International Journal of Research in Marketing 34 (2017) 367–381

Contents lists available at ScienceDirect

IJRM

International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

Assimilation-contrast theory in action: Operationalization and managerial impact in a fundraising context

Arnaud De Bruyn *, Sonja Prokopec

ESSEC Business School, Avenue Bernard Hirsch, 95000 Cergy, France

ARTICLE INFO

Article history:
First received on March 23, 2015 and was under review for 4½ months
Available online 12 November 2016

Guest Area Editor: Ralf van der Lans

Keywords:
Assimilation-contrast
Anchors
Appeals scales
Fundraising
Direct marketing

ABSTRACT

Charities often suggest specific donation amounts in their fundraising appeals, and the assimilation-contrast theory has been well-established as the explanation behind the impact of such anchors on donors' behavior. Yet, researchers who tested it in field studies have reported contradictory findings, and despite its proven reliability in the labs, this theory has had limited impact on managerial practice. Drawing on multiple streams of research, we develop a multi-step strategy to operationalize the assimilation-contrast theory in a fundraising context, and report the results of a large field experiment in which a charity used anchors to influence the behavior of 23,500 of its donors. We found that average donation amount increased by 22% and net margins increased by 36%. We report as one of the key managerial implications that the effects of the assimilation-contrast theory are largely asymmetric, implying that it is far easier for a firm to nudge customers in the direction of increased losses than in the direction of increased profit. We conclude by offering decision heuristics to those managers who do not have the resources to build an econometric model, but wish nonetheless to benefit from our findings.

1. Introduction

Firms regularly use anchors, or external reference points, to influence individuals' judgments and nudge customers into desired behaviors. For instance, realtors communicate on prices of expensive houses sold in the neighborhood to increase the perceived value of a listing and obtain higher bids; TV commercials feature consumers buying or using large quantities of their product; and salespeople announce expensive catalog prices to make the “deal” they are offering appear more attractive.

Assimilation-contrast theory (Sherif, Taub, & Hovland, 1958) well explains the underlying psychological mechanisms at play. This theory states that a consumer's current belief serves as an internal reference point (IRP hereafter) to which the persuasion attempt is compared. When a customer is presented with an anchor –be it an amount, a price, or a quantity– that is credible and congruent enough with his own beliefs, it is assimilated and contributes to adjust his attitudes, which in turn influence his behavior. If the discrepancy is too significant to be assimilated, the anchor is rejected (contrasted), and fails to influence either.

Although assimilation-contrast is a long-standing theory confirmed many times over in the labs, successful, large-scale managerial applications are rare. An illustration of that discrepancy between theory and practice can be found in the fundraising industry. The general practice is to suggest a donation range (e.g., \$20–\$30–\$50–\$100), which we refer to as an appeals scale or donation grid. It is a well-established finding that the values on the appeals scale influence both the likelihood of donation and

* Corresponding author.
E-mail addresses: debruyn@essec.edu (A. De Bruyn), prokopec@essec.edu (S. Prokopec).

http://dx.doi.org/10.1016/j.ijresmar.2016.11.002

the amount donated (e.g., Verhaert & Van den Poel, 2011), causing many charitable organizations to experiment with scales in search of an overall optimum.

Fundraisers in search for specific guidance will find little help in the literature, however, as it does not report the extent to which an organization might benefit from optimizing its appeals scales; offers no guidelines on how to perform such optimization; and even reports contradictory findings. While Schibrowsky and Peltier (1995) report that both likelihood of donation and donation amounts are influenced by appeals scales in the expected (opposite) directions, other authors have reported that only donation amounts were altered by suggested donation amounts, leaving response rates unchanged (Berger & Smith, 1997; Brockner, Guzzi, Kane, Levine, & Shaplen, 1984; Reingen, 1978; Weyant & Smith, 1987). Others report opposite findings, where likelihood of donation was indeed influenced, but where amounts were left unaffected (Desmet, 1999; Desmet & Feinberg, 2003; Doob & McLaughlin, 1989; Fraser, Hite, & Sauer, 1988). Fraser, Hite, and Sauer (1988) report that large anchors increase donation amounts, and legitimization of small donations increase likelihood, but that these effects cancel out when both tactics are used in conjunction. A final group of authors, finally, found no impact whatsoever (Abrahams & Bell, 1994; DeJong & Oopik, 1992).

In such context, our research objective is to bridge the gap between theory and practice, offer a clear pathway to fundraisers who wish to optimize their appeal scales, and demonstrate the potential managerial benefits of doing so.

To that avail, we first build on multiple streams of research (assimilation-contrast theory, prospect theory, configural weight theory, etc.) to list a set of modeling guidelines that practitioners should follow in their efforts to econometrically capture the influence of anchors on donors' behaviors –and more importantly, to harvest such influence to maximize charity's profits.

Second, we illustrate how these modeling guidelines can be implemented in practice by reporting a major empirical application in a fundraising context, and by showing the extent of financial improvement. We develop an econometric model that follows the above guidelines, calibrate the model on data collected to that avail from 50,200 donors, and capture the impact of appeals scales on their response rates and donation amounts. We then use the model in a subsequent fundraising campaign to develop recommendations for 23,500 additional donors, such that suggested donation amounts nudge behavior in the direction of optimized profits, and test the managerial impact of these individually-tailored appeals scales. We find an average donation amount increase of 22%, which in turn increases net margins by 36% compared with the baseline.

The manuscript proceeds with 7 modeling guidelines, followed by an empirical test of their application. We conclude with managerial implications and directions for future research.

2. Modeling guidelines

2.1. Modeling anchor influence on donation amount

In a fundraising context, the influence of a suggested donation amount is ultimately measured by the extent to which it nudges donation behavior away from what donors would have done, absent of any external influence. Without outside influences, a donor's behavior is determined by his internal reference point (IRP), his current attitudes toward the charity, past experience, values, etc. Assimilation-contrast theory (Sherif et al., 1958) suggests that an anchor's influence is conditional on how it compares to IRP.

Kalyanaram and Winer (1995, p.161), addressing price anchors specifically, observe "there is a significant body of literature to support the notion that individuals make judgments and choices based on comparisons of observed phenomena to an internal reference price" (see also Janiszewski & Lichtenstein, 1999; Mazumdar & Papatla, 2000). Extending their inquiry beyond prices, Hardie, Johnson, and Fader (1993) demonstrate that consumers' preferences are also reference-dependent when it comes to multi-attribute and brand choices.

Therefore, to integrate donors' internal reference point (IRP) in the model is essential not only to gauge the extent of anchor's influence, but also to model it. Thus, we refer to what the donor deems appropriate to pay as IRP location and integrate it into our first modeling principle below:

P1: The model should explicitly capture the heterogeneity in customers' IRP *location*.

Though rarely measurable, donors' IRP can often be inferred from past individual data. In the pricing literature, IRP has been estimated as the most recent price paid, the weighted mean of the logarithms of past prices, or as an exponential smoothing of past prices (Briesch, Krishnamurthi, Mazumdar, & Raj, 1997; Janiszewski & Lichtenstein, 1999; Kalyanaram & Winer, 1995; see Mazumdar, Raj, & Sinha, 2005, for a review). For a charity, a donor's last or average donation amount offer good proxies.

The target behavior is the visible resolution of a potential inconsistency between donors' internal beliefs (IRP) and external anchors (manipulated by the charity). For instance, if a donor plans to make a $50 donation, and the charity suggests $100 instead as a more appropriate amount, the final donation will be a reflection of the relative strength of each of the reference points. If the individual eventually gives $50, the external anchor would have had no influence whatsoever; if she gives $90, the external anchor would have all but supplanted the donor's initial intents.

Birnbaum's configural weight theory (Birnbaum, Coffrey, Mellers, & Weiss, 1992; Birnbaum & Zimmermann, 1998) states that in the presence of multiple reference points (e.g., internal vs. external), an individual places different weights on each, and the influence of a reference point is determined by its relative weight (see Mazumdar & Papatla, 2000, for an application of this competing principle in pricing research with multiple external anchors). We therefore suggest:

P2: The target behavior should be modeled as a weighted average of all reference points, either internal (IRP) or external (anchors). The larger the weight of a reference point, the more influence it has on the target behavior.

A corollary is that the stronger customers' IRP and the more ingrained customers' beliefs are, the less influence external anchors will have on their behavior.

Birbaum's theory also accounts for the possibility that some external anchors may bear more influence than others. For instance, some anchors are more prominently displayed, or are more memorable because they are presented first to the customer (primacy effect), or last (recency effect). Birnbaum et al. (1992) have shown that the weight of an anchor depends on its absolute value and rank among other anchors. If the suggested donation amount is portrayed as being suggested by a specific individual, such as the President of the University in a letter to its alumni, the source credibility or expertise will also influence weights (Birnbaum & Stegner, 1979; Birnbaum & Zimmermann, 1998).

