–CS relationship quality	0.34 (0.09)***
CS brand image	0.12 (0.13)
DV: ROS growth (indirect effects through OEMs' value perceptions of CS)	
OEM–CS relationship quality	0.61 (0.24)**
CS brand image	0.22 (0.30)
N	98

: Significant at $p < .01$; *: significant at $p < .001$ (directional one-tailed test). All parameter estimates are unstandardized estimates. Standard errors are in parentheses.

[a] Variable instrumented: CS brand image; Instruments: CS brand awareness, CS direct communication intensity, CS joint communication intensity.

[b] Variable industry-adjusted.

value perceptions of CS is not significant ($\beta = 0.12$, n.s.). The results for the indirect effects show that OEMs' value perceptions of CS mediate the effect of OEM–CS relationship quality on CS profitability growth ($\beta = 0.61$, $p < .01$). This result explains why we were unable to find the direct effect of OEM–CS relationship quality on profitability growth formulated in Hypothesis 1. We also do not find a significant mediated effect of CS brand image on CS profitability growth ($\beta = 0.22$, n.s.). The latter finding helps rule out an alternative mechanism for the financial impact of CS brand image, which would occur if OEMs perceived enhanced value in procuring components from strong CS brands (e.g., Cretu & Brodie, 2007; Webster, 2000). This enhances our confidence in the power-based mechanism put forward in the rationale for Hypothesis 2.

5.4. Robustness checks

We tested whether the results are sensitive to the inclusion of additional interaction terms of the moderators as well as to the operationalization and selection of the dependent variable. First, we added interaction terms of the four moderators with OEMs' value perceptions of CS and OEM–CS relationship quality, respectively, as controls to the model. We do so to ensure that the estimated moderated effects of CS brand image on financial performance go beyond what could have been explained by these well-established metrics from the B2B literature. The first two columns in Appendix A show that the results are robust to the inclusion of these controls and that some of the coefficient estimates for the hypothesized effects even increase slightly. Stepwise regression indicates that the additional interaction terms do not explain a significant amount of variance in ROS growth.

Second, we re-ran the model using ROA growth as an alternative operationalization of profitability growth (third column in Appendix A). The ROA model reported here also includes the interaction terms of OEMs' value perceptions of CS with the moderators, which slightly increases coefficient estimates for the hypothesized effects. A comparison of the ROS and ROA growth models yields several insights: the pattern of coefficient estimates for the hypothesized effects is largely similar for ROA.[10] However, stepwise regression reveals that the moderated effects of CS brand image do not explain a significant amount of variance in ROA growth. Also, the negative interactions of CS brand image with importance of OEM–CS relationships and OEM brands are much less pronounced. ROA ($= profit\ /\ assets$) is the product of ROS ($= profit\ /\ sales$) and asset turnover ($= sales\ /\ assets$)—that is, $ROA = $

profits / sales × sales / assets. A potential explanation for the results obtained is that CS brand image does not affect asset turnover, which means that the outcomes of CS brand image are diluted in the ROA model. One reason is that OEMs do not want to become overly dependent on strong CSs and procure less volume from them. Alternatively, weaker CS brands may be able to make up for lower ROS growth through enhanced asset turnover by increasing asset efficiency.

Third, we estimate the model using downside risk as the dependent variable (i.e., the negative semi-variance of profitability; Ang, Chen, & Yuhang, 2006).[11] In doing so, we respond to recent calls in the literature to examine marketing's impact on financial risk (Tuli & Bharadwaj, 2009). The results in columns 4, 5, and 6 in Appendix A are generally consistent with the expectation that strong CS brand image reduces downside risk under conditions that are favorable to CS branding. We find a negative interaction of CS brand image with CS industry product differentiation. CS brand image also interacts positively with both OEM–end customer relationship importance and brand importance in the OEM industry. An important finding lies in the non-significant interaction effect of CS brand image on CS industry technology intensity. This result could be due to the continuous investment that both R&D and branding require, even during times of economic downturn, causing stronger volatility in profitability. According to stepwise regression, the hypothesized moderating effects do not explain a significant amount of variance for these models. This finding could be due to the fact that, while strong CS brands command higher ROS growth on average, they also have greater flexibility to temporarily lower prices when necessary (e.g., to stabilize sales during an economic downturn), which would increase downside risk. Another notable finding is that enhanced OEM–CS relationship quality is associated with slightly higher ROS-based CS risk. CSs that are highly committed to their OEMs might sacrifice short-term ROS for the benefit of continued, stable relationships and long-term profits.

The results of the control variables and main effects of moderator variables in Table 5 and Appendix A are relatively consistent with expectations and thus provide further support for the robustness of our results. OEMs' value perceptions of CS enhance profitability growth and reduce ROA-based risk. In addition, most of the interaction terms of our moderators with OEMs' value perceptions (which are not displayed in Appendix A due to space constraints) are not significant, indicating that the impact of OEMs' value perceptions is largely universal. Exceptions to this are two interactions of CS industry product differentiation (negative) and importance of OEM–end customer relationships (positive), respectively, with OEMs' value perceptions on ROA growth. The former is in line with the commonly held assumption that value creation in relationships matters even more in less differentiated industries (e.g., Ulaga & Eggert, 2006), and the latter suggests that push strategies are more important when OEMs have strong ties with end customers. In addition, profitability growth is higher for CSs in technology-intensive industries. Larger CSs show better financial performance, while larger size of OEMs affects CS financial performance unfavorably.

---

[10] In addition to the results shown in Appendix A, we obtained a consistent pattern of coefficients using incremental earnings before interest and taxes (EBIT) margin growth as well as non-industry-adjusted ROS and ROA growth as dependent variables.

[11] We thank an anonymous reviewer and the area editor for suggesting risk as an alternative dependent variable. Following the literature (Ang et al., 2006; Tuli & Bharadwaj, 2009), we operationalize downside CS risk as the downside variance of profitability. Research shows that managers characterize risk in terms of the failure to meet an aspired performance level (Miller & Leiblein, 1996). Because investors and managers are loss averse, we assume that they strive for non-negative profitability, and thus we compute risk on the basis of the negative deviation of ROS and ROA from zero (Kahneman & Tversky, 1979; Miller & Leiblein, 1996) for CS firm i and years 1 to T as $RISKD_{iT} = \frac{1}{T}$

$\sqrt{\sum_{t=1}^{T}\left[ROS_{it}^{2} \mid (ROS_{it} < 0)\right]}$. We also compute an alternative risk measure based on year-over-year decrease in ROS: $RISKD_{iT} = \frac{1}{T}\sqrt{\sum_{t=1}^{T}\left[(ROS_{it} - ROS_{i(t-1)})^{2} \mid (ROS_{it} - ROS_{i(t-1)}) < 0\right]}$.

6. Discussion

6.1. Research contributions

Extant research provides little evidence of the financial payoff that CSs derive from strong CS brand image among end customers, despite increasing interest in the branding of industrial components from B2B marketing professionals (ISBM/BMA, 2005; Wiersema, 2012). We questioned whether it makes sense for CSs to build strong CS brand image among their customers' customers. More specifically, we examined how CS brand image helps CSs sustain or grow profitability and assessed the extent to which environmental characteristics moderate this impact.

6.1.1. Financial impact of CS brand image

We drew on the RBV to examine how CS brand image affects CS financial performance. The study's results support the expectation that CS brands affect CS profitability growth when CSs can capitalize on OEMs' dependence on strong CS brand image in their negotiations with OEMs (e.g., Shervani et al., 2007). Similar to brand mindset metrics in consumer markets, strong CS brand image is associated with enhanced CS profits and provides protection against competitive pressure on profitability (Bharadwaj, Tuli, & Bonfrer, 2011; Johansson et al., 2012).

Meanwhile, our supplementary mediation tests do not provide support for an indirect effect formulated in the literature (e.g., Webster, 2000), which postulates that CS brand image enhances OEMs' value perceptions of CSs, which in turn enhance CS profitability. One explanation for the lack of support for the indirect path might be that the majority of CS brands are not available exclusively to OEMs and thus do not create a sustainable competitive advantage for OEMs (Srivastava, Fahey, & Christensen, 2001). Another potential explanation is that widespread quality certification could have eliminated low-quality CS brands from OEMs' consideration sets, reducing variance on our independent variable.

6.1.2. Contingency of financial impact of CS brand image

We find that CSs' ability to leverage strong CS brand image is contingent on certain industry conditions. As such, the study contributes to the emerging contingency perspective on the marketing–finance interface

that examines the conditions under which certain market-based assets create a competitive advantage (Bharadwaj et al., 2011). This finding is in line both with anecdotal evidence for mixed success of CS brand strategies and with previous research that finds that outcomes of B2B brands depend on context (Homburg et al., 2011; Zablah et al., 2010). Our results provide support for the notion that the overall financial outcome of CS brand strategies is positive only when the financial gains from leveraging the CS brand outweigh the costs. However, building and sustaining a strong CS brand image require substantial and continuous allocation of financial resources and commitment and may not be worthwhile in industries with limited differentiation or when embedded in strong OEM brands (see Section 3.3.2). Thus, financial outcomes of CS brand strategies can also be negative in contexts in which CSs are unable to extract higher prices from OEMs because of increased power. We propose two sets of moderators that affect OEMs' ability to leverage CS brand image: the interaction plots in Fig. 3a, b, c, and d illustrate how characteristics of the CS and OEM industries influence the CS brand image–ROS growth relationship. The plots show the simple slope whThe subsequent changes have been suggested by the AEen a single moderator is at the level of its mean plus or minus one standard deviation (SD). Note that three of the four positive simple slopes are not significant. We therefore report threshold values required for each of these three moderators for a significant positive simple slope (the reported threshold values are observed in the sample):

(a) When product differentiation is low in the CS industry, strong CS brand image reduces CS ROS growth (negative simple slope significant at $p < .1$); thus, CS branding is ineffective in CS industries with low product differentiation (Fig. 3a). Conversely, when product differentiation is very high in the CS industry, strong CS brand image enhances CS ROS growth (positive simple slope significant at $p < .05$ for $moderator > mean + 1.48SD$).

(b) When R&D intensity is low, strong CS brand image is associated with reduced CS ROS growth ($p < .1$) (Fig. 3b). However, when R&D intensity is very high in the CS industry, CS ROS growth increases with CS brand image (positive simple slope significant at $p < .05$ for $moderator > mean + 2.25SD$). This finding is in line with Aaker and Jacobson (2001), who find a significant impact of B2B brand image for high-technology products
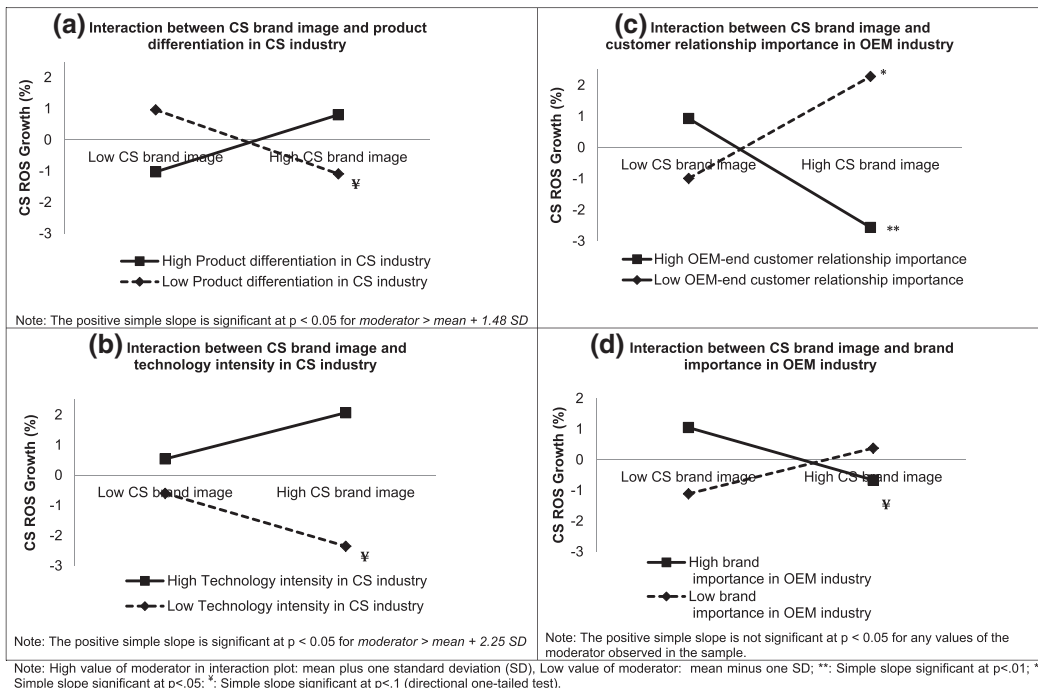

(a) Interaction between CS brand image and product differentiation in CS industry

— High Product differentiation in CS industry
- - - Low Product differentiation in CS industry

Note: The positive simple slope is significant at p < 0.05 for *moderator > mean + 1.48 SD*

(b) Interaction between CS brand image and technology intensity in CS industry

— High Technology intensity in CS industry
- - - Low Technology intensity in CS industry

Note: The positive simple slope is significant at p < 0.05 for *moderator > mean + 2.25 SD*

(c) Interaction between CS brand image and customer relationship importance in OEM industry

— High OEM-end customer relationship importance
- - - Low OEM-end customer relationship importance

(d) Interaction between CS brand image and brand importance in OEM industry

— High brand importance in OEM industry
- - - Low brand importance in OEM industry

Note: The positive simple slope is not significant at p < 0.05 for any values of the moderator observed in the sample.

Note: High value of moderator in interaction plot: mean plus one standard deviation (SD), Low value of moderator: mean minus one SD; **: Simple slope significant at p<.01; *: Simple slope significant at p<.05; ¥: Simple slope significant at p<.1 (directional one-tailed test).

Fig. 3. Moderating effects of CS and OEM industry characteristics.

(e.g., semi-conductors).

(c) Furthermore, strong CS brand image reduces CS ROS growth when OEM–end customer relationships are important in an OEM industry (Fig. 3c) ($p < .01$). Thus, CSs do not benefit from branding when OEMs have strong relationships with their B2B customers. Conversely, when OEM–end customer relationship importance is low, strong CS brand image enhances CS ROS growth ($p < .05$).

(d) Finally, as Fig. 3d shows, when brand importance is high in an OEM industry, strong CS brand image reduces CS ROS growth ($p < .1$). However, we do not observe the very low levels of brand importance in the OEM industry that would be required for the corresponding positive single slope to be significant in our sample.

Given that the positive effects of CS brand image are only observed at rather extreme values of (some of) the moderators, we conducted an additional analysis to explore situations where two moderators are simultaneously fixed at reasonable values *(i.e., mean +/- 1 SD)* that favour a positive effect of CS brand image. Table 7 shows the positive simple slopes for paired combinations of the moderator values. These simple slopes apply to industry contexts in which two moderators are simultaneously one standard deviation above (i.e., product differentiation and technology intensity) or below (i.e., relationship and brand importance) the mean. In these situations, the simple slope takes values between 1.33 and 2.36. CS brand image exerts a significant positive effect for all paired combinations of the moderators shown in the table. This means that many situations (defined by paired combinations of moderator values) exist in which CS brand image enhances CS profitability significantly. Note that the effects observed for moderator pairs that include brand importance are generally weaker than the effects involving the other moderators.

6.1.3. Impact of CS brand image versus that of OEM-CS relationships

Given that the B2B literature has traditionally focused on building strong relationships with direct customers (e.g., Palmatier et al., 2006), an important contribution of our study is the examination of the financial impact of CS brands in the context of the OEM–CS relationship. In contrast with CS brand image, relationship quality has a universally favorable impact on OEMs' value perceptions, holding across contexts. Thus, our results lend support to the established chain of effects that links OEM–CS relationship quality to OEMs' value perceptions and CS financial performance (Anderson & Narus, 2004; Cannon & Homburg, 2001; Ulaga & Eggert, 2006). However, the study clearly demonstrates that CS brand image affects CS performance beyond the impact of these established measures in the B2B marketing literature. According to the hierarchical regression analysis results in Table 5, CS brand image explains 15% of the variance in ROS growth (increase in R^2 from 0.32 to 0.47 = 0.15), while OEMs' value perceptions explain only 6% ($\Delta R^2 = 0.06$).

6.2. Managerial implications

6.2.1. Does CS branding pay off?

This study offers important implications for the growing numbers of B2B CSs "experimenting" with CS brand strategies, according to

Wiersema (2012, p. 51). A key takeaway from our study is that CS brand image matters for financial performance, but only in very specific industry settings. The impact of a CS branding strategy on B2B supplier firm performance has been the subject of frequent debate among B2B marketers. Managers often question whether the principle of ingredient branding can be transferred from consumer markets to their industry context, in which end users are B2B customers (ISBM/BMA, 2005). By showing how CS brand strategies affect firms' financial performance, this study sheds light on this crucial question, enabling marketers to better understand and communicate the outcomes of their CS brand investments to investors and senior management. The predicted impact of a one standard deviation increase in CS brand image on ROS is 2.25 percentage points ($p < .05$) in favorable conditions (i.e., CS product differentiation and technology intensity at one SD above the mean), against an average sample ROS level of 6%. Although this may appear to be a small absolute gain in profitability, a gain of 2.25 percentage points on a base of 6 percentage points is a 38% gain in profitability. This impact is somewhat larger to that of a similar increase in perceived value for OEMs (1.5 percentage points) and substantially larger than that of an increase in OEM–CS relationship quality (0.6 percentage points).

6.2.2. When is a CS brand strategy viable?

Given the high degree of variance in CS branding outcomes, our study helps marketers in CS firms better understand the well-defined conditions in which CS branding investments yield positive payoffs. It makes more sense to invest in CS brands in CS industries in which at least two of the industry moderators are favorable (for example, when product differentiation in the CS industry is high and OEMs' relationships with end customers are unimportant). Furthermore, the study's results clarify the common misperceptions among managers that "As products become increasingly similar, companies are turning to branding as a way to create preference for their offerings" (Kotler & Pfoertsch, 2006, book cover). As we show, CSs in industries with vanishing product differentiation and technology intensity might be better off concentrating their investments on value creation in the OEM–CS relationship. Conversely, CSs in differentiated and technology-intensive industries, which could reap the largest benefits from CS brands, have a tendency to stay away from branding, assuming that "good products need no marketing."

6.3. Limitations and future research directions

The limitations of this study offer fruitful avenues for further research. First, although we took considerable care to address the concerns of endogeneity, common method variance, variable operationalization, and omitted variables, the results could suffer from some of the general limitations of survey research, such as informant bias and perceptual measures. While responses from multiple informants were available for 42% of the sample, the use of single informants for the rest of the sample was less than ideal. However, high inter-rater reliabilities for the multiple informant data enhance our confidence in single-informant measures. In addition, because we match the survey measures with archival data, it seems unlikely that the observed effects are driven by common method variance (Podsakoff et al., 2003). Our use of a differential, industry-adjusted financial measure eliminates unobserved, firm-level, and industry-level variables as competing explanations for the effects observed.

A second potential limitation is the indirect measurement of CS brand image, obtained from carefully selected informants in charge of marketing and sales at OEM firms. We are confident, however, that this measure is valid because B2B marketing and sales managers interact frequently and intensively with their direct customers. Therefore, OEMs' assessments of end customers' mindsets should be accurate. High confidence of informants in the accuracy of their responses and prior research findings that B2B marketing

Table 7
Simple slope of CS brand image as a function of combinations of the moderator values.

