
# ERASMUS SCHOOL OF ECONOMICS

**Bachelor Thesis – International Economics and Business Economics (IBEB)**

# CHARITY SELECTION BASED ON TRUSTWORTHINESS: A DISCRETE CHOICE EXPERIMENT


| | |
|---|---|
| **Name Student** | *Justus Katerberg* |
| **Student Number** | *428677* |
| **Supervisor** | *Prof. Dr. Bas Donkers* |
| **Second Assessor** | *Dr. Arie Barendregt* |
| **Date Final Version** | *21-06-2018* |

## ABSTRACT

This thesis studies the importance of information that is proposed to communicate trustworthiness to mitigate the principal-agent problem and the relative effect of different information on charity choice. The paper starts by conceptualizing cognitive trust and the importance of information in charity marketing. Six hypotheses are developed regarding relative importance of information, both dependent on and independent of experience, which are then tested through a discrete choice experiment. Results showcasing the importance of trustworthiness based on the hypotheses are presented, while additional and equally noteworthy results illustrate the significant effects of information for different demographics. Lastly, marketing implications are presented.


## TABLE OF CONTENTS

# 1. INTRODUCTION

Charities regularly practice direct marketing in an effort to recruit donors. A large share occurs through direct marketing, which involves charities directly approaching decision-makers by making an appeal by presenting information. Charities aim to optimize their direct marketing strategies to generate a sufficient return on investment, which can be achieved by tuning different facets of direct mailings. In the past, several aspects of direct mailings and associated effects on return on investment have been evaluated. These can be divided into two categories: studies focused on the optimal frequency with which direct marketing is practiced and those focused on optimizing presented information. Regarding the former, for example, overall quantity and periodic frequency are two determinants of the size of a charity's donation pool (Piersma & Jonker, 2004). Others found that decision-makers' reactions are more defensive when confronted with direct mailings on a more frequent basis (Diamond & Noble, 2001). Van Diepen, Donkers, and Franses (2009) have extended this analysis by concluding that, at the individual level, stated irritation induced by increased frequency of mailings has no significant effect on actual donations. They continue by suggesting that the content of mailings – which constitutes the second category of studies surrounding charitable direct marketing – may evoke stronger positive feelings that decision-makers are susceptible to. Most certainly, it is true that the type of information provided can have significant influence on response rates and average donations (Smith & Berger, 1996).

In accordance with this suggested importance of optimizing information contained in appeals, charities take considerable effort to communicate trust. Doing so generates a feeling of confidence that funds are spent appropriately, i.e., to actually pursue the cause the charity was founded for. Trust problems are very persistent; charity marketing generally only has the ability to offset a sector-wide decline in trust partly (Gaskin, 1999), which could, for example, be attributed to insufficient self-regulation of scandalized charities (Cordery & Baskerville, 2011). The presence of such issues indicates that trust problems are prominent and can impose consequences on charity success: communicating information effectively is key to a sustainable donor-charity relationship.

Charities are aware of the importance of communicating trustworthiness. Consistent with this statement, they tend to provide more information to donors that show a willingness to receive information than those who have not. Handy (2000) found solicited direct mailings – mail that has been requested by the donor – often provide more information and include more informational indicators of trustworthiness than unsolicited mail. Apparently, when potential donors request mail, they are considered more promising prospects, thereby justifying the extra effort that goes into establishing a trust relationship. Unfortunately, there is a gap in the literature with respect to the effect and relative importance of indicators of trustworthiness in direct marketing, despite the fact that such

knowledge is valuable for creating persuasive pieces of text without mentioning redundant information.

This thesis examines the significance of indicators of trustworthiness through a pair-wise choice experiment. More formally, the experimental design comprises a discrete choice experiment (DCE). DCEs have been applied to reveal consumer preferences in a wide array of areas, including product development of clean-fuel vehicles (Ewing & Sarigöllü, 2000), primary care consultations (Cheraghi-Sohi, et al., 2008), and healthcare products (Ryan, Bate, Eastmond, & Ludbrook, 2001). To my knowledge, this is the first time a DCE is used to elicit consumer preferences with respect to charities and, more specifically, trustworthiness information in charity marketing. By pioneering the use of DCEs to investigate the desirability of particular types of information, charities can develop a more thorough understanding of which content should be integral to their direct marketing policy. Ergo, the primary goal of this thesis is to minimize the holdup that trust issues cause for value creation in the market for charitable donations. It does so by providing an indication of which information facilitates this.

The thesis continues by stipulating the theoretical framework, after which a research question and hypotheses are formulated. After that, the employed methodology is explained and acquired data is described. At the end, results and respective implications for marketing strategy are discussed, after which the paper is concluded with limitations and recommendations.

## 2. THEORETICAL FRAMEWORK

### 2.1. The Agency Problem

The most convenient way of modeling a trust problem in the donor-charity interaction is a principal-agent problem, as per Ross (1973), for which utility of the principal depends on the actions of the agent. In this case, a problem arises once the utility of the agent comes from different courses of action than those which the principal prefers. In the agency relationship between donors and charities, the 'principal' is the donor, while the charity represents the 'agent'. A relationship as such is unique, because, unlike a classical corporate principal-agent problem, stakeholders cannot align the incentive scheme of the agent, so that the latter is encouraged to comply with the former's wishes, through an incentive scheme. Once a donation has been made, the donor surrenders all forms of control to the charity. As a result, the donor risks the possibility of an imperfect alignment of incentives with defalcation being the ultimate adverse consequence.

The problem specific to the non-profit sector can be evaluated in conjunction with dynamic game theory by considering the charity a player who has a binary choice. After a donor has contributed a

donation, the charity can cooperate by designating the money for the intended cause or defect by using the money to enhance an employee's private benefit beyond agreed upon wage benefits. The donor, on the other hand, can sequentially punish a charity by not offering donations on any upcoming occasions, granted he or she possesses perfect information on charity activities. This is where the problem may arise. Charities have an information advantage warranted by the fact they do not publicize all activities they undertake – especially non-ethical ones, quite logically. In line with this proposition, a few scandals can have dramatic spill-over effects to trust attitudes towards the entire voluntary sector (Gibelman & Gelman, 2012). In other words, such adverse events lead perceived risk of donating to increase.

In addition to potential scandalization of the charity-donor trust relationship, a donor engages in another agency relationship resulting from the donation. Donated funds are allocated to vulnerable members of society or other noble causes alike. However, since the donated funds are paid to different organizations and individuals, the donor implicitly trusts the recipient to allocate the resources efficiently as well (Katz, 2000). This second agency relationship renders the charity as the mediating factor between the donor and the final recipient. Consequently, the stability of the donor-charity agency relationship is not merely a question of ethics, but also of ability. The charity is expected to, in turn, be able to find reliable candidates for receipt of funds from the charity. Given the undeniable weight donor trust carries, the long-run equilibrium has donors converging to the charity that can provide most of it. Of course, this is a *ceteris paribus* prediction, but it does intuitively illustrate the information problem at hand.

### 2.2. Trust: A Mediating Factor

Additionally, it is of great importance to clearly define the type of trust relationship that is studied. Johnson and Grayson (2005) distinguish between cognitive trust, which means consciously being reliant on a third party's ability to complete a task, and affective trust, which implies relying on a third party through emotional attachment. In the context of the non-profit sector, cognitive trust is driven by knowledge on a charity and its intentions with acquired funds. On the contrary, affective trust could be driven by emotional appeals, such as guilt appeals (Chang, 2011). Regarding this thesis, the setting surrounds cognitive trust, since the experiment provides hypothetical information without the charity's name, brand, reputation, or prior relation with the donor. The indicators in the design, as the section on methodology will illustrate, are objective. One might argue that such a focus on cognitive trust is myopic, for it does not reflect the trust formation process in its entirety. This claim has substance, since charity marketing may contain particular elements that elicit effective emotional reactions, as per Burt and Strongman (2004). Nevertheless, Lewis and Weigert (1985) help to refute this concern by stating cognitive trust is an initial requirement for affective trust to develop. In a different paper, Johnson and Grayson (2005) reiterate this by finding a significant relationship from cognitive to

affective trust. With respect to this thesis, the conducted experiment is exclusively focused on testing relative importance of non-qualitative information. Because it is unlikely affective associations are made as a result, cognitive trust, based on the aforementioned definition, is tested in isolation.

### 2.3. Marketing in the Non-Profit Sector

All in all, given the paragraphs above, trust is a mediator that drives donations from the agent to the principal. In order to establish a trust relationship, it is imperative charities center their strategies on providing information that proves their trustworthiness. As Handy (2000) confirms, they try to accomplish this by mentioning factual information or what I commonly refer to as 'indicators of trustworthiness'. For the purpose of the rest of the paper, an indicator of trustworthiness is defined as a piece of factual information in a direct mailing from an organization with the ultimate purpose being that the recipient trusts the organization. As such, trust resulting from information moderates the adverse implications of the principal-agent problem.

Sargeant (2010) models the behavioral process surrounding charity appeals as an interaction between charity inputs and perceptual reactions from donors. Factual information, in this model, is one of the inputs which aim to elicit such reaction, which can take the form of a positive portrayal of the charity or increased perceived fit of the charities' values with one's own values. Whereas Sargeant himself does not explicitly relate such reactions to the trust relationship, Morgan and Hunt (1994) do integrate trust into their key mediator variable (KMV) model by labeling it as the product of shared values and frequent communication. Charities should apparently seek prevalent values in their target groups and explicitly communicate how the charity conforms to these through information. This suggests indicators of trustworthiness can play an important role in satisfying the informational needs of a charity's target group. These needs are analogous to the importance attached to different pieces of information provided in the DCE.

### 2.4. Providing Relevant Information

Lastly, for charitable direct marketing, a focus on reporting relevant information is important. To persuade prospective donors, it is helpful to elaborate on points they value and provide less relevant information in a more simplified fashion. This prevents a possible information overload for the recipient (Mitchell & Papavassiliou, 1999). Since information load is related to decision accuracy, it is vital that organizations focus on providing an optimal amount of information (Eppler & Mengis, 2004). All in all, return on investment can be optimized by preventing the possibility of information overload.

## 3. RESEARCH QUESTION AND FORMULATED HYPOTHESES

Because the thesis examines indicators of trustworthiness and resulting donating intention, the research question is the following:

*What information from charities do potential donors value with respect to different communicable indicators of trustworthiness?*

I formulate hypotheses that will treat indicators of trustworthiness individually. These indicators have been extracted from Handy (2000). A requirement for selection is that the indicator is reliably quantifiable in a DCE. Furthermore, it has to be specifically reported as a communicable characteristic that can moderate the principal-agent problem.