2.2. Modeling IRP strength

The first modeling difficulty to overcome is the strength of the belief: donors vary not only on their IRP *location* (e.g., what they believe to be an appropriate donation amount), but also on the rigidity or *strength* of their beliefs[1] (Schibrowsky & Peltier, 1995). The weight associated with donors' IRP is heterogeneous (Mazumdar & Papatla, 2000). Some donors hold very strong beliefs about what is an acceptable amount, be that because of their experience, habits, or values, and external anchors will have very little influence in nudging them away from their IRP (De Bruyn, 2013; Verhaert & Van den Poel, 2011; Desmet, 1999). Other donors are more easily influenced. Since the effectiveness of an external anchor is inversely proportional to the strength of a donor's current beliefs, we suggest:

P3: The model should explicitly capture the heterogeneity in donors' IRP *strength*.

Heterogeneity in IRP strength can be explicitly modeled as a function of donors' characteristics, such as loyalty, expertise, recency, frequency, past experience, or psychological traits (Desmet, 1999; De Bruyn & Prokopec, 2009; De Bruyn, 2013; Verhaert & Van den Poel, 2011; Kalyanaram & Little, 1994). For instance, an alumnus who has made many donations to his university in the past will hold stronger beliefs about what constitutes an acceptable amount than a less-experienced alumnus, and the latter will be more easily influenced by manipulations of external anchors. In such case, "donation frequency" offers a proxy to explicitly model heterogeneity in donors' IRP.

If the researcher has access to a rich data set in which donors have been repeatedly exposed to various anchors, IRP strength can also be estimated at the individual level. In the best case scenario, both donors' characteristics and past data are available, the researcher can build a Bayesian hierarchical model that measures how donors' characteristics predict the strength of their IRPs (through the hyperparameter structure), while allowing additional heterogeneity to be captured by individual random effects. The best way to capture heterogeneity is ultimately determined by the richness of the available data.

2.3. Modeling anchor congruency

Assimilation-contrast theory suggests that an anchor's influence is moderated by its congruency, or credibility. If a suggested donation amount is close to donors' IRP, the anchor is assimilated, contributes to update donors' mental model, and eventually influences behavior. Alternatively, if it is largely incongruent with donors' IRP, it is contrasted, rejected, and bears no influence. Given the theoretical and practical importance to capture anchor's influence:

P4: The model should explicitly specify anchors' congruency. The weight of an anchor and its congruency are conceptually distinct and should be modeled separately, the latter being individual specific, and a moderator of the former.[2]

Assimilation-contrast theory (Sherif et al., 1958) suggests that anchor's congruency is a function of how distant it is from donors' IRP: the closer an external anchor is to IRP, the more likely it will be assimilated (and bear full impact); while the more distant it is from IRP, the more likely it will be rejected and fail to influence behavior (Berger & Smith, 1997). Hence:

P5: Anchor congruency should be maximum around customer's IRP, and monotically decrease as the anchor departs from it.

Congruency can be seen as a multiplicative factor that leaves the influence of an anchor unchanged if it is fully congruent (multiplies its weight by 1), and makes it vanish if it is incongruent and fully rejected (multiplies it by 0), with any possible value in between.

[1] Some authors argue that IRP should be seen as a point estimate which is more or less malleable, while others argue that IRP should rather be modeled as a range of values (e.g., Kalyanaram & Little, 1994). This distinction is akin to the difference between a frequentist and a Bayesian approach to parameter estimation: either the IRP can be seen as a point estimate that can be either strong or weak (equivalent to a small or large standard deviation), or the IRP can be seen as a distribution of points within either a narrow or wide range (equivalent to seeing the posterior distribution itself as the parameter estimate). While this distinction is conceptually interesting, we suspect they both lead to the same empirical insights, and we adopted the first approach in this paper.

[2] A pictorial illustration would be to compare the modeling of an anchor's impact to how loud a sound is heard by bystanders. The loudness of the sound is a function of how loud the sound is at the origin (i.e., weight of an anchor), which is fixed regardless of the listener; while how loud it is heard is individual specific, and depends among other factors on the distance (i.e., congruence) that separates the origin of the sound and the bystander. A sound might not be heard (an anchor might fail to influence behavior) either because it is not loud enough (the relative weight is too small), or because it is too distant (it is too incongruent), and both explanations are conceptually distinct enough to warrant two separate modeling processes.

The previous criteria leads to an interesting corollary: if an anchor is too close from IRP, it fails to nudge behavior in any meaningful way; if it is too distant, it becomes incongruent, is contrasted, and fails to influence behavior. Hence, there is a "sweet spot", an optimal distance where anchor's influence on donation behavior reaches its maximum.

Anchors nudge donors to move away from their IRP, but upward or downward deviations are not psychologically equivalent (Anderson, 1973; Hovland, Harvey, & Sherif, 1957; Sherif, 1963; Sherif et al., 1958), and such asymmetry may impact the optimal strategy of the charity or firm (Park, MacLachlan, & Love, 2011). One obvious candidate that predicts this asymmetry is loss aversion theory (Kahneman & Tversky, 1979), referring to people's tendency to strongly prefer avoiding losses to acquiring gains.

Suppose an individual receives a donation request from a charity he has never supported in the past. He decides to make a donation, but is unsure about the amount he should commit; he believes an amount of $100 or so may be appropriate. If the charity calls for a donation of a lesser amount, the prospective donor might feel a sense of relief, since complying with the suggestion would represent a psychological gain (Kahneman & Tversky, 1979; and Thaler, 1985, define "psychological gain" as the difference between an internal reference amount and a smaller anchor.) The external anchor might be easy to assimilate, since it is easy to comply with. If the requested donation amount is greater than the prospect planned initially, complying with the external anchor would represent a psychological loss (Kahneman & Tversky, 1979 and Thaler, 1985), and the donor might experience greater psychological resistance (i.e., stronger contrast), leading the external anchor to have a lesser influence.

This is equivalent to a central notion in behavioral pricing literature that "defining p^0 to be the observed retail price and p^r to be the individual's internal reference price, the underlying assumption […] is that positive values of (p^0-p^r) are perceived negatively, […] while negative values of (p^0-p^r) are viewed positively" (Winer, 1988, p.35; Kalyanaram & Winer, 1995). Erdem, Mayhew, and Sun (2001) have shown that sensitivity to losses with respect to IRP is greater than sensitivity to gains, and in the pricing literature in particular, considerable effort has been devoted to understanding the stability of the IRP and the factors that can alter it (Kalyanaram & Winer, 1995; Lichtenstein, Burton, & Karson, 1991; Mazumdar & Jun, 1992; Urbany, Bearden, & Weilbaker, 1988).

This psychological asymmetry is especially important from the charity's point of view since, in most situations, a donor's loss coincides with a charity's gain (e.g., a more generous donation); and if it is easier to nudge donors' behavior in the direction of increased losses for the charity than in the direction of increased profit, it is essential for the charity to model this asymmetry explicitly.

P6: Anchor congruency should be allowed to decrease asymmetrically around customers' IRP location.

More specifically, one would expect congruency to decrease faster in the direction of perceived losses than in the direction of perceived gains, which implies that anchors will be rejected (assimilated) faster if they imply psychological losses (gains). This hypothesis needs to be validated, however, and is not as straightforward in the fundraising literature as it is in the pricing literature, since a smaller amount also means a more limited contribution to the charity one wishes to support.

2.4. Modeling anchor influence on behavior occurrence

To nudge donors' IRP in the direction of higher donation amount can still end up decreasing the charity's profit if such manipulation advertently impacts donors' likelihood of donation. Therefore, modeling the influence of anchor on behavior occurrence is critical in the modeling process.

Some anchors merely represent upgrading options, while others constitute *psychological entry barriers*, that we define as the minimum amount signaled by the firm to be meaningful to engage in the behavior of interest. If suggested donation amounts are too high, donors might be concerned that their paltry donations will be negatively perceived, causing them not to donate at all (Gneezy, Gneezy, Nelson, & Brown, 2010). In gauging anchors' effectiveness, it is therefore critical to measure their influence not only on the behavior *quantity* (e.g., amount, price, volume), but also on its *occurrence*, since a positive impact on the first (e.g., revising a potential donation amount upward) may adversely impact the second (e.g., not making any donation).

P7: Beyond the impact of anchors on the *quantity* of interest, the model should capture its impact on behavior *occurrence*.

The criteria relevant to model behavior quantity are equally relevant to model behavior occurrence, such as the need to incorporate donors' IRP, to take into account donors' heterogeneity, to allow for asymmetry, or to model congruency.

