University of Basel
Center for Philanthropy Studies

CEPS Working Paper Series
No. 21, April 2023

Compassion for all: Real-world online donations contradict compassion fade

Dominik S. Meier

Center for Philanthropy Studies (CEPS), University of Basel
dominik.meier@unibas.ch

Citation:
Meier, Dominik. (2023). Compassion for all: Real-world online donations contradict compassion fade. CEPS Working Paper Series, No. 21, Basel: CEPS.

ISSN: 2296-7516 (Print)
2296-7524 (Online)

Abstract

Research has shown that people are more likely to donate money to help a single victim rather than a group of victims. However, recent studies have been able to reverse this compassion fade effect by presenting people with multiple donation appeals with different victim group sizes (joint evaluation) instead of just one donation appeal (separate evaluation). The reversal of this effect when people evaluate multiple donation requests at once has important implications for fundraising. This study tests whether this effect can be replicated in the field by using data from GoFundMe, the world's largest crowdfunding platform. When browsing projects on GoFundMe, people see multiple projects displayed at once, placing them in a joint evaluation context. Using the project campaign category and description to control for confounding, I find that there is indeed a positive effect of the perceived victim group size on the amount of funds raised by a project.

Keywords: Crowdfunding, compassion fade, fundraising, donations, identifiable victim effect

Disclosure Statement: The author(s) declare that there are no conflicts of interest that relate to the research, authorship, or publication of this article

Content

1 Introduction

Online fundraising has become increasingly popular, as it provides access to a huge donor pool at very low costs (Hart, 2002). The promises of online fundraising have been successfully exploited by crowdfunding sites such as GoFundMe.com which allow people to set up crowdfunding campaigns on their online platform with only a few clicks. The $15 billion that has been collectively raised since 2010 on GoFundMe (GoFundMe, 2022), which is the largest donation-based crowdfunding website, speaks for the success of these platforms. Although the amount of funds raised on GoFundMe is impressive, most campaigns fall short of their fundraising targets (Kenworthy & Igra, 2022). This low success rate might be explained by the fact that traditional strategies intended to boost donations (e.g., Ruehle et al., 2021) might not work or even backfire on such platforms. This is because in contrast to more traditional means of fundraising, such as mail solicitation where potential donors often receive only one donation request at a time, potential donors on crowdfunding platforms can choose from a large number of projects to donate to. As noted by Erlandsson (2021), whether people evaluate one option separately or multiple options jointly has been very influential for research on judgment and decision making. For example, Erlandsson (2021) quotes evidence that shows that emotional reactions are more predictive of attitudes toward policies in separate evaluations (Ritov & Baron, 2011), while efficiency-related attributes are more predictive in joint evaluations (Bazerman et al., 2011; Caviola et al., 2014). This evidence led Erlandsson (2021) to test seven helping effects (i.e., strategies that fundraisers can use to boost donations) both when people evaluated multiple donation requests at once (joint evaluation) and when they only evaluated one donation request (separate evaluation).

Erlandsson (2021) found that potential donors indeed prefer projects with different attributes depending on whether they only evaluate one project or multiple projects at once. For example, while research using separate evaluation found that donors prefer projects with a single identified victim to projects with multiple unidentified victims (Lee & Feeley, 2016), Erlandsson (2021) was able to reverse this effect in the joint evaluation condition. This outcome is in line with another recent experimental study by Garinther et al. (2022). In contrast to previous studies that used a separate evaluation design, Garinther et al. (2022) found that people donated more to donation requests depicting larger victim groups than to donation requests depicting smaller victim groups when participants evaluated multiple donation requests at once. According to Garinther et al. (2022), it is the comparison of multiple donation requests with different depicted group sizes that leads to the positive effect of depicted group size on giving. This result has important consequences for fundraising, since studies have traditionally concluded that larger victim groups attract smaller donations (see Butts et al. (2019) for a meta-analysis).

This study examines whether fundraisers that use crowdfunding can leverage the results of Erlandsson (2021) and Garinther et al. (2022) by manipulating the *perceived* victim group size. In lab studies that test the effect of victim group size on giving, the perceived victim group size (i.e., how many people are depicted on the picture) usually corresponds with the real victim group size (i.e., the size of the group that will receive the donations) e.g., Garinther et al. (2022). However, in real-life donation requests, there is usually no direct correspondence between the size of the depicted victim group (e.g., a poor child from Sudan) and the size of the group that benefits from the donation (poor Sudanese children, in this example). Thus, in this work, I attempt to test whether fundraisers can raise more funds by manipulating the size of the perceived victim group in a joint evaluation context (i.e., crowdfunding).

To test this hypothesis, I estimate the effect of the depicted victim group size on giving in a real-world setting where people usually see multiple donation requests at once. I use data from more than 60,000 crowdfunding projects from GoFundMe, the world's largest social fundraising platform. When browsing fundraising projects on GoFundMe.com, people see multiple fundraising projects displayed in a grid (see Figure A1 in the appendix), which places them into a joint evaluation framework. According to Erlandsson

(2021) and Garinther et al. (2022), we should thus observe a positive effect of the number of persons depicted on a project's project profile picture on the funds acquired by the project.

Given the observational setting, I first need to identify the effect of the perceived victim group size on the funds raised. To identify this effect, I need to account for all confounders between the number of people depicted on a project profile picture and the amount of funds raised. The topic of a fundraising project is such a confounder. Whether the funds are raised for a sick child or a college football team likely influences both how many people are depicted on the profile picture and how much people will donate to the project. Fortunately, fundraising projects on GoFundMe must be assigned to one of 18 predefined categories (e.g., "medical", "sports"). Indeed, the category of a project correlates with both the amount of funds raised and the number of persons depicted on a project profile picture (see Figure 1).

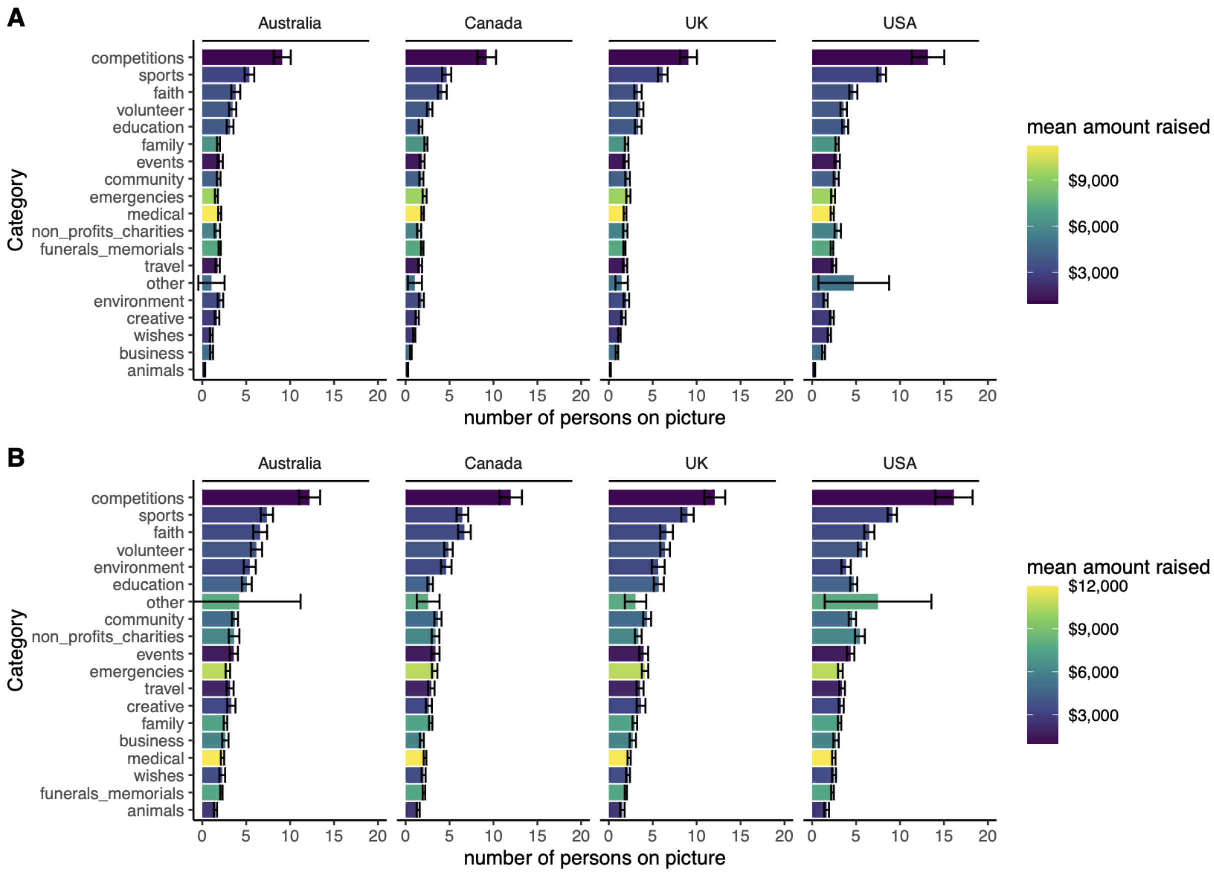
To assess how robustly the category of a project controls for confounding, I also use the campaign description text to additionally control for the topic of a project. The campaign description is free text provided on the project's profile page that fundraisers can use to describe their project. Campaign descriptions have been shown to influence the success of crowdfunding projects (Kuo et al., 2022). I use document embeddings (Le & Mikolov, 2014; Reimers & Gurevych, 2019) and topic models (Blei, 2012) to encode this text into numbers that can then be used as controls.