First moderator	Second moderator		
	1	2	3
1. CS industry product differentiation = mean + 1 * SD			
2. CS industry technology intensity = mean + 1 * SD	2.25*		
3. Importance of OEM–end customer relationships = mean − 1 * SD	2.36**	2.22**	
4. Brand importance in OEM industry = mean − 1 * SD	1.62¥	1.33¥	2.19*

**: Significant at $p < .01$; *: significant at $p < .05$; ¥: significant at $p < .1$ (directional one-tailed test). All parameter estimates are unstandardized estimates.

managers give highly accurate measures of their customers' satisfaction provide support for this expectation (Homburg et al., 2008). By including OEMs' perceptions of value of CS and relationship quality in the model, we take further precaution to rule out that the observed effects are biased by OEMs' own perceptions of the CS. The difference between the pattern of main and moderating effects of OEMs' value perceptions (reported as robustness check) and that of CS brand image lends nomological validity to our CS brand measure.

Third, we focus on industry-level moderators of the relationship between CS brand image and CS performance. Future studies could examine firm- and dyad-level moderators such as OEM brand image and characteristics of the OEM–CS relationship.

Fourth, we examine the impact of the level of CS brand image on firms' ability to sustain or grow profitability in a context of increasingly powerful and demanding OEMs. Although previous research has taken a similar approach by examining the role of the level of consumer brand quality perceptions in stabilizing financial returns and reducing their volatility (Johansson et al., 2012), future studies could explore the impact of changes in CS brand image over time (e.g., Bharadwaj et al., 2011). However, such panel data are difficult to obtain for the given B2B setting.

Acknowledgments

Appendix A. Robustness check using additional interactions and alternative dependent variables

Predictors	Instrumental variable regression models using a GMM estimator[a] (clustered robust standard errors)					
	(1) DV: CS ROS growth[b]	(2) DV: CS ROS growth[b]	(3) DV: CS ROA growth[b]	(4) DV: Downside risk (ROS)[c]	(5) DV: Downside risk (ROA)[c]	(6) DV: Downside risk (ROS)[d]
Main effects						
OEM–CS relationship quality (RELQ)	−0.13	0.12	−0.82	0.55*	0.30	0.38¥
CS brand image	−0.50	−0.31	−0.59	0.38	0.39	0.29
Hypothesized interactions						
CS industry product differentiation × CS brand image	1.15**	1.06**	1.31**	−0.83**	−1.15**	−0.83***
CS industry technology intensity × CS brand image	0.88**	1.07**	0.78*	0.53	0.37	−0.10
Importance of OEM–end customer relationships × CS brand image	−1.63**	−1.99***	−1.37¥	0.79*	1.15**	0.93***
Brand importance in OEM industry × CS brand image	−1.03*	−0.89*	0.015	0.86*	1.15*	1.05**
Selected controls						
OEMs' value perceptions of CS (OEVP)	1.71***	1.41***	1.77*	−0.81*	−1.22**	−0.77*
Controlling for additional interactions of moderators with RELQ or OEVP	OEVP	RELQ	OEVP	N.A.	N.A.	N.A.
N	98	98	101	106	106	99
R²	0.48	0.51	0.11	0.21	0.20	0.15
R² of alternative model excluding interactions of CS brand image (F-test statistic)	0.35 (4.69**)	0.33 (7.06***)	0.08 (0.58)	0.16 (0.85)	0.15 (1.18)	0.13 (0.52)
R² of alternative model excluding additional interactions of moderators with RELQ or OEVP and the moderated main effect (F-test statistic)	0.42 (1.83)	0.47 (1.15)	0.06 (0.82)	N.A.	N.A.	N.A.

***: Significant at $p < .001$; **: Significant at $p < .01$; *: Significant at $p < .05$; ¥: Significant at $p < .1$ (directional one-tailed test). All parameter estimates are unstandardized estimates. [a]Variable instrumented: CS brand image; Instruments: CS brand awareness, CS direct communication intensity, CS joint communication intensity. [b]Variable industry-adjusted. [c]Risk computed on the basis of negative deviation from zero. [d]Risk computed on the basis of year-over-year decrease.

Appendix B. Verbatim description of Datamonitor measures

Measure	Verbatim	Value on seven-point scale
CS industry product differentiation	"Market is characterized by low product differentiation."	2
	"Market is characterized by relatively low product differentiation."	3
	"There is some degree of product differentiation."	4
	"Players can differentiate themselves by the specification of their products."	5
	"It is possible to differentiate products to a relatively high degree."	6
Importance of OEM–end customer relationships	"While some work may be done on a regular basis for large customers under long-term contracts, most work is done on a project basis."	3
	"Buyers typically enter into long-term contracts with players."	6
Brand importance in OEM industry	"There is little brand loyalty."	2
	"There is a great deal of brand loyalty."	6

References

Aaker, D. A., & Jacobson, R. (2001). The value relevance of brand attitude in high-technology markets. *Journal of Marketing Research, 38*(4), 485–493.

Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology, 76,* 732–740.

Anderson, J. C., & Narus, J. A. (1990). A model of distributor firm and manufacturer firm working partnerships. *Journal of Marketing, 54*(1), 42–58.

Anderson, J. C., & Narus, J. A. (1991). Partnering as a focused market strategy. *California Management Review, 33*(3), 95–113.

Anderson, J. C., & Narus, J. A. (2004). *Business market management: Understanding, creating and delivering value.* Upper Saddle River, NJ: Prentice Hall.

Ang, A., Chen, J., & Yuhang, X. (2006). Downside risk. *Review of Financial Studies, 19*(4), 1191–1239.

Antia, K. D., & Frazier, G. L. (2001). The severity of contract enforcement in interfirm channel relationships. *Journal of Marketing, 65*(4), 67–81.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies, 58,* 194–277.

Armstrong, J. S., & Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *Journal of Marketing Research, 14*(3), 396–402.

Bagozzi, R. P. (1974). Marketing as an organized behavioral system of exchange. *Journal of Marketing*, *38*(4), 77–81.

Bagozzi, R. P., & Baumgartner, H. (1994). The evaluation of structural equation models and hypothesis testing. In R. P. Bagozzi (Ed.), *Principles of marketing research* (pp. 386–422). Cambridge: Blackwell.

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74–94.

Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing construct validity in organizational research. *Administrative Science Quarterly*, *36*, 421–458.

Bahadir, S.C., Bharadwaj, S. G., & Srivastava, R. K. (2008). Financial value of brands in mergers and acquisitions: Is value in the eye of the beholder? *Journal of Marketing*, *72*(6), 49–64.

Barney, J. B. (2001). Is the resource-based "view" a useful perspective for strategic management research? Yes. *Academy of Management Review*, *26*(1), 41–56.

Beier, F., & Stern, L. W. (1969). Power in the channel of distribution. In L. W. Stern (Ed.), *Distribution channels: Behavioral dimensions* (pp. 92–116). Boston: Houghton Mifflin.

Bharadwaj, S., Bharadwaj, A., & Bendoly, E. (2007). The performance effects of complementarities between information systems, marketing, manufacturing, and supply chain processes. *Information Systems Research*, *18*(4), 437–453.

Bharadwaj, S. G., Tuli, K. R., & Bonfrer, A. (2011). The impact of brand quality on shareholder wealth. *Journal of Marketing*, *75*(5), 88–104.

Black, J. A., & Boal, K. B. (1994). Strategic resources: Traits, configurations and paths to sustainable competitive advantage. *Strategic Management Journal*, *15*, 131–148.

Buchanan, L. (1992). Vertical trade relationships: The role of dependence and symmetry in attaining organizational goals. *Journal of Marketing Research*, *29*(1), 65–75.

Cameron, C. A., & Trivedi, P. (2009). *Microeconometrics using stata.* College Station, TX: Stata Press.

Cannon, J. P., & Homburg, C. (2001). Buyer–supplier relationships and customer firm costs. *Journal of Marketing*, *65*(1), 29–43.

Cicchetti, D.V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290.

Cohen, J. P., Cohen, S. G., West, L., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Cooper, R., & Slagmulder, R. (1999). Develop profitable new products with target costing. *Sloan Management Review*, *40*(4), 23–33.

Cretu, A. E., & Brodie, R. J. (2007). The influence of brand image and company reputation where manufacturers market to small firms: A customer value perspective. *Industrial Marketing Management*, *36*(2), 230–240.

Davis, D. F., Golicic, S. L., & Marquardt, A. J. (2008). Branding a B2B service: Does a brand differentiate a logistics service provider? *Industrial Marketing Management*, *37*(2), 218–227.

Desai, K. K., & Keller, K. L. (2002). The effects of ingredient branding strategies on host brand extendibility. *Journal of Marketing*, *66*(1), 73–93.

Doney, P.M., & Cannon, J. P. (1997, April). An examination of the nature of trust in buyer–seller relationships. *Journal of Marketing*, *61*, 35–51.

Dyer, J. H., & Singh, H. (1998). The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of Management Review*, *23*(4), 660–679.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: What are they? *Strategic Management Journal*, *21*(10/11), 1105–1121.

Erdem, T., & Swait, J. (1998). Brand equity as a signaling phenomenon. *Journal of Consumer Psychology*, *7*(2), 131–157.

Erdem, T., Swait, J., & Valenzuela, A. (2006). Brands as signals: A cross-country validation study. *Journal of Marketing*, *70*(1), 34–49.

Fang, E., Palmatier, R. W., & Grewal, R. (2011). Effects of customer and innovation asset configuration strategies on firm performance. *Journal of Marketing Research*, *48*(3), 587–602.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50.

Frazier, G. L., Gill, J.D., & Kale, S. H. (1989). Dealer dependence levels and reciprocal actions in a channel of distribution in a developing country. *Journal of Marketing*, *53*(1), 50–69.

Frazier, G. L., Spekman, R. E., & O'Neal, C. R. (1988). Just-in-time exchange relationships in industrial markets. *Journal of Marketing*, *52*(4), 52–67.

Ghosh, M., & John, G. (1999). Governance value analysis and marketing strategy. *Journal of Marketing*, *63*(4), 131–145.

Ghosh, M., & John, G. (2005). Strategic fit in industrial alliances: An empirical test of governance value analysis. *Journal of Marketing Research*, *42*(3), 346–357.

Ghosh, M., & John, G. (2009, October). When should original equipment manufacturers use branded component contracts with suppliers? *Journal of Marketing Research*, *46*, 597–611.

Gundlach, G. T., & Cadotte, E. R. (1994). Exchange interdependence and interfirm interaction: Research in a simulated channel setting. *Journal of Marketing Research*, *31*(4), 516–532.

Hansen, B. E., & West, K. D. (2002). Generalized method of moments and macroeconomics. *Journal of Business & Economic Statistics*, *20*(4), 460–469.

Hayes, A. F., & Preacher, K. J. (2013). Statistical mediation analysis with a multicategorical independent variable. White paper. Retrieved from. http://www.afhayes.com/

Homburg, C., Artz, M., & Wieseke, J. (2012). Marketing performance measurement systems: Does comprehensiveness really improve performance? *Journal of Marketing*, *76*(3), 56–77.

Homburg, C., Droll, M., & Totzek, D. (2008, September). Customer prioritization: Does it pay off, and how should it be implemented? *Journal of Marketing*, *72*, 110–130.

Homburg, C., Klarmann, M., Reimann, M., & Schilke, O. (2012). What drives key informant accuracy? *Journal of Marketing Research*, *49*(4), 594–608.

Homburg, C., Klarmann, M., & Schmitt, J. (2011). Brand awareness in business markets: When is it related to firm performance? *International Journal of Research in Marketing*, *27*(3), 201–212.

Homburg, C., & Pflesser, C. (2000). A multiple-layer model of market-oriented organizational culture: Measurement issues and performance outcomes. *Journal of Marketing Research*, *37*(4), 449–462.

Hutton, J. G. (1997). A study of brand equity in an organizational-buying context. *Journal of Product & Brand Management*, *6*(6), 428–439.

ISBM (2010). *B-To-B marketing trends 2010 study.* College Park, PA: ISBM.

ISBM/BMA (2005). *Summary of roundtable discussion: ISBM/BMA business marketing brand consortium.* Chicago: ISBM.

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, *69*(1), 85–98.

James, L. R., Demaree, R. G., & Wolf, G. (1993). An assessment of within-group interrater agreement. *Journal of Applied Psychology*, *78*(2), 306–309.

Jaworski, B. J., & Kohli, A. K. (1993). Market orientation: Antecedents and consequences. *Journal of Marketing*, *57*(3), 53–70.

Johansson, J. K., Dimofte, C. V., & Mazvancheryl, S. K. (2012). The performance of global brands in the 2008 financial crisis: A test of two brand value measures. *International Journal of Research in Marketing*, *29*(3), 235–245.

Johnston, W. J., & Lewin, J. E. (1996). Organizational buying behavior: Toward an integrative framework. *Journal of Business Research*, *35*(1), 1–15.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *Journal of Marketing*, *57*(1), 1–22.

Kelley, H. H., & Thibaut, J. W. (1978). *Interpersonal relations: A theory of interdependence.* New York: Wiley.

Kirmani, A., & Rao, A.R. (2000). No pain, no gain: A critical review of the literature on signaling unobservable product quality. *Journal of Marketing*, *64*(2), 66–79.

Kohli, A. K., & Jaworski, B. J. (1990). Market orientation: The construct, research propositions, and managerial implications. *Journal of Marketing*, *54*(2), 1–18.

Kotler, P., & Pfoertsch, W. (2006). *B2B brand management.* Heidelberg: Springer.

Kotler, P., & Pfoertsch, W. (2010). *Ingredient branding: Making the invisible visible* (1st ed.). Berlin: Springer-Verlag.

Kumar, N., Scheer, L. K., & Steenkamp, J. -B. E. M. (1995). The effects of supplier fairness on vulnerable resellers. *Journal of Marketing Research*, *32*(1), 54–65.

Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-selectional research designs. *Journal of Applied Psychology*, *86*(1), 114–121.

Marketing Science Institute (2008). *2008–2010 research priorities.* Cambridge, MA: Marketing Science Institute.

Menon, A., Homburg, C., & Beutin, N. (2005). Understanding customer value in business-to-business relationships. *Journal of Business-to-Business Marketing*, *12*(2), 1–35.

Miller, K. D., & Leiblein, M. J. (1996). Corporate risk-return relations: Returns variability versus downside risk. *Academy of Management Journal*, *39*(1), 91–122.

Mintz, O., & Currim, I. S. (2013). What drives managerial use of marketing and financial metrics and does metric use affect performance of marketing-mix activities? *Journal of Marketing*, *77*(2), 17–40.

Morgan, R. M., & Hunt, S. D. (1994, July). The commitment–trust theory of relationship marketing. *Journal of Marketing*, *58*, 20–38.

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1995). *Applied linear statistical models* (4th ed.). New York: Irwin/McGraw-Hill.

Noordewier, T. G., John, G., & Nevin, J. R. (1990). Performance outcomes of purchasing arrangements in industrial buyer–vendor relationships. *Journal of Marketing*, *54*(4), 80–93.

Palmatier, R. W., Dant, R. P., & Grewal, D. (2007). A comparative longitudinal analysis of theoretical perspectives of interorganizational relationship performance. *Journal of Marketing*, *71*(4), 172–194.

Palmatier, R. W., Dant, R. P., Grewal, D., & Evans, K. R. (2006). Factors influencing the effectiveness of relationship marketing: A meta-analysis. *Journal of Marketing*, *70*(4), 136–153.

Palmatier, R. W., Scheer, L. K., & Steenkamp, J. -B. E. M. (2007, May). Customer loyalty to whom? Managing the benefits and risks of salesperson-owned loyalty. *Journal of Marketing Research*, *44*, 185–199.

Park, C. W., Jun, S. Y., & Shocker, A.D. (1996, November). Composite branding alliances: An investigation of extension and feedback effects. *Journal of Marketing Research*, *33*, 453–466.

Payne, A., & Frow, P. (2005). A strategic framework for customer relationship management. *Journal of Marketing*, *69*(4), 167–176.

Phillips, L. W. (1981). Assessing measurement error in key informant reports: A methodological note on organizational analysis in marketing. *Journal of Marketing Research*, *18*(4), 395–415.

Podsakoff, P.M., MacKenzie, S. B., Lee, J. -Y., & Podsakoff, N.P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903.

Porter, M. E. (1980). *Competitive strategy.* New York: Free Press.

Priem, R. L., & Butler, J. E. (2001). Is the resource-based "view" a useful perspective for strategic management research? *Academy of Management Review*, *26*(1), 22–40.

Rao, A.R., Qu, L., & Ruekert, R. (1999). Signaling unobservable product quality through a brand ally. *Journal of Marketing Research*, *36*(2), 258–268.

Reibstein, D. J., Day, G., & Wind, J. (2009, July). Guest editorial: Is marketing academia losing its way? *Journal of Marketing*, *73*, 1–3.

Reimann, M., Schilke, O., & Thomas, J. S. (2010). Toward an understanding of industry commoditization: Its nature and role in evolving marketing competition. *International Journal of Research in Marketing*, *27*(2), 188–197.

Reinartz, W., Krafft, M., & Hoyer, W. D. (2004, August). The customer relationship management process: Its measurement and impact on performance. *Journal of Marketing Research, 41*, 293–305.

Shervani, T. A., Frazier, G., & Challagalla, G. (2007). The moderating influence of firm market power on the transaction cost economics model: An empirical test in a forward channel integration context. *Strategic Management Journal, 28*, 635–652.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428.

Simonin, B.L., & Ruth, J. A. (1998). Is a company known by the company it keeps? Assessing the spillover effects of brand alliances on consumer brand attitudes. *Journal of Marketing Research, 35*(1), 30–42.

Smith, W. R. (1956). Product differentiation and market segmentation as alternative marketing strategies. *Journal of Marketing, 21*(1), 3–8.

Srivastava, R. K., Fahey, L., & Christensen, H. K. (2001). The resource-based view and marketing: The role of market-based assets in gaining competitive advantage. *Journal of Management, 27*(6), 777–802.

Srivastava, R. K., Shervani, T. A., & Fahey, L. (1998). Market-based assets and shareholder value: A framework for analysis. *Journal of Marketing, 62*(1), 2–18.

Srivastava, R. K., Shervani, T. A., & Fahey, L. (1999). Marketing, business processes, and shareholder value: An organizationally embedded view of marketing activities and the discipline of marketing. *Journal of Marketing, 63*(4), 168–179.

Stock, J., & Yogo, M. (2005). Testing for weak instruments in linear IV regressions. In D. W. K. Andrews, & J. Stock (Eds.), *Identification and inference for econometric models* (pp. 109–120). Cambridge, MA: Cambridge University Press.

Swaminathan, V., Reddy, S., & Dommer, S. (2012). Spillover effects of ingredient branded strategies on brand choice: A field study. *Marketing Letters, 23*(1), 237–251.

Teece, D. J. (1988). Capturing value from technological innovation: Integration, strategic partnering, and licensing decisions. *Interfaces, 18*(3), 46–61.

Tuli, K. R., & Bharadwaj, S. G. (2009). Customer satisfaction and stock returns risk. *Journal of Marketing, 73*(6), 184–197.

Tuli, K. R., Bharadwaj, S. G., & Kohli, A. K. (2010). Ties that bind: The impact of multiple types of ties with a customer on sales growth and sales volatility. *Journal of Marketing Research, 47*(1), 36–50.