3. Empirical application

3.1. Experimental procedure

We now report an empirical application of the above modeling principles to a fundraising context. Namely, we develop, calibrate and deploy an econometric model to manipulate appeals scales used in a direct marketing campaign, in order to nudge donor behavior in the direction of increased profit for the charity. This application followed 5 steps:

1. *Data generation*. Since the charity was not used to manipulate appeals scales, past data did not contain enough variance to reliably calibrate a model. We therefore submit 50,200 donors to randomly manipulated appeals scales, and report their behavior.
2. *Model development*. We develop an econometric model to capture the impact of appeal scales manipulations on donation behavior, following the 7 model guidelines presented earlier.

3. *Model calibration*. We calibrate the model on the data generated in step 1.
4. *Appeal scale optimization*. Using the calibrated model and a search algorithm, we find for 23,500 donors solicited for a subsequent fundraising campaign the amounts on the appeals scales that nudge their behavior in the direction of maximized profit.
5. *Empirical test*. We report the actual impact of these individually tailored appeals scales in a field experiment. We find that applying this econometric model increased net margins by 36% compared with the baseline.

We develop these five steps in detail hereafter.

3.2. Step 1: data generation

A large European nonprofit organization sent a solicitation letter to its existing donors. All solicited donors had made at least one donation in the past; prospective donors were excluded. Each solicitation included a letter, a stamped envelope, and a reply coupon. Typically, if a donor was willing to make a donation, he or she would use the preaddressed envelope and send a check together with the reply coupon. The coupon was personalized with the donor's name, address and ID, as well as a personalized appeals scale that contained a varying list of four suggested donation amounts (e.g., 100 €, 120 €, 150 €, 200 €) and an “Other” option.

The nonprofit organization and research team manipulated these appeals scales to obtain a data set with sufficient variability to quantify the impact of anchor manipulations on both compliance and generosity. The study design proceeded as follows.

First, a few days before the solicitation letters were printed, the research team obtained the charity's donor database, extracted the last donation amount for all donors, and used it as a reference point to generate appeal scales randomly. We used the last donation amount, rather than the average amount or other proxy of IRP, because some donors had been in the database for decades, and their average donation amounts were not representative of their more recent behavior. Besides, because the purpose of this first experiment was to generate data suitable for further econometric calibrations, using the last donation amount also created more wide-ranging data points than other metrics would have.

Second, after extracting their last donation, we randomly assigned each donor to one of the 3 × 3 experimental conditions. The first condition referred to how the donor's last donation amount appeared in the appeals scale. Traditionally solicitations have offered a first suggested amount very close to the donor's most likely behavior, and the other amounts represented upgrading options. We manipulated this first, left-end amount such as it would be lower, equal, or higher than donors' last donation amount. The other suggested amounts then increased monotically, and we further manipulated the steepness of this increase, such that each subsequent amount was 20%, 50%, or 80% higher than the preceding amount on the grid. We used stratified sampling to ensure that donors in all charity-defined segments were equally represented across conditions. Each experimental cell was tested with approximately 5578 randomly selected donors.

Third, we constructed a customized appeals scale for each donor, tailored to both their last donation amount and the assigned experimental condition. Two donors in the same experimental condition could be exposed to different scales, depending on their last donation amount, used as the basis of the appeal scale design. In Table 1, we report the specific proportions used for the four suggested amounts for all 3 × 3 experimental conditions. In Table 2, we detail the left and right ends of the appeals scales (after rounding) for two hypothetical donors with different last donation amounts.

Six months after the launch of the campaign, we obtained a list of donors who had made a donation, along with their donation amounts. For completeness, we report average response rates and donation amounts for all 9 conditions in Table 3 and Table 4.

We now develop a model built on the 7 aforementioned guidelines to fit the data.

3.3. Step 2: model development

3.3.1. Donation amount model

When deciding how much to contribute, a potential donor will have to chose between various amount options: what he believes to be appropriate on the one hand (his IRP, which following *P*1 should be explicitly modeled and individual specific), and what the charity signals is expected from him on the other hand (the anchors on the appeals scale). The effectiveness of

Table 1
Ratio of the four suggested amounts to donors' last donation amounts for all experimental conditions.
Notes: For a donor in the Lower/Steepest condition, the first suggested amount is 56% of his last donation; the second amount equals his last donation, and the last suggested amount is more than three times larger his last donation.

		First amount		
		Lower	Equal	Higher
Steepness	Steep	0.83; 1.00; 1.20; 1.44	1.00; 1.20; 1.44; 1.73	1.20; 1.44; 1.73; 2.07
	Steeper	0.67; 1.00; 1.50; 2.25	1.00; 1.50; 2.25; 3.38	1.50; 2.25; 3.38; 5.06
	Steepest	0.56; 1.00; 1.80; 3.24	1.00; 1.80; 3.24; 5.83	1.80; 3.24; 5.83; 10.50

Table 2
Low and high ends of suggested appeals scales (rounded), for two typical donors with last donation amount of 20 € (Left) and 100 € (Right).
Notes: Because the suggested amounts are proportional to the last donation amount, two donors in the same experimental condition may be exposed to different appeals scales in absolute values. Compared to their last donation amount, however, they are proportionally identical.

		First amount					First amount		
		Lower	Equal	Higher			Lower	Equal	Higher
Steepness	Steep	18...30	20...35	25...40	Steepness	Steep	85…140	100...170	120...200
	Steeper	14...45	20...70	30…100		Steeper	70...230	100...350	150...500
	Steepest	10...65	20...120	35…200		Steepest	55...320	100...580	180...1000

an anchor, either internal or external, is a function of its strength relative to the strength of all other competing anchors. As per P2, stemming from Birnbaum et al. (1992); Birnbaum and Zimmermann (1998), we model donation amount as:

$$m_i = \frac{\sum_{j=0}^{J} \omega_{i,j} \cdot r_{i,j}}{\sum_{j=0}^{J} \omega_{i,j}} \tag{4}$$

where
i Index referring to the *ith* donor;
j Index referring to the *jth* reference amount. To reduce notational clutter, we refer to $j = 0$ as the donor's IRP, and $j = 1, \ldots, J$ as the *jth* amount suggested on the appeals scale ($J = 4$ in our context);
m_i Amount donated by the *ith* donor;
$r_{i,j}$ Reference amounts, either internal ($j = 0$) or external ($j = 1, \ldots, J$);
$\omega_{i,j}$ Relative influence of each reference amount on the target behavior, where $\omega_{i,j} \geq 0 : \forall i,j$.

This model stipulates that various reference amounts ($r_{i,j}$) compete to influence donation amounts (m_i). These reference amounts exert varying influences ($\omega_{i,j}$); the larger the $\omega_{i,j}$, the more the donation amount of the *ith* donor is influenced by the *jth* reference point. We now discuss how we modeled IRP ($j = 0$) and external anchors' ($j \geq 1$) weights in this application.

3.3.2. IRP strength ($w_{i,0}$)

The influence of the *ith* donor's IRP on his decision is noted $w_{i,0}$. As $w_{i,0} \to \infty$, IRP exerts (relatively) more influence on the donor's behavior; that is, the higher the value of $w_{i,0}$, the more salient the IRP will be in deciding how much to give, and $m_i \to r_{i,0}$. In contrast, a lower $w_{i,0}$ implies that donor's behavior is more malleable, and that the appeals scale exerts a greater influence in nudging him away from his IRP ($m_i \neq r_{i,0}$).

Following *P*3, and consistent with Mazumdar and Papatla (2000) who show that IRP's salience might vary across individuals, we stipulate that the strength of a donor's IRP is individual specific. Among potential indicators, the number of donations made in the past is a good indicator of IRP strength (De Bruyn & Prokopec, 2009; Verhaert & Van den Poel, 2011; Desmet, 1999), and we model it as:

$$\omega_{i,0} = \delta_0 + \delta_1 \ln(f_i) \tag{5}$$

where
$\delta.$ Two parameters to be estimated;
f_i Frequency of the *ith* donor (total number of donations made in the past).

As $f_i \to \infty$, $\omega_{i,0} \to \infty$, hence $m_i \to r_{i,0}$; that is, as the donor's frequency increases, the strength of his IRP increases. Making many donations contributes to form a stronger and more rigid idea of what constitutes an appropriate amount, and appeal scale manipulations exert less influence on frequent donors. We note that the data set at hand is not rich enough (no panel structure) to allow the estimation of truly individually-estimated coefficients $\omega_{i,0}$. We discuss potential model refinements in the conclusions.

Table 3
Average return rates (donations) of the solicitation campaign.