Controlling for the category and the campaign description allows me to identify the effect of the perceived victim group size on donations for projects that belong to the same category and have similar campaign descriptions. This places us close to an experimental design where we could vary the number of people depicted on the profile project picture while keeping the description of the project constant.

I use regression models and double machine learning (Chernozhukov et al., 2017) as a robustness check to estimate the effect of the number of people on the profile picture on the amount of funds raised. Double machine learning (Chernozhukov et al., 2017) uses off-the-shelf machine learning algorithms to estimate causal effects in the presence of potentially high-dimensional confounders. Double machine learning allows us to control for confounders (e.g., document embeddings) without making strong assumptions about the functional form of our model.

In contrast to the majority of the extant research that has mostly found a negative effect on the victim group size on funds raised (Butts et al., 2019), I find no such negative effect, or even a positive effect, of the perceived victim group size on the amount of funds raised by crowdfunding projects. In contrast to past findings and in line with recent evidence from laboratory studies, it thus seems beneficial to increase the perceived victim group size in settings where potential donors evaluate multiple fundraisers at once.

Figure 1 Mean number of persons on the project profile picture by project category.



Notice: Error bars represent 95% confidence intervals. Figure A) with all projects, Figure B) only with projects that have at least one person on the project profile picture.

2 Literature review

2.1 Effect of victim group size on donations

There are two well-known effects of victim group size on donations, namely, the identifiable victim effect (Jenni & Loewenstein, 1997) and the compassion fade effect (Västfjäll et al., 2014). For both effects, the effect of the victim group size on the funds raised is negative. The identifiable victim effect refers to the tendency of individuals to provide more help to specific, identifiable victims than to anonymous (statistical) victims (Jenni & Loewenstein, 1997). For example, Kogut and Ritov (2005a) found that when asked to help sick children who need costly life-saving treatment, participants were more willing to contribute to a single child identified by age, name, and picture than to a single unidentified child or a group of unidentified children. While this effect results in people donating less to larger victim groups, it mainly operates, as the name implies, through the identification of the victims (Lee & Feeley, 2016). Indeed, a meta-analysis by Lee and Feeley (2016) found that this effect only works for a single victim and not for a group of victims. This finding is in line with the results of Kogut and Ritov (2005a) and Kogut and Ritov (2005b), who found that people donated more to a single identified victim than to a nonidentified victim, while there was no significant difference between donations made to a group of identified victims and those made to a group of nonidentified victims.

In contrast to the identifiable victim effect, the compassion fade effect specifies a negative effect of the depicted victim group size on willingness to donate that is directly caused by the size of the victim group. As mentioned by Butts et al. (2019), the compassion fade effect has also been referred to as compassion fatigue (Figley, 1995), compassion collapse (Cameron, 2017), and psychic numbing (Slovic, 2007). Butts et al. (2019) also highlighted that it is important to note that compassion here refers to compassionate behavior (e.g., donating) and not to the emotion of feeling compassion. Butts et al. (2019) analyzed 41 studies in a meta-analysis and found that victim group size negatively affected both helping intent and behavior (e.g., donations). They also found that anticipated positive affect and perceived impact, which were negatively associated with victim group size, mediated this effect.

One prominent explanation of the compassion fade effect is that it is caused by numeracy limitations and biases in the basic affective processing underlying the decision to help (Hamilton & Sherman, 1996; Slovic, 2007). Butts et al. (2019) called this the affective bias perspective. This explanation is related to the identifiable victim effect; it postulates that a single victim is depicted in more detail (i.e., with more information) than groups, which elicits stronger emotional reactions (Hamilton & Sherman, 1996). In contrast to a single victim, a group of victims constrains people's capacity for attention and imagery, which results in a fragmented representation of the victims and thus a weaker affective response (Dickert & Slovic, 2009; Kogut & Ritov, 2005a).

The other prominent explanation postulates that people expect the needs of large groups to be potentially overwhelming and therefore engage in emotion regulation to prevent themselves from experiencing these overwhelming emotions (Cameron & Payne, 2011). According to this explanation, people regulate their emotions to maximize their personal goals while minimizing potential costs that may seem overwhelming (Butts et al., 2019). Butts et al. (2019) called this the motivated choice perspective and noted that this explanation aligns with the cost-reward model of helping (Dovidio et al., 1991) and past work on empathy avoidance (Shaw et al., 1994). The abovementioned work demonstrated that potential donors regulate their emotions to avoid feelings that will compel them to help when helping is foreseen as being too costly.

To explain the reversion of the identifiable victim effect in joint evaluation, Erlandsson (2021) referred to X. Li and Hsee (2019), who posited that attributes in decision situations can differ in both justifiability (whether people think the attribute should affect decisions) and in evaluability (how easily the attribute in itself can be understood). Erlandsson (2021) noted that the size of the victim group is a prime example of an attribute with a high level of justifiability (most people would agree that helping more people is better than helping fewer) but a low level of evaluability (without any comparison, it is difficult to judge whether four victims are few or many). However, evaluability is higher in joint evaluations than in separate evaluations (Erlandsson, 2021; Hsee, 1996). Hsee (1996) noted that there might be a greater level of uncertainty in judging the value of a hard-to-evaluate attribute (e.g., victim group size) in separate evaluations than in joint evaluations. Therefore, these factors could have less impact in separate evaluations than in joint evaluations (Hsee, 1996). In line with this, Hsee, Zhang, Wang, et al. (2013) showed that willingness to donate when one could save 200 rather than 100 polar bears was twice as high in joint evaluation, while there was no difference in separate evaluation.

While the evidence for the compassion fade effect is substantial (Butts et al., 2019), this evidence rests on some limitations. Mainly, as mentioned by Garinther et al. (2022), only a few studies have used designs that required participants to jointly evaluate donation requests with different victim group sizes. The meta-analysis from Butts et al. (2019) explicitly excluded such studies. The authors acknowledged this shortcoming by stating that designs with "separate evaluations do not adequately reflect the realistic settings in which people make donation decisions" (p. 27). This leads to the second limitation, namely, to the best of my knowledge, this effect has never been tested using real-world donation data.

Regarding the first limitation, the few studies that have used joint evaluations to study the compassion fade effect have found inconsistent results (Butts et al., 2019). As mentioned by Garinther et al. (2022), only one of the studies replicated the compassion fade effect (Västfjäll et al., 2014, study 2). The other studies either found comparable donations to different victim group sizes (Kogut & Ritov, 2005b) or were even able to reverse the compassion fade effect (Hsee, Zhang, Lu, et al., 2013; Kogut & Ritov, 2005b). Garinther et al. (2022) took the existing inconsistencies in the design and results of these studies as motivation to systematically study the compassion fade effect in joint evaluation conditions. Over multiple studies, the authors showed that when subjects saw multiple donation requests at once, either simultaneously or sequentially, they donated more to larger victim groups.

This finding has important practical implications since it contradicts the compassion fade effect literature in a setting that "better mirror[s] real charitable giving contexts" (Butts et al., 2019, p. 27). Thus, in contrast to the majority of the extant literature on the relationship between victim group size and donations (Butts et al., 2019), fundraisers whose solicitations are evaluated jointly with other solicitations can potentially attract more donations by increasing the (perceived) victim group size. If this recommendation is externally valid, we should observe the following:

> *Hypothesis 1: There is a positive effect of perceived victim group size (i.e., number of persons on the project profile picture) on the amount of funds raised by the project.*