Tuli, K. R., Kohli, A. K., & Bharadwaj, S. G. (2007). Rethinking customer solutions: From product bundles to relational processes. *Journal of Marketing, 71*(3), 1–17.

Ulaga, W., & Eggert, A. (2006). Value-based differentiation in business relationships: Gaining and sustaining key supplier status. *Journal of Marketing, 70*(1), 119–136.

Van Bruggen, G. H., Lilien, G. L., & Kacker, M. (2002). Informants in organizational marketing research: Why use multiple informants and how to aggregate responses. *Journal of Marketing Research, 39*(4), 469–478.

Van Osselaer, S. M. J., & Janiszewski, C. (2001). Two ways of learning brand associations. *Journal of Consumer Research, 28*(2), 202–223.

Venkatraman, N. (1989). The concept of fit in strategy research: Toward verbal and statistical correspondence. *Academy of Management Review, 14*(3), 423–444.

Venkatraman, N., & Ramanujam, V. (1986). Measurement of business performance in strategy research: A comparison of approaches. *Academy of Management Review, 11*(4), 801–814.

Verbeek, M. (2008). *A guide to modern econometrics.* New York: John Wiley & Sons.

Webster, F. E. J. (2000). Understanding the relationships among brands, consumers, and resellers. *Journal of the Academy of Marketing Science, 28*(1), 17–23.

Webster, F. E., Jr., & Keller, K. L. (2004). A roadmap for branding in industrial markets. *Journal of Brand Management, 11*(5), 388–402.

Wernerfelt, B. (1988). Umbrella branding as a signal of new product quality: An example of signalling by posting a bond. *RAND Journal of Economics, 19*(3), 458–466.

Wiersema, F. (2012). *The B2B agenda: The current state of B2B marketing and a look ahead.* University Park, PA: ISBM.

Wise, R., & Zednickova, J. (2009). The rise and rise of the B2B brand. *Journal of Business Strategy, 30*(1), 4–13.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Boston: The MIT Press.

Wuyts, S., Verhoef, P. C., & Prins, R. (2009). Partner selection in B2B information service markets. *International Journal of Research in Marketing, 26*(1), 41–51.

Zablah, A.R., Brown, B. P., & Donthu, N. (2010). The relative importance of brands in modified rebuy purchase situations. *International Journal of Research in Marketing, 27*(3), 248–260.

Zandan, P., & Clark, F. S. (1987). FUD is one reason why research for high-tech products is different. *Marketing News, 21*(18), 16.

Full Length Article

Billboard and cinema advertising: Missed opportunity or spoiled arms? ☆

Steffi Frison [a,*], Marnik G. Dekimpe [b,a,1], Christophe Croux [a,2], Peter De Maeyer [c,3]

[a] KU Leuven, University of Leuven, Naamsestraat 69, 3000 Leuven, Belgium
[b] Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands
[c] Mahidol University, 69 Vipawadee Rangsit Road, Bangkok 10400, Thailand

ARTICLE INFO

ABSTRACT

Advertising remains one of the most popular marketing instruments, and many studies have studied its sales effectiveness. However, prior research has either looked at the total spending of a brand/firm, or has focused on the most popular media, especially TV advertising. Even though huge amounts are also spent on "smaller" media such as billboards and cinema, little is known on their effectiveness.

While many brands never use them (which could be a missed opportunity), others allocate a substantial part of their advertising budget to these media (which could represent spoiled arms in case of insufficient effectiveness). In this study, we conduct a large-scale empirical investigation, using close to 7 years of monthly data on over 250 brands of consumer packaged goods, to quantify the sales elasticity of these often-neglected media.

Even though a significant long-run elasticity is found for a number of brands, we obtain a substantially lower proportion of significant effects for billboard and cinema advertising than for the more popular TV medium. Also meta-analytically, and after correcting for the brands' self-selection of media on which to spend their advertising budget, no evidence of a significant short- or long-run sales elasticity is found for these two media, while significant effects are obtained for both TV and magazine advertising. In addition, little evidence of systematic synergy effects with the more traditional media is found. Hence, from a sales-response point of view, investments in billboard and cinema advertising appear to act as spoiled arms for most mature CPG brands.

1. Introduction

Advertising remains one of the most important marketing-mix instruments. Total global adspend in 2012 has been estimated at nearly US$ 500 billion, or 0.7% of the world's global GDP (Barnard, 2012). At the company level, advertising represents a substantial investment as well. A recent compilation across more than 5000 companies from 300+ industries reports an average advertising-to-sales ratio of 3.1% (AdAge, 2012). Within the consumer packaged goods (CPG) sector, ratios close to 10% (food products: 9.2%) or higher (e.g., soaps and detergents: 12.1%; cosmetics: 20%) are not uncommon. In an effort to justify these expenditures, a large literature has emerged that quantifies the effectiveness of advertising in terms of its sales elasticity. Recent reviews include, among others, Allenby and Hanssens (2005) and Sethuraman, Tellis, and Briesch (2011). The latter compiled a meta-analytic data set of 751 (402) short-run (long-run) advertising elasticities from 56 (38) publications, and report an average elasticity of 0.12 (0.24). Within a CPG setting, Van Heerde, Gijsenberg, Dekimpe, and Steenkamp (2013) studied advertising's sales effectiveness across 150 brands in the U.K., and report average short- and long-run elasticities of 0.002 and 0.013. Srinivasan, Vanhuele, and Pauwels (2010), looking at 74 brands across four categories in France, obtained values of 0.020 and 0.036, respectively.

The aforementioned numbers do not recognize, however, that different media (television, radio, magazines, newspapers, …) can differ widely in their short- and/or long-run effectiveness. Still, that information is critical when deciding on brands' budget allocation across different media (Fisher, Albers, Wagner, & Frie, 2011). To that extent, Sethuraman et al. (2011) distinguished in their meta-analysis between studies reporting, respectively, elasticities for (i) television, (ii) print media, and (iii) an aggregation across multiple media. After accounting for a variety of other factors, they found a significantly higher long-run elasticity for print advertising than for television advertising. For the short-run elasticity, this order was reversed. Other media were not considered separately, given the more limited number of studies that report their elasticities.

Recently, some studies have started to look at a wider variety of media. Naik and Peters (2009), for example, analyzed an advertising campaign for cars involving six media (television, magazines, newspapers, radio, internet banners and sponsored search). They found radio advertising to be most cost effective, followed by newspapers, TV and magazines. Danaher and Dagger (2013), in turn, report the short-run elasticities for 10 media in the context of a blitz (1-month) media campaign by an up-market Australian department store. However, given (i) the limited number of such studies, and (ii) their rather unique character (e.g., the blitz-advertising setting in the latter study acted more like a sales-promotion tool, making potential carryover or long-run effects less relevant), no empirical generalizations on those other media are available yet.

This is especially the case for some of the so-called "smaller" media, such as outdoor (billboard) and cinema advertising, which are typically excluded from consideration. Deleersnyder, Dekimpe, Steenkamp, and Leeflang (2009), for example, report across 37 countries the proportion of total advertising spent on four "key media" (p. 628), television, radio, magazines and newspapers, for which they analyze the cyclical sensitivity. However, as most other studies, they exclude (see their Table 2, footnote b) the amounts spent on cinema and outdoor media. As their share tends to be smaller, very few studies have explicitly considered the effectiveness of billboard and/or cinema advertising. Two notable exceptions are Berkowitz, Allaway, and D'Souza (2001) and Naik, Peters, and Raman (2008). The former analyzed 1 year of weekly sales, radio advertising and billboard advertising for three stores of a single regional retailer, and concluded (without reporting specific elasticities) that "radio advertising is anywhere from three to seven times (depending on the store) more effective than billboard advertising" (p. 64). Naik et al. (2008) considered a multimedia campaign for soft drinks, and reported a greater impact of TV advertising (elasticity of 0.32) relative to print (0.02), outdoor (0.06) and cinema (0.06). Again, it is hard to interpret this lower effectiveness of billboard and/or cinema advertising as an empirical regularity on the basis of just two studies, especially since the reported elasticity of TV advertising in the Naik et al. (2008) application appears to be much higher than the average value reported in the aforementioned meta-analyses, which could be attributed to their less conventional dependent variable (i.e., medium-specific advertising awareness levels rather than sales or market share).

Against this background, our study offers new substantive and managerial insights. Substantively, we contribute novel empirical generalizations on the effectiveness of advertising in two media that have been largely ignored in prior research by systematically analyzing the short- and long-run advertising effectiveness of both billboard spending and cinema advertising for a wide variety of CPG brands (40 + for cinema advertising; 100 + for billboard advertising). While the proportion spent worldwide on those media (estimated at 6.6% for outdoor and 0.6% for cinema in 2012; Barnard, 2012) may be smaller than for the more often studied television (40.2%) and print (27.7% for newspapers and magazines combined) media, the absolute spending levels remain very large (32.3 US$ billion and 2.7 US$ billion, respectively), and warrant more research attention. For the more traditional media, this paper is the first to provide empirical generalizations on sales, as well as market-share, elasticities that have been corrected for the brands' self-selection in media usage. Managerially, a better understanding of their relative effectiveness will help managers make better media-allocation decisions, not only managers who currently abstain from using those media, but also those who already allocate a substantial portion of their media budget to them.

2. Data

Through GfK Benelux (monthly revenue) and the Belgian Centre for Information about Media (advertising expenditures across media), we obtained monthly sales and advertising-spending information on 261 leading brands in the consumer packaged goods (CPG) market. The

Table 1
Data coverage.

Product class	Number of categories/brands	Examples of included categories	Examples of included brands
Food	55/132	Biscuits and cookies	Delacre
		Canned vegetables	Bonduelle
		Yoghurt	Danone
Beverages	16/50	Beers	Stella Artois
		Carbonated soft drinks	Fanta
		Mineral water	Evian
Household care	13/40	Air refreshing products	Febreze
		General cleaners	Mr. Propre
		Fabrics cleaning	Dash
Personal care	10/31	Deodorants	Axe
		Depilatory & shaving	Gillette
		Haircare	Head & Shoulders
Pet food	2/8	Dog & cat food	Sheba
		Other pet food	Vitakraft
Total	96/261		

brands cover a wide variety of 96 categories, involving food, beverage, household-care, personal-care and pet-food products (a summary is provided in Table 1). The revenue data consist of the aggregated data across all SKUs of the brand within the category at hand (e.g., combined sales across all Axe deodorants). Special care was taken that the advertising data were compiled according to the same category classification. All brands advertised at least two times during the observation period (January 2004–August 2010) in at least one of the following six advertising media: television, radio, newspapers, magazines, billboards and cinema.[4] Sales revenues and advertising expenditures were deflated based on the relative Consumer Price Index (CPI).

As shown in Tables 2 and 3, there is considerable variability in media usage. Table 2 illustrates how the different media differ in both (i) their number of users[5] and (ii) their share of the total advertising budget (adspend share). In line with previous research (see, e.g., Deleersnyder et al., 2009; Sethuraman et al., 2011), TV advertising is by far the most popular medium: 231 out of the 261 brands use TV advertising, resulting in a combined spending share of 82.9%. All other media are considerably less popular. Billboard (cinema) advertising, for example, has an adspend share of only 7.2 (1.8)%, and is used by only 117 (43) brands. Among those users, the corresponding adspend shares are considerably higher, however (column 4 of Table 2), and even exceed (with a share of 11.9 and 6.9%) the proportion spent on more traditional media such as radio (5.7%), newspapers (3.1%) and magazines (4.4%). Interestingly, brands differ considerably in the number of media they use, as summarized in Table 3. While 52 (20%) brands use only a single medium, other brands (18) use all six media.

Besides sales revenues,[6] GfK provided information on the brands' penetration level and market share. These variables will be used to address potential sample-selection issues (we refer to the Methodology section for details), and to explore differences in both media usage and media effectiveness.

---

[4] No internet spending was considered. The data collection on this medium was started two years after that on the other media, which would have reduced considerably the length of the time series.

[5] A brand is considered a user of a certain advertising medium if it uses the medium at least two times during the observation period (a similar decision rule was recently used by Van Heerde et al., 2013).

[6] No price information was available. However, cross-sectional differences in average price level across brands in a given category are captured through our brand-specific intercepts, as explained in the Methodology section. Moreover, Sethuraman et al. (2011), Table 2) show in their meta-analysis on brand-level advertising elasticities that the absence of a price variable in the model has no significant impact on the resulting advertising-effectiveness estimates, neither in the short run nor in the long run.

Table 2
Media usage.

Medium	Adspend share sample-wide	Number of users	Adspend share among users
TV	82.9%	231	83.6%
Radio	2.6%	81	5.7%
Newspaper	1.9%	123	3.1%
Magazine	3.6%	185	4.4%
Billboard	7.2%	117	11.9%
Cinema	1.8%	43	6.9%

Table 3
Number of media used.

Number of media used	Number of brands
1	52
2	57
3	60
4	44
5	30
6	18
Total	261

3. Methodology

Our modeling approach consists of two steps: (i) an analysis at the brand level to estimate the advertising elasticities for the different media used by each individual brand, and (ii) a meta-analysis to combine these results into empirical generalizations. In our first step, i.e. the brand-specific analysis, we closely follow Van Heerde et al. (2013), and use an error-correction model to estimate each medium's short- and long-run advertising elasticity, while we correct for endogeneity (and allow for correlated error terms within a given product category) through a 3SLS estimation procedure. However, unlike Van Heerde et al. (2013), where *all* brands made use of the (aggregated across media) advertising instrument, we are confronted with an intrinsic selection problem. Indeed, many brands make no use of certain advertising channels (as documented in Table 3). We explicitly account for this selection issue in our second-stage meta-analysis.

3.1. Brand-specific analysis

An error-correction model is used to estimate both the short- and long-run advertising elasticities of each medium used by each individual brand. If i represents the brand (i.e., $i = 1,...,B_c$) and c the category the brand belongs to ($c = 1,...,96$), the error-correction model is written as:

$$\Delta \ln \text{Revenue}_t^{ic} = \alpha^{ic} + \sum_{j=1}^{J^{ic}} \beta_j^{ic} \Delta \ln \text{Advertising}_t^{icj} + \beta_{J^{ic}+1}^{ic} \Delta \ln \text{Other}_t^{ic}$$
$$+ \varphi^{ic} \left[\ln \text{Revenue}_{t-1}^{ic} - \sum_{j=1}^{J^{ic}} \gamma_j^{ic} \ln \text{Advertising}_{t-1}^{icj} - \gamma_{J^{ic}+1}^{ic} \ln \text{Other}_{t-1}^{ic} - \delta^{ic} \text{trend} \right] + \varepsilon_t^{ic},$$
(1)

where J^{ic} denotes the number of media for which brand i from category c had (in line with the decision rule of Van Heerde et al., 2013) at least two non-zero spending levels. The β_j^{ic} coefficients represent the short-run (same period) advertising-to-sales elasticities of TV, radio, newspaper, magazine, billboard or cinema, as $\Delta \ln \text{Advertising}_t^{icj}$ gives the first difference of the log-transformed advertising expenditures. The γ_j^{ic} parameters represent their long-run counterparts, and reflect the cumulative effect (same period + future periods) of a one-period shock to the medium at hand.[7] $\beta_{J^{ic}+1}^{ic}$ and $\gamma_{J^{ic}+1}^{ic}$ capture the combined short-run and long-run advertising elasticity across media that were used only once in the observation period. The φ^{ic} parameter reflects the speed of adjustment towards the underlying long-run equilibrium, and the trend variable serves as a proxy for all other variables that have gradually changed over the sample period (cf. Dekimpe & Hanssens, 1995). This trend variable ranges from -1 to $+1$ in order to interpret the elasticities mid-observation period. An in-depth discussion of the error-correcting specification is provided in Fok, Horváth, Paap, and Franses (2006). Its use is well-established in the marketing literature (see, e.g., Horváth & Fok, 2013; Van Heerde et al., 2013, Van Heerde, Helsen, & Dekimpe, 2007 for other applications), and is particularly suited for our research

setting, as it provides direct estimates of the different media's short- and long-run elasticities, without the need to impose a common carry-over coefficient (as is often done in Koyck-type and partial-adjustment-type models involving multiple media; see Naert & Leeflang, 1978, pp. 94–97).[8] 3SLS estimation is used to account for the possible endogeneity of the $\Delta \ln \text{Advertising}_t^{icj}$ variables.[9] The averaged values of the lagged log advertising expenditures from the other product classes, disaggregated across the six advertising media, serve as instruments (see Van Heerde et al., 2013 or Lamey, Deleersnyder, Steenkamp, & Dekimpe, 2012 for a similar practice). As we have 24 instruments (six media across four non-focal product classes) for a maximum of six endogenous variables, the model is overidentified. The model is estimated jointly across the B_c brands within a given category c, allowing their error terms to be correlated.

3.2. Meta-analysis

In the second step of our analysis, we meta-analytically combine the brand-specific parameter estimates, and derive empirical generalizations on the relative effectiveness of the different media. Unlike Van Heerde et al. (2013), we cannot rely on Rosenthal's method of added Zs (Rosenthal, 1991). As we estimate the elasticities at a more disaggregate level (i.e., at the level of the individual medium, rather than summed across all media—in which case all included brands are a user), and given that not all media are used by every brand, there is a potential issue of sample selection. As brands self-select which media to use, only those brands for which a medium is likely to be (more) effective may do so, and their estimates may not generalize to the other brands. As such, we have to test (and correct) for this potential self-selection bias. In addition, we need to account for the fact that (i) the elasticities are estimated quantities, and (ii) that estimates from brands within the same category may not be independent of one another.

Two meta-regressions are constructed for, respectively, the stacked short-run elasticity estimates (2a) and the stacked long-run estimates (2b) from the brand-specific analyses:

$$\beta_m^{ic} = \sum_{k=1}^{6} \beta_k D_k^{ic}(m) + \sum_{k=1}^{6} \rho_k^{SR} \lambda_k^{ic} D_k^{ic}(m) + u_m^{ic},$$
(2a)

$$\gamma_m^{ic} = \sum_{k=1}^{6} \gamma_k D_k^{ic}(m) + \sum_{k=1}^{6} \rho_k^{LR} \lambda_k^{ic} D_k^{ic}(m) + v_m^{ic}.$$
(2b)

The index m denotes the medium at hand (TV, radio, newspapers, magazines, billboards or cinema), with each brand contributing J^{ic} elasticity estimates. $D_k^{ic}(m)$ is an indicator variable, taking on the value 1 when $k = m$. As such, β_k and γ_k provide the expected (short- and long-run) effectiveness of each medium k.

However, their estimates may not represent the expected short- and long-run advertising elasticities for a randomly-selected CPG brand because of the aforementioned self-selection into the sample. In fact, they

---

[7] This error-correction interpretation presumes stationarity of the (log) sales series, which was confirmed through the Levin, Lin, and Chu (2002) panel unit-root test. A similar conclusion was obtained on the basis of brand-specific Phillips and Perron (1988) tests (detailed results are available from the first author upon request).

[8] Note that the autoregressive partial-adjustment model will be applied as a robustness check in Section 4.4.