		First amount			
		Lower	Equal	Higher	Average
Steepness	Steep	9.8%	9.2%	8.5%	9.2%
	Steeper	9.4%	9.1%	8.1%	8.9%
	Steepest	10.1%	9.3%	7.9%	9.1%
	Average	9.8%	9.2%	8.2%	9.0%

Table 4
Donation amounts as deviations from Base of 1.
Notes: we report the donation amounts as a proportion of the donors' most recent donation, such that a high value (>1) indicates a donation amount above what the charity likely expected from that donor, whereas a low value (<1) indicates a smaller donation amount. Namely, we report (donor i's donation amount for this campaign/donor i's last donation) × scaling factor, where the scaling factor is computed across all donors who made a donation for this campaign, equal to (average of last donations/donation amounts for this campaign), which ensures that the grand average is equal to 1.

		First amount			
		Lower	Equal	Higher	Average
Steepness	Steep	.949	.960	1.037	.980
	Steeper	.895	1.000	1.085	.989
	Steepest	.937	.982	1.212	1.031
	Average	.927	.980	1.109	1.000

3.3.3. Influence of external anchors ($w_{i,j}$)

Whereas $\omega_{i,0}$ indicates the strength of donor's *internal* reference point, $\omega_{i,j}$: $j \geq 1$ indicate the influence of *external* anchors. Following *P*4, we model the latter as a weight, moderated by an individual-specific congruency indicator:

$$\omega_{i,j} = c_{i,j} \cdot \gamma_j \tag{6}$$

where
γ_j Attraction weight to be estimated at the population level; and
$c_{i,j}$ Congruency of the *jth* reference amount for the *ith* consumer.

Note that in this model, external anchors' attraction weights are homogeneous across donors, for two reasons. First, while there could be situations where donors could be exposed to different kinds of anchors (e.g., different anchor presentations, repeated exposures), this is not the case in our specific experimental setting. Second, following Birnbaum's configural weight theory, anchors' weights are relative to one another. Consequently, arguing that an external anchor may exert less influence on some donors is structurally equivalent to arguing that those same donors are less susceptible to be influenced by say external anchor. In other words, the heterogeneity introduced in IRP strength partly captures the potential heterogeneity in external anchors' attraction weights. Different experimental contexts or access to richer data (e.g., panel structure with repeated exposures) would call for a more explicit introduction of heterogeneity.

Anchor congruency is a pivotal concept, since both the assimilation-and-contrast theory (Sherif et al., 1958) and adaptation-level theory (Helson, 1964) predict that external reference points that are highly incongruent with the IRP are likely to be rejected or ignored, hence bear no influence. Greater believability allows for greater influence on consumer behavior (Compeau & Grewal, 1998). Following *P*4, any model that aims at applying the assimilation-contrast theory to real-life problems should explicitly calibrate anchor congruency at the individual level. For this specific application, we modeled the degree of congruence of an anchor as a normal-shaped function, centered around the donor's IRP:

$$\mathrm{c}_{\mathrm{i,j}} = \mathrm{e}^{\frac{\mathrm{d}^2_{\mathrm{i,j}}}{2\sigma^{*2}}} \tag{7}$$

where

$$\mathrm{d}_{\mathrm{i,j}} = \frac{\mathrm{r}_{\mathrm{i,j}} - \mathrm{r}_{\mathrm{i,0}}}{\mathrm{r}_{\mathrm{i,0}}} \tag{8}$$

$$\sigma^* = \begin{cases} \sigma_1 & \text{if } \mathrm{d}_{\mathrm{i,j}} \leq 0 \\ \sigma_2 & \text{if } \mathrm{d}_{\mathrm{i,j}} > 0 \end{cases} \tag{9}$$

and
$d_{i,j}$ A measure of relative distance between external anchor $r_{i,j}$ and IRP $r_{i,0}$; and
$\sigma_.$ Congruency parameters (>0) to be estimated.

Following P5, congruence is maximum and equal to 1 when $d_{i,j} = 0$, that is, when $r_{i,j} = r_{i,0}$. When a suggested donation amount equals the donor's IRP, it is fully congruent and assimilated. Conversely, as $|r_{i,j}\text{-}r_{i,0} \to \infty$, the external anchor becomes incongruent ($c_{i,j} \to 0$ at a rate determined by σ^*), hence it exerts less influence ($\omega_{i,j} \to 0$). Consistent with *P*6, we also allow the congruency parameter σ^* to be asymmetric around the IRP (Eq. (9)). We expect donation amounts to appear more incongruent if they are larger than $r_{i,0}$, since larger amounts represent psychological losses (being asked to give more than originally planned), so we expect $\sigma_1 > \sigma_2$.

Note that most donors use the lowest value on the appeals scale as the anchor for judging their contribution (Schibrowsky & Peltier, 1995). The first amount is the most salient and should bear greater influence than subsequent suggested amounts. Because the donor's IRP likely has the greatest influence (even for infrequent donors), we expect $\omega_{i,0} > \gamma_1 > \gamma_j{:}\forall j \geq 2$.

3.3.4. Donation likelihood model

Beyond anchor influence on behavior *quantity*, it is crucial to model its influence on behavior *occurrence*. Per P7, we argue that some anchors play the role of psychological entry barriers. In our fundraising application, following Schibrowsky and Peltier (1995) and De Bruyn and Prokopec (2009), we argue that the first suggested amount on the appeals scale plays such role, since it indicates the minimum amount the charity deems acceptable for a donation. It serves as a frame of reference that should reduce donation likelihood if it is higher than the donor's IRP and increase it if lower, though to a lesser extent. Consistent with these propositions, we model the donor's likelihood of donation using a classic logit specification as:

$$p_i = \text{logit}\left(\alpha_0 + \alpha_1 y_i + \alpha_2 \ln(y_i) + \alpha_3 \ln(f_i) + \alpha_4 \ln(m_i) + \beta^* d_{i,1} c_{i,1}\right) \tag{11}$$

$$\beta^* = \begin{cases} \beta_1 & \text{if } d_{i,1} \leq 0 \\ \beta_2 & \text{if } d_{i,1} > 0 \end{cases} \tag{12}$$

and

p_i Likelihood that the *ith* donor will make a donation;
y_i Recency of the *ith* donor: number of years elapsed since the last donation;
f_i Frequency: total number of donations made by the *ith* donor to date;
m_i Average donation amount; and
α., β. Parameters to be estimated.

We approximate the donor's likelihood of donation with a logit model (Eq. (10)), using a simple model in which recency, frequency, and monetary values (RFM) are independent variables. Since, in a simple logit model, multicollinearity can quickly become an issue, we do not introduce additional covariates (such as last donation amount, variability in past donations, etc.), but note that they might have additional predictive power. Recency relies on a nonlinear formulation; donors who have made a donation very recently are unlikely to support the charity again in the immediate future, as are those who have lapsed from donating for a long period. We therefore expect $\alpha_1 < 0$ and $\alpha_2 > 0$. Donation behavior often follows an annual rhythm, and the nonlinear formulation captures this inverted U-shaped relationship between recency and donation likelihood. The logarithmic transformation of frequency is common practice: The marginal likelihood increases tremendously from one to two donations, when new donors confirm their willingness to support the charity, a little less from two to three donations, and then keeps decreasing as f_i increases. We expect $\alpha_3 > 0$. The logarithmic term for monetary value indicates that donors who contribute very little tend to be less loyal to the charity, whereas this effect tends to disappear for the more generous donors, so we expect $\alpha_4 > 0$.

We augmented the classic RFM model with the impact of the first amount on the appeals scale, expressed according to a relative deviation from the donor's IRP (Eq. (8)), weighted by its congruency (Eq. (7)). The interplay of these two parameters capture the fact that an anchor might have limited impact on donation likelihood, either because it is very close to the donor's IRP (inviting donors to contribute an amount they would have given anyway), or very distant (hence contrasted, rejected, and failing to influence the donor's frame of reference).

Consistent with *P*6, its impact is captured by β and allowed to be asymmetric (Eq. (12)). We expect $\beta_1 < 0$ and $\beta_2 < 0$; when the first suggested donation amount is smaller than the donor's IRP, $d_{i,1}$ is negative, and $\beta_1 < 0$ captures a positive impact on donation likelihood. Based on the arguments leading to *P*6, the overall impact should be stronger when $d_{i,1} > 0$ than when $d_{i,1} < 0$. This formulation is consistent with prior literature (Helson, 1964; Sherif et al., 1958): To have an impact, the anchor must push the donor away from his or her IRP ($d_{i,1} > 0$), but it might lose impact if it becomes too extreme ($c_{i,1} \rightarrow 0$). To our knowledge, this moderating role has been theorized, but never econometrically calibrated.