2.2 Success Factors of Donation based Crowdfunding Campaigns

To understand what drives a crowdfunding campaign's success we need to understand what motivates people to give to those campaigns and how crowdfunding campaigns can tap into these motivations. This short review largely draws on two excellent reviews by van Teunenbroek and Dalla Chiesa (2022) and van Teunenbroek et al. (2023). van Teunenbroek and Dalla Chiesa (2022) summarise motives that lead people to donate to crowdufunding campaigns. Many of the mechanisms that affect charitable giving in traditional contexts are also likely to play a role in crowdfunding campaigns (van Teunenbroek et al., 2023). Among these are altruism (Fehr & Fischbacher, 2003), the joy of giving (i.e., warm glow Andreoni, 1990) (i.e., warm glow) and solicitation (van Teunenbroek & Hasanefendic, 2023). van Teunenbroek and Dalla Chiesa (2022) also mention feeling part of a community as a motive. Project backers cannot only donate, but also comment on the project and share it on social media. Thus, by donating, donors can become part of a community (Josefy et al., 2017). The narrative of a project, communicated for example through the project description, can also motivate donors to contribute (van Teunenbroek & Dalla Chiesa, 2022). For example, Wang et al. (2022) found that a guilt-evoking project description positively influenced the willingness to donate. Campaign pictures also influence potential donors. As mentioned by van Teunenbroek and Dalla Chiesa (2022), Rhue and Robert (2018) found that campaigns that depict people with happy facial expressions raised more money than campaigns depicting people with neutral facial expressions. van Teunenbroek and Dalla Chiesa (2022) note that the crowdfunding environment is characterized by high uncertainty because there exists information asymmetry between the donors and the project initiators. According to van Teunenbroek and Dalla Chiesa (2022), people use quality signals to guide their behavior in such situations (van Teunenbroek et al., 2020). Therefore, the perceived quality of a project is positively related to its funding success (Mollick, 2014). Similarly, the number of campaign updates is also positively related to campaign success (Mollick, 2014). As mentioned by van Teunenbroek and Dalla Chiesa (2022), the perceived credibility of the project initiator also feeds into the perceived quality of a project. For example, Hörisch (2015) found that projects initiated by an officially recognised non-profit organization tend to be more successful. With crowdfunding being an inherently online based fundraising channel, social media plays an important role. Sharing a project on social media increases its visibility and therefore solicits potential donors (Bhati & McDonnell, 2020; Priante et al., 2022). Unsurprisingly, the number of social media shares is positively related to a project's success (Kubo et al., 2021).

van Teunenbroek et al. (2023) conducted a literature review about mechanisms that affect giving via philanthropic crowdfunding. Based on the review, they developed a conceputal model that specifies how these mechanisms mediate the effect of crowdfunding features (e.g., project description) on giving beahvior. The crowdfunding features they study are the project creator, social information, project description and rewards. A few of the mediating mechanisms were already mentioned above, namely the perceived credibility of the project initiator, the perceived quality of the project, the emotional reaction elicited by the projects as well as the identification with a community. van Teunenbroek et al. (2023) also specify the strength of the tie a donor has with the project initiator as a mechanism. I do not summarize the reward specific mechanisms since the crowdfunding platform studied in this study is not reward-based. Knowing how crowdfunding campaign factors affect the success of a campaign, we can now go on to discuss the identification strategy that we use to identify the effect of the perceived victim group size on the amount of funds raised by a project.

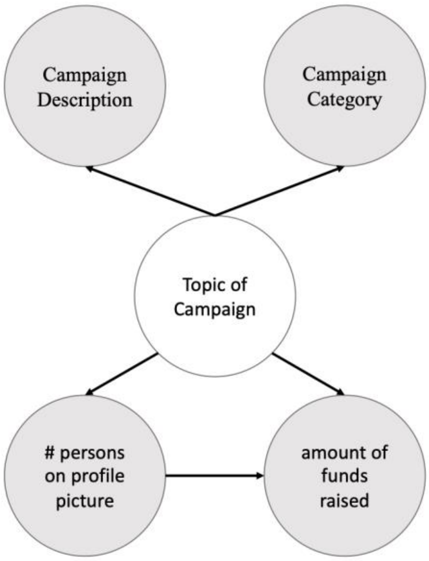
3 Methods

3.1 Data and Identification Strategy

I use data downloaded from GoFundMe to test the compassion fade effect in a real-life donation setting. Data from GoFundMe have been successfully used by researchers to study nonexperimenter solicited charitable contributions in a real-world setting (Sisco & Weber, 2019). In March 2022, I downloaded more than 60,000 fundraising projects from four countries (the US, the UK, Australia and Canada). When visiting GoFundMe.com, people see a grid of fundraising projects (see Figure A1 in the appendix). This grid displays the most important information for each project. One can see the location of the project, the project profile picture, the title of the project and the first few words of the description, the target amount to be raised, the funds already raised and when the last donation was made. Fundraising progress is visualized by a green progress bar. Although projects on GoFundMe have a target amount, GoFundMe follows a direct donation structure (van Teunenbroek et al., 2023) that allows initiators to keep the donated money, regardless of whether the target was reached or not. Raising more than the target amount is also possible. By clicking on a project from the project overview page, one is forwarded to the profile page of the project. On this page, project creators have the opportunity to display more photos and videos and to provide a detailed textual description of the project. From the project overview page and the profile page, I download the data that I need to test hypothesis 1 (see Table 1). Each project belongs to a category (see Figure 1), and up to 1,000 projects per category can be downloaded.

My identification strategy relies on the backdoor criterion. A set of variables Z satisfies the backdoor criterion relative to an ordered pair of variables (x, y) if (1) Z blocks every path between x and y that contains an arrow into x and (2) no node in Z is a descendant of x. Given our treatment variable (number of persons on a project profile picture) and our dependent variable (amount of funds raised), I assume that the topic of a fundraising campaign, which I measure with the campaign description text and the campaign category, satisfies the backdoor criterion. As already mentioned, the topic of a campaign likely influences both how many people are shown on the project profile picture and how much people will donate to the project. The directed acyclic graph (DAG) (Rohrer, 2018) that visualizes this assumption is shown in Figure 2. We can directly condition on the category of a campaign since each campaign is assigned to a category. To control for the campaign description text, I use natural language processing (NLP) methods to convert the text data to a numerical representation. There is a relatively new but growing stream of literature on using textual data as controls in statistical models (Keith et al., 2020). As mentioned by (Keith et al., 2020), there exist multiple ways of measuring confounders from text, such as lexicons, supervised classifiers, topic models and embeddings. As I want to control for the overall topic of a project, I use the latter two methods. These methods inductively learn confounding factors to (ideally) account for all known and unknown aspects of the text (Keith et al., 2020). Topic models are generative probabilistic models that represent text as a mixture of latent topics (Roberts et al., 2014). Embeddings represent text as low-dimensional, dense vectors that encode the meaning of the text. These numerical text representations are then used in place of the confounder (the topic of the campaign) in a causal adjustment method (e.g., linear regression) (Keith et al., 2020). As our identification strategy crucially depends on our ability to measure the confounders from text, I use multiple text representation methods. Namely, I use topic models (Roberts et al., 2014) and two state-of-the-art document embedding techniques (Le & Mikolov, 2014; Reimers & Gurevych, 2019). This approach allows us to see how sensitive our estimates are to different text representations. I describe these methods in more detail below.

Figure 2 Directed acyclic graph (DAG)



Notice: Directed acyclic graph (DAG) showing how I use the campaign category and campaign description to control for confounding between the number of persons on a project profile picture and the amount of funds raised by the project.

3.2 Control Variables

As just mentioned, I use the category of a fundraising project and the project description to identify the effect of the perceived victim group size on the amount of funds raised (i.e., to control for confounding). However, there are other variables that, although they are not confounders, can increase the precision of the estimate of interest (Cinelli et al., 2022). The section on the "success factors of donation based crowdfunding campaigns" motivates and informs the selection of these variables. Although the effects of these variables on campaign success were already discussed, I still briefly state the reason for inclusion when presenting the control variables. I include the total photos of a fundraising project, the number of updates posted and the length of the project description as control variables because studies have found that better documented projects raise more money (Wu et al., 2022). I also control for whether the fundraiser is organized by an organization or by people, whether it is a team fundraiser or not, and how many people are organizing the fundraiser. I include these controls because these variables likely affect the sharing of the fundraiser, which increases the visibility of the fundraiser and thus likely also the amount of funds raised (Kubo et al., 2021). For the same reason, I also created variables that control for whether the fundraiser was organized for anyone or not and if so whether it was organized for another person or to benefit an organization. I use the state-of-the-art named entity recognition model "ner-english" provided by the flair Python library to do this (Akbik et al., 2018).

To control for the popularity of the fundraiser, I include the number of times the fundraiser was shared on social media, the number of hearts (i.e., likes) the fundraiser collected and the number of comments that were made on the fundraiser project page as control variables. All of these variables likely affect the visibility of the fundraiser, which in turn should affect the amount of funds raised. For the same reason, I also control for the page position of the fundraiser in the category project overview page, as projects that appear on top of the page should receive more attention. Many of these control variables are also included to ensure that I control for the number of people who directly visit a fundraising project without browsing other projects beforehand. This approach is crucial since my hypothesis rests on the assumption that potential donors are in a joint evaluation context when they decide on which project they should donate to. Controlling for variables that affect the number of people who directly visit a fundraising project without browsing other projects before or after should ensure that donations that were made in a separate evaluation context do not affect the results. The target amount is also included as a control as it has been shown to be associated with campaign success (Mollick, 2014). For reasons that are obvious I also control for the days that passed since the project launched.