[9] $\ln \text{Advertising}_{t-1}^{icj}$ and $\ln \text{Other}_{t-1}^{ic}$ are predetermined variables, and are thus not endogenous. We treat $\Delta \ln \text{Other}_t^{ic}$ as exogenous, as too few non-zero observations might prevent a reliable auxiliary-regression estimation (Steenkamp et al., 2005).

would only pertain to those brands that actually use the medium. To correct our meta-analytic results for this potential selection bias, we include in Eqs. (2a) and (2b) λ_k^{ic} variables resulting from Heckman's two-stage method (Greene, 2000). For each of the six media, a probit model is estimated to quantify the probability that the medium is used by brands with certain characteristics:

$$\Pr\left(s_m^{ic} = 1 \mid z_m^{ic}\right) = \Phi\left(z_m^{ic}\,\xi_m\right). \tag{3}$$

The dummy variable s_m^{ic} is a selection indicator, which takes on the value 1 when brand i from CPG category c is a user of medium m. z_m^{ic} is a 1×12 vector of variables that might have explanatory power over the decision whether or not a brand is a user of a given medium. This vector includes (i) an intercept, (ii) four product-class dummies indicating whether or not the brand is a beverage, food, personal-care or household-care brand (with pet food as reference category), (iii) the average values of the log advertising expenditures for each of the five other media, (iv) the average market share of the brand during the observation period, and (v) its average penetration level. Using the fitted values of this probit model, the corresponding inverse Mills ratio is derived as:

$$\lambda_m^{ic} = \frac{\phi\left(z_m^{ic}\hat{\xi}_m\right)}{\Phi\left(z_m^{ic}\hat{\xi}_m\right)}, \tag{4}$$

which is added to the meta-regression (2) as an additional explanatory variable.

When estimating (2), one needs to account for the fact that the dependent variables β_m^{ic} and γ_m^{ic} are estimated quantities, as are the inverse Mills ratios. Moreover, the error terms corresponding to brands from the same category may not be independent of one another. To accommodate the first issue (estimated dependent variables), we use weighted least squares (WLS) to obtain parameter estimates in the meta-regressions. The weights are set to equal the inverse of the standard errors of the β_j^{ic} and γ_j^{ic} estimates from the 3SLS estimation (see Saxonhouse (1976) for a formal motivation, or Nijs, Srinivasan, and Pauwels (2007) for another marketing application). To address the two other issues (estimated explanatory variables and correlated error terms), the standard errors for the meta-analytic parameters are derived by means of a bootstrap procedure, where we also account for clustered error terms between brands of the same product category (Field & Welsh, 2007).

4. Results

4.1. Small-medium usage

All six probit selection models (3) resulted in a high hit rate (ranging between 71 and 89%), which was substantially higher than the number expected by chance (Morrison, 1969). Even though not our primary research interest, we obtained several interesting insights (detailed results are provided in Online Appendix A) into which brands/categories are most likely to *use* the less popular billboard and cinema media. First, brands with a higher penetration level (two-sided $p < 0.10$) are more likely to use the billboard medium.[10] In contrast, high-share brands are not more/less likely ($p > 0.10$) to use billboard or cinema advertising than smaller brands. Interestingly, brands that make more extensive use of the TV medium are also more likely (two-sided $p < 0.05$) to use cinema advertising, while brands that make more extensive use of the magazine medium are also more likely (two-sided $p < 0.05$) to use billboard advertising. These combinations make intuitive sense, as tactical consistency (Sheehan & Doherty, 2001) may be easier to realize

[10] This is also the case for television advertising.

between TV and cinema advertising on the one hand, and between magazine and billboard advertising on the other hand, which should facilitate consumers' information processing when these media pairs are used together in a coordinated media campaign (Edell & Keller, 1989). Finally, significant differences between the different category types are present as well. Billboard advertising, for example, is more likely to be used by brands in the food and beverage category, and less likely in the household-care, personal-care and pet-food category, while beverage brands are typically more inclined to use cinema advertising compared to food brands.

4.2. Brand-specific findings

Table 4 summarizes the individual-level elasticity estimates. For seven brands, preliminary single-equation estimations of the sales equation resulted in a variance inflation factor (VIF) in excess of 10, indicating serious multicollinearity issues (Hair, Anderson, Tatham, & Black, 1995). We omit these brands when reporting the proportion of significant effects (and from our subsequent meta-analysis).[11] Focusing on the long-run parameter estimates for the remaining 254 brands, we find that 141 (18.7%) of the 755 estimated advertising elasticities are positive and significant ($p < 0.10$, one-sided).[12] This proportion is comparable to the one obtained in other CPG-based studies (see, e.g., Van Heerde et al., 2013 for the U.K., and Steenkamp, Nijs, Hanssens, & Dekimpe, 2005 for the Netherlands). Interestingly, the proportion of significant estimates for TV (28%) is significantly higher ($p < 0.05$) than for the other media, offering a first justification for its more frequent use. Still, even for that medium, no significant effect is found for a substantial number of brands, in line with earlier findings (see, e.g., Lodish et al., 1995; Allenby & Hanssens, 2005) that advertising for mature products often fails to induce significant sales increases. Moreover, a large number of brands (143 or 56.3%) does not obtain a significant effect with any medium, and only 26 brands (out of the 202 that use at least two media) have a significant long-run elasticity for more than one (with a maximum of three) medium. This again suggests that, at least from a sales-response perspective, advertising spending can often be perceived, in the terminology of Leeflang and Wittink (1996) and Steenkamp et al. (2005), as spoiled arms.

Considering our two focal media, billboards (12.9%) and cinema (15.0%), their proportion of significant effects is not significantly lower ($p > 0.10$) than for the more traditional radio (18.4%), newspaper (12.8%) and magazine (15.5%) media. These numbers exceed (albeit only marginally) the percentage that would be expected by chance (10%), but are (as indicated before) significantly smaller than the proportion obtained for TV advertising.

Only 26 of the 202 brands that actively use more than one medium obtain a significant long-run elasticity for multiple media. This raises serious concerns about the optimality of the allocation rules used by many brands. Indeed, almost all brands allocate resources to one or more non-effective (from a sales-response point of view) media. For only 11 of these 26 brands, one of the significant long-run elasticities involved billboard (8) and/or cinema advertising (4). Comparing the relative share allocated to these media (e.g., spending on billboard relative to the spending on all "effective" media) with the near-optimal allocation heuristic developed in Fisher et al. (2011),[13] we note (i) that among the eight billboard users, five brands are "on target." They allocate a relative adspend share to billboard advertising (relative referring to the shares allocated to the other significant media) that lies within the 90% confidence interval for the near-optimum. The other brands significantly

[11] We do so because the resulting inflation of the standard errors would render the estimation results too imprecise.
[12] In Online Appendix B, we also report the proportion of positive (not necessarily significant) estimates.
[13] According to Fisher et al.'s heuristic, this proportion should be equal to the ratio between that medium's own long-run elasticity and the sum of all the significant long-run elasticities.

Table 4
Brand-specific findings.

Medium	Number of users	Share of significant[a] positive short-run elasticities (β_j^{ic})	Share of significant[a] positive long-run elasticities (γ_j^{ic})
TV	225	19.6%	28.0%
Radio	76	14.5%	18.4%
Newspaper	117	12.0%	12.8%
Magazine	181	9.9%	15.5%
Billboard	116	16.4%	12.9%
Cinema	40	7.5%	15.0%

(a) One-sided, 10% level.

underspend on this type of advertising, with an average deficit in relative budget share of 39.6 percentage points. For cinema advertising, the results on adspend share allocation are not as conclusive (one brand is on target, while one (two) brand(s) significantly overspends (underspend) on cinema advertising).

4.3. Meta-analytic findings

One could argue that the rather bleak picture on advertising's effectiveness may be attributed to the low power of the statistical tests, as "only" 80 observations are available for each brand. To increase the power of our inference, we meta-analytically combine the evidence across all brands that use a particular medium (see, e.g., Deleersnyder, Dekimpe, Sarvary, & Parker, 2004, Deleersnyder et al., 2009; Lamey et al., 2012; Van Heerde et al., 2013 for a similar reasoning). In both the short- (television, radio and newspapers) and the long-run (television, newspapers, billboards) equation, several of the ρ_k estimates are significant ($p < 0.10$, one-sided), indicating the need to control for self-selection. This was also confirmed when looking at the combined evidence across the six ρ parameters by means of the Strube (1985) test (see Deleersnyder et al., 2004 for a marketing application), which was highly significant in both instances (one-sided $p = 0.005$ and 0.013 for Eqs. (2a) and (2b), respectively).

The resulting elasticity estimates are summarized in Table 5. In both the short and long run, the overall advertising elasticity is only significant for TV and magazines. In other words, for a random CPG brand, TV and magazines are the only media that are able to significantly affect the brands' sales revenues, with television the most effective ($p < 0.001$) of the two. For the other media, including billboards and cinema, no meta-analytic evidence of their effectiveness is found (short-run p-values of 0.152 (billboards) and 0.317 (cinema); long-run p-values of 0.451 and 0.148, respectively).

Comparing the meta-analytic results *with* and *without* correction for self-selection, we find that in the latter case, the estimate for the long-run TV elasticity is biased upwards (0.0070 versus 0.0043). Moreover, for the other media, a significant long-run effectiveness is also found for radio advertising and billboard advertising (in terms of short-run effectiveness, three additional effects become significant, i.e. radio, newspapers and billboards). Hence, not accounting for the fact that the observed set of users of a given medium may not be fully random leads one to seriously overestimate its effectiveness when generalizing to the overall population of brands.

4.4. Robustness checks

4.4.1. Alternative performance metric: market share

The earlier discussion focused on the impact of the different media on the brands' sales revenues, which can reflect both competitive gains and gains because of primary market expansion (Sethuraman et al., 2011). In a first robustness check, we replaced the dependent variable $\Delta \ln \text{Revenue}_t^{ic}$ in the brand-specific Eq. (1) by $\Delta \ln \text{MarketShare}_t^{ic}$, which only captures the competitive gains. All other steps of the analytic

approach remained unchanged,[14] except that we now used the brand's revenues (rather than its share) as an explanatory variable in the probit selection models for media usage. The results were very robust across both dependent measures: (i) no significant effect was found for the two focal small media, i.e. billboard and cinema, neither in the short run (Table 6a) nor in the long run (Table 6b), (ii) this was also the case for radio and newspaper advertising, while (iii) a significant effect was again found for TV advertising. The only exception was the magazine medium, where the significant effect in terms of sales revenue appears to be attributable to a primary-demand, rather than a competitive-gain, effect (as the impact becomes insignificant in the market-share model).

4.4.2. Alternative model specification: autoregressive partial-adjustment model

To ensure that the lack of empirical evidence on the short-run and long-run effectiveness of the small advertising media is not idiosyncratic to the error-correction specification used in Eq. (1), we re-estimated the brand-specific advertising elasticities using the autoregressive partial-adjustment model (APAM) discussed in Hanssens, Parsons, and Schultz (2001, p. 147) and Leeflang, Wittink, Wedel, and Naert (2000, p. 489). The APAM specification is very similar to the Koyck model, but is somewhat more flexible, in that it does not impose the restriction that the autocorrelation parameter in the error term equals the coefficient associated with the lagged dependent variable. Specifically:

$$\ln \text{Revenue}_t^{ic} = \left(1 - \lambda^{ic}\right)\left[\alpha^{ic} + \sum_{j=1}^{J^c}\beta_j^{ic}\ln \text{Advertising}_t^{icj} + \beta_{J^c+1}^{ic}\ln \text{Other}_t^{ic}\right] + \lambda^{ic}\ln \text{Revenue}_{t-1}^{ic} + \varepsilon_t^{ic},$$

(5)

with $\varepsilon_t^{ic} = \tau^{ic}\varepsilon_{t-1}^{ic} + w_t^{ic}$, denoting that w_t^{ic} and ε_t^{ic} follow a white-noise respectively AR(1) process. In this specification (see also Bass & Clarke, 1972), the β_j^{ic} coefficients represent the long-run advertising-to-sales elasticities of TV, radio, newspaper, magazine, billboard or cinema, while the short-run elasticities can be calculated as $(1 - \lambda^{ic})\beta_j^{ic}$. λ^{ic} gives the brand-specific carryover parameter that is instrumental in linking the short- and long-run elasticities. Again, similar conclusions were obtained for the two focal small media (with no significant meta-analytic effect in the short or long run), for radio and magazine advertising (with no significant effect either), and for television (with a significant short- and long-run effect ($p < 0.01$)). For magazine advertising, a comparable short- (0.0024 versus 0.0019) and long-run (0.0041 versus 0.0059) elasticity as before was obtained, even though the latter failed to reach statistical significance in the APAM specification.

4.4.3. Alternative diminishing returns to scale

The error-correction model was specified in the popular log-log format, which offers direct estimates of the respective elasticities, and which captures diminishing returns to scale. To verify that our main insights remain unchanged with a different rate of diminishing returns to scale, we adopted the specification of Naik and Raman (2003), where we model the dependent variable (sales revenues) as a function of the square root of the media expenditures (see also Leeflang et al., 2000, p. 69). Across all six media, the same conclusions[15] as in our focal

---

[14] As Leeflang and Wittink (1996), we did not consider the logical-consistency requirements in our market-share validation, for two reasons. First, empirical evidence on the relative performance of multiplicative versus attraction specifications remains inconclusive (we refer to their footnote 10 for an extensive literature review on the issue). Second, for several categories, information on only a small subset of the brands in the category was available, making a full attraction specification infeasible. Also, given that our data set contains a large number (96) of micro-categories, the approach used in Fok, Paap, and Franses (2003), where one would use one brand per (micro-) category as reference brand, would lead to a substantial reduction in the number of elasticities that could be obtained per medium, making also that approach less appealing.

[15] Given the different scaling of both the dependent (log-transformed or not) and independent (log-transformed or square root) variables, the parameter estimates are not directly comparable in magnitude across both specifications. As such, we focus on the sign and significance in this comparison.

Table 5
Meta-analytic findings.

| | | Without self-selection correction | | | | With self-selection correction | | | |
| | | Short-run elasticity (β_k) | | Long-run elasticity (γ_k) | | Short-run elasticity (β_k) | | Long-run elasticity (γ_k) | |
	Number of contributing brands	Estimate[a]	One-sided p-value	Estimate[a]	One-sided p-value	Estimate[a]	One-sided p-value	Estimate[a]	One-sided p-value
TV	225	**0.0037**	0.000	**0.0070**	0.000	**0.0026**	0.002	**0.0043**	0.000
Radio	76	**0.0042**	0.002	**0.0045**	0.019	−0.0001	0.513	0.0024	0.249
Newspaper	117	**0.0028**	0.034	0.0009	0.356	0.0006	0.372	−0.0022	0.759
Magazine	181	**0.0033**	0.002	**0.0042**	0.000	**0.0024**	0.040	**0.0041**	0.002
Billboard	116	**0.0027**	0.020	**0.0037**	0.049	0.0017	0.152	0.0004	0.451
Cinema	40	0.0017	0.269	0.0017	0.245	0.0028	0.317	0.0059	0.148

(a) Effects significant at $p < 0.10$ (one-sided) are put in bold.

model were obtained, i.e. insignificant meta-analytic effects for billboard, cinema, newspaper and radio, and a significant short- and long-run effect for TV and magazines. This was also the case when we varied, in the spirit of Lamey, Barbara, Dekimpe, and Steenkamp (2007), Table 4), the rate of diminishing returns further, and used Advertising$^{0.4}$ and Advertising$^{0.6}$ instead of Advertising$^{0.5}$.

4.4.4. Stability

As a final robustness check, we tested the stability of the obtained insights on media effectiveness through a split-half analysis. To that extent, Eq. (1) was augmented as follows:

$$\Delta \ln \text{Revenue}_t^{ic} = \alpha^{ic} + \sum_{j_1=1}^{J_1^{ic}} \beta_{1,j}^{ic} \Delta \ln \text{Advertising}_t^{icj} S_t^1 + \sum_{j_2=1}^{J_2^{ic}} \beta_{2,j}^{ic} \Delta \ln \text{Advertising}_t^{icj} S_t^2$$
$$+ \beta_{1,J_1^{ic}+1}^{ic} \Delta \ln \text{Other}_t^{ic} S_t^1 + \beta_{2,J_2^{ic}+1}^{ic} \Delta \ln \text{Other}_t^{ic} S_t^2$$
$$+ \varphi^{ic} \left[\ln \text{Revenue}_{t-1}^{ic} - \sum_{j_1=1}^{J_1^{ic}} \gamma_{1,j}^{ic} \ln \text{Advertising}_{t-1}^{icj} S_t^1 - \sum_{j_2=1}^{J_2^{ic}} \gamma_{2,j}^{ic} \ln \text{Advertising}_{t-1}^{icj} \right.$$
$$\left. S_t^2 - \gamma_{1,J_1^{ic}+1}^{ic} \ln \text{Other}_{t-1}^{ic} S_t^1 - \gamma_{1,J_1^{ic}+1}^{ic} \ln \text{Other}_{t-1}^{ic} S_t^2 - \delta^{ic} \right] + \varepsilon_t^{ic}, \quad (6)$$

with S_t^1 and S_t^2 dummy variables representing the two subsamples.[16] J_1^{ic} and J_2^{ic} denote the number of media for which brand i from category c had at least two non-zero spending levels in, respectively, subsample 1 and 2. Subsequently, we estimated subsample-specific (sample-selection-corrected) weighted averages of the brand-specific elasticities per medium:

$$\beta_{m,s}^{ic} = \sum_{l=1}^{2} \sum_{k=1}^{6} \beta_{k,l} D_k^{ic}(m) D_l^{ic}(s) + \sum_{k=1}^{6} \rho_k^{SR} \lambda_k^{ic} D_k^{ic}(m) + u_m^{ic}, \quad (7a)$$

$$\gamma_{m,s}^{ic} = \sum_{l=1}^{2} \sum_{k=1}^{6} \gamma_{k,l} D_k^{ic}(m) D_l^{ic}(s) + \sum_{k=1}^{6} \rho_k^{SR} \lambda_k^{ic} D_k^{ic}(m) + v_m^{ic}, \quad (7b)$$

where $\beta_{m,s}^{ic}$ ($\gamma_{m,s}^{ic}$) consist of the estimated short-run (long-run) elasticities of the different media for both subsamples (i.e., $s = 1, 2$ refers to the first, respectively second, subsample). $D_l^{ic}(s)$ is an indicator variable, taking on the value 1 when $l = s$. The final colucxmns in Table 6 report the respective estimates. As before, similar substantive conclusions were obtained, with insignificant effects (both in the short and in the long run) for billboard, cinema, radio and newspaper advertising (in both subsamples), and a highly significant ($p < 0.01$) effect for TV advertising (again in both subsamples). In addition, three of the four magazine elasticities were found to be significant ($p < 0.10$). Moreover, none of the elasticities found in the second half of the sample was significantly different ($p > 0.10$) from the corresponding value in the first subsample, which again confirms the stability of our inferences over time.

In sum, our results were found to be very robust across all sensitivity analyses. We consistently found no significant meta-analytic effect for the two focal small media, nor for the more traditional radio and newspaper media. In contrast, we each time found a significant short- and long-run effect for TV advertising. The impact of magazine spending,

in turn, while robust in sign and magnitude, sometimes failed to reach statistical significance.

4.5. Heterogeneity across brands and categories

In the spirit of Van Heerde et al. (2013), we also derived the small media's (average) effectiveness in some major product groups (beverages, food, household care, personal care, and pet food for billboards, and beverages, food and personal care for cinema).[17] No significant meta-analytic elasticity was found (all p-values > 0.10) in any of these category types, neither in the short nor in the long run. We also checked (one at a time)[18] for a moderating impact of the brands' (i) average market share, (ii) penetration level, and (iii) advertising intensity (measured as the combined spending across all six media relative to the brands' sales revenue). However, the total (meta-analytic) impact of the two small media (main + moderating effect) never reached statistical significance (i.e., $p > 0.10$) across the range of these moderators observed among their users.