3.3.5. Donor's internal reference point

We have repeatedly referred to $r_{i,0}$ (the *ith* donor's IRP) without specifying its operationalization. Despite the many potential metrics to approximate donors' IRPs, existing theory provides little guidance. We therefore compare model fit with various proxy candidates: first, last (most recent), average, and maximum donation amounts, as well as exponential smoothing averages with different smoothing factors (0.20, 0.50, and 0.80), and keep the one that best explains the data.

Previous research has conceptualized donation behavior as a two-step process (Xu & Wyer, 2007), and since decisions about whether to donate and then how much to donate might tap into different psychological mechanisms, they might also summon different IRPs. More specifically, Adaval and Monroe (2002) have shown that the context in which a judgment is made influences the internal standard that consumers use. When a donor decides *whether or not* she will make a donation, the first donation amount on the appeals scale has been shown to be the most salient and influential (Schibrowsky & Peltier, 1995), and plays the role of a psychological barrier (De Bruyn & Prokopec, 2009). Then, when a donor decides *how much* to give, the full range of anchors come into play, even if they are artificially manipulated or are perceived to be incidental (Nunes & Boatwright, 2004; Schwarz, Hippler, Deutsch, & Strack, 1985). We follow Adaval and Monroe (2002) and posit that different decisional contexts (donate at all? vs. how much?), and focusing on different anchors (the first vs. the whole range) might trigger different IRPs. While we do not necessarily argue that there are two completely different IRPs at work, we cannot rule out the hypothesis that the donor's IRP adjusts during the two stages of the decision-making process. We therefore allow the reference points to differ across models.

3.4. Step 3: model calibration

The data set we used to calibrate the models is the one we described previously ($N = 50.208$), augmented by 50,180 additional solicitations, for a total of 100,388 solicited donors. The 50,180 additional solicitations had been sent during the same campaign to an equivalent population of donors, but with the appeals scales the charity was using at the time.

3.4.1. Donation likelihood model calibration

The donation likelihood model contains nine parameters to estimate: five parameters of the RFM model (α_0 to α_4), two parameters that capture the impact of the first donation amount on the likelihood of donation (β_1 and β_2), and two parameters that capture the speed at which suggested donation amounts became incongruent as they diverge from the donor's IRP (σ_1^p and σ_2^p, where the superscript p refers to the donation model). In line with our theoretical developments, we expect $\alpha_1 < 0$ and $\alpha_2 > 0$ (inverted U-shaped relationship between recency and likelihood of donation); $\alpha_3 > 0$ (frequency); and $\alpha_4 > 0$ (monetary). We also expected $\beta_1 < 0$ and $\beta_2 < 0$. We fit the model to the entire data set ($N = 100.388$), using maximum likelihood estimation, and report results in Table 5.

The IRP with the best fit (i.e., highest likelihood) is the donor's last donation amount (Table 6); the data suggest that when judging the acceptability of donation requests, donors compare them with their most recent (and salient) donation amount. Changing the reference point does not affect the log-likelihood much; another popular metric, the model lift at 10% points to the average donation amount, while the last donation amount comes a close second. All in all, the best IRP operationalization is not clear. Of the 100,388 donors in the data set, 20,403 made only one donation, and 5275 made multiple donations of equal amounts, leading to identical IRPs, regardless of the measure used. Another 26,000 donors reveal a history of donations within a tight range, and such reduced variability leads to small variations in the log-likelihood values. Due to the lack of data with more variability, and without stronger theory to guide our choices, we simply retain the measure that leads to the highest likelihood.

The parameter estimates of the impact and congruence parameters also tell an interesting story. A high value of σ_2^p indicates that asking for more than the donor's IRP keeps exerting a negative impact on donation likelihood, however extreme the request. An unrealistic request, regardless of the extent to which it seems out of line, upsets donors. The reverse is untrue though; a small value for σ_1^p means very small requests are promptly deemed incongruent and ignored, in line with DeJong and Oopik's (1992) finding that "even a penny will help" donation tactics are ineffective in direct marketing. In Fig. 2, we illustrate a hypothetical donor's response likelihood for various values of the first anchor; it strongly supports the asymmetry of the congruence parameters (t-test, $p < 0.001$).

3.4.2. Donation amount model calibration

In Birnbaum's configural weight theory, the weights of the anchors, either internal or external (γ_j), are only relative to one another. Consequently, for identification purpose, it is necessary to fix one of them to a constant (i.e., 1). In this fundraising context, however, the first donation amount is under the research team's control, and each subsequent anchor is a multiple of the preceding one, at a rate that has been experimentally manipulated as well. It ensues that the model has fewer free parameters than is apparent, and more parameters need to be fixed to take into account the inherent correlation across suggested donation amounts. Since our theoretical framework predicts that the first donation amount will be the most influential, we freely estimate γ_1 and fix the other weights to $\gamma_j = 1{:}\forall_j > 1$.

The donation amount model contains five parameters to estimate: γ_1 (weight of the first suggested donation amount in the attraction model); δ_0 (intercept of the weight of the donor's IRP, equivalent to γ_0 for donors who made only 1 donation); δ_1

Table 5
Parameter estimates and key statistics: Donation likelihood model.
Notes: All parameters are significant and have the expected signs (except for α_4).

Parameter	Estimate	Standard deviation	p-Value	Description
α_0	−1.578	0.261	0.000	RFM model intercept
α_1	−1.044	0.108	0.000	Recency
α_2	0.232	0.045	0.000	Recency (log)
α_3	0.502	0.012	0.000	Frequency (log)
α_4	−0.285	0.045	0.000	Amount (log)
β_1	−0.577	0.289	0.023	Anchor impact, below reference point
β_2	−0.226	0.039	0.000	Anchor impact, above reference point
σ_1^p	0.275	0.084	0.001	Anchor congruence, below reference point
σ_2^p	6.835	3.924	0.041	Anchor congruence, above reference point

Table 6
Candidate metrics to capture donors' internal reference points.
Notes: For the donation amount model, errors and r^2 are computed on log(amounts). The last column reports the deviance of an alternative model fit to maximize the log-likelihood (as opposed to minimized the sum-of-square errors), and lead to identical results.

Donation likelihood model	−2 LogL	Lift 10%	Donation amount model	Sq. errors	R^2	−2 LogL
Last donation amount	58,248	2.328	Exp. average (0.5)	928.9	0.848	5223
Exponential average (0.2)	58,271	2.310	Exp. average (0.8)	940.2	0.846	5311
Maximum	58,272	2.074	Last donation	996.0	0.836	5818
Average	58,280	2.371	Exp. average (0.2)	1062.8	0.828	6429
Exponential average (0.5)	58,285	2.262	Average	1180.3	0.811	7364
Exponential average (0.8)	58,287	2.254	Maximum	1300.2	0.735	8253

(influence of the donor's frequency on the strength of the IRP); and σ_1^m and σ_2^m (congruency parameters, with superscript m for this model). Following our theoretical developments, we expect $\delta_0 > \gamma_1 > 1$, $\delta_1 > 0$, and $\sigma_1^m > \sigma_2^m$. We fit the donation amount model to the subset of donors who made a donation ($N = 9010$) and performed a squared error minimization of the difference of the logs between observed and predicted donation amounts. We report parameter estimates in Table 7.

As we expected, $\delta_0 > \gamma_1$ ($p = 0.008$) and $\gamma_1 > 1$ ($p = 0.051$). The donor's IRP is much more important than external reference points in determining actual behavior (especially among frequent donors, with $\delta_1 > 0$), and the first suggested amount exerts an impact three times greater than the other amounts on the appeals scale.

Our results also confirm the asymmetry hypothesis ($\sigma_1^m > \sigma_2^m$, $p < 0.001$). With a high value for σ_1^m, suggested amounts below the donor's IRP never appear incongruent, regardless of how small they are. A donor who is used to giving 100 € might find it perfectly reasonable to be asked to give 10 € and may comply. However, a much smaller value for σ_2^m means suggested donation amounts larger than the donor's IRP are quickly deemed incongruent and have no impact on donation amounts, as we illustrate in Fig. 3.

The IRP that provides the best fit is not the last donation amount (as in our previous model) but an exponential smoothing average with a smoothing factor of 0.5 (Table 6), regardless of the success metric used. More recent donation amounts contribute greatly to the reference point used in the amount model, but the smoothing average captures a natural, well-known reversion to the mean phenomenon.

While model parameters are consistent with theory, significant and in the expected directions, it remains to be seen if this effort can have a significant managerial impact. With that in mind, we next run these models to identify the profit-maximizing appeals scales for each donor solicited during a second fundraising campaign and measure the impact of individually tailored appeals scales on donation likelihoods, donation amounts, and profitability.