Finally, I also control for the emotions displayed by the people who are depicted on the project profile photo. I do this because the facial expressions of victims have been shown to affect giving behavior (Rhue & Robert, 2018; Small & Verrochi, 2009). Since this approach is only possible for projects that depict at least one face on the project profile page, I fit all models once without controlling for depicted facial emotions and once with controlling for depicted facial emotions.

When thinking about which controls to include, one must make sure that no "bad" controls are included (Cinelli et al., 2022). Colliders are an example of such bad controls. Conditioning on a collider, i.e., a common effect of the exposure and outcome, leads to a noncausal association between the exposure and the outcome (Cinelli et al., 2022; Hünermund et al., 2021). In my case, what I call the "social" variables (number of social media shares, number of comments, and number of campaign hearts) could potentially be such collider variables. People could share the fundraiser on social media, like, or comment because they are depicted on the fundraiser's project profile picture. Making a donation could also lead people to do these same things, which could make these variables serve as colliders between the number of people depicted on the project profile picture and the amount of funds raised. I therefore also fit the regression models without these three variables as controls (see appendix). However, I think that including these variables for the reasons stated above is more important, which is why these variables are included by default. However, the results for the models without these variables are very similar to those with these variables. In fact, the effect of the number of persons on the project profile picture is even stronger when not controlling for these variables. Because the estimated effects of these control variables are unlikely to have a causal interpretation, I follow previously made recommendations to not report them (Hünermund & Louw, 2020; Westreich & Greenland, 2013).

3.3 Natural Language Processing Methods

Embeddings embed text into a dense numerical space. There exist several methods to do this. These methods can roughly be divided into contextualized and noncontextualized methods. Word2vec (Mikolov et al., 2013) is a popular method used to obtain noncontextualized embeddings. Word2vec encodes words into numerical vectors by predicting a target word by its context words (or vice versa) with a shallow neural network. This prediction task is only a means to an end to obtain the weights of the neural network that are then used as the numerical vectors that represent a given target word. Using such a prediction task to obtain numerical representations has the advantage that words that are used in similar contexts end up having similar numerical representations (Camacho-Collados & Pilehvar, 2018). This process builds on the distributional hypothesis, which assumes that words that occur in similar contexts have similar meanings (Firth, 1957). After training word2vec on a (preferably large) corpus, one ends up with a vector of typically approximately 100-300 dimensions for each word of the corpus. To obtain a document vector, one can use the mean of all the word vectors that make up a document (Lau & Baldwin, 2016). Although word2vec uses the context of words to compute the word vectors, it does not assign different representations to the same word used in different contexts. For example, the word “bank” has the same numerical representation regardless of whether it is used in a financial context or not.

Contextualized models such as BERT (Devlin et al., 2018) alleviate this shortcoming by producing context-dependent embeddings for each word. These methods thus produce one word embedding for each unique word and each unique context the word appears in (Liu et al., 2020). Due to this approach, among other things, these models have achieved groundbreaking results in natural language understanding tasks (Rogers et al., 2020). Since BERT produces one embedding per word, we also need a way to aggregate those embeddings over the course of a document. I use Sentence BERT (SBERT) to do this (Reimers & Gurevych, 2019). Similar to averaging the word2vec embeddings per document, SBERT adds a pooling operation (mean) to the output of BERT to derive a fixed-sized document embedding. While contextualized models objectively perform better in most areas, they are not without limitations. First, because these models are very memory intensive, there is a limit to the length of text they can process. Second, since these methods are pretrained, they might fail to capture peculiarities of the text data at hand. I therefore use both of these methods to obtain embeddings of the project descriptions since word2vec can be trained on the data at hand and does not have a text length limit.

In addition to embeddings, I also use topic models to operationalize the topic of a campaign. Topic models are a popular method to detect latent topics in a collection of texts (Roberts et al., 2014). Topic models treat each document as a mixture of topics and each topic as a mixture of words. Latent Dirichlet allocation (LDA) is a popular method for fitting such topic models. I use the implementation provided by Roberts et al. (2014) to fit the topic model and use the method developed by Mimno and Lee (2014) to decide on the number of topics per topic model.

As evidenced by this short presentation of different methods for encoding the topic of a campaign from its campaign description text, these methods have differing strengths and weaknesses. For example, while the strength of embeddings is that they promise to encode all aspects of a text (e.g., meaning, affect, topic), their high level of dimensionality could complicate inference. The reverse is true for topic models; they only encode the topic(s) of the text but are often lower in dimension (i.e., number of topics) than embeddings. By using these different methods with differing strengths and weaknesses, we can verify how sensitive our estimates are to the type of text encoding.

3.4 Inference Methods

To estimate the effect of the number of persons on a campaign profile picture, I mainly rely on regression models and use double machine learning to assess the robustness of the results. I use double machine learning (Chernozhukov et al., 2017) as a robustness check because it allows us to control for confounders in a flexible (i.e., nonlinear) way. Double machine learning can be illustrated with a partially linear model in the following form:

$$Y = D\theta_0 + g_0(X) + \zeta,\ \mathbb{E}(\zeta \mid D, X) = 0, \tag{1}$$

$$D = m_0(X) + V,\ \mathbb{E}(V \mid X) = 0, \tag{2}$$

where Y is the outcome variable (i.e., the amount of funds raised by a project) and D is the treatment variable (i.e., the number of persons on a projects profile picture). The (potentially) high-dimensional vector X contains the confounding and control variables, and ζ and V are stochastic errors. Equation (1) is the equation of interest, and θ_0 is the main regression coefficient that we would like to infer. Assuming that D is conditionally exogenous, θ_0 has the interpretation of a structural or causal parameter. Equation (2) keeps track of confounding, i.e., the dependence of D on covariates. These covariates X affect the treatment variable D via the function $m_0(X)$ and the outcome variable via the function $g_0(X)$.

Applying machine learning methods directly to Equations (1) and (2) may have a very high level of bias, which is caused by the regularization properties of machine learning algorithms (Bach et al., 2021; Chernozhukov et al., 2017). Double machine learning uses orthogonalization to overcome this regularization bias. To illustrate this, we rewrite the abovementioned PLR model in the following residualized form:

$$W = V\theta_0 + \zeta,\ \ \mathbb{E}(\zeta \mid D, X) = 0, \tag{3}$$

$$W = \big(Y - l_0(X)\big), \quad l_0(X) = \mathbb{E}(Y \mid X) = 0, \tag{4}$$

$$V = \big(D - m_0(X)\big), \quad m_0(X) = \mathbb{E}(D \mid X) = 0, \tag{5}$$

Given identification, double machine learning then estimates l_0 and m_0 by $\widehat{l_0}$ and $\widehat{m_0}$ by solving the two problems of predicting Y and D using X. These prediction problems can be solved by using any off-the-shelf machine learning method. This gives us the following estimated residuals:

$$\widehat{W} = Y - \widehat{l_0}(X), \tag{6}$$

$$\hat{V} = D - \widehat{m_0}(X). \tag{7}$$

To avoid overfitting, these residuals are of a cross-validated form. We can then finally estimate θ_0 by regressing the residual $\widehat{W}$ on $\hat{V}$. Conventional inference is used for this final regression estimator. Double machine learning uses a method-of-moments estimator for θ_0 with a Neyman-orthogonal score function. This approach ensures that the moment condition used to identify and estimate θ_0 is insensitive to small perturbations of the nuisance functions (i.e., $\widehat{l_0}(X)$ and $\widehat{m_0}(X)$) estimated by the machine learning models. Although this ensures some robustness, a good approximation of the nuisance functions is still crucial. I therefore use three different machine learning algorithms, namely, regression trees (Therneau et al., 2015), random forests (Breiman, 2001) and XGBoost (T. Chen & Guestrin, 2016). I refer interested readers to Chernozhukov et al. (2017) for a detailed treatment of double machine learning.

To ensure that the machine learning methods can well approximate the nuisance functions, I train the methods via random search (Bergstra & Bengio, 2012) for 20 iterations each. To set the tuning space of the hyperparameters, I rely on the current advice from the literature (Bischl et al., 2023).

3.5 Person and facial emotion detection

Given the large number of project profile pictures, it is not feasible to detect the number of persons in a project profile picture and the facial expressions shown by these people by hand. I therefore rely on machine learning algorithms for this task. The state-of-the-art algorithms for these tasks are as good or even better than human raters while being significantly faster. To detect the number of persons in a project profile picture, I use a Faster R-CNN model (Ren et al., 2015) that was pretrained to detect persons in the COCO dataset (Lin et al., 2014). I use the mmdetection Python library (K. Chen et al., 2019) to implement the model. More information regarding this model is provided in the appendix.

To detect the facial emotions expressed by the people on the project profile picture, I use a model that is considered state of the art in facial expression detection at the time of performing the analysis (Savchenko, 2021). I use the hsemotion Python library to detect expressed facial emotions with this model. For each detected face, this model returns the probability that this face shows emotion x for a total of seven emotions (angry, disgust, fear, happiness, sadness, surprise, and neutral). To obtain one value per project profile picture, I take the mean per emotion when multiple faces are detected.