First of all, charities may be deemed more reliable if they have existed for a longer period of time. Although, Handy (2000) found solicited mail contains a message about longevity significantly more often than unsolicited mail, literature delving into the importance of longevity for charity choice is scarce. If one were to draw an analogy to other consumer situations, it is evident that longevity of an institution communicates the fact that it has not publicly failed to conform to the wishes of the principal in the agency relationship. This effect is likely even stronger when longevity is specifically emphasized in a direct mailing, for example. Therefore, the first hypothesis is as follows:

*$H_1$: Information on increased longevity increases charity preference.*

The second indicator of trustworthiness surrounds the percentage of the donation pool that is spent on overhead costs or, phrased more omni-inclusively, all expenditures made not in direct pursuit of the mission of the charity. Explicitly stating an acceptable percentage can increase perceived efficiency of the charity and, possibly, trust in an ethical allocation of resources. This reasoning makes sense logically, but only a relatively small percentage of donors stated overhead costs represent a major consideration for them (Iwaarden, Wiele, Williams, & Moxham, 2009). Then again, donors have been found to donate significantly more when it was explicitly mentioned that overhead costs had already been covered for by previous donations (Gneezy, Keenan, & Gneezy, 2014). This statement is intriguing, because, when analyzing this idea based on rational choice, the donation pool is the same regardless of whether this is mentioned. Apparently, people have been found to express they are insensitive to an increase in the *perceived* price of donating, even when they may behave in opposite fashion. Bowman (2006) confirms that a higher percentage of overhead costs negatively affects donations, but that other considerations might be substantially more important. In this article, it is unclear whether this is due to a higher perceived price of donating or a more pessimistic appraisal of

ethical resource allocation. In response to the unclarities above, it is necessary to study whether the overhead percentage is interpreted as a cost imposed on the donor and/or whether reporting it communicates trust. Therefore, it is imperative to disentangle the trust effect by comparing situations when the amount is not stated at all to when different percentages are reported. If people, indeed, do trust a charity that reports an overhead percentage and value efficiency at the same time, the following hypotheses should hold:

*$H_{2a}$: Information on a higher percentage of overhead costs decreases charity preference.*
*$H_{2b}$: Information on a sector-average percentage of overhead costs increases charity preference compared to omitting such information.*

Moving on, another possible indicator of trustworthiness is the amount of funding received from the government. Referring back to the principal-agent problem, when there is information asymmetry, principals wish to remove this imbalance. Because governments can be expected to have more insight into the activities of charities, communicating the fact that the charity is receiving subsidies stresses the charity's reliability. Indirectly, government funding reduces the information imbalance, as governments may be expected to have the ability to see to it that funds are used ethically. In prior literature, however, government funding has mainly been inspected to prove the existence of impure altruism (Andreoni & Payne, 2003; 2011). This was studied by proving the existence of imperfect crowd-out by studying donations subject to different levels of government spending. In direct marketing, nevertheless, the sole purpose of mentioning government support would be to establish trust. With any crowd-out whatsoever, making possible donors aware of government subsidies is counter-productive. In fact, if increased government spending is indeed a reason to donate more, this entails not only the absence of perfect crowd-out, but a negative crowd-out rate. Of course, this would be very context-dependent, but, in the context of charity marketing, solicited mail actually does contain information on government spending significantly more frequently than unsolicited mail (Handy, 2000). Since government spending is apparently used as a signal of reliability by charities, the hypotheses testing this aspect are as follows:

*$H_{3a}$: Information on higher government funding increases charity preference.*
*$H_{3b}$: Information on a sector-average of government funding increases charity preference compared to omitting such information.*

Accreditation of activities by third parties, as stressed in Bekkers (2003), is essential to promote trust. Agents that suffer from asymmetric information can be assisted by third parties. While the government can moderate information asymmetry, its own reliability is, in turn, determined by a range of political matters (Chanley, Rudolph, & Rahn, 2000). It is likely that, because of this, skepticism is much higher

towards government spending than towards an independent consumer interest group whose reputation is exclusively dependent on developments within the organization. Because of the fact that such monitoring can be beneficial for decision-makers in streamlining the decision, the two following hypotheses are formulated:

*$H_{4a}$: Information on accreditation of activities by an external consumer interest group increases charity preference.*
*$H_{4b}$: Information on accreditation of activities by an external consumer interest group is more important than reporting a sector average of government funding.*

Regarding the relative importance of trustworthiness factors, it is vital to distinguish between different levels of experience with charitable marketing. If more severe implications regarding required information present itself for more dedicated donors, for example, charities' direct marketing should adapt to facilitate a more intensive trust relationship with this group. Obermiller and Spangenberg (2008) recognize the effect of age on consumer skepticism towards advertising, which they claim to be the result of greater experience with manipulative intent. Because such skepticism can negatively alter attitudes towards advertising (Cotte, Coulter, & Moore, 2005; Hibbert, Smith, Davies, & Ireland, 2007), it is reasonable to think that sufficient information on trustworthiness should be provided in order to mitigate concerns. Therefore, the fifth hypothesis is as stated below:

*$H_5$: More frequent exposure to direct mailings increases importance of at least one of the proposed indicators of trustworthiness relative to information that is not of this kind.*

The above hypothesis defines experience with charitable marketing as the result of unsolicited exposure to marketing. This ignores that exposure is not solely dependent on the effort the charity invests in approaching its (prospective) donors. Some donors value donations more and, thus, experience can be a consequence of preference, i.e., donors' experience increases as a result of enhanced familiarity with the actual product, consistent with Johnson and Russo (1984). Since such experience may pertain to the same mechanism that is relevant for $H_5$, the sixth hypothesis is the following:

*$H_6$: More frequent donations increase importance of at least one of the proposed indicators of trustworthiness relative to information that is not of this kind.*

Because of practical constraints regarding the collectible number of participants, no hypotheses were developed regarding socio-demographic influences, as I desired to run the same model for each

hypothesis. However, significant effects of socio-demographic characteristics are reported in the section '6.3. Socio-Demographics'.

# 4. METHODOLOGY

## 4.1. Discrete Choice Experiment

In order to test the aforementioned hypotheses, a discrete choice experiment (DCE) is used. A DCE is one of the most practical applications of modern-day consumer choice theory. Underlying such theory is the assumption that consumers gain utility from an alternative in a choice set because of its attributes (Fader & Hardie, 1996). They do not gain utility from the concept in its entirety; they do so from the levels of the attributes encompassed in the concept. This distinction is important, because it allows utility profiles to be portrayed as an additive function of the utility from each individual attribute. Attributes of a charity $j$ are summarized in a vector $x_j$ with a preference vector $\beta$ with an equal number of factors, thereby weighing each attribute against preference through a linear combination:

$$(\mathbf{1}) \quad U_j = x_j \times \beta, with\ x_j = \{x_1, x_2, \dots, x_J\}\ and\ \beta = \{\beta_1, \beta_2, \dots, \beta_J\}.$$

In a DCE, the participant's task is straightforward: she has to indicate a preference for any of the alternatives or profiles – used interchangeably hereafter – within a choice set. All included alternatives have their attributes in common, while each attribute's levels may vary. By having a sample of people answer a survey on multiple choice sets, average preference for each attribute is revealed. In the case of this thesis, choice sets include two profiles of charities with varying levels of the aforementioned indicators. Formally, alternative $j$ will be chosen over the alternative $j'$ when $U_{js} + \varepsilon_{js} > U_{j's} + \varepsilon_{j's}$. In this inequality, the error terms account for measurement error. This entails that one cannot always incontestably conclude which alternative will be selected. Nonetheless, the probability that an alternative is chosen can be determined if an assumption is made regarding the distribution of its error term. More formally, consistent with Lancsar and Louviere (2008), the probability of choosing a charity $j$ over a different charity $j'$ within choice set $s$:

$$(\mathbf{2}) \quad P_{js} = P\left(U_{js} + \varepsilon_{js} > U_{j's} + \varepsilon_{j's}\right)$$
$$= P(U_{js} - U_{j's} > \varepsilon_{j's} - \varepsilon_{js})$$
$$(\mathbf{3}) \quad = F_{\varepsilon_{j's} - \varepsilon_{js}}(U_{js} - U_{j's}).$$

The transition from (2) to (3) shows that the probability that the utility differential between the alternatives $j$ and $j'$ exceeds the difference in error terms is described along a cumulative distribution

function. The transition from (3) to (4), is depicted below, can be made under the assumption that the error terms are distributed independently and identically along a Gumbel distribution. This is, because the difference between two variables with such a distribution is logistically distributed. As two charities are evaluated per choice set, the probability that charity $j$ is chosen follows a binomial logistic distribution:

$$(\mathbf{4}) \quad P_{js} = \frac{e^{U_{js}}}{e^{U_{js}} + e^{U_{j's}}}.$$

This probability depends on $U_{js}$ and $U_{j's}$ and increases when the former increases but decreases when the latter increases. Therefore, a DCE revolves around discovering the average utility of each of the alternatives by examining the choice patterns and choice probabilities that people show.

Following the derivation of the appropriate probability distribution as in Equation 4, a fractional factorial choice design is generated. A design as such uses only a fraction of the possible alternatives from the full factorial design. By specifying prior estimates of the preference distribution, utility balance can be attained (Huber & Zwerina, 1996). This implies choice designs are generated that are expected to give participants similar utility a priori. Doing so optimizes the amount of information each choice gives, because dominated alternatives – alternatives that are inferior regarding all attributes in a choice set – and alternatives that are extremely unlikely to be selected are excluded from the choice design. It uses prior estimates for the distribution of the preference vector, which is assumed to follow a multivariate Normal distribution $N(\beta_0, \Sigma_0)$. Here, $\beta_0$ represents the prior mean of the preference distribution, whereas $\Sigma_0$ denotes the prior variance-covariance matrix. The latter indicates how the distribution is expected to vary across the population and how preferences over one attribute co-vary with a different one.

In *Appendix B,* the prior mean estimates are specified. Based on the hypothesized relationships, superior levels are assigned a prior mean of 1, while inferior levels are assigned a prior mean of -1. The final estimates of main effects will be generated based on effects-coding, which is why the priors add up to zero. Effects-coding is useful, because the final estimates of the coefficients show deviation from the "average level" of that attribute. More formally, the reference or base level is coded as -1, while this would be 0 for dummy coding (Bech & Gyrd- Hansen, 2005). To compute the utility differential between a particular level of interest and a reference level, the coefficient on the reference level should be multiplied with -1 and added to the coefficient on the level of interest. Incidentally, this is also how coefficients would be interpreted if one were to use dummy coding: coefficients show the deviation from a reference level. Lastly, the prior specification of the variance-covariance matrix is an identity matrix, because there are no prior studies that allow me to be more certain of particular

prior estimates. Also, no substantial reason has been identified to reliably assume that the values will co-vary one way or another.