4.6. How about synergy effects?

Even in the absence of significant "own" effects, allocating a portion of one's advertising budget to a certain medium may be justified when it enhances the effectiveness of another medium (Naik, 2007, p. 44; Naik & Raman, 2003). To this end, we investigated to what extent billboard and cinema advertising have a positive synergistic effect with television, the most frequently used (Table 2) and most effective (Tables 4 and 5) medium.

Specifically, we augmented Eq. (1) with an interaction term between both media, both in the short-run (i.e., between the first difference terms) and the long-run (i.e., between the lagged level terms) part. Seventy brands were a joint user of both TV and billboard (i.e., had at least two joint spending occurrences), of which 41 had a VIF value smaller than 10 after the inclusion of the relevant interaction terms. Of these 41 brands, 3 (6) experienced a significantly positive ($p < 0.10$, one-sided) short-run (long-run) synergy effect from this joint spending, a proportion not exceeding what could be expected by chance (to correct for potential self-selection, we estimated, in a similar spirit as before, a probit model on an indicator variable taking the value of 1 if the brand was a user of both media, and zero otherwise, to derive the corresponding inverse Mills ratio). Also meta-analytically, no significant synergy effect was found in the short or long run ($p > 0.10$). As for cinema advertising, only 16 brands passed both criteria (at least two joint occurrences, and no evidence of serious multicollinearity). In this case, only 2 (0) brands experience a significant long-run (short-run) synergy effect. Across the 16 brands, no significant meta-analytic effect was found ($p > 0.10$). Hence, little evidence is found in support of the synergy claim that is often raised (see, e.g., Bhargava & Donthu, 1999,

---

[16] $S_t^1 = 1$ when $1 \le t \le 40$ while $S_t^2 = 1$ when $40 < t \le 80$.

[17] No pet-food or household-care brand in our sample made use of cinema advertising.
[18] Specifically, we added interaction terms between a given moderator and the six media dummies to Eq. (2a,2b).

Table 6
Robustness checks on the short- (a) and long-run (b)meta-analytic results.

	Focal model	Alternative performance metric: market share	Alternative model specification: APAM[i]	Alternative DRS[ii]			Stability	
				$\eta = 0.4$	$\eta = 0.5$	$\eta = 0.6$	First 50% of the data	Last 50% of the data
(a)								
TV	**0.0026**	**0.0009**	**0.0025**	358.64	99.48	27.58	0.0023	0.0024
Radio	−0.0001	0.0007	−0.0022	−468.19	−160.24	−52.63	−0.0032	0.0023
Newspaper	0.0006	−0.0006	−0.0013	−604.93	−165.06	−41.15	−0.0050	−0.0028
Magazine	**0.0024**	0.0003	**0.0019**	581.47	197.10	64.14	0.0023	**0.0058**
Billboard	0.0017	0.0003	0.0004	330.60	82.89	24.39	0.0077	0.0033
Cinema	0.0028	0.0017	0.0014	−786.19	−292.04	−89.69	−0.0004	−0.0024
(b)								
TV	**0.0043**	**0.0017**	**0.0191**	554.15	155.25	42.68	0.0036	0.0055
Radio	0.0024	0.0004	−0.0282	−605.71	−223.02	−74.15	−0.0011	0.0047
Newspaper	−0.0022	−0.0021	0.0117	−846.13	−220.28	−53.32	−0.0081	−0.0039
Magazine	**0.0041**	0.0009	0.0059	992.53	329.15	106.32	0.0054	0.0058
Billboard	0.0004	−0.0014	−0.0411	540.85	147.45	42.79	0.0124	0.0069
Cinema	0.0059	0.0023	−0.0092	526.73	120.01	34.96	0.0011	0.0031

Effects significant at $p < 0.10$ (one-sided) are put in bold.
(i) APAM = autoregressive partial-adjustment model.
(ii) DRS = diminishing returns to scale, with η denoting the according rate, i.e. Advertising$^\eta$.

Table 7
Tests for synergy effects with the small media.

Synergy between	Number of common users[a]	Share (%) of significant[b] positive estimates (short | long run)	Meta-analytic estimate[c] (short | long run)	Meta-analytic one-sided p-value (short | long run)
Billboard & TV	41	7 | 15	−0.0005 | 0.0005	0.998 | 0.232
Billboard & Radio	24	8 | 17	−0.0010 | **0.0045**	0.919 | 0.011
Billboard & Newspaper	31	6 | 10	−0.0005 | 0.0011	0.953 | 0.161
Billboard & Magazine	42	5 | 12	−0.0001 | 0.0006	0.579 | 0.304
Billboard & Cinema	13	8 | 8	0.0009 | 0.0024	0.231 | 0.369
Cinema & TV	16	0 | 13	0.0008 | −0.0019	0.191 | 0.776
Cinema & Radio	8	25 | 13	0.0018 | −0.0067	0.198 | 0.982
Cinema & Newspaper	9	22 | 22	−0.0010 | 0.0026	0.666 | 0.254
Cinema & Magazine	12	0 | 0	0.0007 | 0.0025	0.347 | 0.236

(a) This number only includes those brands with VIF values smaller than 10.
(b) One-sided, 10% level.
(c) Effects significant at $p < 0.10$ (one-sided) are put in bold.

p. 8; Ewing, du Plessis, & Foster, 2001, p. 78) to motivate a more extensive use of these smaller media.

Following a similar procedure, we checked (one media combination at a time)[19] for a positive synergy effect between our two focal media (billboard and cinema) and the three other media (radio, newspapers, and magazines). Table 7 summarizes the various synergy tests. For cinema advertising, no significant synergy effect was obtained (neither in the short nor in the long run) with any of these media ($p > 0.10$ in all instances). This was also the case for billboard advertising when combined with newspaper and magazine advertising. However, a positive long-run synergy effect, estimated at 0.005, was found between billboard and radio advertising. This effect was estimated on 24 brands that used both media simultaneously in at least two instances (and for which the VIF values stayed below 10 after the inclusion of the relevant interaction terms). The synergy effect was significant ($p < 0.10$) for four individual brands, as well as meta-analytically ($p < 0.05$). Even though the long-run main effect stayed insignificant,[20] the *total* impact of billboard spending (consisting of the main and moderated impact) was significant ($p < 0.10$) in case of positive spending levels on radio. Still, brands often seem to ignore this opportunity. Indeed, many brands

(89) used billboard advertising without ever using it simultaneously with radio spending. Across the 24 brands that did use both media jointly, billboard advertising coincided with radio spending in only 27% of the time periods that the billboard medium was used (and in only 4% of all possible calendar months). Hence, in many instances, brands fail to capitalize on the potential synergy effects that a joint use of both media could entail.

Finally, we also checked for a synergy effect between both small media, but no such evidence ($p > 0.10$) was found, neither in the short run nor in the long run.

5. Conclusion

Even though numerous studies have quantified the sales effectiveness of advertising spending, most have either focused on *aggregated* spending (i.e., summed across media), or on the most popular (*television*) medium (e.g., Sethuraman et al., 2011). In this study, we provide some new empirical generalizations on the relative effectiveness of less-frequently used, and definitely much less studied, media such as billboard and cinema advertising. Importantly, we show the need to correct for self-selection when making inferences, not only on the effectiveness of these smaller media, but also on the more traditional media.

Using a rich data set on over 250 popular CPG brands, we find very little evidence in support of the small media's sales effectiveness: (i) a significant short- and/or long-run elasticity is found for only a small fraction (16.4 and 12.9% for billboards; 7.5 and 15.0% for cinema advertising) of brands, and (ii) also meta-analytically, no significant positive

---

[19] Retaining all possible interaction terms jointly would lead to excessive multicollinearity and unstable results. In the spirit of Bijmolt, Van Heerde, and Pieters (2005), we opted instead to add one media combination at a time to our base model (1). We refer to Van Heerde et al., 2013; Steenkamp & Geyskens, 2014, or Yeung, Ramasamy, Chen, & Paliwoda, 2013 for a similar practice.

[20] The short-run main effect of billboard, in contrast, became significant ($p < 0.05$) after allowing for an interaction effect for those 24 brands.

effect ($p > 0.10$) was obtained. Moreover, no evidence of synergy effects with the most popular and most effective medium (TV) was found—precluding another justification for their use. Only with one other medium (radio) did we find evidence of a positive (long-run) synergy effect. However, only in very few instances (both in terms of the number of brands and in terms of the number of time periods) did brands try to capitalize on this opportunity.

These results may sound discouraging. Indeed, both managers and academics prefer significant results with large effect sizes. However, as emphasized in Hubbard and Armstrong (1992) and Sawyer and Peter (1983), there is clear value in the identification of a null result, especially when this result is obtained in many replications (in our case, across many brands and categories), and when this also holds up in the much more powerful meta-analytic tests (see also Fagley, 1985). Moreover, it is important to realize that our results (both in terms of the proportion of significant effects and in terms of the average effect size) are in line with many of the earlier findings on advertising's limited sales effectiveness in mature CPG categories (e.g., Srinivasan et al., 2010; Van Heerde et al., 2013). This could be interpreted as evidence that advertising spending in general, and in the smaller billboard and cinema media in particular, reflects ineffective/spoiled arms in those markets.[21] However, it is interesting to note that, without self-selection correction, the meta-analytic effect for billboards was significant (cf. Table 5). Hence, the (small) subset of brands that makes use of this medium appears well-informed (on average) that this medium works for them.[22] Also, one should not forget that advertising (also in those smaller media) may well have other benefits, such as a reduced vulnerability to competitive actions (Van Heerde et al., 2007), a lower sensitivity to cyclical fluctuations (Deleersnyder et al., 2009), a lower private-label growth (Lamey et al., 2012), an enhanced visibility to financial stakeholders (Joshi & Hanssens, 2010), or the stimulation of word-of-mouth communication among potential customers, which is an essential element in any successful grassroots marketing campaign (Stephen & Galak, 2012), to name just a few (see also Allenby & Hanssens, 2005 or Dekimpe & Hanssens, 2011 for a similar reasoning). Also, we focused on the sales (and market-share) effectiveness of the different media. It would be interesting to also study, in the spirit of Srinivasan et al. (2010) and Pauwels, Erguncu, and Yildirim (2013), how they influence intermediate mind-set metrics, such as brand recognition or brand liking.

Also within the sales-responsiveness domain, multiple avenues for future research remain. First, it would be useful to control for other marketing-mix instruments, such as distribution and salesforce, when estimating the brand-specific advertising elasticities in Eq. (1). However, these are likely to be positively correlated with current-period sales and with current-period advertising. As such, one can expect the omission of distribution (salesforce) *to bias the advertising-elasticity estimate positively* (Sethuraman et al., 2011, p. 461, italics added). This was indeed found in the meta-analysis of Sethuraman et al. (2011; p. 462, p. 468). Given this result, the absence of any significant effect in our model without these two measures makes our "spoiled arms" conclusion a conservative one.[23] Still, we agree that it would be useful to add specific (also longitudinal) information on those variables. Second, and in spite of their lower overall effectiveness, significant billboard or cinema effects are obtained for certain brands. Indeed, 20.7% (17.5%) of the users obtained either a significant short- or long-run response from their investments in billboard (cinema) advertising. In several instances, this effect was even larger than for any of the other media used

by the brand. Hence, more research is warranted not only to identify (beyond the potential moderators studied in Section 4.5) for what brands, categories and/or market conditions the occurrence of such positive effects becomes more likely, but also to identify what creative (more qualitative) aspects of the campaign make a significant main and/or synergy effect for a particular medium more pronounced. Third, we focused on established brands in mature CPG categories. More research is needed whether stronger effects for billboard and/or cinema advertising are found for newer brands, or with non-CPG categories.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ijresmar.2014.05.004.

References

AdAge (2012). Advertising to sales ratios by industry. Posted June 28, 2012. Available at. http://adage.com/article/datacenter-advertising-spending/advertising-sales-ratios-industry/106575/

Allenby, Greg, & Hanssens, Dominique M. (2005). Advertising response. *Special Report, (05-200)* (pp. 8). Marketing Science Institute.

Barnard, Jonathan (2012). ZenithOptimedia forecasts 4.1% growth in global adspend in 2013. Posted December 3, 2012. Available at. http://www.zenithoptimedia.com/zenith/zenithoptimedia-forecasts-4-1-growth-in-global-adspend-in-2013/

Bass, Frank M., & Clarke, Darral G. (1972). Testing distributed lag models of advertising effect. *Journal of Marketing Research*, 9(3), 298–308.

Berkowitz, David, Allaway, Arthur, & D'Souza, Giles (2001). The impact of differential lag effects on the allocation of advertising budgets across media. *Journal of Advertising Research*, 41(2), 27–38.

Bhargava, Mukesh, & Donthu, Naveen (1999). Sales response to outdoor advertising. *Canadian Journal of Marketing Research*, 12(1), 87–97.

Bijmolt, Tammo H. A., Van Heerde, Harald J., & Pieters, Rik G. M. (2005). New empirical generalizations on the determinants of price elasticity. *Journal of Marketing Research*, 42(2), 141–156.

Danaher, Peter J., & Dagger, Tracey S. (2013). Comparing the relative effectiveness of advertising channels: A case study of a multimedia blitz campaign. *Journal of Marketing Research*, 50(4), 517–534.

Dekimpe, Marnik G., & Hanssens, Dominique M. (1995). Empirical generalizations about market evolution and stationarity. *Marketing Science*, 14(3), G109–G121.

Dekimpe, Marnik G., & Hanssens, Dominique M. (2011). *The hidden powers of advertising investments. Liber Amicorum of Peter Leeflang, 209–223.*

Deleersnyder, Barbara, Dekimpe, Marnik G., Sarvary, Miklos, & Parker, Philip M. (2004). Weathering tight economic times: The sales evolution of consumer durables over the business cycle. *Quantitative Marketing and Economics*, 2(4), 347–383.

Deleersnyder, Barbara, Dekimpe, Marnik G., Steenkamp, Jan-Benedict E. M., & Leeflang, Peter S. H. (2009). The role of national culture in advertising's sensitivity to business cycles: An investigation across continents. *Journal of Marketing Research*, 46(5), 623–636.

Edell, Julie A., & Keller, Kevin L. (1989). The information processing of coordinated media campaigns. *Journal of Marketing Research*, 26(2), 149–163.

Ewing, Michael T., du Plessis, Erik, & Foster, Charles (2001). Cinema advertising reconsidered. *Journal of Advertising Research*, 41(1), 78–85.

Fagley, Nancy S. (1985). Applied statistical power analysis and the interpretation of nonsignificant results by research consumers. *Journal of Consulting Psychology*, 32(2), 391–396.

Field, Chris A., & Welsh, Alan H. (2007). Bootstrapping clustered data. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, 69(3), 369–390.

Fisher, Marc, Albers, Sönke, Wagner, Niks, & Frie, Monica (2011). Dynamic marketing budget allocation across countries, products, and marketing activities. *Marketing Science*, 30(4), 568–585.

Fok, Dennis, Horváth, Csilla, Paap, Richard, & Franses, Philip H. (2006). A hierarchical Bayes error-correction model to explain dynamic effects of price changes. *Journal of Marketing Research*, 43(3), 443–462.

Fok, Dennis, Paap, Richard, & Franses, Philip H. (2003). Modeling dynamic effects of the marketing mix on market shares. *ERIM Report Series Research in Management* (ERS-2003-044-MKT).

Greene, William H. (2000). *Econometric analysis.* New Jersey: Pearson Education Inc.

Hair, Joseph F., Jr., Anderson, Rolph E., Tatham, Ronald L., & Black, William C. (1995). *Multivariate data analysis.* New York: Macmillan.

---

[21] Strictly speaking, we were not able (in spite of many replications and robustness checks) to prove the null hypothesis of no sales impact wrong, which does *not* imply, however, that the null hypothesis has been proven true. Still, we follow common practice (see, for example, Leeflang & Wittink, 1996; Steenkamp et al., 2005) of calling a marketing instrument ineffective (or spoiled arms) if no significant sales elasticity is found.

[22] We would like to thank the Associate Editor for this observation.

[23] Also, in line with our earlier footnote 3, cross-sectional differences in average distribution/salesforce support across brands in a given category are already reflected in the brand-specific intercepts.

Hanssens, Dominique M., Parsons, Leonard J., & Schultz, Randall L. (2001). *Market response models: Econometric and time series analysis.* Massachusetts: Springer.

Horváth, Csilla, & Fok, Dennis (2013). Moderating factors of immediate, gross, and net cross-brand effects of price promotions. *Marketing Science, 32*(1), 127–152.

Hubbard, Raymond, & Armstrong, Scott (1992). Are null results becoming an endangered species in marketing? *Marketing Letters, 3*(2), 127–136.

Joshi, Amit, & Hanssens, Dominique M. (2010). The direct and indirect effects of advertising spending on firm value. *Journal of Marketing, 74*(1), 20–33.

Lamey, Lien, Barbara, Deleersnyder, Dekimpe, Marnik G., & Steenkamp, Jan-Benedict E. M. (2007). How business cycles contribute to private-label success: Evidence from the United States and Europe. *Journal of Marketing, 71*(1), 1–15.

Lamey, Lien, Deleersnyder, Barbara, Steenkamp, Jan-Benedict E. M., & Dekimpe, Marnik G. (2012). The effect of business-cycle fluctuations on private-label share: What has marketing conduct got to do with it? *Journal of Marketing, 76*(1), 1–19.

Leeflang, Peter S. H., & Wittink, Dick R. (1996). Competitive reaction versus consumer response: Do managers overreact? *International Journal of Research in Marketing, 13*(2), 103–119.

Leeflang, Peter S. H., Wittink, Dick R., Wedel, Michel, & Naert, Philippe A. (2000). *Building models for marketing decisions.* Dordrecht: Kluwer.

Levin, Andrew, Lin, Chien-Fu., & Chu, Chia-Shang J. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. *Journal of Econometrics, 108*(1), 1–24.

Lodish, Leonard M., Abraham, Magid, Kalmenson, Stuart, Livelsberger, Jeanne, Lubetkin, Beth, Richardson, Bruce, et al. (1995). How TV advertising works: A meta-analysis of 389 real world split cable TV advertising experiments. *Journal of Marketing Research, 32*(2), 125–139.

Morrison, Donald G. (1969). On the interpretation of discriminant analysis. *Journal of Marketing Research, 6*(2), 156–163.

Naert, Philippe A., & Leeflang, Peter S. H. (1978). *Building implementable marketing models.* New York City: Springer.

Naik, Prasad A. (2007). Integrated marketing communications: Provenance, practice and principles. In Gerard J. Tellis, & Tim Ambler (Eds.), *The SAGE Handbook of Advertising* (pp. 35–53). London: Sage Publications.

Naik, Prasad A., & Peters, Kay (2009). A hierarchical marketing communications model of online and offline media synergies. *Journal of Interactive Marketing, 23*(4), 288–299.

Naik, Prasad A., Peters, Kay, & Raman, Kaylan (2008). Heterogeneous response functions in advertising. *Arbeitspapiere des Lehrstuhls für Innovation.* Neue Medien und Marketing.