3.6 Fitted models and preprocessing

To ensure that the results are not driven by outliers, I follow previous literature that used data from GoFundMe (Sisco & Weber, 2019) and removed projects that raised more than the mean plus 3 standard deviations per country. The analyses conducted on the full sample are reported in the appendix. The results for the models with outliers are similar to those without outliers but tend to have larger standard errors. To account for heteroscedasticity, heteroskedasticity-consistent standard errors (HC1) are used. To ensure comparability to Garinther et al. (2022), I run the analysis once for all projects and once only for projects that have twelve or fewer people on the project profile picture. This approach also increases the comparability with other studies; M.-R. Li and Yin (2022) found that in studies that showed participants pictures of beneficiaries, showing eight beneficiaries was the most frequent approach. Since studies conducted in the lab usually have at least one person in the solicitation picture, I also run the analysis for all projects and only for projects with at least one person in the project profile picture.

4 Results

4.1 Descriptive Results

The data without outliers contain a total of 64,024 fundraising projects without any missing data in the variables listed in Table 1. This table also contains summary statistics for all of these variables. On average, these projects raised $4,390, with a standard deviation of $11,586. There are on average 2.6 persons depicted on the project profile picture, with a standard deviation of 5.5 persons. The majority of the projects (59.5%) have at least one person on the project profile picture. Only a few projects (4.8%) have more than twelve persons on the project profile picture. The majority of projects are thus comparable to the experimental literature with regard to the number of persons shown on the donation request (M.-R. Li & Yin, 2022).

The subset of projects with at least one detected face on the project profile picture contains 29,446 projects (Table 2). The mean facial emotions by project profile picture are mostly happy and neutral. The appendix shows the values given in Tables 1 and 2 for the full sample (i.e., with outliers). As I consider projects that raised more than the mean plus three standard deviations per country as outliers, 824 (1.27%) of the projects were excluded from the sample with all projects, and 464 (1.56%) projects were excluded from the sample where at least one face was detected on the project profile page.

Table 1 Summary Statistics of all projects without outliers.

Variable	NotNA	Mean	Median	Sd	Min	Max
Number of persons on campaign picture	64024	2.596	1	5.46	0	84
At least one person on campaign picture	64024					
... No	25927	40.5%				
... Yes	38097	59.5%				
Max. 12 people on campaign picture	64024					
... No	3060	4.8%				
... Yes	60964	95.2%				
Amount raised	64024	4389.74	768.416	11585.774	0	137690
Target amount	64024	11344.629	3475.118	35071.549	0.695	1108890
Created x days ago	64024	63.625	45	78.796	1	3089
Length of description (words)	64024	205.693	152	208.625	0	8256
Total updates of campaign	64024	0.634	0	2.303	0	117
Total photos of campaign	64024	1.846	1	3.254	0	218
Number of social media shares	64024	95.81	11	305.226	0	18445
Number of campaign hearts	64024	46.253	17	164.025	0	21028
Number of comments	64024	1.901	0	9.252	0	1074
Organized by	64024					
... an organization	3590	5.6%				
... a person	60434	94.4%				
Number of people organizing	64024	1.117	1	1.464	0	138
Team fundraiser	64024					
... No	58913	92%				
... Yes	5111	8%				
Organized for	64024					
... not organized for anyone	50801	79.3%				
... an organization	4785	7.5%				
... a person	8438	13.2%				
Country	64024					
... Australia	16114	25.2%				
... Canada	14574	22.8%				
... UK	17004	26.6%				
... USA	16332	25.5%				

Table 2 Summary Statistics of projects with at least one detected face without outliers.

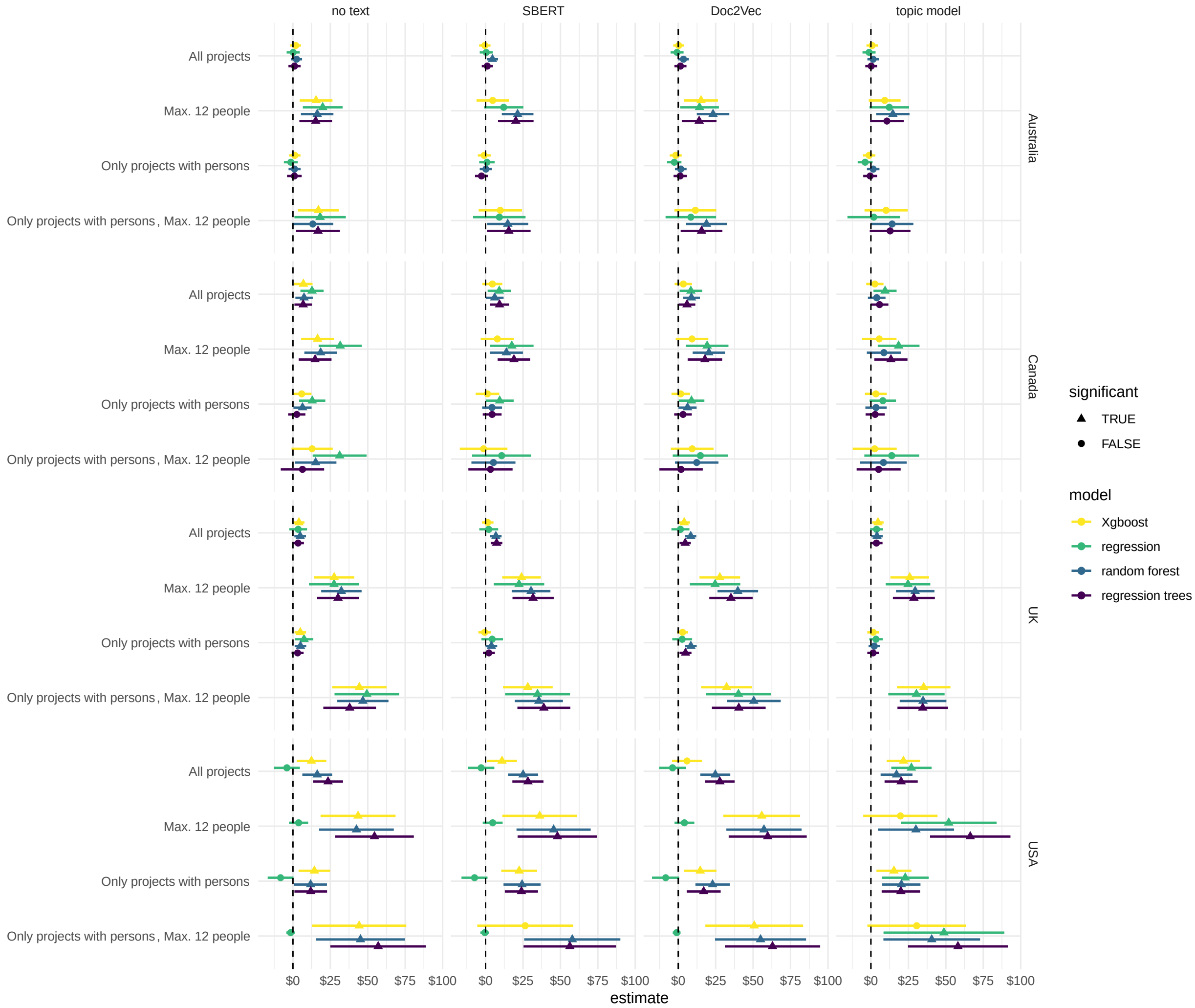
Variable	NotNA	Mean	Median	Sd	Min	Max
Number of persons on campaign picture	29446	4.533	2	6.754	0	80
At least one person on campaign picture	29446					
... No	811	2.8%				
... Yes	28635	97.2%				
Max. 12 people on campaign picture	29446					
... No	2536	8.6%				
... Yes	26910	91.4%				
Mean facial emotion: angry	29446	0.062	0.019	0.113	0	0.99
Mean facial emotion: disgust	29446	0.048	0.015	0.096	0	0.981
Mean facial emotion: fear	29446	0.025	0.003	0.064	0	0.964
Mean facial emotion: happy	29446	0.559	0.608	0.371	0	1
Mean facial emotion: sad	29446	0.063	0.021	0.108	0	0.98
Mean facial emotion: surprise	29446	0.047	0.011	0.09	0	0.965
Mean facial emotion: neutral	29446	0.195	0.119	0.212	0	0.972
Amount raised	29446	3405.149	997.359	6919.25	0	72677
Target amount	29446	11562.584	3890.6	32217.524	0.695	800865
Created x days ago	29446	60.297	42	80.774	1	3089
Length of description (words)	29446	223.752	168	220.584	1	8256
Total updates of campaign	29446	0.69	0	2.413	0	117
Total photos of campaign	29446	1.878	1	3.424	0	218
Number of social media shares	29446	132.291	36	351.934	0	16599
Number of campaign hearts	29446	57.962	24	115.497	0	4972
Number of comments	29446	2.39	1	5.691	0	214
Organized by	29446					
... an organization	1229	4.2%				
... a person	28217	95.8%				
Number of people organizing	29446	1.132	1	1.273	0	105
Team fundraiser	29446					
... No	26871	91.3%				
... Yes	2575	8.7%				
Organized for	29446					
... not organized for anyone	22742	77.2%				
... an organization	1612	5.5%				
... a person	5092	17.3%				
Country	29446					
... Australia	6826	23.2%				
... Canada	6397	21.7%				
... UK	6989	23.7%				
... USA	9234	31.4%				

4.2 Effect of perceived victim group size on funds raised

Figure 3 shows the effect of the number of persons on a project's profile picture on the amount of funds raised by a project. For each country, text control method, and subset of data, four models are fitted (three double machine learning models and one regression model). I first present the results that consider all projects (top row for each country panel). For projects in Australia, the number of people on the picture has no significant effect on the amount of funds raised. For Canada, there is a small but significant effect for most of the models. For the UK, most models indicate that there is no significant effect. For US-based projects, most models indicate a positive effect; however, compared to the other countries, there is more variance in the estimates between the different models. In contrast to the double machine learning results, the regression results indicate no positive effect in all but the topic model estimate.