4.2. Experimental Design

The DCE is assigned to participants in an online setting. All are asked to choose between two different sets of information provided by the charity with attributes and levels as described in *Appendix A*. This procedure is repeated for ten different choice sets. To limit task complexity, each participant is asked to review a random subset, comprising ten choice sets, from the sixteen provided in *Appendix A*. Caussade et al (2005) have found that, indeed, ten choice sets is optimal with respect to the inevitable trade-off between the amount and reliability of acquired information. The latter can be jeopardized by effects from continuing exposure to abstract choices, such as tiredness and boredom (see '4.3. Reliability and Validity'). Furthermore, varying the number of levels per attribute did not cause significant changes in the variance of responses in the heteroskedastic model they estimated.

Moreover, each alternative consists of seven different attributes (*Appendix B*). One attribute, 'Sector', holds the same level throughout the experiment and merely serves the purpose of providing context. Given the fixed level of this attribute, it can be computed that the full factorial design has a total of 324 alternatives ($3^4 \times 2^2$). For these alternatives, the numeric values of the levels of 'Donation Pool', 'Overhead Costs', and 'Money from Government' have been adapted from a branch report to preserve realism (Goede Doelen Nederland, 2016). This report is based on data supplied by the Dutch Centraal Bureau Fondsenwerving (CBF), an independent consumer organization that certifies charities for ethical behavior. The reasons for selecting the levels are shown in *Appendix A* in the rightmost column.

Below, each attribute and its relationship with the hypotheses in section '3. Research Question and Formulated Hypotheses' is explained:

- 'Founded in': This attribute shows in which year the charity was founded, thereby testing the effect of longevity on charity preference.
- 'Donation Pool': This attribute shows the size of the aggregate donation pool a charity receives from their donors on a yearly basis on average. It acts as a benchmark attribute to limit effects arising from contingency awareness, provide context and provide a benchmark for information on trustworthiness.
- 'Overhead Costs': This attribute shows the percentage of the aggregate donation pool that gets allocated to resources other than any that directly pursue the cause the charity stands for, thereby testing whether efficiency and transparency signal trustworthiness.

- ‘Money from Government’: This attribute shows the amount of financial aid the charity receives from the government on a yearly basis on average, thereby testing whether reporting higher government spending signals trustworthiness.
- ‘Method of Collection’: This attribute shows the method in which donations will be collected as announced by the charity. It acts as a benchmark attribute to limit effects arising from contingency awareness, provide context and provide a benchmark for information on trustworthiness.
- ‘CBF label of Quality’: This attribute shows whether a charity has received official accreditation for ethical behavior, thereby testing whether accreditation by a third party is important to consumers in a trust relationship.

Using these attributes, combined with the prior means as in *Appendix A* and the aforementioned identity matrix as variance-covariance matrix, the choice design provided in *Appendix B* is generated. In order to study segment-specific preferences, questions related to socio-demographic attributes and donating behavior are posed to participants. Key socio-demographic variables include participants’ gender, age, and education, all of which have been found to be predictors of charitable giving (Bekkers & Wiepking, 2011; 2012) (see *Appendix C*). In addition, donating behavior is described by familiarity with direct mails, yearly donation amount and donation frequency as in *Appendix D*. Donation behavior likely affects the way in which information about a charity is perceived. As an example, greater experience with marketing stimuli leads to increased skepticism of the techniques used (Hibbert, Smith, Davies, & Ireland, 2007). In addition, trust attitudes differ between people who are experienced donors and people who have never donated or do so rarely (Sargeant & Lee, 2002).

### 4.3. Reliability and Validity

This section treats reliability and validity of the used methodological approach.

#### 4.3.1. Reliability

For any experiment, reliability is of high priority, which may be compromised if measurement error is inconsistent. Obtrusiveness and scrutiny are at the forefront when it comes to risks that make true behavior deviate from stated behavior during an experiment. Fortunately, the nature of the experiment is not sensitive and the survey is completely anonymous. Additionally, incentives in pair-wise choice experiments do not significantly alter results (Beattie & Loomes, 1997). This means reliability is not likely to be hampered by the lack of a monetary incentive. Also, task complexity is a very relevant concern with DCEs, which is influenced by the number of attributes included in the profiles. With regards to the DCE in this thesis, there are six varying attributes that define the information supplied by a charity. These attributes are consistently reported in the same order allowing for side-by-side

comparison. Despite the fact this refutes most concerns regarding the measurements' reliability, the results of a pilot study are described in section '5.1. Pilot Study'.

Another valid concern regarding the reliability of results from a DCE are order effects, i.e., response behavior dependent on the order in which treatment conditions are administered during an experiment. For example, at later stages in an experiment, participants will exert less effort to answer questions in comparison to early in an experiment (Galesic & Bosnjak, 2009). Despite the fact that the experiment, as stated before, is not very complex, randomization is applied to the block of choice sets, so that fatigue/boredom effects are limited. On top of this, I randomize the order in which the alternatives per choice set are presented. In spite of the fact that each set only encompasses two alternatives, people have been found to show an unconscious preference for the first answer to multiple choice questions (Biswas, Labcrecque, Lehmann, & Markos, 2014). Therefore, the order in which the alternatives are presented is randomized as well.

#### 4.3.2. Validity

Regarding the context, it may be argued that repeated pair-wise comparison of information cannot be generalized to a setting outside of the experiment. Moreover, as, for example, direct mailings contain more information than what is distributed to participants in the experiment, this obviously does not mirror reality. Even though, I acknowledge the fact that no participants will encounter a setting like this, I would also like to emphasize the thesis' focus on cognitive trust, which is knowledge-based. As a result, the factual nature of the information provided in the experiment reveals that which potential donors deem valuable for trusting charities through cognitive processing. On top of this, DCEs yield externally valid results in other sectors (Telser & Zweifel, 2007), in spite of the lack of context engrained in its practice. Lastly, as mentioned in the 'Experimental Design' section, the majority of attribute levels have been extracted from the 2016 branch report in order to simulate realism.

## 5. DATA

### 5.1. Pilot Study

To further mitigate any remaining concerns regarding task complexity and potential obtrusiveness, a small pilot study was conducted, comprising eight participants ($M_{age}$ = 34.5 years, $SD_{age}$ = 18.21 years, female = 62.5%). To warrant a pilot sample somewhat representative of the eventual sample, participants from each of the education levels, except for 'No Schooling', were included through a judgmental sampling method.

5.2. Final Study

The final version of the online experiment was distributed to 103 participants ($M_{age}$ = 36.97 years, $SD_{age}$ = 17.44, female = 58.25%), all of whom were living in the Netherlands at the time. This sample size meets the minimum requirement $N \geq \frac{500c}{t \times a}$ for estimating main effects in choice experiments, according to Orme (1998). In this inequality, $c$ denotes the highest number of levels for a single attribute, $t$ the number of choice tasks and $a$ the number of possible choices. Since $c$ = 3, $t$ = 10, and $a$ = 2, this implies$N \geq 75$.

To prevent the results from becoming confounded through learning effects, participants in the pilot study were excluded from participation. A quota sampling technique was used to accumulate a sample that is evenly distributed over the three socio-demographic variables in the study. Although, admittedly, probability sampling allows for theoretically sound generalizations, quota sampling can yield sampling variance close to probability sampling (Mose & Stuart, 1953). *Appendix F* reports the targeted quotas for each of the demographic indicators and the corresponding sample characteristics.

The categorical variables in the study were coded as in Table 2, which explains the coded values of the interaction terms in the appendices. Also, the table shows that, for 'Mailing Frequency' and 'Donation Frequency' categories were merged to alleviate sample size constraints. Lastly, this table also shows the sample composition per variable category.

| Variable | Participant's Answer | Coding | % of Participants |
|---|---|---|---|
| **Gender** | Male | 1 | 41.75 |
| | Female | 0 | 58.25 |
| **Education Level** | University/vwo | 1 | 27.18 |
| | Applied Sciences/havo | 2 | 38.33 |
| | Practical Education/mavo | 3 | 33.98 |
| **Mailing Frequency** | Daily<br>Weekly | 1 | 20.39 |
| | Monthly | 2 | 21.36 |
| | Less than Monthly<br>Never | 3 | 54.37 |
| **Donation Frequency** | Weekly<br>Monthly | 1 | 37.86 |
| | Quarterly | 2 | 27.18 |
| | Yearly<br>Never | 3 | 31.07 |

***Table 2**: Categorization and coding process of categorical variables.*

# 6. RESULTS

In order to answer the research question, the hypotheses are studied, different models are estimated to verify the robustness of obtained results, and results based on socio-demographic characteristics are reported. First, a short discussion of the results from the pilot study concerning task complexity is provided.

## 6.1. Pilot Study

After they completed their participation in the experiment, participants from the pilot study were tasked with rating four statements on a 5-point Likert scale, as per *Appendix E*. Participants were given the option to report any motivations they had for the allotted number of points. Table 1 shows the average scores with standard deviations for each of the questions asked.

| | 1. Clear Expectations | 2. Appropriate Length | 3. Sufficient Information | 4. Invasive Questions |
|---|---|---|---|---|
| **Average** | *4.50* | *4.13* | *4.25* | *1.63* |
| **Standard Deviation** | *0.53* | *0.35* | *0.71* | *0.74* |

***Table 1:*** *Descriptive statistics of the pilot study.*

The Likert scale was interpreted as an interval scale from 5 ('Fully Agree') to 1 ('Fully Disagree'). The first three constructs limit task complexity; each of these was rated in excess of 4.0 on average. The final construct, on the other hand, may increase perceived obtrusiveness for participants, but no participant expressed any noticeable concerns regarding this matter.

Lastly, participants notified me of several practical errors with the survey regarding phrasing of questions/statements and the manner in which alternatives were displayed. These issues were resolved before the final questionnaire rolled out.

## 6.2. Hypotheses

For testing each hypothesis, a consistent model is used, so that none of the results are subject to the addition of supplementary explanatory parameters in the model. All attributes are included as profile effects, with mailing and donation frequency being subject effects to distinguish between different levels of experience. Furthermore, it is important to note that all reported *p*-values have been computed using likelihood ratio tests. Here, significance implies that the difference in parameter estimates between different levels of the same attribute can be meaningfully interpreted. These differences show the change in utility when moving from one level to the other. Finally, all results reported for testing the hypotheses are in reference to *Appendix G*.

The first hypothesis concerns information on the longevity of a charity. The various levels of longevity ($p < 0.05$) follow the hypothesized order when examining the effect marginals. Therefore, $H_1$ is supported. Older charities are generally preferred, but the relationship seems to be non-linear. The utility differential between 2010 and 1970 far exceeds the utility differential between 1970 and 1930. Apparently, there seems to be a certain longevity after which additional years of existence do not contribute to relative trust.

The second pair of hypotheses concerns the importance of information on overhead costs ($p < 0.0001$). Unsurprisingly, $H_{2a}$ is supported, as an overhead percentage of 12% yields negative utility, while 6% yields positive utility. $H_{2b}$ is also supported. The experiment used 12% as the highest value, which is even above the sector average of 9%, but this is still preferred to not disclosing information regarding overhead percentages. To emphasize the size of the effect, the marginal probability of choosing an alternative with '12%' doubles in comparison to 'Not Mentioned' *ceteris paribus*. This suggests there is a preference for transparency with respect to this matter, regardless of the cost percentage.