3.5. Step 4: appeal scale optimization

A year after the first campaign used to generate data and calibrate the model, the sponsor charity sent approximately 57,000 solicitations to existing donors who had made at least one donation in the past two years. These donors were randomly assigned, by the research team, to one of two conditions:

Control 33,302 donors exposed to the charity's traditional appeals scales.
Optimized 23,552 donors exposed to an optimized appeals scale, specifically designed to maximize their expected contribution (see below).

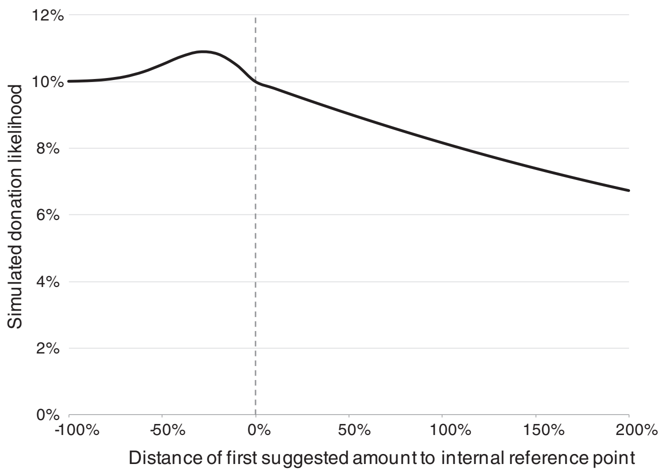

Fig. 2. Impact of first suggested donation amount on donation likelihood: hypothetical donor.

Table 7
Parameter estimates and key statistics: Donation amount model.
Notes: All parameters are significant and have the expected signs.

Parameter	Estimate	Standard deviation	*p*-Value	Description
γ_1	3.405	0.859	0.000	Attraction weight, first external ref. point. Subsequent anchors have a weight of 1.
δ_0	12.748	3.899	0.001	Attraction weight, internal reference point
δ_1	5.646	2.297	0.007	Influence of frequency on internal reference point's strength
σ_1^m	10.008	1.028	0.000	Anchor congruence, below reference point
σ_2^m	1.844	0.359	0.000	Anchor congruence, above reference point

Prior to the campaign, the charity provided the donation history of all solicited donors, and we used the two preceding models to design individually tailored appeals scales for all donors in the OPTIMIZED group. However, these optimized appeals scales were subject to two managerial constraints. First, donors of this charity tend to expect a certain pattern in appeals scales, with suggested amounts that increase monotonically, so $(r_4\text{-}r_3) \geq (r_3\text{-}r_2) \geq (r_2\text{-}r_1)$. We decided that violating this expectation could create undesired perturbations. Moreover, the original experiment used to calibrate the models followed the same convention, so we judged it safer to keep this pattern unchanged rather than to stretch the models into uncharted territories. We therefore only optimized the first suggested donation amount $r_{i,1}$, and the rate of increase, which was constant:

$$\Delta_i = \frac{r_{i,j}}{r_{i,j-1}} = \frac{r_{i,j+1}}{r_{i,j}} : \forall j > 1 \tag{10}$$

Second, the management team requested rounded suggested amounts, which is consistent with Desmet and Feinberg's (2003) report that rounded suggested amounts exert a greater influence on donors' behavior.

For each individual in the OPTIMIZED group, we determined the appeals scale as follows:

1. Determine the suggested amount that will appear first on the appeals scale ($r_{i,1}$) and the rate of increase of further suggested amounts (Δ_i);
2. Given the two above values, obtain the entire appeals scale ($r_{i,1}$ $r_{i,2}$ $r_{i,3}$ $r_{i,4}$);
3. Based on the appeals scales the donor will be exposed to and the donor's individual variables (e.g., recency, frequency, last, average and smoothed donation amount, etc.), apply the models described previously to estimate the likelihood of donation (Eq. (10)), the most likely donation amount (Eq. (4)), and consequently the expected profit;
4. Find the combination of values $r_{i,1}$ and Δ_i that maximize donor's expected profit. To achieve that, we applied a nonlinear search algorithm where the starting values were initialized using simulated annealing to avoid local minima, and the solution was refined using a gradient search;
5. Once the optimal values are found for this particular donor, round the amounts appearing on the appeals scales.

For instance, consider a hypothetical donor with a last donation amount of 80 €, a smoothed average donation amount of 95 €, and a likelihood of donation (absent of any anchor influence) of 8.5%. Plugging these figures into our model, the combination of the first suggested donation amount and rate of increase that maximizes expected net margins is 122.5 € and 27.7%. After rounding, the optimal appeals scale for this donor is [120 €, 160 €, 200 €, 260 €]. Such anchor combination would decrease donation likelihood by 0.6%, but increase expected donation amount by 12 €, for an expected net gain of 45 cents (or +4.6%).

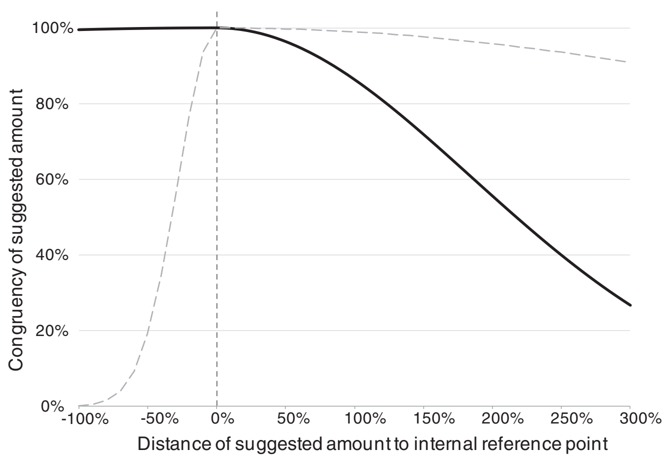

Fig. 3. Congruence of suggested donation amount based on distance from donors' internal reference point. Notes: The bold line depicts the donation amount model; the dashed line is the donation likelihood model.

Table 8
Financial results for comparison of traditional appeals scales (CONTROL condition) against individually tailored, optimized appeals scales (OPTIMIZED condition).

Condition	N	Donations	Return rate	Average donation amount	Average contribution	Average net margin
CONTROL	33,302	1484	4.46%	49.2 €	2.19 €	1.54 €
OPTIMIZED	23,552	1077	4.57%	60.2 €	2.72 €	2.10 €
Difference			+2.5%	+22.4%	+24.2%	+36.4%

3.6. Step 5: empirical test

3.6.1. Financial results

We report the financial results of this second field study in Table 8. Donors did not find the theme of that specific fundraising campaign particularly appealing, so the results overall were poor. Specifically, in the CONTROL condition, the return rate was 4.46%, with an average donation amount of 49.2 € (close to the historical average of 52 €). The average contribution margin was 2.19 €. With a cost of 65 cents per mailing, the average net margin was 1.54 € per donor (the exchange rate between dollar and euro at the time was 0.74).

In the OPTIMIZED condition, the return rate was virtually unchanged (4.57%, n.s.), but the average donation amount improved by more than 22% (60.2 €, $p < 0.01$), which increased the average total contribution by 24%. With an unchanged cost structure, the average net margin per donor moved from 1.54 € to 2.10 €, a 36% improvement.

3.6.2. Robustness tests

Noting the amplitude of the observed improvements, we ran further robustness tests to confirm the validity and reliability of our results. Specifically, we explored whether improvements in the results were widespread in the tested population, or if they were driven primarily by a few individual outliers or a few very responsive segments.

To explore whether extremely generous donation amounts (e.g., outliers) have caused the striking results, we replaced the top 5% of the most generous donations (150 € or greater) by the ceiling value associated with the 95th percentile of donation amounts. Therefore, a donation of 5000 € would instead be accounted for 150 €, which provides a very stringent test. The differences remained significant: The average donation amount increased by +12.3%, and the net margin increased by +22.0%, which ruled out the outlier hypothesis.

Second, we checked whether the results improved uniformly across segments. For confidentiality reasons, we cannot report the exact segmentation structure used by the charity, but the data suggested improvements across all segments, with varying amplitude. The average donation amount increased between +8% and +44%, depending on the segment, and the most impressive improvements appeared among the most generous donors.

Although it is tempting to assume that generous donors were more susceptible to appeals scale manipulations, analyzing the charity's default strategy draws a different picture. In the *CONTROL* condition, the charity adapted appeals scales to segment membership. Donors were divided into 18 segments, each segment being assigned to one of four appeals scale, and the brackets grew wider with the shift toward more generous (and more rare) donors. For example, the first bracket clustered donors with average generosity between 1 and 50 €; the last bracket grouped donors with an average donation amount between 750 € and 7500 €—a much wider range. Because the charity had been suggesting identical appeals scales to all donors in the same bracket, the segments of generous donors offered greater room for improvement.