Restricting the analysis to projects that show no more than twelve people on the project profile picture increases the effect of the number of people on the picture on the amount of funds raised across all countries. For most countries and models, the estimates show that an additional person on the project profile picture leads to an increase in the amount of funds raised of approximately $25. The models presented in the last two rows for each country only consider projects with at least one person on the project profile picture, namely, once for all projects where this condition is met and once further restricted to those with a maximum of twelve people on the picture. The results are very similar to those models where all projects are considered. Overall, these results suggest that there is a small, but not consistent, effect of the number of persons on the project profile picture on the amount of funds raised. The positive effect is most consistent for the subset of projects with at most twelve people on the project profile picture. Most importantly, not one of the estimates of the 256 fitted models shown in Figure 3 indicates a significant negative effect of the number of persons on the project profile picture on the amount of funds raised.

Figure 3 Effect of the number of persons on a project's profile picture on the amount of funds raised.

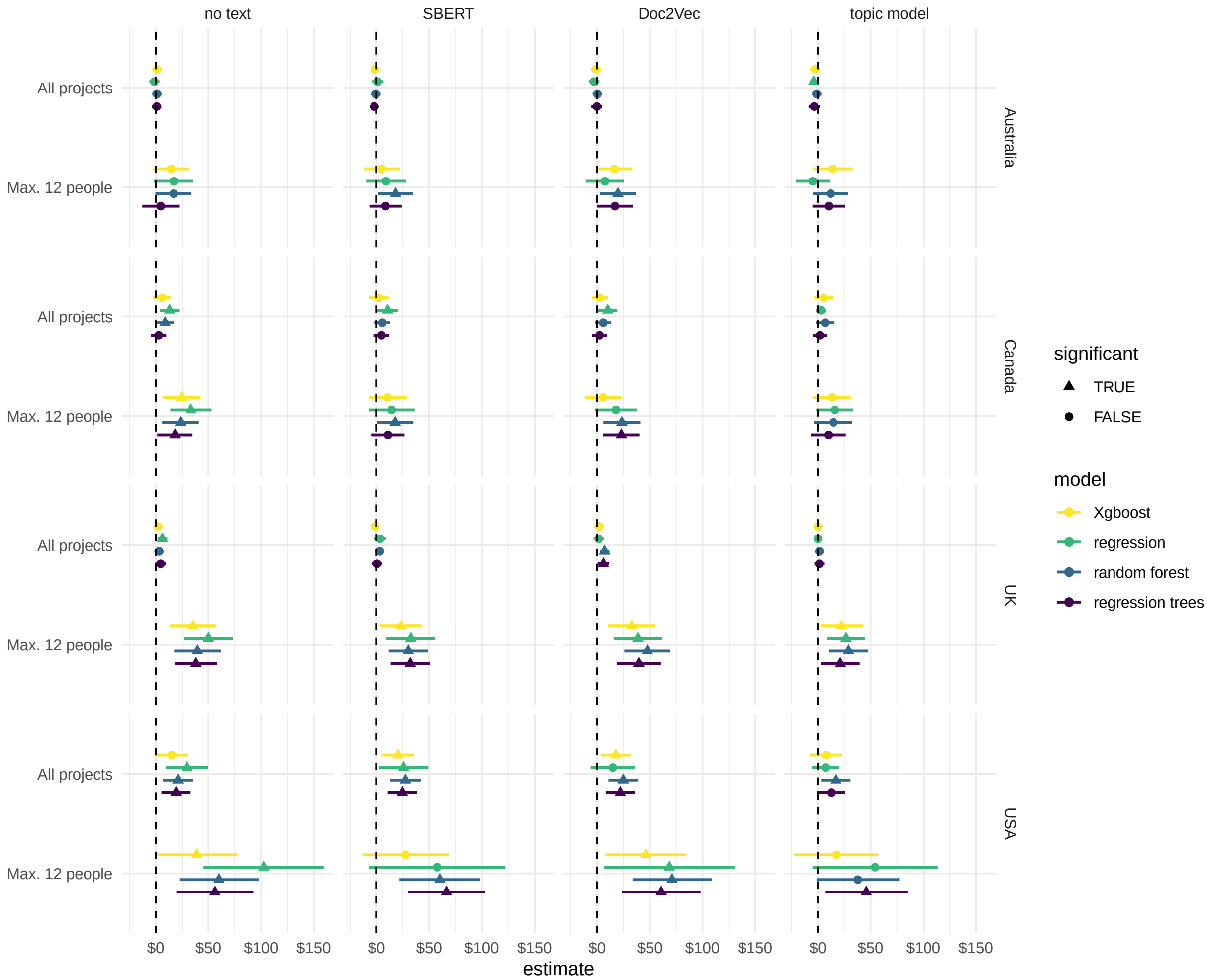


Notice: Error bars denote 95% confidence intervals.

Figure 4 shows the same analysis as that shown in Figure 3, but it includes models that also account for the facial emotion expressed by the people on the project profile picture. The results are very similar to the results reported in Figure 3. The models that consider all projects within this subsample mostly indicate no significant effect, except for projects posted in the US, where the majority of models indicate a positive effect. As in Figure 3, the effects become larger when further restricting the sample to projects with a maximum of twelve people on the project profile picture. Most models still indicate no effect for projects posted in Australia and Canada. However, for the UK, there is now a consistent positive effect and a mostly consistent positive effect for the US.

The fact that models with projects that have at most twelve people on them show larger effects could indicate a nonlinear (i.e., concave) effect of the number of people on a project profile picture on the amount of funds raised. I therefore also fit the models shown in Figures 3 and 4 with a quadratic effect of the number of persons on a project's profile picture. I only use ordinary least squares regression to fit these models because the results of the regression and double machine learning estimates are very similar.

Figure 4 Effect of the number of persons on a project's profile picture on the amount of funds raised for models that control for the facial emotions expressed by the persons.



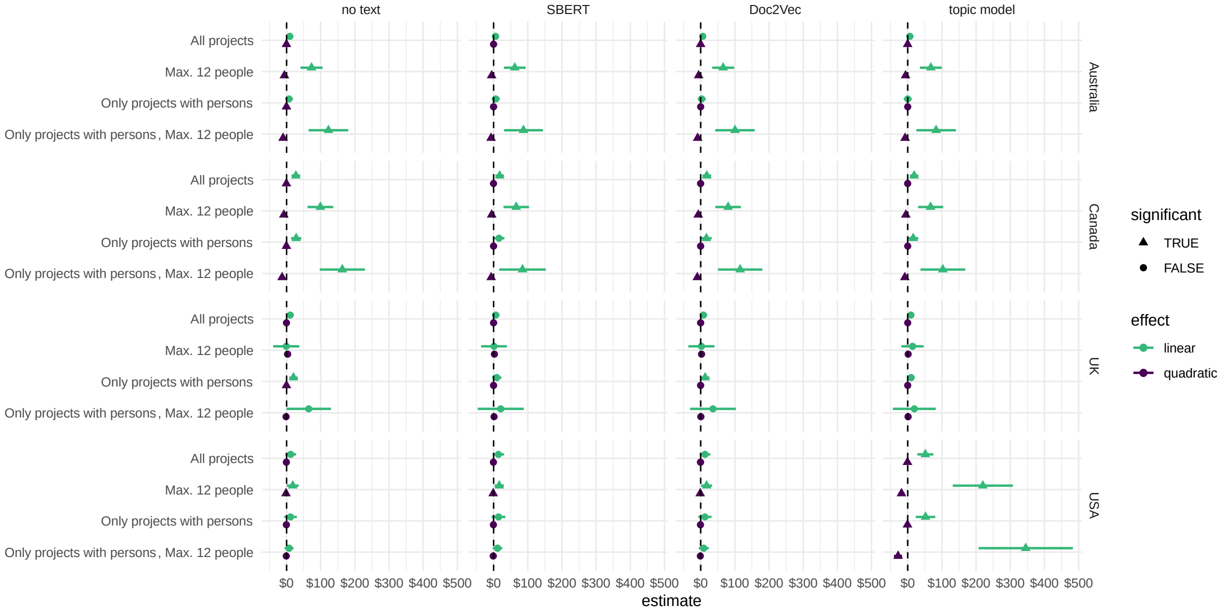
Notice: Error bars denote 95% confidence intervals.