The third set of hypotheses addresses information on government funding ($p < 0.0001$). In contrast to the previously introduced notion of crowd-out, $H_{3a}$ is supported, implying higher levels of government funding are preferred to lower levels. Since the highest percentage of 30% is most preferred and exceeds the sector's average of 24%, it is unlikely that crowding out is an issue in direct marketing in the non-profit sector. As for $H_{3b}$, which concerns the fact that charity preference decreases when omitting information government funding compared to reporting a sector average, this hypothesis is also supported. Charity preference decreases more when omitting information on government funding than when a lower than average level of government funding (15%) is reported. The findings hint at the existence of a trust effect of reported government subsidies and refutes concerns of crowd-out when reporting above average amounts of public funding.

The fourth hypothesis addresses the importance of independent accreditation ($p < 0.0001$). Unsurprisingly, $H_{4a}$ is supported. More interesting is the support found consistent with $H_{4b}$, which compares the importance of average government accreditation to independent CBF accreditation. The utility range between independent accreditation and the absence of such information was compared to the utility range between the average level of government support – in the form of subsidies – and the absence of any such information. As the sector-average level of governmental support was not included in the choice design, the range between highest level of government support and no reported support was used to warrant robustness. The difference in range between the highest and lowest level is almost 0.25 higher in favor of accreditation by CBF. This evidence confirms $H_{4b}$.

For $H_5$, concerned with mailing frequency, the significance of interaction terms is studied. Out of these, only the interaction between mailing frequency and independent accreditation ($p < 0.05$) is significant. Participants who have been have been exposed on a more frequent basis show a relatively higher adverse reaction to omitting information on independent accreditation by CBF in comparison to people who receive email less frequently. For this to be evidence in support of the hypothesis, the utility range of indicators of benchmark attributes ('Donation Pool' and 'Method of Collection') should a) not be significant and, if it is, b) have a smaller range than that of the interaction term between CBF accreditation and highest and lowest level mailing frequency. The interaction with 'Donation Pool' is insignificant, but the one with 'Method of Collection' ($p < 0.05$) is significant. Therefore, the utility range of the latter variable's interaction with mailing frequency should be compared, so as to study whether the utility range of the interaction between mailing frequency and CBF accreditation is indeed greater. Since this is the case (0.443 > 0.273), CBF accreditation becomes relatively more important for people exposed to greater mailing frequency in comparison to both 'Method of Collection' and 'Donation Pool'. To clarify this result, utility ranges represent the range between the most and least preferred level of an attribute. This illustrates how important the attribute is, subject of course to the best and worst levels included in the experimental design. For interactions, utility ranges show how much more or less utility a participant gets from various levels of a subject effect. Ergo, utility ranges of coefficients on interaction terms – between one subject effect level and different attribute levels – show how much more or less important an attribute is for a person for the relevant subject effect level. Therefore, support is found of increased relative importance of at least one proposed indicator of trustworthiness resulting from higher mailing frequency. Lastly, it is noteworthy that the interaction between overhead costs and mailing frequency ($p < 0.10$) is weakly significant. However, the relationship is such that increased mailing frequency leads to higher utility of the level 'Not Mentioned' for people who receive mailings 'more than weekly' in comparison to 'monthly' mail recipients. This seems to contradict the notion of an increase in required trustworthiness, but is not adequately significant to re-evaluate the conclusion drawn regarding $H_5$.

$H_6$, which focuses on donation frequency, finds no support in the model. The interaction between this factor and the attribute 'Method of Collection' ($p < 0.10$) is weakly significant, but this is not a proposed indicator of trustworthiness. Incidentally, the stand-alone effect of 'Method of Collection' is not significant at all, so it seems that people who donate more might care relatively more about practicalities surrounding collection than trustworthiness. This suggestion would make sense, but indicates donation frequency is not a determinant of increased relative importance of trustworthiness indicators.

Besides the hypotheses, this model, in its particular setting, revealed that the order of importance is independent accreditation (by CBF), overhead costs, government accreditation, and longevity.

6.3. Robustness checks

Because the above conclusions are subject to the variables incorporated, I, first of all, estimate a model neglecting any subject effects, as per *Appendix H*. The signs and significance of all four sets of hypotheses are the same and, in fact, effect marginal estimates are remarkably similar. As a result, no reason is found to argue against the first four hypotheses.

Furthermore, I estimate two models that separate the subject effects. The main motivation is a Pearson correlation coefficient of approximately 0.352 between the participants' exposure to mailing frequency and their expressed donation frequency. This could have led to inaccuracy, as the causal effects may not be appropriately disentangled by the model. The first model, as in *Appendix I*, exclusively incorporates mailing frequency as a subject effect. In this version, the interaction with CBF accreditation ($p < 0.05$) remains significant, but a major distinction from the version used to test the hypotheses is the significance of the interaction between mailing frequency and percentage of overhead costs. This points to additional importance attached to overhead costs subject to different levels of mailing frequency when ignoring donor frequency. Nevertheless, there is no consistent relationship between increased mailing frequency and different levels of overhead costs, as the sign is positive for some and negative for other interaction terms. There is no substantial reason to challenge the support of $H_5$ – because of the inconsistency in the relationship – but it could raise questions regarding the evaluation of overhead costs for different levels of mailing frequency. Moving on, the second model, as in *Appendix J*, uses donation frequency as the sole subject effect. No different conclusions regarding any of the hypotheses can be drawn, as significance and sign of the utility estimates remains unchanged. Lastly, neither model challenges the evidence found in support of $H_1$ to $H_4$.

To check the final claim from '6.1. Hypotheses' – regarding the order of importance of trustworthiness indicators – an output grid with different utility profiles of the model in *Appendix H* was generated. This table illustrated that, when sorted by utility in descending order, the top profile maintained the most positive value of each of the most important attributes for longest. To illustrate, the 38$^{th}$ alternative from this grid was the first not to have a CBF label of quality, while the 2$^{nd}$ alternative already showed a sub-optimal level for 'Longevity'.

6.4. Socio-demographics

Beyond the hypotheses, it is meaningful to observe the influence of demographic variables on the parameter estimates, as per *Appendix K*. First of all, the main effect marginals of all attributes remain

similar. The coefficients on overhead costs and CBF accreditation preserve significance, while government funding ($p = 0.0563 < 0.10$) is almost significant and longevity loses significance.

6.4.1. Age

All interactions between participants' age and hypothesized indicators of trustworthiness are significant at 1%, while none of the interactions with the benchmark attributes are significant. This implies that age has an effect on perception of information on trustworthiness. What is interesting, is that age strengthens the effect of each of the trustworthiness indicators, with the signs on the interaction terms between age and each attribute variable exactly in accordance with those of the main marginal effect estimators. This implies that older people a) attach more value to information on trustworthiness, b) are able to more accurately determine what shows trustworthiness, or c) both.

6.4.2. Gender

The interactions between gender and the method of collection ($p < 0.0001$) and between gender and the importance of a CBF label of quality ($p < 0.05$) are significant. For women in the sample, accreditation by CBF is significantly more important than it is for men. Since accreditation by third parties is more impactful for decisions than government backing, according to evidence in support of $H_4$, women supposedly require more information on trustworthiness.

6.4.3. Education

With respect to hypothesized indicators of trustworthiness, interactions between education and longevity ($p < 0.05$), overhead costs ($p < 0.01$), and CBF label of quality ($p < 0.01$) are all significant. For longevity, it seems participants that are higher educated generally have a greater dislike of the charities founded in 2010, whereas there is only a slight change for 1970. In addition, there is a clear negative preference for omitting information on overhead costs among participants that are higher educated. Lastly, CBF accreditation is also more important to higher educated than lower educated people. When evaluating these three statements regarding trustworthiness, it is straightforward to conclude that better educated people attach greater utility to them.

To verify the validity of this last statement, I estimated three more models, one for each level of education separately. These are summarized in *Appendix L*. For every level, all trust indicators were significant. The highest level of education had greater utility ranges than the two lower ones, leading me to believe that participants with this level of education make more informed choices on the basis of trustworthiness.

A summary of all results is given in *Appendix M*.

# 7. DISCUSSION, MARKETING IMPLICATIONS AND ANSWER TO RESEARCH QUESTION

## 7.1. Content of Mailings

In order to overcome the principal-agent problem that hampers value creation in the market for charitable donations, charities should and can take action to mitigate these. Regarding the indicators of trustworthiness tested in the experiment, it is vital to note the importance of providing information on independent accreditation. As mentioned before, this presented the most significant consideration for many participants. Interestingly, the use of CBF accreditation was only explained to participants in a single sentence. Although, none of the participants were asked to evaluate their familiarity with this particular organization, it could imply that it is sufficient for charities to state why having particular accreditation is important quite briefly.

For overhead costs, it is evident that charities should report a cost as low as possible. In the context of this thesis, this was also found to be the second most important consideration for donations. Also, ambiguous overhead costs were least preferred. When information is not disclosed, potential donors apparently do not trust a charity to a) have a low overhead percentage or to b) make an attempt at being trustworthy by disclosing such information. Although, the implications of both types of trust are likely similar to a large extent, they deserve a slightly different explanation. The first suggestion would be consistent with an aversion to ambiguity specific to the number of the overhead percentage. On the other hand, the second suggestion implies that not reporting this percentage deteriorates the trust image of the charity in its entirety. Unfortunately, since such motivations may be tested through qualitative interviews, this is beyond the scope of collected data. What can be said, is that it is much more preferred to report overhead costs above average than not disclosing information on this matter. This finding does suggest there is some sort of trust issue looming when not reporting overhead costs, regardless of how people evaluate it. Essentially, a similar principle applies to government funding. Despite the fact that this was not valued to be of equal importance as overhead costs, acquiring government funds and distributing information emphasizing this will boost perceived trustworthiness. This latter finding also challenges the existence of crowd-out, especially when a charity pro-actively reports on its own government grants.

Finally, information on longevity, in a relative sense, was the least important to consider out of all the trustworthiness indicators. Still, it is highly significant. For reasonably new charities, it is important to refrain from focusing on longevity. If a charity has existed for longer, however, additional focus on longevity becomes redundant at a certain point. It seems donors have a preference for charities that have existed in the recent past. Having a longer history and actively proclaiming its importance, on the other hand, does not seem to contribute to utility and may be a useless investment.

7.2. Adjusting Marketing Policy to Experience

In spite of the fact that it is very beneficial to identify prospective donors that have particular levels of exposure to direct mailings, this is also incredibly challenging. As I found, experienced people will generally be more critical of accreditation by independent organizations, such as CBF. This may be the result of an enhanced ability to digest factual information, which is typical of consumers with greater experience (Maheswaran & Sternthal, 1990). Therefore, charities' market research should reveal whether its donors can be classified as such. If it turns out less experienced people are part of the target demographic, trust information will not be as important, which is why such demographics may be better targeted through relatively less extensive information on independent accreditation. Additionally, charities do not need to alter provided information based on the frequency at which the target demographic donates.