Third, a small improvement in the average donation amount of a segment can lead to significant financial results. The segment for which our proposed model achieved the smallest improvement contained occasional and not-so-generous donors, whose average donation amount was 29.2 € in the CONTROL condition, with a return rate of 3.12%. In the OPTIMIZED condition, the average donation amount improved by only +7.90% (lowest improvement across all segments), with a return rate of 3.64%. Because the donors in this segment were among the least generous, they were also the least profitable, and their average expected contribution was only 91 cents per donor, for a profit of 26 cents. Any cent beyond mailing costs represents a pure net margin, and the latter improved to 50 cents per donor in the OPTIMIZED condition, a 90% improvement.

3.6.3. Linking financial results and segment heterogeneity

The increase in average donation amounts surprisingly was not accompanied by a corresponding decrease in return rates; literature suggests that one occurs only at the expense of the other (e.g., Desmet & Feinberg, 2003; Verhaert & Van den Poel, 2011). The reason lies in the heterogeneity of donors within each segment. Suppose, for example, that a charity clusters all donors with an average donation amount of 100–500 € and solicits them with an appeals scale beginning at 350 €. For a donor with a usual contribution of 100 €, the lowest amount of the appeals scale seems outrageous, because it suggests a contribution 3.5 times larger than her previous donations. Therefore, based on our model predictions, this person's donation likelihood will fall greatly, but the donation amount will be largely unaffected (due to the incongruence of the appeals scale). On the other hand, for a more generous donor used to donating 500 €, the first suggested donation amount is 30% smaller than a usual contribution, which facilitates further donations but also diminishes his contribution.

Suggesting optimized appeals scales to these two extreme donors will increase the donation likelihood of the first while reducing the donation likelihood of the second (as suggested by theory and our models), leaving the aggregate return rate almost unchanged. In terms of expected amounts though, both will improve: the first because the suggested contributions will be more congruent and exert a greater positive impact, the second because the appeals scale becomes more in line with potential and stops pushing generosity downward. The overall effect is an average return rate left unchanged but a higher average donation amount.

4. Conclusions

4.1. Key contribution summary

Researchers have developed the assimilation-and-contrast theory decades ago, and have demonstrated its principles many times in the lab. Yet, despite wide agreement on its reliability, practical applications are lacking. This paper aims at bridging this gap between theory and practice; providing practitioners with a list of sound and research-supported modeling guidelines; and demonstrating the potential financial impact of an effective implementation of the assimilation-and-contrast theory on a charity's performance.

We build on well-established streams of literature to conjecture that multiple moving parts are required to predict—and influence—donors' behavior. We identify seven modeling guidelines: the need to capture heterogenity in donors' IRP location (*P*1) and strength (*P*3); the competitive dynamics of internal and external anchors in influencing donors' behavior (P2); and the moderating role of external anchors' congruency (*P*4), allowed to monotically (P5) and asymmetrically (*P*6) decrease around donors' IRP. Finally, we note that anchor may not only influence behavior amount, but also behavior occurrence, and highlight the need to model that as well (P7).

We then develop a multi-step strategy to put these principles in practice. We (1) collect data suitable to our model calibration on 50,200 donors, (2) develop two econometric models of donation likelihood and donation amount following the seven aforementioned model guidelines, (3) calibrate the models on actual data, (4) use these models to individually tailor 23,500 donation grids such that recommended anchors are expected to maximize charity's revenues, and (5) empirically test the model recommendations. We find that applying these econometric models increased average donation amount by 22% and net margins by 36%.

4.2. Directions for research

This study offers several avenues for further research. First, additional exploratory analyses—not reported here—suggest that many factors, beyond frequency, might contribute to an understanding of donors' IRPs. For example, donors with a very stable pattern of donations (e.g., small median average deviation) seem to have stronger IRPs than donors who demonstrate greater variability in their generosity; they also are less susceptible to appeals scale manipulations. Donors who have given very recently are less susceptible to appeals scale manipulations that attempt to influence their generosity. They likely have a vivid memory of their most recent donation amount, so their IRP is accessible and strong. Even two donors with identical donation histories might have reacted differently to appeals scale manipulations in the past, which could offer a means to attain an individual estimation of the malleability of their IRPs. Others factors, such as the nature of the call or of the charity, donor's involvement, age, gender, or other psychographics might explain how malleable a donor's IRP is. A traditional frequentist approach –such as the one we embraced to facilitate further adoption by the sponsor organization– is fairly unsuited to capturing these phenomena (mainly due to multicollinearity and data scarcity at the individual level). A Bayesian approach, such as a hierarchical model with a Tobit II specification, would offer a solution (Lee & Feinberg, 2012). Our sponsor organization's suspicion toward complex ("black box") models led us to embrace a simpler modeling approach, but we hope the first achievements reported in this paper will pave the way to models that will better account for heterogeneity.

Second, the models we have developed treat donation likelihood and donation amount as two separate decision processes occurring independently. This approach was easier to replicate by the sponsor organization, hence facilitating its further adoption. A discrete-continuous model (e.g., Hanemann, 1984), where both phenomena are modeled jointly, might exploit data more efficiently and improve parameter estimations. Although the financial impact we have reported is unquestionable, it should be seen as a lower bound of its potential. For completeness, we estimated a discrete-continuous model, and it provided strikingly similar results: the fit of the donation model was not improved in any significant way by including the residuals of the likelihood model as independent variable, and the substantive results (i.e., IRP specifications leading to the best model fit) remained unaffected. Since this latter model variation has not been tested in a field study, we do not detail it here.

Third, in our experimental setting, since anchors are presented in increasing order, the first anchor is also the lowest suggested amount, and it is therefore a prime candidate against which donors judge the acceptability of their IRP. If anchors were presented in reverse order, however, where the highest suggested amounts would appear first and then decrease, the psychological entry barrier might be a combination of anchors, such as the first they see (primacy effect) and the last (lowest possible suggested amount). We believe that this point in particular, on what constitutes a psychological entry barrier warrants further investigation.

Fourth, the financial results we have reported are impressive, but it is not clear whether they are sustainable. Many donors exhibit rather stable behavior and remain in the same segments over time. In the context of the studied charity, donors have been exposed to the same traditional appeals scales, year after year. Our manipulations introduced a "shock" that significantly affected donors' behavior, but this shock might not be repeated easily. If the results are not sustainable, charities will benefit from a more subtle laddering strategy that encourages donation amounts to increase incrementally over time (Lee & Feinberg, 2012).

Finally, the modeling guidelines we have summoned from the literature are easily transferable to contexts outside the realm of fundraising. It remains to be seen whether their implementation to different contexts could lead to equally impressive financial results.

4.3. Managerial implications

As predicted by theory, using anchors larger than customers' IRP tends to increase behavior amount, but to decrease its likelihood of occurrence. In our fundraising context, a more aggressive donation grid shifts donor's frame of reference upwards, which increases their contributed amounts in case of donation. At the same time, it creates a stronger psychological entry barrier and makes it harder for potential donors to comply with the charity's request, which negatively affects donation likelihood. Managers rightfully conjecture that there must be an anchors combination that balances these two opposing effects, and maximizes the charity's expected revenues.

While simple in principle, the problem becomes excruciatingly complex in practice, for two reasons. First, an anchor's influence is moderated by its congruence, or believability, which monotically decreases as it departs from a donor's IRP. Since the latter is individual specific, both in terms of location and strength, so is the anchor that maximizes the charity's expected revenues. The same anchor might increase expected contribution for one donor, decrease it for another, and be so unrealistic that it would fail to influence behavior in any meaningful way for a third. To be optimal, anchor manipulation must therefore be individually tailored. Failure to do so likely explains many contradictory findings reported in the literature.

Second, anchor congruence does not decrease symmetrically around customers' IRP (Fig. 3). In fact, in our fundraising application, we have shown that the assimilation-and-contrast theory predictions only hold in the direction of psychological losses, whereas anchors in the direction of psychological gains, however unrealistic they might be, keep influencing behavior.[3] While this asymmetry might be less extreme in other contexts, we can conjecture that an anchor will be deemed incongruent and fail to influence behavior faster in the direction of psychological loss, where compliance is more costly (e.g., giving more money to a charity), than in the direction of psychological gain (e.g., giving less money, or not giving at all).

Unfortunately, what constitutes a gain for the customer often constitutes a loss for the firm. Thus, for a charity, it is far easier to nudge donors in the direction of decreased profit than to nudge them in the direction of increased gains. Consequently, specific knowledge of targeted customers, and careful econometric estimation—as depicted in this research—, are key to improving the effectiveness of the marketing tactics.

Amid recommending caution, this research clearly demonstrates the need to (and feasibility of) optimizing anchor manipulation at the individual level in a fundraising context. We hope these modeling guidelines could be successfully applied in other contexts, with replicable results.