Figure 5 shows the results of the same models as those fitted in Figure 3, but it includes an additional quadratic term for the number of persons on a project profile picture. In the majority of the models, the quadratic term has a small but significant negative effect. Compared to Figure 3, the linear effect of the number of persons on the project profile picture seems to be larger.

Looking at the models that also control for the emotions expressed by the people on the project profile picture (Figure 6), we see a similar result. In most of the models, the quadratic effect is significantly negative. Except for projects from the UK, there is a significant positive linear effect of the number of people on the project profile picture on the amount of funds raised. The effect of an additional person on the project profile picture on the amount of funds raised is considerable (approximately $100 for projects in Australia and Canada and around $500 for projects in the US for models fitted on the subgroup of projects with a maximum of twelve people on the project profile picture).

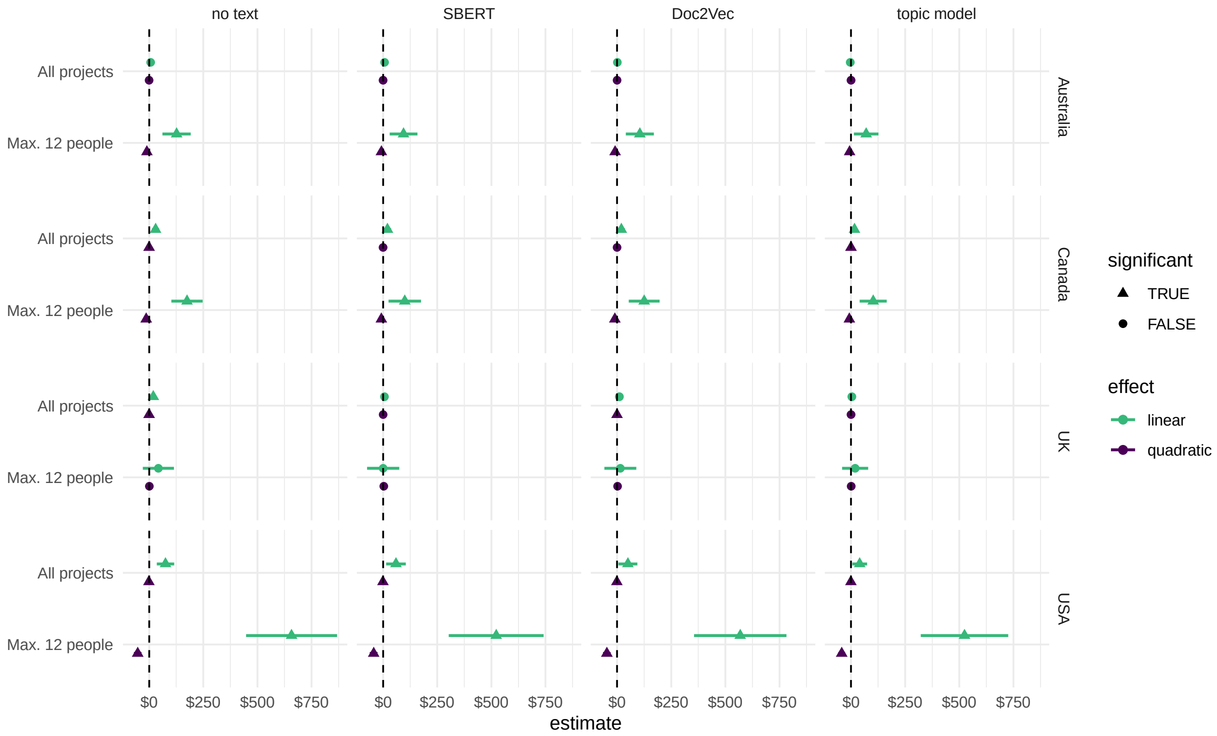
Looking at the results with regards to hypothesis 1, I thus find that the postulated positive effect of the perceived victim group size (i.e., number of persons on the project profile picture) on the amount of funds raised by the project is mixed. I found the most consistent positive effect for models that were fitted on the subset of projects with a maximum of twelve people on the project profile picture. In the estimates obtained with the double machine learning estimators, projects in the UK and the US show the most consistent positive effect. Interestingly, the effect vanishes for projects from the UK in the models where a nonlinear effect was included and the facial emotions were controlled for (Figure 6). However, all other models with this specification for the other three countries show a significant positive effect of the number of people on the project profile picture on the amount of funds raised. I can thus confirm hypothesis 1 for all but projects from the UK.

Figure 5 Linear and nonlinear effect of the number of persons on a project's profile picture on the amount of funds raised.



Notice: Error bars denote 95% confidence intervals.

Figure 6 Linear and nonlinear effect of the number of persons on a project's profile picture on the amount of funds raised for models that do control for the facial emotions expressed by the persons.



Notice: Error bars denote 95% confidence intervals.

5 Discussion

Crowdfunding has great promise both for individual and institutional fundraisers (Alexiou et al., 2020). However, most fundraising projects fail to reach their targets (Kenworthy & Igra, 2022). In this study I explored whether this lack of success could in part be explained by the different characteristics of crowdfunding and more traditional fundraising (e.g., mail solicitation). Among other things, crowdfunding platforms differ from more traditional means of fundraising in that potential donors are in a joint evaluation context, i.e., they can choose among a large number of projects to donate to. In contrast, traditional fundraising often occurs in a separate evaluation context, where potential donors often face only one donation request at a time. Importantly for fundraisers, laboratory studies have shown that the effect of the victim group size on donations reverses when going from a separate to a joint evaluation context (Erlandsson, 2021; Garinther et al., 2022). In contrast to results obtained in a separate evaluation context, people in a joint evaluation context donate more to larger victim groups (Erlandsson, 2021; Garinther et al., 2022). Using data from over 60,000 GoFundMe crowdfunding projects from four countries, I tested whether this effect generalizes to a real-world setting (i.e., crowdfunding). I did this by testing the effect of the number of people on a project's profile picture (i.e., perceived victim group size) on the amount of funds raised by a project. Consistent with recent evidence from the lab that tested this effect in a joint evaluation condition (Garinther et al., 2022), I found a mostly significant positive effect of the number of people depicted on the project profile picture on the amount of funds raised by the project for the subgroup of projects that is most similar to the stimuli used in the laboratory (i.e., maximum twelve people on the picture). The results of the models fitted on the full sample were more mixed, and only a few of them indicated a significant positive effect. This discrepancy calls for more experimental studies that vary the range of victims beyond what has mostly been done in the literature (e.g., a maximum of twelve victims).

Most of the models that include the quadratic effect indicate a concave effect of the perceived victim group size on the amount of funds raised. This has not yet been found in experimental studies. However, the nonlinear effect is rather small; therefore, laboratory studies might lack the power to detect this effect. This nonlinear effect might indicate that even in a joint evaluation setting, people might be prone to affective biases. The affective bias perspective denotes that in separate evaluations, people's numeracy limitations and biases in affective processing (Hamilton & Sherman, 1996; Slovic, 2007) might be responsible for the compassion fade effect (Butts et al., 2019). My results show that these biases could still affect decision-making in a joint evaluation setting but are trumped by attributes that have a high level of justifiability (i.e., victim group size) (Erlandsson, 2021). Thus, to attract more donations from people who are browsing projects on GoFundMe, it is still beneficial to include more rather than fewer people on the project profile picture. This advice is relative to the other projects that share the same category (e.g., medical, sports). As shown in Figure 1, the variance of the number of people depicted on the profile picture is rather small within projects that share the same category. My results hold for the typical range of depicted persons per category and do not necessarily extrapolate beyond that range. I thus advise people who want to raise money in a joint evaluation context to increase the victim group size by a sensible amount. For example, when raising money for a sick child, show the whole family (but not, e.g., the whole child's school class) instead of only the child. Assuming that people use the (perceived) victim group size (high justifiability) (Erlandsson, 2021) to choose from among similar fundraising projects but are also affected by affective biases to some extent (Butts et al., 2019), having a marginally higher victim size than that of other similar projects should be most effective. The assumptions used to derive this advice are consistent with my results and with the evidence quoted by Erlandsson (2021) that shows that emotional reactions are more predictive of attitudes toward policies in separate evaluations (Ritov & Baron, 2011), while efficiency-related attributes are more predictive in joint evaluations (Bazerman et al., 2011; Caviola et al., 2014). Future research could test whether there are other fundraising related attributes that have differing effects depending on whether people evaluate them in a separate or joint evaluation context. For example, the overhead ratio is an efficiency-related attribute that is difficult to evaluate in separate evaluation and should thus receive more weight in a joint evaluation context.