Naik, Prasad A., & Raman, Kaylan (2003). Understanding the impact of synergy in multi-media communications. *Journal of Marketing Research, 40*(4), 375–388.

Nijs, Vincent R., Srinivasan, Shuba, & Pauwels, Koen (2007). Retail-price drivers and retailer profits. *Marketing Science, 26*(4), 473–487.

Pauwels, Koen, Erguncu, Selin, & Yildirim, Gokhan (2013). Winning hearts, minds and sales: How marketing communication enters the purchase process in emerging and mature markets. *International Journal of Research in Marketing, 30*(1), 57–68.

Phillips, Peter C. B., & Perron, Pierre (1988). Testing for a unit root in time series regression. *Biometrika, 75*(2), 335–346.

Rosenthal, Robert (1991). *Meta-analytic procedures for social research.* Newbury Park: Sage Publications.

Sawyer, Alan G., & Peter, J. Paul (1983). The significance of statistical significance tests in marketing research. *Journal of Marketing Research, 20*(2), 122–133.

Saxonhouse, Gary R. (1976). Estimated parameters as dependent variables. *The American Economic Review, 66*(1), 178–183.

Sethuraman, Raj, Tellis, Gerard J., & Briesch, Richard A. (2011). How well does advertising work? Generalizations from meta-analysis of brand advertising elasticities. *Journal of Marketing Research, 48*(3), 457–471.

Sheehan, Kim B., & Doherty, Caitlin (2001). Re-weaving the web: Integrating print and online communications. *Journal of Interactive Marketing, 15*(2), 47–59.

Srinivasan, Shuba, Vanhuele, Marc, & Pauwels, Koen (2010). Mind-set metrics in market response models: An integrative approach. *Journal of Marketing Research, 47*(4), 672–684.

Steenkamp, Jan-Benedict E. M., & Geyskens, Inge (2014). Manufacturer and retailer strategies to impact store brand share: Global integration, local adaptation, and worldwide learning. *Marketing Science, 33*(1), 6–26.

Steenkamp, Jan-Benedict E. M., Nijs, Vincent R., Hanssens, Dominique M., & Dekimpe, Marnik G. (2005). Competitive reactions to advertising and promotion attacks. *Marketing Science, 24*(1), 35–54.

Stephen, Andrew T., & Galak, Jeff (2012). The effects of traditional and social earned media on sales: A study of a micro lending marketplace. *Journal of Marketing Research, 49*(5), 624–639.

Strube, Michael J. (1985). Combining and comparing significance levels from nonindependent hypothesis tests. *Psychological Bulletin, 97*(2), 334–341.

Van Heerde, Harald J., Gijsenberg, Maarten J., Dekimpe, Marnik G., & Steenkamp, Jan-Benedict E. M. (2013). Price and advertising effectiveness over the business cycle. *Journal of Marketing Research, 50*(2), 177–193.

Van Heerde, Harald J., Helsen, Kristiaan, & Dekimpe, Marnik G. (2007). The impact of a product-harm crisis on marketing effectiveness. *Marketing Science, 26*(2), 230–245.

Yeung, Matthew C. H., Ramasamy, Bala, Chen, Junsong, & Paliwoda, Stan (2013). Customer satisfaction and consumer expenditure in selected European countries. *International Journal of Research in Marketing, 30*(4), 406–416.

Full Length Article

The effects of a "no-haggle" channel on marketing strategies

CrossMark

Xiaohua Zeng [a,*], Srabana Dasgupta [b], Charles B. Weinberg [c]

[a] College of Business, City University of Hong Kong, Hong Kong, China
[b] Beedie School of Business, Simon Fraser University, Canada
[c] Sauder School of Business, University of British Columbia, Vancouver, Canada

ARTICLE INFO

ABSTRACT

As sellers increasingly turn to multi-channel retailing, the opportunity to implement different pricing policies has grown. With the advent of the internet, many traditionally bargained products such as automobiles, jewelry, watches, appliances and furniture are now being offered online at a fixed pre-determined price. We explore the strategy of simultaneously offering two pricing formats (fixed and bargained) via two different channels (on-line and brick and mortar) and find that in a market where there are two types of consumers—those with a high cost of haggling and others with a lower cost—a dual-pricing strategy is optimal only when there are enough high haggling-cost consumers, but not too many, and when the haggling costs between the two types of consumers are sufficiently different. We also find that it is optimal for the seller to specify a higher-than-cost minimum acceptable price as the price floor of bargaining. By doing so, the seller increases the bargained price by complementing the salesperson's bargaining ability, and also softens the internal competition between the two channels. Finally, we find that, surprisingly, the dual-pricing strategy may serve fewer customers while still being more profitable than a single price structure. The implications for consumer surplus are also explored.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

In many markets, bargaining is the norm. In the automobile market, consumers only infrequently pay the sticker price for a car. For products such as electronics, jewelry and furniture, while bargaining is not as overt as in the car market, consumers still expect to be able to haggle with salespeople, either directly on the sales price of the product or on service-related costs: chain retailers such as Best Buy and Sleep Country routinely bargain with in-store customers by offering them in-store discounts as well as additional services such as free delivery and extended warranties.

With the advent of the internet and the growing popularity of online buying, however, many manufacturers and retailers are now offering their products at fixed prices either through their own websites or third party sites, ostensibly addressing some consumers' dissatisfaction with bargaining and time spent visiting the physical store (Business Week, 2007). In the automobile market, third-party websites such as www.CarsDirect.com, www.Autobytel.com and the Canadian website www.unhaggle.com allow consumers to obtain price quotes (typically provided by several competing dealers) for the car of their choice. Consumers simply review the price and, if acceptable, the car is shipped to them directly. Best Buy and other large retailers continue to allow bargaining on the shop floor even though the prices on their websites

are fixed.[1] High end stores, such as Cartier and Zales for jewelry and Ethan Allen for furniture, have recently introduced online shopping that, like the online auto-buying websites, allow consumers to avoid haggling and visiting the physical store. In some cultures, such as in Asia, where haggling is traditional even for small-ticket items including clothing, food and home appliances, the growing use of the internet has led to many retailers launching their own web-stores or joining online aggregators such as Taobao (China's leader in e-commerce), where typically prices are fixed and cannot be bargained over.

Despite the growing opportunity for sellers to use multi-channel settings to simultaneously implement different pricing policies, there is significant variation across and within industries in the extent to which this strategy has been adopted, for which the extant literature does not provide a satisfactory explanation. There have been numerous studies examining a seller's choice between a fixed-price format and a bargaining format (e.g., Riley & Zeckhauser, 1983; Wang, 1995; Arnold & Lippman, 1998), all of which focus on a seller's choice of one pricing format over another and do not consider the possibility that the seller may want to offer both simultaneously. In all of these studies, in choosing a fixed, no-bargain price, a seller must weigh the cost of giving up the ability to discriminate through bargaining in favor of the higher prices it is able to charge consumers who can no longer haggle. In these studies, offering a fixed price is an equilibrium strategy under

---

[1] According to www.slybaldguys.com, "…managers (of Best Buy) have goals that their teams have to meet and managers that manage the slower times have a harder time of meeting these goals, thus they are more willing to negotiate."
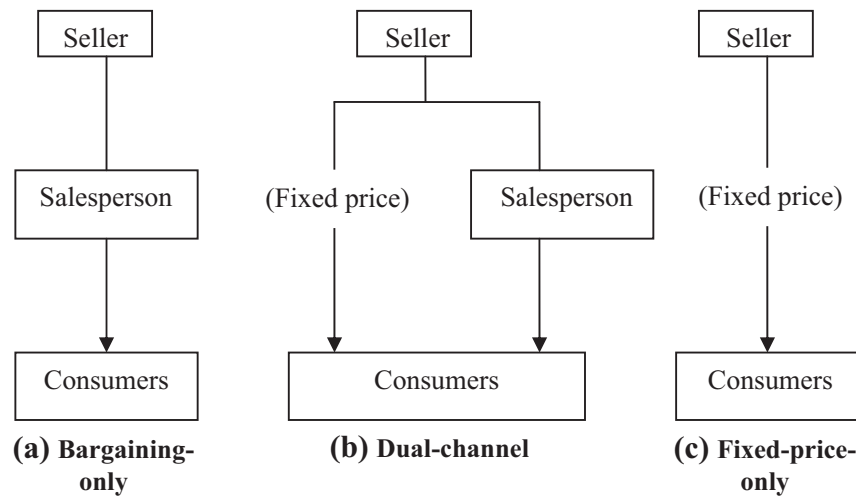
Fig. 1. Three channel structures.

such conditions as the seller being able to make a credible commitment to a fixed price strategy (Riley & Zeckhauser, 1983), or the buyers' bargaining abilities being, on average, sufficiently high (Arnold & Lippman, 1998), or the operating cost of implementing a bargaining strategy being too high (Wang, 1995). While these findings give us some insights into the benefits of fixed pricing over bargaining, this is different from a situation where consumers have the option to *choose* between the two different pricing formats. As a result, we do not have a clear understanding of why and when a strategy of simultaneously offering bargained and fixed prices is optimal.

Our objective is therefore to answer the following questions. When is it optimal for a seller to bargain, offer a fixed price, or to use a mix of the two via two different channels and given the optimal choice, what prices should the seller set in each channel? To answer these questions, we develop a model where we diverge from the existing literature to allow both pricing formats (bargained and fixed prices) to be offered simultaneously via a dual distribution system so that consumers can self-select into a channel that maximizes their utility. We model the interaction of three parties: (1) a seller that can sell via bargaining in a brick and mortar store, or at a fixed-price online, or both, (2) a salesperson who bargains over price in the physical, brick and mortar channel on behalf of the seller, and (3) the consumer who incurs a "haggling cost" if she decides to bargain. Thus, we consider three potential channel structures: the conventional bargaining channel that allows for face-to-face haggling with the consumer (Fig. 1(a)), a "dual channel" that offers consumers a choice between a fixed price and a bargained one (Fig. 1(b)), and a fixed-price-only channel (Fig. 1(c)).

One distinct feature of our model is that it allows the salesperson's commission to be based on the difference between the sales price and a seller determined "minimum acceptable price".[2] This is in contrast to the existing literature (Basu, Lal, Srinivasan, & Staelin, 1985; Misra, Coughlan, & Narasimhan, 2005) where the commission is based on the difference between the sales price and the marginal cost of the product. Thus, rather than imposing the constraint that the marginal cost of the product represent the lowest price the seller is willing to accept, we treat the bottom line of bargaining as a strategic variable that

may be equal to or higher than marginal cost. This flexibility that the seller now has in setting the lowest bargaining point for the salesperson serves two purposes: first, it raises the salesperson's threat point and allows the sales representative to commit to a higher price during bargaining (Cai & Cont, 2004; Gatehouse, 2007), and second, it controls the cost information on which the bargaining is based and helps the seller reach a more favorable bargaining outcome (Wilken, Cornelißen, Backhaus, & Schmitz, 2010).

Our model yields several interesting findings. First, a dual channel is optimal if there are (i) two types of consumers in the market – those with a high cost of haggling and those with a lower cost – and (ii) a high enough proportion of high haggling cost consumers whose cost of haggling is sufficiently different from the low haggling cost customers. Second, we find that it is optimal for the seller to specify a higher-than-cost minimum acceptable price above which it pays the salesperson a commission. While a higher price floor means that the salesperson fails to reach agreements with more buyers (e.g., those with valuations above marginal cost but below minimum acceptable price), the seller still finds it optimal to do so. The lower the salesperson's bargaining ability, the greater the seller's incentive to set a higher minimum acceptable price. The minimum acceptable price also serves to soften the internal competition between the two channels. Third, surprisingly, under certain conditions a dual-channel seller may serve fewer customers while still making a higher profit than under a single-channel structure, i.e., either a bargaining-only or fixed-price-only channel. This is because the minimum acceptable price set in a dual channel is higher, allowing the seller to charge a higher fixed price, which in turn helps the salesperson achieve a higher price in the bargaining channel. Finally, we find that no one channel structure is ideal for every customer: the bargaining-only channel generates the greatest surplus for low haggling-cost and low-valuation consumers, while the fixed-price-only channel generates the greatest surplus for high haggling-cost consumers. Overall, the fixed-price-only channel generates the highest surplus, the bargaining-only channel the lowest, while the dual channel stands between the two.

The contribution of our study lies in two domains. First, we contribute to the channels literature by identifying the conditions under which we would observe a dual channel structure in a market where bargaining is the norm. This is distinct from the existing dual-distribution literature where the addition of a channel does not involve implementing a different pricing format from the original channel (e.g., Moriarty & Moran, 1990; Chiang, Chhajed, & Hess, 2003; Kumar & Ruan, 2006).

A second contribution of our research is to the pricing and bargaining literature, where we explore a means by which the seller

---

[2] Empirical evidence of a minimum acceptable price can be found in the automobile industry, where the invoice price for vehicles effectively plays the role of a minimum acceptable price because consumers typically observe a vehicle's invoice price through websites such as www.edmunds.com, www.autobytel.com and the Kelly Blue Book, and they view this as the dealer's true cost. In reality, the invoice price is different from the actual cost the dealer incurs. This is because of hidden incentives offered to the dealer by the manufacturer that are not reflected in the invoice price. This includes incentives such as special allowances, dealer cash, dealer holdback and discounts, all of which the dealer may choose not to pass on to the consumer (Besanko et al., 2005).

can limit the freedom of the salesperson in order to influence the outcome of the bargaining process and consequently soften competition between the two channels. In this, our model is similar to that of Thanassoulis and Gill (2010) who explore how limiting the sales person's ability to offer discounts to matching a rival's posted price can soften competition between sellers leading to higher prices. Instead of price matching, however, our model uses a different mechanism that is internal to the seller, namely the imposition of a price floor, to manipulate the outcome of the bargaining between the consumer and the salesperson.

Finally, it is important to note that the role that the salesperson plays in our model is that of a price delegate in that he simply performs the bargaining on behalf of the seller. It is necessary to have the salesperson in our model for the following reason: the salesperson acts as a commitment device — by appropriately designing the delegation contract (i.e., specifying the minimum acceptable price), the seller commits his delegate (the salesperson) to tougher negotiations with the buyer. Thus, our study is also related to the price delegation literature, which examines delegation contracts ranging from decisions to give full pricing authority, limited pricing authority (i.e., pricing latitude was limited to pre-specified ranges), or no pricing authority (i.e., salesperson is not allowed to deviate from list prices). While most of the studies are concerned with designing delegation contracts that influence the amount of effort the delegate or sales person makes (Lal, 1986; Joseph, 2001), our focus is on understanding the role that the minimum acceptable price plays in influencing the final bargaining outcome in a dual channel context. In order to do this and to make the analysis tractable we make two simplifying assumptions. First, we assume that the salesperson is risk neutral and second, we assume that the seller has complete information about the salesperson's capabilities and effort level so that moral hazard and adverse selection problems are abstracted away from. These assumptions allow us to focus on the issue of alternative pricing strategies and the role of the new instrument, namely the minimum acceptable price, in a multi-channel context. We discuss the limitations of making these assumptions in Section 4.

The rest of the paper is organized as follows. In Section 2, we develop a model of dual channels for a monopolist seller. This is followed by analyses of the model in Section 3. The overall conclusion of our study and directions for future research are outlined in Section 4.

2. A model of dual channels

In this section, we discuss the basic setup of a dual-channel model as shown in Fig. 1(b). Due to space limitations, we omit the discussion of the bargaining-only and fixed-price-only channels, as they can be easily derived from the dual-channel model. For clarity, Table 1 provides a list of the notations that we use in the model.

2.1. Consumers

Let V be consumers' valuation of a product (e.g., electronics, car, furniture), which we assume follows a uniform distribution, $U(0,1)$. Consumers can either buy the product at a fixed price online, p^f, or they can bargain with the salesperson (in the physical store) to reach a price, p^b, and by doing so, incur a *haggling cost*, hc, that is incurred from the time and effort spent in (and the inherent aversion to) negotiation, as well as the time spent visiting the physical store. The notion of a haggling cost has been reported in the press (Business Week, 2007) and considered in several theoretical studies of bargaining (Desai & Purohit, 2004; Terwiesch, Savin, & Hann, 2005; Gill & Thanassoulis, 2009). Surveys of automobile buyers and buyers of other consumer products indicate that as many as two thirds of them have a strong aversion to the negotiation process, not only because of the time and effort required, but also because of a fear of being taken advantage of by the salesperson. For simplicity, we assume that β proportion of consumers

Table 1
Model notation.

Notation	Explanation
p^f	No-haggle fixed price
p^b	Bargained price
M	Minimum acceptable price
B	Salesperson's commission rate, $0 \le B \le 1$
C	The seller's true cost of the vehicle
q^f	Unit sales in the fixed-price channel
U_S	Salesperson's outside option, $U_S \ge 0$
α	Salesperson's bargaining power, $\alpha \in (0, 1)$
V	Consumer's valuation of the product, $V \sim U(0, 1)$
hc	Consumer's haggling cost, $hc \in \{hc_l, hc_h\}$, where $hc_l < hc_h < 1 - C$
β	The proportion of high haggling-cost consumers, where $0 \le \beta \le 1$
U_C	Consumer's outside option, $U_C \ge 0$

have a high cost of haggling, hc_h, while the remaining $1 - \beta$ consumers incur a lower cost, hc_l.

We assume that consumers make decisions using the following steps: first, they obtain the fixed price, p^f, from the seller's website or a third-party website. Then, based on p^f and other information, the consumer estimates the price she expects to pay if she bargains with the salesperson. If buying the product through either channel does not generate sufficiently high utility, the consumer opts for an outside option, $U_C \ge 0$. The outside option can be thought of as the consumer's status quo, i.e., not buying a new product. Once the consumer has determined p^f and p^b, she computes her utility from the three options and chooses the one that maximizes her utility. We define the consumer's utility as:

$$U = \begin{cases} V - p^f & \text{if the product is bought at the no-haggle fixed price} \\ V - p^b - hc & \text{if the product is bought at the bargaining price; } hc \in \{hc_l, hc_h\} \cdot \\ U_C & \text{if the outside option is exercised} \end{cases}$$

(1)

Note that the presence of the outside option, U_C, allows us to focus on a single brand while still ensuring reasonable pricing behavior by the seller. Because our research focuses on the vertical relationship among the seller, salesperson and consumer, this simplification seems reasonable.

2.2. The salesperson

Consumers who buy from the bargaining channel bargain over the price with a salesperson employed by the seller. As discussed earlier, the salesperson's role is that of a delegate who can credibly commit to a mutually agreed price. Since our objective is to understand how the seller can best design a contract that can influence the final bargaining price, we abstract away from issues such as moral hazard and adverse selection by assuming the salesperson to be risk neutral and that the seller has full information. This assumption also maintains analytical tractability in our multichannel pricing strategy setting. The salesperson receives a commission from the seller (Srinivasan, 1981), which we assume follows a linear form[3]:

$$\pi^S = B \left[\beta \int_{V \in F_{V1}} \left(p^b(V, hc_h) - M \right) dV + (1 - \beta) \int_{V \in F_{V2}} \left(p^b(V, hc_l) - M \right) dV \right].$$

(2)

The commission rate is denoted by B, and the minimum acceptable price is denoted by M. We allow the seller to specify an $M \ge C$ instead of restricting $M = C$. We also assume that consumers have information

---

[3] We restrict our attention to a commission-only plan. A more general compensation plan contains a fixed salary and a commission (e.g., Basu et al., 1985). Interestingly, compensation schemes in the automobile industry tend to be heavily dependent on commissions, the salary component being very small (Winter, 2004).

about M but not about C. As Eq. (5) below shows, the bargaining price, p^b, turns out to be a function of both consumers' valuation and haggling cost. F_{V1} (F_{V2}) are the set of valuations of high haggling-cost consumers (low haggling-cost consumers) who buy from the bargaining channel. As Appendix I shows, $F_{V1} = \varnothing$ and $F_{V2} = [hc_l + M + U_C, 1]$, that is, only some low haggling-cost consumers buy from the bargaining channel under a dual-channel and none of the high haggling-cost consumers do so. β is the proportion of high haggling-cost consumers. Let U_S be the salesperson's reservation utility that he can obtain from alternative employment. The salesperson receives a commission if and only if there is a sale. Thus, for the salesperson to work, we require that $\pi^S \geq U_S$.