7.3. Adjusting Marketing Policy to Socio-Demographics

Regarding the key socio-demographic characteristics, several vital marketing implications can be proposed.

Age, out of the three key socio-demographic characteristics, is most important in determining the importance of trust information. Older people attach more value to each of the trustworthiness indicators relative to supplementary information. Therefore, if a target demographic is relatively old, marketing policy should focus on creating a reliable image through objective information. For younger people, trustworthiness may be heightened through an image that promotes change, which is more consistent with their values (Williams & Page, 2011). Also, this article, in accordance with my conclusions, advocates the use of information to persuade Generation X. Likewise, it stresses the importance of presenting simple and clear information to Baby Boomers. Because of the fact that information has been found to be so important for older people in determining charity choice, charities that have an older target demographic should factor in this group's reduced cognitive ability. Provided information should be adjusted accordingly.

Since women find independent accreditation significantly more important than men, charities should attempt to adapt to these informational needs. For this finding, several explanations are offered by literature. Schwieren and Sutter (2008) explain that women are significantly less eager to trust someone else's ability, which seems to be a very feasible analogy to the nature of the principal-agent problem between charities and donors. Also, Andreoni and Vesterlund (2001) find a similar tendency in the realm of altruism, i.e., women are less altruistic than men when it becomes expensive to be altruistic. Because of the importance of altruism in driving charitable donations, the finding of increased preference for reliable accreditation seems sensible. Nonetheless, it is challenging to exploit

this. When a charity's pool of donors is fairly mixed in terms of gender, it may be detrimental to a charity's image to practice gender-based marketing in its information provision. This would be inappropriate, given the fact that a donation to a charity is the exact same good, regardless of gender. Marketing such a product in different ways can have extremely undesirable effects on the trust relationship between charities and donors, let alone other forms of appeals as well. As mentioned before, if donors perceive an organization's marketing practice to be manipulative, its consequences can be persistent. Nevertheless, if charities know they can effectively target a certain group made up of either predominantly male or female donors, they could adjust their investment in attaining and promoting independent accreditation according to the findings in this paper.

For education, implications are simple. For higher educated audiences, similar to older people, providing factual information is critical to alleviate the principal-agent problem. A feasible explanation would be that increased cognitive effort is inversely related to analytical ability and memory (Alba & Hutchinson, 1987), i.e., processing information is costlier for people with lower education. When generalizing this statement to the context of charities and, especially, the DCE, the amount of information processed accurately is lower, because the disutility from doing so is greater. Still, factual information on trustworthiness should not at all be omitted for lower educated segments, as each indicator of trustworthiness did significantly alter decision-making. It could perhaps be presented in a simplified form. That is, of course, if a particular donor segment can be identified as being primarily lower educated and targeted easily.

### 7.4. Answer to Research Question

*What information from charities do potential donors value with respect to different communicable indicators of trustworthiness?*

Although this question has been elaborately addressed in prior paragraphs, it is obvious that donors very much value transparency for a trust relationship. For example, firms are better off reporting sub-par information on government grants and overhead costs than not mentioning these facets at all. Also, independent accreditation communicates superior trustworthiness than average governmental accreditation. This effect is stronger for donors with more experience and female donors. Finally, with respect to other socio-demographic characteristics, age and education are determinants of importance attached to information intended to communicate trustworthiness. This implies marketing for older and better educated segments should elaborate on such information, since these segments possess the cognitive ability to process it and/or value a carefully considered donation more. On the other hand, younger and worse educated segments should still receive such information, but in simplified form and supplemented with other forms of stimuli.

## 8. LIMITATIONS AND RECOMMENDATIONS FOR FURTHER RESEARCH

One of the limitations of the study is encompassed in the sample composition. Because of the absence of a usable sampling frame, the sample is composed of participants accumulated through a quota sampling method. Despite the fact that, in comparison to convenience sampling, quota sampling warrants some representativeness regarding key socio-demographic variables, generalizations towards a sampling frame cannot be made. Another limitation of DCEs is potential experimenter bias, since attributes and corresponding levels are selected by the person conducting the experiment. Whereas levels can be made fairly representative and, thus, externally valid by using realistic levels, attributes are a lot more subjective. Fortunately, Handy (2000) gives a sound rationale for selecting the attributes, as these are used by charities in practice. At the same time, none of the participants from the pilot study expressed concerns regarding assignment clarity (see '6.1. Pilot Study'). Still, it is unlikely that people's impression of charities is reflected in the attributes given to them during the experiment, despite the focus on valuation of information. Therefore, it is certainly valuable to conduct research on what people explicitly state they value in the trust relationship.

As mentioned in the section '4.3. Reliability and Validity', the decision-making process for donors is likely different than described. First of all, a consideration set composed of different charities possibly includes more charities than the binary choice presented here. Also, when confronted with marketing stimuli from a specific charity, the short-term consideration set may only include that charity, thereby generating a binary decision problem between donating and not doing so.

To mitigate the concerns above and to build onto the results from this thesis, the study into trustworthiness indicators should be elaborated. For example, field experiments are a viable option for testing the findings from this study. One of the main advantages is obviously task realism. The effectiveness of direct mailings could be compared by tracking the recipients and their donation decisions in a pre-defined period of time after receipt of mail. Furthermore, as section '2. Theoretical Framework' stipulated, cognitive trust is the subject of the study. However, it would be interesting to compare the importance of cognitive trust versus affective trust by sending several types of mail containing different indicators of trustworthiness. This shows how the presence of affective trust may yield different implications.

Furthermore, because of the significant findings from this thesis' application of a DCE, the use of such experimental designs should have a prominent role in academia in the non-profit sector. Whereas field experiments may be fruitful from a broader perspective, DCEs offer a relatively inexpensive alternative for testing simpler hypotheses. In a way, they can act as a reliable basis for testing presumptions, after which larger and more expensive field experiments can be adapted to suit the goal

of obtaining an optimal amount of information. Whereas the DCE in this paper focused on the effects of trustworthiness, preference for complete charity profiles could also be tested. This has the potential to yield implications for marketing policy and other policy issues.

## 9. CONCLUSION

To conclude my thesis, I truly recommend charities to incorporate the relative effects of trust information into direct marketing to provide convincing, but simple information to potential donors to prevent information overload. The findings on socio-demographic characteristics are especially well-implementable, assuming charities know which are present within their target group. Furthermore, whereas the most effective type of direct mail is probably partly subject to the sector and charity-specific characteristics, it could prove to be efficient to incorporate the information from this paper as a fundament.

## REFERENCES


Alba, J. W., & Hutchinson, J. W. (1987). Dimensions of Consumer Expertise. *Journal of Consumer Research, 13*(4), 411-454.

Andreoni, J., & Payne, A. A. (2003). Do Government Grants to Private Charities Crowd Out Giving or Fund-raising? *American Economic Review, 93*(3), 792-812.

Andreoni, J., & Payne, A. A. (2011). Is crowding out due entirely to fundraising? Evidence from a panel of charities. *Journal of Public Economics, 95*(5-6), 334-343.

Andreoni, J., & Vesterlund, L. (2001). Which is the Fair Sex? Gender Differences in Altruism. *The Quarterly Journal of Economics, 116*(1), 293-312.

Beattie, J., & Loomes, G. (1997). Journal of Risk and Uncertainty. *The Impact of Incentives Upon Risky Choice Experiments, 14*(2), 155-168.

Bech, M., & Gyrd- Hansen, D. (2005). Effects coding in discrete choice experiments. *Health Economics, 14*(10), 1079-1083.

Bekkers, R. (2003). Trust, Accreditation, and Philantropy in the Netherlands. *Nonprofit and Voluntary Sector Quarterly, 32*(4), 596-615.

Bekkers, R., & Wiepking, P. (2011). Who gives? Aliterature review of predictors of charitable giving Part One: Religion, education, age and socialisation. *Voluntary Sector Review, 2*(3), 337-365.

Bekkers, R., & Wiepking, P. (2012). Who gives? A literature review of predictors of charitable giving. Part Two: Gender, familiy composition and income. *Voluntary Sector Review*, 217-245.

Biswas, D., Labcrecque, L., Lehmann, D. R., & Markos, E. (2014). Making Choices While Smelling, Tasting, and Listening: The Role of Sensory (Dis)Similarity When Sequentially Sampling Products. *Journal of Marketing, 78*(1), 112-116.

Bowman, W. (2006). Should Donors Care About Overhead Costs? Do They Care? *Nonprofit and Voluntary Sector Quarterly, 35*(2), 288-310.

Burt, C. D., & Strongman, K. (2004). Use of images in charity advertising: Improving donations and compliance rates. *International Journal of Organisational Behaviour, 8*(8), 571-580.

Caussade, S., Ortúzara, J. d., Rizzi, L. I., & Hensher, D. A. (2005). Assessing the influence of design dimensions on stated choice experiment studies. *Transportation Research Part B: Methodological, 39*(7), 621-640.

Chang, C. (2011). Guilt appeals in cause-related marketing. *International Journal of Advertising, 30*(4), 587-616.

Chanley, V. A., Rudolph, T. J., & Rahn, W. M. (2000). The Origins and Consequences of Public Trust in Government: A Time Series Analysis. *Public Opinion Quarterly, 64*(3), 239–256.

Cheraghi-Sohi, S., Hole, A. R., Mead, N., McDonald, R., Whalley, D., Bower, P., & Roland, M. (2008). What Patients Want From Primary Care Consultations: A Discrete Choice Experiment to Identify Patients' Priorities. *Annals of Family Medicine, 6*(2), 107-115.

Cordery, C. J., & Baskerville, R. F. (2011). Charity Transgressions, Trust and Accountability. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations, 22*(2), 197-213.

Cotte, J., Coulter, R. A., & Moore, M. (2005). Enhancing or disrupting guilt: the role of ad credibility and perceived manipulative intent. *Journal of Business Research, 58*(3), 361-368.

Diamond, W. D., & Noble, S. M. (2001). Defensive responses to charitable direct mail solicitations. *Journal of Interactive Marketing, 15*(3), 2-12.

Eppler, M. J., & Mengis, J. (2004). The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines. *The Information Society, 20*(5), 325-344.

Ewing, G., & Sarigöllü, E. (2000). Assessing Consumer Preferences for Clean-Fuel Vehicles: A Discrete Choice Experiment. *Journal of Public Policy and Marketing, 19*(1), 106-118.

Fader, P. S., & Hardie, B. G. (1996). Modeling consumer choice among SKUs. *Journal of Marketing Research, 33*(4), 442-452.

Galesic, M., & Bosnjak, M. (2009). Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Surve. *Public Opinion Quarterly, 73*(2), 349-360.