References

Abrahams, M. F., & Bell, R. (1994). Encouraging charitable contributions — An examination of three models of door-in-the-face compliance. *Communication Research*, *21*(2), 131–153.

Adaval, R., & Monroe, K. B. (2002). Automatic construction and use of contextual information for product and price evaluations. *Journal of Consumer Research*, *28*(4), 572–588.

Anderson, R. E. (1973). Consumer dissatisfaction: The effect of disconfirmed expectancy on perceived product performance. *Journal of Marking Research*, *10*, 38–44.

Bell, D. R., & Lattin, J. M. (2000). Looking for loss aversion in scanner panel data: The confounding effect of price response heterogeneity. *Marketing Science*, *19*(2), 185–200.

Berger, P. D., & Smith, G. E. (1997). The effect of direct mail framing strategies and segmentation variables on university fundraising performance. *Journal of Interactive Marketing*, *11*(1), 30–43.

Birnbaum, M. H., & Stegner, S. E. (1979). Source credibility in social judgment: Bias, expertise, and the Judge's point of view. *Journal of Personality and Social Psychology*, *37*(1), 48–74.

Birnbaum, M. H., & Zimmermann, J. M. (1998). Buying and selling prices of investments: Configural weight model of interactions predicts violations in joint independence. *Organizational Behavior and Human Decision Processes*, *74*(2), 145–187.

Birnbaum, M. H., Coffrey, G., Mellers, B. A., & Weiss, R. (1992). Utility measurement: Configural-weight theory and the judge's point of view. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 331–346.

Briesch, R. A., Krishnamurthi, L., Mazumdar, T., & Raj, S. P. (1997). A comparative analysis of reference price models. *Journal of Consumer Research*, *24*(2), 202–214.

Brockner, J., Guzzi, B., Kane, J., Levine, E., & Shaplen, K. (1984). Organizational fundraising: Further evidence on the effect of legitimizing small donations. *Journal of Consumer Research*, *11*(1), 611–614.

Compeau, L. D., & Grewal, D. (1998, Fall). Comparative advertising: An integrative review. *Journal of Public Policy & Marketing*, *17*, 257–273.

De Bruyn, A. (2013). Opening a donor's wallet: The influence of appeal scales on likelihood and magnitude of donation. *Journal of Consumer Psychology*, *23*(4), 496–502.

De Bruyn, A., & Prokopec, S. (2009). In J. -P. Helfer, & J. -L. Nicolas (Eds.), *When asking for more leads to getting nothing: The impact of anchors on donor's behavior*. Nantes: European Marketing Academy.

DeJong, W., & Oopik, A. J. (1992). Effect of legitimizing small contributions and labeling potential donors as helpers on responses to a direct mail solicitation for charity. *Psychological Reports*, *71*(1), 923–928.

Desmet, P. (1999). Asking for less to obtain more. *Journal of Interactive Marketing*, *13*(3), 55–65.

Desmet, P., & Feinberg, F. M. (2003). Ask and ye shall receive: The effects of the appeals scale on consumers' donation behavior. *Journal of Economic Psychology*, *24*(3), 349–376.

Doob, A. N., & McLaughlin, D. S. (1989). Ask and you shall be given: Request size and donations to a good cause. *Journal of Applied Social Psychology*, *19*(12), 1049–1056.

Erdem, T., Mayhew, G., & Sun, B. (2001). Understanding reference-price shoppers: A within- and cross-category analysis. *Journal of Marketing Research*, *38*(4), 445–457.

Fraser, C., Hite, R. E., & Sauer, P. L. (1988, September). Increasing contributions in solicitation campaigns: The use of large and small anchor points. *Journal of Consumer Research*, *15*, 284–287.

Gneezy, A., Gneezy, U., Nelson, L. D., & Brown, A. (2010). Shared social responsibility: A field experiment in pay-what-you-want pricing and charitable giving. *Science*, *329*, 325–327.

[3] Bell and Lattin (2000) question the amplitude of loss aversion usually found in scanner panel data, arguing that loss sensitivity is often overestimated because of the inherent heterogeneity found in the marketplace. When all consumers are exposed to identical anchors, they rightfully reason, the more price-responsive consumers tend to have lower reference prices than the rest of the population, leading to an estimation bias. Though definitely a valid concern with panel data, we note that donors in our sample have been exposed to anchors designed relative to their own individual IRP, hence the heterogeneity issue mentioned by Bell and Lattin has been addressed and cannot, in and by itself, explain our results.

Hanemann, W. M. (1984). Discrete/continuous models of consumer demand. *Econometrica*, *52*, 541–561.
Hardie, B. G. S., Johnson, E. J., & Fader, P. S. (1993). Modeling loss aversion and reference dependence effects on brand choice. *Marketing Science*, *12*(4), 378–394.
Helson, H. (1964). *Adaptation-level theory: An experimental and systematic approach to behavior.* New York: Harper & Row.
Hovland, C. I., Harvey, O. J., & Sherif, M. (1957). Assimilation and contrast effects in reactions to communication and attitude change. *The Journal of Abnormal and Social Psychology*, *55*(2), 244–252.
Janiszewski, C., & Lichtenstein, D. R. (1999). A range theory account of price perception. *Journal of Consumer Research*, *25*(4), 353–368.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.
Kalyanaram, G., & Little, J. D. C. (1994). An empirical analysis of latitude of price acceptance in consumer package goods. *Journal of Consumer Research*, *21*(3), 408–418.
Kalyanaram, G., & Winer, R. S. (1995). Empirical generalizations from reference price research. *Marketing Science*, *14*(Special Issue), 161–169.
Lee, K. Y., & Feinberg, F. M. (2012). Modeling scale attraction effects: An application to charitable donations and optimal laddering. *Working paper*.
Lichtenstein, D. R., Burton, S., & Karson, E. J. (1991). The effect of semantic cues on consumer perceptions of reference price ads. *Journal of Consumer Research*, *18*(3), 380–391.
Mazumdar, T., & Jun, S. Y. (1992, October). Effects of price uncertainty on consumer purchase budget and price thresholds. *Marketing Letters*, *3*(4), 323–329.
Mazumdar, T., & Papatla, P. (2000). An investigation of reference price segments. *Journal of Marketing Research*, *37*(2), 246–258.
Mazumdar, T., Raj, S. P., & Sinha, I. (2005). Reference price research: Review and propositions. *Journal of Marketing*, *69*(4), 84–102.
Nunes, J. C., & Boatwright, P. (2004). Incidental prices and their effect on willingness to pay. *Journal of Marketing Research*, *41*(4), 457–466.
Park, J. H., MacLachlan, D. L., & Love, E. (2011). New product pricing strategy under customer asymmetric anchoring. *International Journal of Research in Marketing*, *28*, 309–318.
Reingen, P. H. (1978). On inducing compliance with requests. *Journal of Consumer Research*, *5*(2), 96–102.
Schibrowsky, J. A., & Peltier, J. E. (1995, Winter). Decision frames and direct marketing offers: A field study in a fundraising context. *Journal of Direct Marketing*, *9*, 8–16.
Schwarz, N., Hippler, H. -J., Deutsch, B., & Strack, F. (1985). Response scales: Effects of category range on reported behavior and comparative judgments. *Public Opinion Quarterly*, *49*(3), 388–395.
Sherif, M. A. (1963). Social categorization as a function of latitude of acceptance and series range. *Journal of Abnormal and Social Psychology*, *67*, 148–156.
Sherif, M., Taub, D., & Hovland, C. (1958). Assimilation and contrast effects of anchoring stimuli on judgments. *Journal of Experimental Psychology*, *55*(2), 150–156.
Thaler, R. (1985). Mental accounting and consumer choice. *Marketing Science*, *4*(3), 199–214.
Urbany, J. E., Bearden, W. O., & Weilbaker, D. C. (1988). The effect of plausible and exaggerated reference prices on consumer perceptions and price search. *Journal of Consumer Research*, *15*, 95–110.
Verhaert, G. A., & Van den Poel, D. (2011). Improving campaign success rate by tailoring donation requests along the donor lifecycle. *Journal of Interactive Marketing*, *25*(1), 51–63.
Weyant, J. M., & Smith, S. L. (1987). Getting more by asking for less: The effects of request size on donations of charity. *Journal of Applied Social Psychology*, *17*(4), 392–400.
Winer, R. S. (1988). Behavioral perspective on pricing: Buyers' subjective perceptions of price revisited. In T. M. Devinney (Ed.), *Issues in pricing: Theory and research* (pp. 35–57). Lexington, MA: Lexington Books.
Xu, A. J., & Wyer, R. S., Jr. (2007). The effect of mind-sets on consumer decision strategies. *Journal of Consumer Research*, *34*, 556–566.