2.3. Determination of the bargaining price

Next we outline the way in which the bargaining price, p^b, is determined. We follow the Nash axiomatic approach (Nash, 1950; Roth, 1979), a bargaining mechanism by which each party receives its reservation utility while any remaining surplus is split depending on the relative bargaining power of the two parties. Thus, the party with either the higher bargaining power or a more appealing outside option, i.e., reservation utility, is able to extract a larger proportion of the total surplus. Besides being widely used by previous researchers (Dukes & Gal-Or, 2003; Desai & Purohit, 2004; Guo & Iyer, 2013) and being a general and intuitive method by which to capture the bargaining process, the Nash axiomatic approach also allows us to incorporate competition between the two channels in a straightforward manner, something that could become quite intractable with an alternative approach (such as Rubenstein's (1982) sequential bargaining model).

Let α be the salesperson's bargaining power relative to the consumer, where $\alpha \in (0, 1)$. The Nash solution to the bargaining process maximizes the following expression:

$$Max \left[V - p^b - hc - D_C\right]^{1-\alpha} \times \left[B\left(p^b - M\right) - D_S\right]^{\alpha}. \tag{3}$$

$V - p^b - hc$ and $B(p^b - M)$ represent the consumer's and salesperson's respective gains from the transaction. D_C and D_S are the utilities for the consumer and the salesperson, respectively, from their best alternative in case of a disagreement.

The bargainers will reach an agreement if and only if it makes both parties better off, i.e., $V - p^b - hc \geq D_C$ and $B(p^b - M) \geq D_S$. For the salesperson, we assume that $D_S = 0$, as he makes no money if no agreement is reached in the bargaining. Note that M determines the disagreement point of bargaining because once the price is below M, bargaining breaks down. The consumer, however, can either buy from the fixed-price channel at p^f or resort to her outside option, U_C, depending on which alternative generates a higher surplus, i.e.,

$$D_C = \max\left\{V - p^f, U_C\right\}. \tag{4}$$

Intuitively, if negotiations are unsuccessful, Eq. (4) implies that consumers with a high valuation would prefer to buy from the fixed-price channel, while low-valuation consumers would choose their outside option, U_C.

We assume that consumer characteristics $\{V, hc\}$ are known to the salesperson. Then, depending on the consumer's best alternative, the bargaining price takes the following form:

$$p^b = \begin{cases} \alpha\left(p^f - hc\right) + (1-\alpha)M & \text{if } p^f + U_C \leq V \leq 1 \\ \alpha(V - hc - U_C) + (1-\alpha)M & \text{if } hc + U_C + M \leq V < p^f + U_C \end{cases}. \tag{5}$$

2.4. The seller

The objective of a dual-channel seller is to maximize the joint profit from the two channels by setting (i) the no-haggle fixed price, p^f, and

(ii) the salesperson's commission rate B and minimum acceptable price M, which determine the price bargained in the bargaining channel, p^b. Formally, the seller's optimal decision will be a solution to the following problem:

$$\underset{(p^f, B, M)}{Max} \; \pi^D = \left(p^f - C\right)q^f \tag{6}$$
$$+ \left[\beta \int_{V \in F_{V1}} \left(p^b(V, hc_h) - C\right)dV + (1-\beta) \int_{V \in F_{V2}} \left(p^b(V, hc_l) - C\right)dV\right]$$
$$- \pi^S \quad \text{s.t. } \pi^S \geq U_S$$

where C is the seller's marginal cost, q^f is the number of units sold in the fixed-price channel (as a function of p^f and M), and the constraint represents the fact that the seller must guarantee the salesperson his minimum payoff, U_S.

3. Analyses and results

We begin the analysis by solving for the conditions under which a dual channel has positive sales in both channels. If both channels are available but one of them has zero sales, we consider it to be a single-channel structure. We then derive the optimal minimum acceptable price for the salesperson and follow this with a series of results related to the seller's pricing strategy, including a comparison of bargaining and fixed prices, both within and across channel structures. Finally, we present results for demand, channel profitability and consumer surplus.

For clarity, we identify the decisions under different channel structures by adding a corresponding subscript, such as M_{dual} and $M_{bargaining - only}$. Without loss of generality, we set consumers' outside option, U_C, and the low haggling cost, hc_l, to be zero. These assumptions simply scale the solutions but do not change our conclusions. We also confine our analysis to the case where $hc_h < 1 - C$ because, if $hc_h \geq 1 - C$, none of the high haggling-cost consumers will buy the product in the bargaining channel even at its cost, C, which is an uninteresting case. The complete analytical solutions for each channel structure are provided in Appendix II.

3.1. Case where dual channel has positive sales in both channels

As shown in Appendix I, for both the fixed-price channel and bargaining channel to have positive sales, two conditions must be satisfied: (a) two types of consumers exist, i.e., $0 < \beta < 1$, and (b) the difference in haggling costs needs to be sufficiently high, $hc_l \leq p^f_{dual} - M_{dual} \leq hc_h$.

Table 2 illustrates the market segments for each channel, conditional on a given fixed price and minimum acceptable price.

3.2. Optimal minimum acceptable price for the salesperson

Eq. (2) describes the salesperson's problem, which is a function of a commission rate, B, and the minimum acceptable price, M. Since the role of B has been studied extensively in the salesforce literature (e.g., Basu et al., 1985; Lal & Srinivasan, 1993; Chen, 2005), our major emphasis is on understanding how the seller sets M. Unlike the previous literature which assumes that the salesperson's commission pay is contracted on the seller's true marginal cost, C, in our model (consistent with certain features of the auto market) M does not necessarily have to be equal to C, and therefore serves as an additional instrument for the seller over B.

The optimal minimum acceptable price for the salesperson is described in the following proposition (see proof in Appendix III):

Table 2
Market segmentation for the three channel structures.

Consumer characteristics $\{V, hc\}$		Choice
Dual channel ($hc_l + M \leq p^f \leq hc_h + M$)		
$hc = hc_h$	$V \in [p^f + U_C, 1]$	Buy in the fixed-price channel
	$V \in [0, p^f + U_C)$	Outside option
$hc = hc_l$	$V \in [hc_l + M + U_C, 1]$	Buy in the bargaining channel
	$V \in [0, hc_l + M + U_C)$	Outside option
If $p^f > hc_h + M$, the channel structure becomes a bargaining-only channel		
$hc = hc_h$	$V \in [hc_h + M + U_C, 1]$	Buy in the bargaining channel
	$V \in [0, hc_h + M + U_C)$	Outside option
$hc = hc_l$	$V \in [hc_l + M + U_C, 1]$	Buy in the bargaining channel
	$V \in [0, hc_l + M + U_C)$	Outside option
If $p^f < hc_l + M$, the channel structure becomes a fixed-price-only channel		
$hc \in \{hc_l, hc_h\}$	$V \in [p^f + U_C, 1]$	Buy in the fixed-price channel
	$V \in [0, p^f + U_C)$	Outside option

Proposition 1(a). *Under both the bargaining-only channel and the dual channel, it is optimal for the seller to specify a minimum acceptable price that is greater than the marginal cost, i.e., $M > C$.*

The intuition behind this is as follows: The minimum acceptable price, M, serves as the salesperson's threat point or the price floor for bargaining, so that a higher M, according to Eq. (5), achieves a higher bargained price, p^b. In this way, a higher M complements the salesperson's bargaining skill by resulting in a higher bargaining price. This can be seen from the expressions $\frac{\partial \left(M_{bargaining-only} - C \right)}{\partial \alpha}$ < 0 and $\frac{\partial (M_{dual} - C)}{\partial \alpha}$ < 0. Thus, the lower the salesperson's bargaining power, the greater the seller's incentive to set a higher M to achieve a higher bargaining price.

In a dual channel, the minimum acceptable price has a further strategic role. To see this we first state the following proposition:

Proposition 1(b). *The minimum acceptable price in a dual channel is higher than that in a bargaining-only channel.*

When the seller offers two channels, there is internal competition between them, as all consumers will be aware of the price in the fixed price channel. To accommodate the no-haggle fixed-price channel, the seller needs to prevent the price in the bargaining channel from being too low and does so by specifying an M that is higher than that under a bargaining-only channel. A higher M effectively increases the bargaining price, so as to lessen the price pressure on the fixed-price channel. In other words, M serves to soften the internal competition between the two channels.

3.3. Bargaining and fixed prices

In this subsection, we derive the optimal fixed and bargaining prices in the dual channel and then compare them to the prices set in the bargaining-only and fixed-price-only channel structures.

We begin by asking whether, in a dual channel, the fixed price is higher or lower than the bargaining prices. We answer this with the following proposition:

Proposition 2. *Under a dual-channel strategy, the no-haggle fixed price is higher than the price bargained in the bargaining channel.*

The intuition for this is fairly straightforward and can be derived directly from the consumer's utility function in Eq. (1): Since consumers incur a haggling cost when they bargain, for them to buy in the bargaining channel instead of the fixed-price channel, the bargained price must be lower than the fixed price. In other words, since the fixed price saves consumers' time and effort, they are willing to pay a premium for this.

Next, we compare the dispersion of bargained prices under the dual channel to that in the bargaining-only channel. We measure dispersion as being the difference between the lowest and highest bargained prices. The dispersion therefore reflects the extent to which sellers are price-discriminating among consumers. We present the following proposition (see Appendix IV for proof):

Proposition 3. *The dispersion of bargaining prices under a dual channel is lower than that in a bargaining-only channel.*

There are two reasons for this. First, the lowest bargaining price in a dual channel is higher than that in a bargaining-only channel, as the dual-channel seller sets a higher minimum acceptable price (Proposition 1(b)). Second, the highest bargaining price in a dual channel is lower than that in a bargaining-only channel because in a bargaining-only channel, consumers with different valuations are charged different prices, while in the dual channel consumers whose valuation exceeds the no-haggle fixed price can cite that price when bargaining with the salesperson to obtain a lower price. As a result, they pay less than they would have had the fixed-price alternative not existed. Empirical evidence supporting this has been found in the auto industry by Zettelmeyer, Scott-Morton, and Silva-risso (2006), who show that buyers who used the fixed-price option (e.g., price quote from internet referral services) tended to pay lower prices for their cars than those who did not use it.

Finally, we compare the no-haggle fixed price in a dual channel with that in a fixed-price-only channel. We present the following proposition:

Proposition 4. *The fixed price in the dual channel is higher than that in the fixed-price-only channel.*

The intuition for this is as follows: As low haggling-cost consumers do not benefit as much from the fixed price as do high haggling-cost consumers, the fixed-price-only channel charges a price that is low enough to attract both types of consumers. In contrast, the dual channel serves the segments via different channels and can price more efficiently. That is, as only the high haggling-cost consumers will buy at the fixed price, the dual-channel seller is able to charge a higher price than the fixed-price-only seller.

Furthermore, the dual-channel seller has an additional incentive to set a higher fixed price because the higher fixed price also increases bargained prices in the bargaining channel. To see why, consider Eq. (5): the bargained price, p^b, is non-decreasing in p^f. This is because as p^f increases, the outside opportunity for certain consumers becomes less attractive, making them more dependent on the bargained outcome. The salesperson can take advantage of them and achieve a higher price.

3.4. Demand

Next we compare the optimal demand levels across all three channel structures. We put forward the following proposition (see Appendix VI for proof):

Proposition 5. *When $M > C$ and when each channel operates optimally, the dual channel does not necessarily generate higher demand than a single-channel structure.*[4]

Specifically, when the high haggling cost is below a certain level, i.e., $hc_h < \frac{\alpha(1-C)}{\alpha(1-\beta)+2\beta}$, and M is a decision variable so that the seller will choose $M > C$ at the optimum, then the demand in the dual channel is lower than in the bargaining-only channel.

To illustrate this proposition more clearly, we present the following numerical examples. In Table 3(a) and (b), we compute the demand levels in each of the three channel structures while we vary β and hc_h.

---

[4] In some cases, despite the lower demand, the dual channel will earn a higher profit than the single channel. We discuss this in Subsection 3.5.

Table 3
Demand comparison.

(a) β varies, $hc_h = 0.065$					
β		**0.15**	**0.35**	**0.55**	**0.75**
Dual	Total demand	0.114	0.111	0.106	0.102
	Demand from hc_l consumers	0.105	0.087	0.060	0.033
	Demand from hc_h consumers	0.009	0.024	0.046	0.069
	M	$17,529	$17,333	$17,333	$17,333
	p^f	$18,829	$18,634	$18,340	$18,154
	π_{dual}	$234	$219	$209	$203
Bargaining-only	Total demand	0.127	0.118	0.110	0.100
	Demand from hc_l consumers	0.116	0.092	0.065	0.037
	Demand from hc_h consumers	0.011	0.026	0.045	0.063
	M	$17,268	$17,182	$17,095	$17,008
	$\pi_{bargaining-only}$	$242	$212	$183	$154
Fixed-price-only	Total demand	0.100	0.100	0.100	0.100
	p^f	$18,000	$18,000	$18,000	$18,000
	$\pi_{fixed-price-only}$	$200	$200	$200	$200
(b) hc_h varies, $\beta = 0.75$					
hc_h		**0.02**	**0.035**	**0.05**	**0.065**
Dual	Total demand	0.075	0.102	0.103	0.103
	Demand from hc_l consumers	0.031	0.032	0.033	0.033
	Demand from hc_h consumers	0.044	0.070	0.070	0.070
	M	$17,675	$17,431	$17,333	$17,334
	p^f	$18,075	$18,131	$18,154	$18,154
	π_{dual}	$202	$203	$203	$203
Bargaining-only	Total demand	0.124	0.116	0.108	0.101
	Demand from hc_l consumers	0.035	0.036	0.036	0.037
	Demand from hc_h consumers	0.089	0.080	0.072	0.064
	M	$17,233	$17,158	$17,083	$17,008
	$\pi_{bargaining-only}$	$227	$200	$176	$154
Fixed-price-only	Total demand	0.100	0.100	0.100	0.100
	p^f	$18,000	$18,000	$18,000	$18,000
	$\pi_{fixed-price-only}$	$200	$200	$200	$200

For other variables in this and all subsequent numerical analyses for profitability and consumer surplus, we set $\alpha = 0.5$ (i.e., the consumer and salesperson have equal bargaining power), $C = 0.8$ and $U_S = 0.0001$. To make the numerical analysis easy to understand, we multiply all monetary amounts by $20,000.

The dual channel versus the bargaining-only channel ($M > C$). As shown in Table 3(a), when $\beta = 0.55$, the total demand in the bargaining-only channel (0.110) is greater than that in the dual channel (0.106). Although generating higher demand is one rationale for why a seller may sell its products through multiple channels (Geyskens, Gielens, & Dekimpe, 2002), for a certain mix of the two types of consumers ($\beta \leq 0.55$ in our numerical example), demand in the dual channel is actually lower than that in the bargaining-only channel. This can be explained if we understand how demand from the two types of consumers is generated. First, the dual-channel seller will always sell to fewer low haggling-cost consumers than the bargaining-only seller. This is because, according to Proposition 1(b), a dual-channel seller will specify a higher M than will a bargaining-only channel seller. Second, the dual-channel seller may or may not sell to more high haggling-cost consumers than its bargaining-only counterpart, depending on the proportion of high haggling-cost consumers, β. When β is low (i.e., the proportion of low haggling-cost consumers is high), the dual-channel seller will have an incentive to charge a higher fixed price, even if it entails sacrificing some demand, as it allows the seller to charge a higher price in the bargaining channel.

Another factor that contributes to the differences in demand between the bargaining-only and dual channel is the high haggling cost, hc_h. As shown in Table 3(b), when $hc_h \leq 0.05$, the total demand in the dual channel is lower than that in the bargaining-only channel. This is because as hc_h decreases, the bargaining-only channel seller has a greater incentive to serve high haggling-cost customers, as it requires less to compensate them for their haggling costs.

However, if hc_h or β is not too low—i.e., as shown in Table 3(a), when $hc_h = 0.065$ and $\beta = 0.35$ or 0.55, or when, as shown in Table 3(b), $hc_h = 0.035$ or 0.05 and $\beta = 0.75$—then despite the

lower demand, the profit in the dual channel is higher than that in the bargaining-only channel. This is because the dual-channel seller is able to raise the fixed price, which in turn increases prices in the bargaining channel, thus compensating for the lower demand.

The dual channel versus the bargaining-only channel ($M = C$). Allowing the seller to set $M > C$ is important for deriving Proposition 5. When M is fixed, i.e., $M = C$, then the total demand in the dual channel is never lower than demand in the bargaining-only channel (see proof in Appendix VI).

The dual channel versus the fixed-price-only channel. We start by noting that demand in the fixed-price-only channel is independent of the values of either β or hc_h. This is because consumers do not haggle, so that haggling costs do not influence the seller's decision and consumers' choices. According to Table 3(a) and (b), demand in the dual channel is higher than that in the fixed-price-only channel in most cases, except for a very low value of high haggling costs, $hc_h = 0.02$.

3.5. Profitability

In this subsection, we compare profits across the three different channel structures in order to determine the optimal conditions for a particular channel structure. We find that, as in the case of demand, haggling costs are critical in determining the relative profitability of each channel.