Gaskin, K. (1999). Blurred vision: public trust in charities. *International Journal of Nonprofit and Voluntary Sector Marketing, 4*(2), 163-178.

Gibelman, M., & Gelman, S. R. (2012). Very Public Scandals: Nongovernmental Organizations in Trouble. *Voluntas: International Journal of Voluntary and Nonprofit Organizations, 12*(1), 49-66.

Gneezy, U., Keenan, E. A., & Gneezy, A. (2014). Avoiding overhead aversion in charity. *Science, 346*(6209), 632-635.

Goede Doelen Nederland. (2016). *Feiten & cijfers goede doelen 2016.* Amsterdam: Goede Doelen Nederland.

Handy, F. (2000). How We Beg: The Analysis of Direct Mail Appeals. *Non-profit and Voluntary Sector Quarterly, 29*(3), 439-454.

Hibbert, S., Smith, A., Davies, A., & Ireland, F. (2007). Guilt appeals: Persuasion knowledge and charitable giving. *Psychology and Marketing, 24*(8), 723-742.

Huber, J., & Zwerina, K. (1996). The Importance of Utility Balance in Efficient Choice Designs. *Journal of Marketing Research, 33*(3), 307-317.

Iwaarden, J. v., Wiele, T. v., Williams, R., & Moxham, C. (2009). Charities: how important is performance to donors? *International Journal of Quality & Reliability Management, 26*(1), 5-22.

Johnson, D., & Grayson, K. (2005). Cognitive and affective trust in service relationships. *Journal of Business Research, 58*(4), 500-507.

Johnson, E. J., & Russo, J. E. (1984). Product Familiarity and Learning New Information. *Journal of Consumer Research, 11*(1), 542-550.

Katz, R. A. (2000). Can Principal-Agent Models Help Explain Charitable Gifts and Organizations? *Wisconsin Law Review, 2000*(1), 1-30.

Lancsar, E., & Louviere, J. (2008). Conduction discrete choice experiments to inform healthcare decision making. *Pharmacoeconomics, 26*(8), 661-677.

Lewis, J. D., & Weigert, A. (1985). Trust as a Social Reality. *Social Forces, 63*(4), 967-985.

Maheswaran, D., & Sternthal, B. (1990). The Effects of Knowledge, Motivation, and Type of Message on Ad Processing and Product Judgments. *Journal of Consumer Research, 17*(1), 66-73.

Mitchell, V.- W., & Papavassiliou, V. (1999). Marketing causes and implications of consumer confusions. *Journal of Product & Brand Management, 8*(4), 319-342.

Morgan, R. M., & Hunt, S. D. (1994). The Commitment-Trust Theory of Relationship Marketing. *Journal of Marketing, 58*(3), 20-38.

Mose, C. A., & Stuart, A. (1953). An Experimental Study of Quota Sampling. *Journal of the Royal Statistical Society. Series A (General), 116*(4), 349-405.

Obermiller, C., & Spangenberg, E. R. (2008). Development of a Scale to Measure Consumer Skepticism Toward Advertising. *Journal of Consumer Psychology, 7*(2), 159-186.

Orne, B. (1998). Sample size issues for conjoint analysis studies. *Sequim: Sawtooth Software Technical Paper*.

Piersma, N., & Jonker, J.-J. (2004). Determining the optimal direct mailing frequency. *European Journal of Operational Research, 158*(1), 173-182.

Ross, S. A. (1973). The Economic Theory of Agency: The Principal's Problem. *American Economic Review, 63*(2), 134-139.

Ryan, M., Bate, A., Eastmond, C., & Ludbrook, A. (2001). Use of Discrete Choice Experiments. *Quality in Health Care, 10*(Suppl. 1), i55-i60.

Sargeant, A. (2010). Charitable Giving: Towards a Model of Donor Behaviour. *Journal of Marketing Management, 15*(4), 215-238.

Sargeant, A., & Lee, S. (2002). Improving public trust in the voluntary sector: An empirical analysis. *International Journal of Nonprofit and Voluntary Sector Marketing, 7*(1), 68-83.

Schwieren, C., & Sutter, M. (2008). Trust in cooperation or ability? An experimental study on gender differences. *Economic Letters, 99*(3), 494-497.

Smith, G. E., & Berger, P. D. (1996). The Impact of Direct Marketing Appeals on Charitable Marketing Effectiveness. *Journal of the Academy of Marketing Science, 24*(3), 219-231.

Telser, H., & Zweifel, P. (2007). Validity of discrete-choice experiments evidence for health risk reduction. *Applied Economics, 39*(1), 69-78.

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

Williams, K. C., & Page, R. A. (2011). Marketing to the Generations. *Journal of Behavioral Studies in Business, 3*, 1-17.

# APPENDICES

| | | LEVELS | | | REASON |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | |
| ATTRIBUTES | **SECTOR** | Against Cancer | Against Cancer | Against Cancer | All equal to prevent bias. |
| | **FOUNDED IN** | 2010 *(-1)* | 1970 *(0)* | 1930 *(1)* | Relatively far apart to accurately determine preference. |
| | **DONATION POOL**[1] | €5M *(0)* | €15M *(0)* | €25M *(0)* | Benchmark attribute to describe size: small, medium, and large.[2] |
| | **OVERHEAD COSTS**[1] | Not mentioned *(-1)* | 12% *(0)* | 6% *(1)* | 9% is average for the health sector. |
| | **MONEY FROM GOVERNMENT** | Not mentioned *(-1)* | 15% *(0)* | 30% *(1)* | 24% is average for the health sector. |
| | **METHOD OF COLLECTION** | Door-to-Door Collection *(0)* | Debit Card *(0)* | | Benchmark attribute indicating the way in which donations are collected.[2] |
| | **CBF LABEL OF QUALITY** | NO *(-1)* | YES *(1)* | | Binary by definition. |

***Appendix A:*** *The attributes used in the DCE with respective possible levels. Prior mean estimates for each level are given between brackets.*

[1] The percentages represent the share of the donation pool relevant to this attribute.

[2] 'Benchmark attributes' are attributes that have been added to mitigate contingency awareness, supplement context, and for testing the findings surrounding indicators of trustworthiness relative to a benchmark.

| Choice Set | Founded in | Donation Pool | Overhead Costs | Government Funding | Method of Collection | CBF Label of Quality |
|---|---|---|---|---|---|---|
| **1** | 1930 | €5M | Not Mentioned | Not Mentioned | Direct Deposit | No |
| **1** | 1970 | €25M | 12% | 15% | Door-to-Door | No |
| **2** | 1970 | €25M | Not Mentioned | Not Mentioned | Door-to-Door | Yes |
| **2** | 1930 | €25M | Not Mentioned | 30% | Direct Deposit | No |
| **3** | 1970 | €15M | 12% | 15% | Door-to-Door | No |
| **3** | 1930 | €25M | 6% | Not Mentioned | Door-to-Door | No |
| **4** | 1970 | €15M | Not Mentioned | Not Mentioned | Door-to-Door | Yes |
| **4** | 2010 | €25M | 12% | 15% | Direct Deposit | Yes |
| **5** | 2010 | €15M | 6% | 15% | Direct Deposit | Yes |
| **5** | 1970 | €25M | 12% | 30% | Door-to-Door | No |
| **6** | 1970 | €15M | 6% | Not Mentioned | Direct Deposit | No |
| **6** | 1930 | €5M | Not Mentioned | 15% | Door-to-Door | No |
| **7** | 2010 | €15M | Not Mentioned | 30% | Door-to-Door | Yes |
| **7** | 1970 | €25M | 12% | 15% | Direct Deposit | No |
| **8** | 1970 | €5M | 6% | 15% | Direct Deposit | No |
| **8** | 2010 | €25M | 12% | Not Mentioned | Direct Deposit | Yes |
| **9** | 1930 | €15M | Not Mentioned | 15% | Direct Deposit | No |
| **9** | 1930 | €5M | 12% | Not Mentioned | Direct Deposit | No |
| **10** | 1970 | €15M | 12% | 30% | Direct Deposit | Yes |
| **10** | 2010 | €5M | 6% | 30% | Door-to-Door | No |
| **11** | 1970 | €25M | Not Mentioned | Not Mentioned | Direct Deposit | No |
| **11** | 2010 | €5M | 12% | 15% | Door-to-Door | No |
| **12** | 1970 | €5M | Not Mentioned | 30% | Direct Deposit | Yes |
| **12** | 2010 | €15M | 6% | 15% | Door-to-Door | Yes |
| **13** | 1970 | €5M | Not Mentioned | 15% | Direct Deposit | Yes |
| **13** | 2010 | €15M | 12% | 30% | Direct Deposit | No |
| **14** | 2010 | €25M | Not Mentioned | 15% | Door-to-Door | Yes |
| **14** | 1970 | €15M | 12% | Not Mentioned | Direct Deposit | No |
| **15** | 1930 | €15M | 12% | 15% | Door-to-Door | No |
| **15** | 1970 | €5M | 6% | 30% | Door-to-Door | No |
| **16** | 1930 | €15M | 12% | Not Mentioned | Door-to-Door | Yes |
| **16** | 2010 | €5M | Not Mentioned | 15% | Direct Deposit | Yes |

***Appendix B:*** *Utility-balanced design generated with prior means specified as in Appendix B and an identity matrix to represent the variance-covariance matrix.*

**1.** What is your age?

*[…]*

**2.** What is your gender?

*[Male] [Female]*

**3.** What is your highest level of education? If you are attending education higher than any other education you have finished, please indicate the level you are currently attending.

*[Vwo/University] [Havo/University of Applied Sciences] [Mavo/Practical Higher Education] [Primary school/No schooling]*

***Appendix C:*** *Questions asked to describe socio-demographic characteristics. The questions that are multiple choice indicate possible choices between square brackets.*

**1.** How often do you receive (e-)mail from a charity, regardless of whether you have requested it or not? Indicate the answer that is most accurate.

*[Daily] [Weekly] [Monthly] [Less often than monthly] [Never]*

**2.** How much do you donate a year? This includes one-time cash donations to collectors, one-time donations via direct deposit, and automatic subscriptions.

*[I do not donate.] [Less than €25] [Between €25 and €50] [More than €50]*

**3.** How often do you donate a year? This includes one-time cash donations to collectors, one-time donations via direct deposit, and automatic subscriptions. Indicate the answer that is most accurate.

*[I do not donate.] [Weekly] [Monthly] [Quarterly] [Yearly]*

***Appendix D:*** *Questions asked to describe donating behavior. The questions are multiple choice with possible choices as indicated between square brackets.*

**1.** It was clear what was expected from me during the choice task.

*[Fully Agree] [Agree] [Neither Agree nor Disagree] [Disagree] [Fully Disagree]*

**2.** The questionnaire was of appropriate length.

*[Fully Agree] [Agree] [Neither Agree nor Disagree] [Disagree] [Fully Disagree]*

**3.** I had sufficient information to make a choice, given the choice tasks.

*[Fully Agree] [Agree] [Neither Agree nor Disagree] [Disagree] [Fully Disagree]*

**4.** The questionnaire asked for information that is confidential.

*[Fully Agree] [Agree] [Neither Agree nor Disagree] [Disagree] [Fully Disagree]*

***Appendix E****: Participants in the pilot study were tasked with rating each of the above statements on a 5-point Likert scale.*

| Statistic | Quota min. | Quota max. | Actual |
|---|---|---|---|
| **Age (mean)** | 30 | 40 | 36.97 |
| **Age (std. dev.)** | 10 | - | 17.44 |
| **Female (% of sample)** | 40 | 60 | 58.25 |
| **Education (% of sample)** | | | |
| **University/vwo** | 20 | 50 | 27.18 |
| **Applied Sciences/havo** | 20 | 50 | 38.83 |
| **Practical Education/mbo** | 20 | 50 | 33.98 |

***Appendix F****: Targeted sample characteristics according to quota minima and maxima and actual value of sample characteristics.*

| Parameter | Estimate | Std. Error | $P(x > \chi^2)$ |
|---|---|---|---|
| Founded In[2010] | -0.47949759 | 0.208516 | 0.0134 |
| Founded In[1970] | 0.25978478 | 0.136759 | |
| Donation Pool[€5M] | 0.47790362 | 0.178235 | 0.0145 |
| Donation Pool[€15M] | -0.15449856 | 0.13958 | |
| Overhead Costs[Not Mentioned] | -1.0406561 | 0.222086 | <0.0001 |
| Overhead Costs[12%] | -0.02355359 | 0.145896 | |
| Government Funding[Not Mentioned] | -0.78982637 | 0.214909 | 0.0003 |
| Government Funding[15%] | -0.03680735 | 0.142273 | |
| Method of Collection[Door-to-Door] | 0.09382447 | 0.106493 | 0.3622 |
| CBF Label of Quality[No] | -0.99573874 | 0.18893 | <0.0001 |
| Mail_Freq[2-1]*Founded In[2010] | -0.12947077 | 0.276603 | 0.8909 |
| Mail_Freq[2-1]*Founded In[1970] | -0.05764643 | 0.171872 | |
| Mail_Freq[2-1]*Donation Pool[€5M] | 0.13978971 | 0.236016 | 0.1189 |
| Mail_Freq[2-1]*Donation Pool[€15M] | -0.07204037 | 0.178104 | |
| Mail_Freq[2-1]*Overhead Costs[Not Mentioned] | -0.16813318 | 0.293673 | 0.0989 |
| Mail_Freq[2-1]*Overhead Costs[12%] | 0.12256869 | 0.192023 | |
| Mail_Freq[2-1]*Government Funding[Not Mentioned] | 0.28102208 | 0.273662 | 0.2414 |
| Mail_Freq[2-1]*Government Funding[15%] | -0.05378418 | 0.178982 | |
| Mail_Freq[2-1]*Method of Collection[Door-to-Door] | 0.24668373 | 0.133594 | 0.0490 |
| Mail_Freq[2-1]*CBF Label of Quality[No] | 0.11593695 | 0.238608 | 0.0388 |
| Mail_Freq[3-2]*Founded In[2010] | 0.15289858 | 0.241998 | 0.8909 |
| Mail_Freq[3-2]*Founded In[1970] | 0.09402788 | 0.144725 | |
| Mail_Freq[3-2]*Donation Pool[€5M] | -0.45936594 | 0.198653 | 0.1189 |
| Mail_Freq[3-2]*Donation Pool[€15M] | 0.18451199 | 0.148043 | |
| Mail_Freq[3-2]*Overhead Costs[Not Mentioned] | 0.57697777 | 0.250742 | 0.0989 |
| Mail_Freq[3-2]*Overhead Costs[12%] | -0.15341772 | 0.161892 | |
| Mail_Freq[3-2]*Government Funding[Not Mentioned] | 0.01948857 | 0.233334 | 0.2414 |
| Mail_Freq[3-2]*Government Funding[15%] | 0.23686128 | 0.14756 | |
| Mail_Freq[3-2]*Method of Collection[Door-to-Door] | 0.02796243 | 0.11406 | 0.0490 |
| Mail_Freq[3-2]*CBF Label of Quality[No] | 0.31888635 | 0.196745 | 0.0388 |
| Don_Freq[2-1]*Founded In[2010] | 0.05252677 | 0.217888 | 0.8186 |
| Don_Freq[2-1]*Founded In[1970] | -0.14625719 | 0.134145 | |
| Don_Freq[2-1]*Donation Pool[€5M] | -0.22918914 | 0.181201 | 0.7848 |
| Don_Freq[2-1]*Donation Pool[€15M] | 0.07293064 | 0.135916 | |
| Don_Freq[2-1]*Overhead Costs[Not Mentioned] | 0.0375871 | 0.220962 | 0.9249 |
| Don_Freq[2-1]*Overhead Costs[12%] | 0.01489531 | 0.148782 | |
| Don_Freq[2-1]*Government Funding[Not Mentioned] | -0.11467431 | 0.217317 | 0.5722 |
| Don_Freq[2-1]*Government Funding[15%] | 0.12513676 | 0.134204 | |
| Don_Freq[2-1]*Method of Collection[Door-to-Door] | -0.10545557 | 0.106246 | 0.0546 |
| Don_Freq[2-1]*CBF Label of Quality[No] | 0.21922987 | 0.184589 | 0.3136 |
| Don_Freq[3-2]*Founded In[2010] | 0.07075672 | 0.234018 | 0.8186 |
| Don_Freq[3-2]*Founded In[1970] | 0.04360989 | 0.145629 | |
| Don_Freq[3-2]*Donation Pool[€5M] | 0.16379147 | 0.188055 | 0.7848 |
| Don_Freq[3-2]*Donation Pool[€15M] | -0.02112286 | 0.148644 | |
| Don_Freq[3-2]*Overhead Costs[Not Mentioned] | 0.02694347 | 0.234822 | 0.9249 |
| Don_Freq[3-2]*Overhead Costs[12%] | -0.13316703 | 0.155101 | |
| Don_Freq[3-2]*Government Funding[Not Mentioned] | 0.10956445 | 0.236258 | 0.5722 |
| Don_Freq[3-2]*Government Funding[15%] | -0.23088451 | 0.151289 | |
| Don_Freq[3-2]*Method of Collection[Door-to-Door] | -0.14873743 | 0.11635 | 0.0546 |
| Don_Freq[3-2]*CBF Label of Quality[No] | -0.28696467 | 0.203873 | 0.3136 |

***Appendix G**: Parameter estimates, standard error, and p-value of the likelihood ratio test of the model used for testing the hypotheses.*

| Parameter | Estimate | Std. Error | $P\ (x > \chi^2)$ |
|---|---|---|---|
| Founded In[1970] | 0.193382929 | 0.053542 | <0.0001 |
| Founded In[2010] | -0.456670656 | 0.085116 | |
| Donation Pool[€5M] | 0.218544686 | 0.068006 | 0.0046 |
| Donation Pool[€15M] | -0.065621448 | 0.053841 | |
| Overhead Costs[Not Mentioned] | -0.80426138 | 0.085761 | <0.0001 |
| Overhead Costs[12%] | -0.063899551 | 0.056798 | |
| Government Funding[Not Mentioned] | -0.600576737 | 0.085955 | <0.0001 |
| Government Funding[15%] | 0.072184249 | 0.054371 | |
| Method of Collection[Door-to-Door] | 0.190422654 | 0.041676 | <.0001 |
| CBF Label of Quality[No] | -0.676769568 | 0.073086 | <0.0001 |

***Appendix H**: Parameter estimates, standard error, and p-value of the likelihood ratio test of the model used as a first robustness check.*

| Parameter | Estimate | Std. Error | $P\ (x > \chi^2)$ |
|---|---|---|---|
| Founded In[2010] | -0.47284993 | 0.182384 | 0.0105 |
| Founded In[1970] | 0.19424776 | 0.121252 | |
| Donation Pool[€5M] | 0.39100716 | 0.161445 | 0.0335 |
| Donation Pool[€15M] | -0.12481378 | 0.126619 | |
| Overhead Costs[Not Mentioned] | -1.03172715 | 0.199932 | <0.0001 |
| Overhead Costs[12%] | -0.02391343 | 0.132769 | |
| Government Funding[Not Mentioned] | -0.8400078 | 0.189709 | <0.0001 |
| Government Funding[15%] | 0.00820241 | 0.127523 | |
| Method of Collection[Door-to-Door] | 0.04108415 | 0.094102 | 0.6655 |
| CBF Label of Quality[No] | -0.9136222 | 0.170187 | <0.0001 |
| Mail_Freq[2-1]*Founded In[2010] | -0.13450465 | 0.273427 | 0.7658 |
| Mail_Freq[2-1]*Founded In[1970] | -0.06172072 | 0.170637 | |
| Mail_Freq[2-1]*Donation Pool[€5M] | 0.1343448 | 0.233644 | 0.0673 |
| Mail_Freq[2-1]*Donation Pool[€15M] | -0.06890558 | 0.175788 | |
| Mail_Freq[2-1]*Overhead Costs[Not Mentioned] | -0.16714292 | 0.291914 | 0.0288 |
| Mail_Freq[2-1]*Overhead Costs[12%] | 0.10135178 | 0.190349 | |
| Mail_Freq[2-1]*Government Funding[Not Mentioned] | 0.2816635 | 0.271358 | 0.2966 |
| Mail_Freq[2-1]*Government Funding[15%] | -0.06480415 | 0.1777 | |
| Mail_Freq[2-1]*Method of Collection[Door-to-Door] | 0.21796851 | 0.131914 | 0.1816 |
| Mail_Freq[2-1]*CBF Label of Quality[No] | 0.07508738 | 0.235913 | 0.0422 |
| Mail_Freq[3-2]*Founded In[2010] | 0.22048401 | 0.231635 | 0.7658 |
| Mail_Freq[3-2]*Founded In[1970] | 0.08121594 | 0.138892 | |
| Mail_Freq[3-2]*Donation Pool[€5M] | -0.46023985 | 0.189728 | 0.0673 |
| Mail_Freq[3-2]*Donation Pool[€15M] | 0.18560755 | 0.14063 | |
| Mail_Freq[3-2]*Overhead Costs[Not Mentioned] | 0.6142305 | 0.239103 | 0.0288 |
| Mail_Freq[3-2]*Overhead Costs[12%] | -0.18610636 | 0.15471 | |
| Mail_Freq[3-2]*Government Funding[Not Mentioned] | 0.05251858 | 0.224537 | 0.2966 |
| Mail_Freq[3-2]*Government Funding[15%] | 0.18737698 | 0.142083 | |
| Mail_Freq[3-2]*Method of Collection[Door-to-Door] | -0.04424383 | 0.107566 | 0.1816 |
| Mail_Freq[3-2]*CBF Label of Quality[No] | 0.30468211 | 0.188971 | 0.0422 |