The expressions for profits from each channel structure are derived in Appendix II. However, due to the complexity of these expressions, we are unable to directly compare profits in closed form. Instead, we do so using a numerical procedure. Specifically, we vary the values of β, hc_h and hc_l in order to capture the mix of high and low haggling-cost consumers and the differences in their haggling costs. For the other parameters, we kept the same set of values as in the demand analyses (i.e., $\alpha = 0.5$, $C = 0.8$ and $U_S = 0.0001$; changing these values will shift the profit numbers but the comparison across the three channel structures follows the same pattern.)
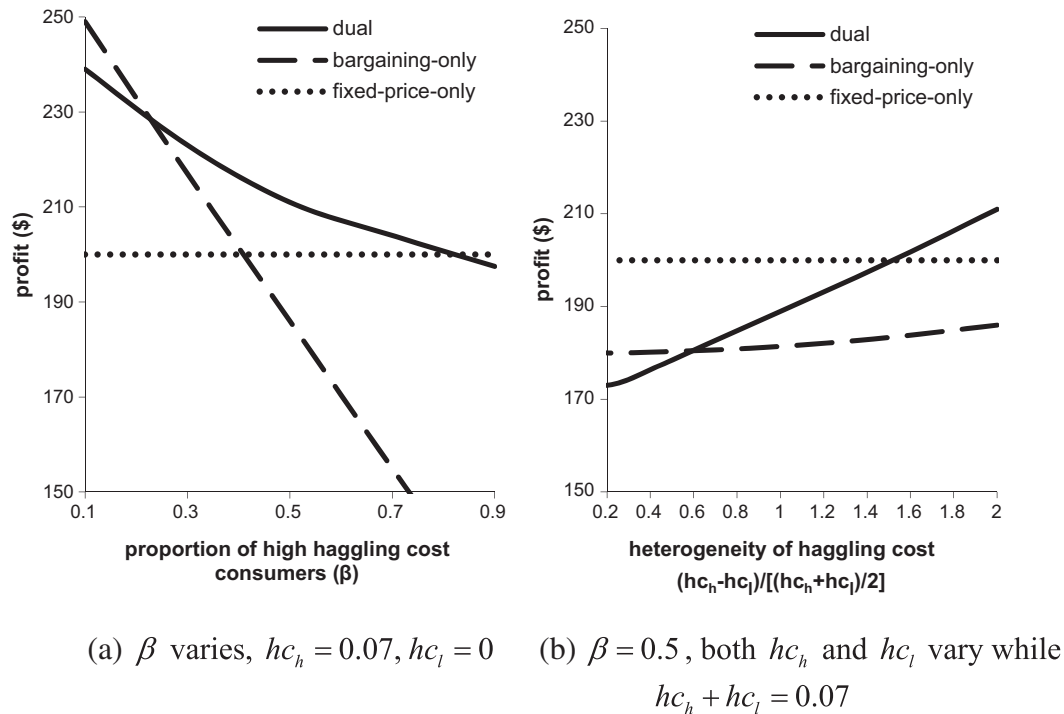
(a) β varies, $hc_h = 0.07$, $hc_l = 0$

(b) $\beta = 0.5$, both hc_h and hc_l vary while $hc_h + hc_l = 0.07$

Fig. 2. Profit comparison across three channels.

We start by varying the proportion of high haggling-cost consumers, β, the effect of which is shown in Fig. 2(a). We follow this with an analysis of the impact of heterogeneity in haggling costs, $hc_h - hc_l$, which is shown in Fig. 2(b). It is important to note that we vary $hc_h - hc_l$, while maintaining the same average haggling cost, i.e., $(hc_h + hc_l)/2$, throughout. This allows us to capture variation in heterogeneity alone and prevents it from being confounded with any changes in the average haggling cost in the market. To facilitate the interpretation, we represent the heterogeneity as a fraction of the average haggling cost, i.e., $(hc_h - hc_l)/[(hc_h + hc_l)/2]$.

The findings from our numerical analysis lead to the following proposition:

Proposition 6. *The dual channel is the most profitable structure when (a) there are enough high haggling-cost consumers but not too many, and (b) when the difference in haggling costs between the two types of consumers is sufficiently high.*

According to Proposition 6, two factors are critical in determining the profitability of the dual channel. First, for the dual channel to be optimal, a certain mix of the two types of consumers is required. For example, if all or nearly all the consumers are high haggling-cost consumers, it is best for the seller to serve consumers only with a fixed price, as the benefits from price discrimination through bargaining are outweighed by the costs of compensating consumers for their haggling costs. In the context of our numerical example, this need for some but not too many high haggling-cost consumers implies that $0.22 \leq \beta \leq 0.87$.

Second, for the dual-channel to be optimal it is also required that there be sufficient heterogeneity in haggling costs. Our numerical analysis shows that when the heterogeneity in haggling costs is $\geq 143\%$, a dual-channel strategy is optimal. Alternatively, when there is insufficient heterogeneity (i.e., $<143\%$), the dual channel is never optimal. In other words, the fact that consumers are differentiated ($\beta > 0$) does not ensure the optimality of using dual channels because, first, although this dual-channel structure allows the seller to discriminate consumers based on their haggling costs, the fixed-price channel eliminates the seller's ability to discriminate based on consumer valuations, something

that is possible in a bargaining-only channel structure. Second, the presence of the fixed-price channel impacts the price bargained in the bargaining channel because it serves as an outside option, resulting in a lower price to some high-valuation consumers than would have been offered if the no-haggle option were absent. These two disadvantages of the fixed-price channel can be mitigated, however, if the fixed price is sufficiently high, which is possible only if some consumers have sufficiently higher haggling costs than others.
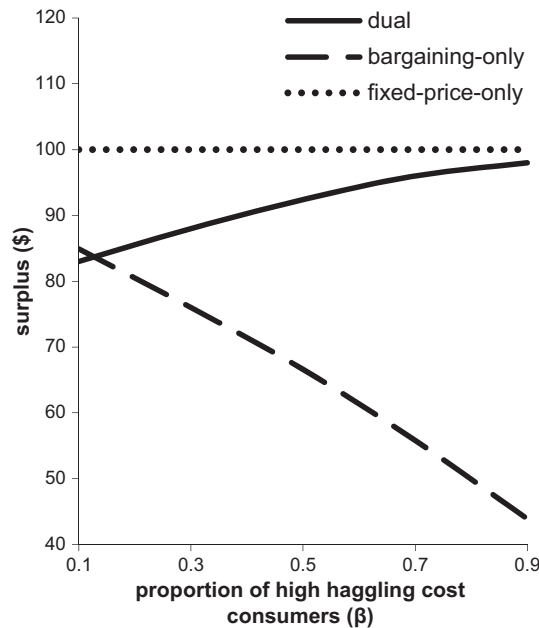
3.6. Consumer surplus

We conclude this section with an examination of consumer surplus under the three channel structures, which we also analyze using a numerical approach. We first analyze the surplus at individual consumer level because each consumer's surplus depends on her valuation and haggling cost. Fig. 3(a) and (b) demonstrates how the surplus is distributed among different consumers.

We find that no single channel structure generates the highest level of consumer surplus for all consumers. Specifically, for low haggling-cost consumers with relatively low valuations (ranging from $17,100 to $18,580 in Fig. 3(a)), the bargaining-only channel generates the highest surplus. This is because in the dual channel, due to a higher minimum acceptable price (Proposition 1(b)), bargained prices for these consumers are higher. For low haggling-cost consumers with high valuations (>$18,580), the dual channel offers the highest consumer surplus, as the presence of the fixed price allows these consumers to bargain a better price. For high haggling-cost consumers, the fixed-price-only channel generates the highest surplus (Fig. 3(b)), as its use of a fixed price format allows these consumers to skip the costly bargaining process. While this is also true for the dual channel, the price paid by these consumers is higher due to the seller's need to soften the internal competition between the two channels (Proposition 4).

Fig. 3(c) summarizes the overall consumer surplus from the three channels. The fixed-price-only channel generates the highest surplus; the bargaining-only channel, while most commonly used in practice, generates the lowest surplus in most cases; and the dual channel stands

(a) Low Haggling-cost Consumers

(b) High Haggling-cost Consumers

$(\beta = 0.5, hc_h = 0.07, hc_l = 0)$
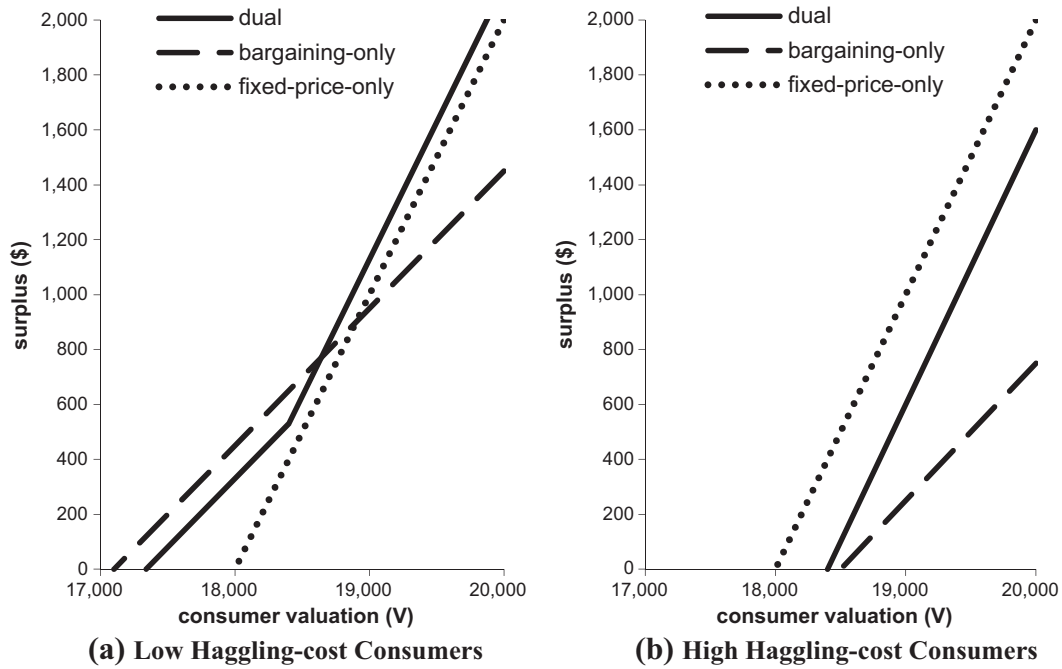
(c) Total Consumer Surplus

$(\beta \text{ varies}, hc_h = 0.07, hc_l = 0)$

Fig. 3. Consumer surplus.

in between. The fixed-price-only channel benefits consumers in two ways. First, the fixed price prevents the seller from engaging in price discrimination among consumers, which benefits, in particular, high-valuation consumers. Second and more importantly, the fixed-price-only channel eliminates the haggling costs that consumers incur when bargaining in the bargaining channel.

The fact that the fixed-price-only channel generates the highest surplus among all channel structures is a somewhat surprising finding, as it seems to contradict the fact that a number of regulatory agencies and consumer groups have argued that a fixed-price policy works against

consumer interests, at least in the auto industry (e.g., Competition Bureau of Canada, 2003; Automobile Consumer Coalition of Canada, 2006). However, it is important to note that our results do not imply that *all* consumers will be better off when prices are fixed. Instead, we claim that only consumers with certain characteristics benefit from it, i.e., low haggling-cost consumers are still better off if they bargain over prices, whether it is the bargaining-only structure or the dual channel. Nevertheless, our results suggest that a fixed-price policy is valuable as a whole, as it eliminates consumer haggling costs and limits the ability of sellers to price-discriminate among consumers.

4. Summary and conclusions

The conventional wisdom in setting prices is that a seller is better off if it is able to price-discriminate among consumers, using mechanisms such as bargaining. Offering a fixed price at the same time appears to reduce the advantage of price discrimination because buyers know the maximum price that can be charged. As a result, the recent emergence of a no-haggle, fixed price in markets that have traditionally relied on bargaining cannot be satisfactorily explained. Our research attempts to explain this phenomenon.

In particular, we explore the strategic implications of offering consumers the choice between bargaining a price and accepting a fixed, no-haggle price through such channels as the internet. We compare the profitability of three channel structures (bargaining only, fixed-price only and a dual-channel structure). Our findings suggest that consumer haggling cost plays a critical role in determining when a particular channel structure is optimal. We find that a dual channel is not always optimal: when either high or low haggling-cost consumers account for a large proportion of the population, or when they do not have very different haggling costs, a single channel is optimal. Our conclusions provide guidance to sellers: no one strategy is always the best, as optimization depends upon the magnitude and dispersion of haggling costs, which in turn may be related to such factors as customer bargaining experience, income and time constraints.

More broadly, we find that when individual-level price discrimination imposes a cost on consumers, as does haggling, it is not always optimal to use that strategy, as consumers attempt to offset that cost by seeking to pay a lower monetary price to the seller. Airlines and other industries that use yield-maximization strategies to change prices dynamically may find that customers, in turn, seek compensation for the time they spend searching for the "best price" by demanding a lower price than if prices were fixed over time.

As the development of new technologies enables sellers to more easily reach customers via multiple channels, allowing different pricing policies across channels, the choice of pricing formats emerges as a strategic consideration for sellers. Our findings generate implications for sellers deciding what pricing format(s) to implement.

Further, we examine the impact of a rarely considered strategic variable, namely the minimum acceptable price, to the pricing and channels literature. We show that the seller may not find it optimal to set the commission based on the true marginal cost of the product. Instead, the seller may be better off by specifying a higher-than-cost minimum acceptable price as the price floor of bargaining. The minimum acceptable price has important implications for both the bargained and no-haggle price in that it affects the outcome of the bargaining. It also leads to cases where the dual channel is more profitable but does not generate higher demand than a single-channel structure.

A limitation of our analyses is that we have not considered the cases in which the salesperson is risk averse and the seller has incomplete information. These factors can play a role in determining the conditions under which different pricing strategies are optimal and the optimal minimum acceptable price. First, increasing risk aversion may reduce a participant's share in the bargaining outcome and increases that of his opponent (Kihlstrom, Romer, & Williams, 1981; Osborne, 1985; Roth, 1989). This is because the more risk-averse participant is relatively more eager to minimize the risk of breakdown. This is exploited by the less risk-averse participant and he or she demands a larger share of the net surplus. Therefore, when the salesperson is more risk averse, we predict that the seller will set a higher price floor to raise the salesperson's threat point. Second, if the seller does not observe the salesperson's efforts and the cost of effort, the contract needs to provide incentives for the salesperson to exert the proper level of effort and to reveal his true type. If a salesperson has high cost of effort, he has a disadvantage in bargaining and thus the seller can set a higher minimum acceptable price. In contrast, for a salesperson with lower cost of effort, higher minimum acceptable is less necessary; rather, a higher

commission rate can motivate him to exercise his greater bargaining power (Cai & Cont, 2004). Therefore, the seller should provide a menu of contracts, such that the higher cost of effort the salesperson has, the higher the minimum acceptable price and the lower the commission rate. In short, in the presence of risk aversion and information asymmetry, we predict that the seller still sets a higher-than-cost minimum acceptable price in most cases but the optimal level may vary. With respect to the optimal channel structure, we believe that our main results remain qualitatively the same, that is, the optimal condition for the dual-channel is that consumers are sufficiently heterogeneous in their haggling costs. However, as risk aversion of the salesperson and the seller's information disadvantage imply a higher operating cost in the haggling channel, the cutoff of consumer heterogeneity may shift. Future research should examine these issues in more details.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ijresmar.2014.06.002.

References

Arnold, M.A., & Lippman, S. A. (1998). Posted prices versus bargaining in markets with asymmetric information. *Economic Inquiry, 36*(3), 450–457.
Automobile Consumer Coalition of Canada (2006). *The effects of "no-haggle" vehicle pricing on Canadian consumers.* Company Report.
Basu, A. K., Lal, R., Srinivasan, V., & Staelin, R. (1985). Salesforce compensation plans: An agency theoretic perspective. *Marketing Science, 4*(4), 267–291.
Besanko, D., Dubé, J. -P., & Gupta, S. (2005). Own-brand and cross-brand retail pass-through. *Marketing Science, 24*(1), 123–137.
Business Week (2007). Haggling starts to go the way of the tail fin, *71*, 4056.
Cai, H., & Cont, W. (2004). Agency problems and commitment in delegated bargaining. *Journal of Economics & Management Strategy, 13*(4), 703–729.
Chen, F. (2005). Salesforce incentives, market information, and production/inventory planning. *Management Science, 51*(1), 60–75.
Chiang, W. K., Chhajed, D., & Hess, J.D. (2003). Direct marketing, indirect profits: A strategic analysis of dual-channel supply-chain design. *Management Science, 49*(1), 1–20.
Competition Bureau of Canada (2003). *Competition Bureau Settles Price Maintenance and Misleading Advertising Case Regarding the Access Toyota Program.*
Desai, P.S., & Purohit, D. (2004). 'Let me talk to my manager': Haggling in a competitive environment. *Marketing Science, 23*(2), 219–233.
Dukes, A., & Gal-Or, E. (2003). Negotiations and exclusivity contracts for advertising. *Marketing Science, 22*(2), 222–245.
Gatehouse, D. (2007). *The Perfect Salesforce: The 6 Best Practices of the World's Best Sales Teams, Portfolio.*
Geyskens, I., Gielens, K., & Dekimpe, M. G. (2002). The market valuation of internet channel additions. *Journal of Marketing, 66*(2), 102–119.
Gill, D., & Thanassoulis, J. (2009). The impact of bargaining on markets with price takers: Too many bargainers spoil the broth. *European Economic Review, 53*(6), 658–674.
Guo, L., & Iyer, G. (2013). Multilateral bargaining and downstream competition. *Marketing Science, 32*(3), 411–430.
Joseph, K. (2001). On the optimality of delegating pricing authority to the sales force. *Journal of Marketing, 62*–70.
Kihlstrom, R. E., Romer, D., & Williams, S. (1981). Risk aversion with random initial wealth. *Econometrica: Journal of the Econometric Society, 911*–920.
Kumar, N., & Ruan, R. (2006). On manufacturers complementing the traditional retail channel with a direct online channel. *Quantitative Marketing and Economics, 4*(3), 289–323.
Lal, R. (1986). Delegating pricing responsibility to the salesforce. *Marketing Science, 5*(2), 159–168.
Lal, R., & Srinivasan, V. (1993). Compensation plans for single-and multi-product salesforces: An application of the Holmstrom–Milgrom model. *Management Science, 39*(7), 777–793.
Misra, S., Coughlan, A. T., & Narasimhan, C. (2005). Salesforce compensation: An analytical and empirical examination of the agency theoretic approach. *Quantitative Marketing and Economics, 3*(1), 5–39.
Moriarty, R. T., & Moran, U. (1990). Managing hybrid marketing systems. *Harvard Business Review, 68*(6), 146.
Nash, J. F. (1950). The bargaining problem. *Econometrica, 18*(2), 155–162.

Osborne, M. J. (1985). The role of risk aversion in a simple bargaining model. *Game Theoretic Models of Bargaining, 181–213*.

Riley, J., & Zeckhauser, R. (1983). Optimal selling strategies: When to haggle, when to hold firm. *Quarterly Journal of Economics*, *98*(2), 267–289.

Roth, A. E. (1979). Proportional solutions to the bargaining problem. *Econometrica*, *47*(3), 775–778.

Roth, A. E. (1989). Risk aversion and the relationship between Nash's solution and subgame perfect equilibrium of sequential bargaining. *Journal of Risk and Uncertainty*, *2*(4), 353–365.

Rubenstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, *50*(1), 97–109.

Srinivasan, V. (1981). An investigation of the equal commission rate policy for a multi-product salesforce. *Management Science*, *27*(7), 731–756.

Terwiesch, C., Savin, S., & Hann, I. H. (2005). Online haggling at a name-your-own-price retailer: Theory and application. *Management Science, 339–351*.

Thanassoulis, J., & Gill, D. (2010). *The optimal marketing mix of posted prices, discounts and bargaining.* Southampton Discussion Paper in Economics and Econometrics 1108.

Wang, R. (1995). Bargaining versus posted-price selling. *European Economic Review*, *39*(9), 1747–1764.

Winter, N. (2004). Rethinking Sales Compensation in the High-Tech Sector. *Compensation and Benefits Review*, *36*(5), 27–32.

Wilken, R., Cornelißen, M., Backhaus, K., & Schmitz, C. (2010). Steering sales reps through cost information: An investigation into the black box of cognitive references and negotiation behavior. *International Journal of Research in Marketing*, *27*(1), 69–82.

Zettelmeyer, F., Scott-Morton, F., & Silva-risso, J. (2006). How the Internet lowers prices: Evidence from matched survey and automobile transaction data. *Journal of Marketing Research*, *43*(2), 168–181.