***Appendix I**: Parameter estimates, standard error, and p-value of the likelihood ratio test of the model including mailing frequency as the only subject effect.*

| Parameter | Estimate | Std. Error | $P\ (x > \chi^2)$ |
|---|---|---|---|
| Founded In[2010] | -0.521168947 | 0.142398 | <0.0001 |
| Founded In[1970] | 0.256414339 | 0.088501 | |
| Donation Pool[€5M] | 0.362659512 | 0.11551 | 0.0040 |
| Donation Pool[€15M] | -0.115116394 | 0.087582 | |
| Overhead Costs[Not Mentioned] | -0.917951285 | 0.145441 | <0.0001 |
| Overhead Costs[12%] | 0.009479505 | 0.096502 | |
| Government Funding[Not Mentioned] | -0.59881347 | 0.142356 | <0.0001 |
| Government Funding[15%] | 0.057567551 | 0.087945 | |
| Method of Collection[Door-to-Door] | 0.290892243 | 0.068664 | <0.0001 |
| CBF Label of Quality[No] | -0.762856823 | 0.116597 | <0.0001 |
| Don_Freq[2-1]*Founded In[2010] | 0.072700952 | 0.215772 | 0.7579 |
| Don_Freq[2-1]*Founded In[1970] | -0.144579944 | 0.132423 | |
| Don_Freq[2-1]*Donation Pool[€5M] | -0.232606588 | 0.176183 | 0.5986 |
| Don_Freq[2-1]*Donation Pool[€15M] | 0.068696297 | 0.132591 | |
| Don_Freq[2-1]*Overhead Costs[Not Mentioned] | 0.08086161 | 0.216345 | 0.4559 |
| Don_Freq[2-1]*Overhead Costs[12%] | -0.020196161 | 0.144558 | |
| Don_Freq[2-1]*Government Funding[Not Mentioned] | -0.097698794 | 0.214662 | 0.7329 |
| Don_Freq[2-1]*Government Funding[15%] | 0.099449974 | 0.132266 | |
| Don_Freq[2-1]*Method of Collection[Door-to-Door] | -0.120797819 | 0.104511 | 0.1614 |
| Don_Freq[2-1]*CBF Label of Quality[No] | 0.214863692 | 0.18127 | 0.5030 |
| Don_Freq[3-2]*Founded In[2010] | 0.098255803 | 0.216551 | 0.7579 |
| Don_Freq[3-2]*Founded In[1970] | 0.067358256 | 0.136443 | |
| Don_Freq[3-2]*Donation Pool[€5M] | 0.013090864 | 0.173636 | 0.5986 |
| Don_Freq[3-2]*Donation Pool[€15M] | 0.045055781 | 0.138775 | |
| Don_Freq[3-2]*Overhead Costs[Not Mentioned] | 0.209110468 | 0.217236 | 0.4559 |
| Don_Freq[3-2]*Overhead Costs[12%] | -0.163139541 | 0.144722 | |
| Don_Freq[3-2]*Government Funding[Not Mentioned] | 0.189689216 | 0.220935 | 0.7329 |
| Don_Freq[3-2]*Government Funding[15%] | -0.150794312 | 0.142621 | |
| Don_Freq[3-2]*Method of Collection[Door-to-Door] | -0.063171089 | 0.107541 | 0.1614 |
| Don_Freq[3-2]*CBF Label of Quality[No] | -0.121858776 | 0.19217 | 0.5030 |

***Appendix J**: Parameter estimates, standard error, and p-value of the likelihood ratio test of the model including donation frequency as the only subject effect.*

| Parameter | Estimate | Std. Error | $P(x > \chi^2)$ |
|---|---|---|---|
| Founded In[2010] | -0.250000982 | 0.231969 | 0.1124 |
| Founded In[1970] | 0.268351581 | 0.145269 | |
| Donation Pool[€5M] | 0.224078443 | 0.191306 | 0.2095 |
| Donation Pool[€15M] | -0.236585197 | 0.152128 | |
| Overhead Costs[Not Mentioned] | -0.758726268 | 0.238684 | 0.0018 |
| Overhead Costs[12%] | -0.022190865 | 0.158483 | |
| Government Funding[Not Mentioned] | -0.378056427 | 0.236225 | 0.0563 |
| Government Funding[15%] | 0.247535202 | 0.154459 | |
| Method of Collection[Door-to-Door] | 0.097081776 | 0.112865 | 0.3714 |
| CBF Label of Quality[No] | -0.596547265 | 0.200261 | 0.0016 |
| Age*Founded In[2010] | -0.017010375 | 0.005858 | 0.0027 |
| Age*Founded In[1970] | -0.002658952 | 0.00347 | |
| Age*Donation Pool[€5M] | 0.00577728 | 0.004453 | 0.3640 |
| Age*Donation Pool[€15M] | -0.003089632 | 0.00352 | |
| Age*Overhead Costs[Not Mentioned] | -0.018270669 | 0.005802 | 0.0024 |
| Age*Overhead Costs[12%] | 0.000481083 | 0.00376 | |
| Age*Government Funding[Not Mentioned] | -0.013235663 | 0.005631 | 0.0043 |
| Age*Government Funding[15%] | -0.007045166 | 0.003558 | |
| Age*Method of Collection[Door-to-Door] | 0.001112314 | 0.002731 | 0.6747 |
| Age*CBF Label of Quality[No] | -0.011529766 | 0.004747 | 0.0093 |
| Gender[0]*Founded In[2010] | -0.096102874 | 0.090702 | 0.2661 |
| Gender[0]*Founded In[1970] | -0.051957523 | 0.056127 | |
| Gender[0]*Donation Pool[€5M] | 0.076574501 | 0.071776 | 0.5573 |
| Gender[0]*Donation Pool[€15M] | -0.024139075 | 0.057138 | |
| Gender[0]*Overhead Costs[Not Mentioned] | -0.148809979 | 0.091587 | 0.2308 |
| Gender[0]*Overhead Costs[12%] | -0.000015288 | 0.060288 | |
| Gender[0]*Government Funding[Not Mentioned] | -0.066948539 | 0.090209 | 0.6618 |
| Gender[0]*Government Funding[15%] | -0.027910464 | 0.057401 | |
| Gender[0]*Method of Collection[Door-to-Door] | 0.163640334 | 0.044077 | <0.0001 |
| Gender[0]*CBF Label of Quality[No] | -0.169549589 | 0.07707 | 0.0224 |
| Education[2-1]*Founded In[2010] | 0.390640764 | 0.233286 | 0.0102 |
| Education[2-1]*Founded In[1970] | 0.063999049 | 0.145358 | |
| Education[2-1]*Donation Pool[€5M] | -0.185946247 | 0.186244 | 0.0284 |
| Education[2-1]*Donation Pool[€15M] | 0.336462057 | 0.149874 | |
| Education[2-1]*Overhead Costs[Not Mentioned] | 0.687421028 | 0.236964 | 0.0004 |
| Education[2-1]*Overhead Costs[12%] | -0.054617545 | 0.156552 | |
| Education[2-1]*Government Funding[Not Mentioned] | 0.295221548 | 0.234334 | 0.3465 |
| Education[2-1]*Government Funding[15%] | 0.099709919 | 0.152324 | |
| Education[2-1]*Method of Collection[Door-to-Door] | 0.09074502 | 0.111153 | 0.5943 |
| Education[2-1]*CBF Label of Quality[No] | 0.302699053 | 0.200344 | 0.0016 |
| Education[3-2]*Founded In[2010] | 0.431517298 | 0.216445 | 0.0102 |
| Education[3-2]*Founded In[1970] | -0.041602825 | 0.133292 | |
| Education[3-2]*Donation Pool[€5M] | -0.226572479 | 0.170136 | 0.0284 |
| Education[3-2]*Donation Pool[€15M] | 0.097161029 | 0.132966 | |
| Education[3-2]*Overhead Costs[Not Mentioned] | 0.375346766 | 0.215941 | 0.0004 |
| Education[3-2]*Overhead Costs[12%] | -0.017420301 | 0.141708 | |
| Education[3-2]*Government Funding[Not Mentioned] | 0.176858418 | 0.21447 | 0.3465 |
| Education[3-2]*Government Funding[15%] | 0.000845523 | 0.134222 | |
| Education[3-2]*Method of Collection[Door-to-Door] | -0.087802872 | 0.105477 | 0.5943 |
| Education[3-2]*CBF Label of Quality[No] | 0.403981132 | 0.181978 | 0.0016 |

***Appendix K**: Parameter estimates, standard error, and p-value of the likelihood ratio test of the model including socio-demographic characteristics.*

| Parameter | Estimate (educ. = 1) | Estimate (educ. = 2) | Estimate(educ. = 3) |
|---|---|---|---|
| Founded In[1970] | -0.72069969**** | -0.463279411*** | -0.266920203** |
| Founded In[2010] | 0.2047181**** | 0.227264433*** | 0.161502457** |
| Donation Pool[€5M] | 0.39929428** | 0.239662227* | 0.139981668 |
| Donation Pool[€15M] | -0.32247465** | -0.029506252* | 0.029705521 |
| Overhead Costs[Not Mentioned] | -1.26245813**** | -0.69594799**** | -0.602263303**** |
| Overhead Costs[12%] | -0.00907095**** | -0.076451483**** | -0.114381054**** |
| Government Funding[Not Mentioned] | -0.73635847**** | -0.54224779*** | -0.564670713**** |
| Government Funding[15%] | 0.05684318**** | 0.09986739*** | 0.008913043**** |
| Method of Collection[Door-to-Door] | 0.12600259 | 0.212025995*** | 0.203303299*** |
| CBF Label of Quality[No] | -0.91076078**** | -0.698795331**** | -0.481908402**** |
| ** = p < 0.10, ** = p < 0.05, *** = p < 0.01, **** = p <0.0001* | | | |

***Appendix L**: Parameter estimates of the three models for different levels of education.*

| [Hypothesis Number]: Effect Tested | Result | Comment |
|---|---|---|
| 1: Longevity | Supported | Preference for charities that have existed for a while, but marginal utility declines for higher longevity. |
| 2a+b: Overhead Costs | Supported | Price and trust effect of higher overhead costs. |
| 3a+b: Government Funding | Supported | Trust effect of government funding and negative crowd-out. |
| 4a+b: Independent Accreditation | Supported | Relatively more important than average government backing. Most important factor in the experiment |
| 5: Mailing Frequency and Importance of Trust | Supported | Increased importance of independent accreditation. |
| 6: Donation Frequency and Importance of Trust | Not Supported | Perhaps greater focus on practicalities surrounding collection. |
| Socio-demographic Characteristics | - | Age and education are most important.<br>Women care more about independent accreditation than men |

***Appendix M**: Summary of the results from testing the hypotheses.*