My results are not without limitations. First, as with any observational study aiming to draw causal conclusions, my results crucially depend on the identifying assumptions. I use the project category and the project description to control for confounders between the number of people shown on the project profile picture and the amount of funds raised. I also use a number of other control variables that should help with identifying the variance needed to draw causal conclusions. However, it is possible that I have omitted a confounder or included a bad control. While this possibility cannot be ruled out, the fact that the results are consistent with results from the laboratory and that the results also hold for the models without the (potentially) bad controls indicates that the results are hopefully not affected by such a possibility. Second, I rely heavily on machine learning algorithms to conduct this study. While these algorithms are already very good, the field is developing rapidly. Methods that use text data to adjust for confounding are constantly evolving (Feder et al., 2022). While I use three different methods to control for the topic of a fundraiser (i.e., campaign description), these methods will soon be surpassed by newer and better methods. The third limitation concerns the data. While it covers countries from three different continents, my analysis is still restricted to a Western sample. Due to cultural differences in how crowdfunding projects are set up (Cho & Kim, 2017), my results do not necessarily generalize to, e.g., more collectivist cultures (Nie et al., 2022). Future research could replicate my results in such cultures.

While experimental studies are immune to some of these limitations, there is only so much we can learn about real-world behavior from laboratory studies (Levitt & List, 2007). Experiments can provide evidence for a causal relationship between X and Y within the constraints of the experimental design, but they are often unable to shed light on the strength of an effect in everyday natural settings (Diener et al., 2022). As argued by Diener et al. (2022), a research program is most successful when experiments are integrated with other methods rather than being considered the sole source of valid information. I follow this perspective and use evidence and theories derived from laboratory studies as the basis for this study to test whether those findings are generalizable to the field. Even though the results from these laboratory studies have direct and easy-to-implement implications for fundraisers, field studies that test whether these effects are actually generalizable are still missing from the literature. This is in contrast to the broader fundraising literature, where field studies are not uncommon (e.g., Alston et al., 2021; Woods et al., 2023). The lack of evidence from field data applies to psychological research more generally (Diener et al., 2022; Grosz et al., 2020). I hope that this study helps to alleviate this shortcoming by showing the usefulness of observational field data to complement and extend findings from the laboratory.

6 References

Akbik, A., Blythe, D., & Vollgraf, R. (2018). Contextual string embeddings for sequence labeling. *COLING 2018, 27th International Conference on Computational Linguistics*, 1638–1649.

Alexiou, K., Wiggins, J., & Preece, S. B. (2020). Crowdfunding acts as a funding substitute and a legitimating signal for nonprofit performing arts organizations. *Nonprofit and Voluntary Sector Quarterly*, *49*(4), 827–848. https://doi.org/10.1177/0899764020908338

Alston, M., Eckel, C., Meer, J., & Zhan, W. (2021). High-capacity donors' preferences for charitable giving. *Nonprofit and Voluntary Sector Quarterly*, *50* (6), 1262–1273. https://doi.org/10.1177/08997640211003251

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The economic journal*, *100* (401), 464–477. https://doi.org/10.2307/2234133

Bach, P., Chernozhukov, V., Kurz, M. S., & Spindler, M. (2021). DoubleML–An Object-Oriented Implementation of Double Machine Learning in R. *arXiv preprint arXiv:2103.09603*. https://doi.org/10.48550/arXiv.2103.09603

Bazerman, M. H., Gino, F., Shu, L. L., & Tsay, C.-J. (2011). Joint evaluation as a real-world tool for managing emotional assessments of morality. *Emotion review*, *3*(3), 290–292. https://doi.org/10.1177/1754073911402370

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, *13*(2).

Bhati, A., & McDonnell, D. (2020). Success in an online giving day: The role of social media in fundraising. *Nonprofit and Voluntary Sector Quarterly*, *49* (1), 74–92. https://doi.org/10.1177/0899764019868849

Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A.-L., Deng, D., & Lindauer, M. (2023). Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. *WIREs Data Mining and Knowledge Discovery*, e1484. https://doi.org/10.1002/widm.1484

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*(4), 77–84. https://doi.org/10.1145/2133806.2133826

Breiman, L. (2001). Random forests. *Machine learning*, *45* (1), 5–32. https://doi.org/10.1023/A:1010933404324

Butts, M. M., Lunt, D. C., Freling, T. L., & Gabriel, A. S. (2019). Helping one or helping many? A theoretical integration and meta-analytic review of the compassion fade literature. *Organizational Behavior and Human Decision Processes*, *151*, 16–33. https://doi.org/10.1016/j.obhdp.2018.12.006

Camacho-Collados, J., & Pilehvar, M. T. (2018). From word to sense embeddings: A survey on vector representations of meaning. *Journal of Artificial Intelligence Research*, *63*, 743–788. https://doi.org/10.1613/jair.1.11259

Cameron, C. D. (2017). *Compassion collapse: Why we are numb to numbers* (Vol. 261, No. 271). Oxford University Press.

Cameron, C. D., & Payne, B. K. (2011). Escaping affect: How motivated emotion regulation creates insensitivity to mass suffering. *Journal of personality and social psychology*, *100*(1), 1. https://doi.org/10.1037/a0021643

Caviola, L., Faulmüller, N., Everett, J. A., Savulescu, J., & Kahane, G. (2014). The evaluability bias in charitable giving: Saving administration costs or saving lives? *Judgment and decision making*, *9*(4), 303–315. https://doi.org/10.1017/S1930297500006185

Chen, K., Wang, J., Pang, J., Cao, Y., Xiong, Y., Li, X., Sun, S., Feng, W., Liu, Z., Xu, J., et al. (2019). Mmdetection: Open mmlab detection toolbox and benchmark. *arXiv preprint arXiv:1906.07155*. https://doi.org/10.48550/arXiv.1906.07155

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., & Newey, W. (2017). Double/debiased/neyman machine learning of treatment effects. *American Economic Review*, *107*(5), 261–65. https://doi.org/10.1257/aer.p20171038

Cho, M., & Kim, G. (2017). A cross-cultural comparative analysis of crowdfunding projects in the United States and South Korea. *Computers in Human Behavior*, *72*, 312–320. https://doi.org/10.1016/j.chb.2017.03.013

Cinelli, C., Forney, A., & Pearl, J. (2022). A crash course in good and bad controls. *Sociological Methods & Research*, 00491241221099552. https://doi.org/10.1177/00491241221099552

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*. https://doi.org/10.48550/arXiv.1810.04805

Dickert, S., & Slovic, P. (2009). Attentional mechanisms in the generation of sympathy. *Judgment and Decision making*, *4*(4), 297–306. https://doi.org/10.1017/S1930297500003879

Diener, E., Northcott, R., Zyphur, M. J., & West, S. G. (2022). Beyond experiments. *Perspectives on Psychological Science*, *17*(4), 1101–1119. https://doi.org/10.1177/17456916211037670

Dovidio, J. F., Piliavin, J. A., Gaertner, S. L., Schroeder, D. A., & Clark III, R. D. (1991). The arousal: Cost-reward model and the process of intervention: A review of the evidence. Sage Publications, Inc.

Erlandsson, A. (2021). Seven (weak and strong) helping effects systematically tested in separate evaluation, joint evaluation and forced choice. *Judgment and decision making*, *16*(5), 1113–1154. https://doi.org/10.1017/S1930297500008378

Feder, A., Keith, K. A., Manzoor, E., Pryzant, R., Sridhar, D., Wood-Doughty, Z., Eisenstein, J., Grimmer, J., Reichart, R., Roberts, M. E., et al. (2022). Causal inference in natural language processing: Estimation, prediction, interpretation and beyond. *Transactions of the Association for Computational Linguistics*, *10*, 1138–1158. https://doi.org/10.1162/tacl_a_00511

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature*, *425* (6960), 785–791. https://doi.org/10.1038/nature02043

Figley, C. R. (1995). Compassion fatigue: Toward a new understanding of the costs of caring. In B. H. Stamm (Ed.), *Secondary traumatic stress: Self-care issues for clinicians, researchers, and educators* (pp. 3–28). The Sidran Press.

Firth, J. (1957). A synopsis of linguistic theory, 1930-1955. *Studies in linguistic analysis*, 10-32.

Garinther, A., Arrow, H., & Razavi, P. (2022). Victim number effects in charitable giving: Joint evaluations promote egalitarian decisions. *Personality and Social Psychology Bulletin*, *48*(1), 34–48. https://doi.org/10.1177/0146167220982734

GoFundMe. (2022). *Gofundme's 2021 giving report*. Retrieved December 20, 2022, from https://www.gofundme.com/c/gofundme-giving-report-2021

Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The taboo against explicit causal inference in nonexperimental psychology. *Perspectives on Psychological Science*, *15*(5), 1243–1255. https://doi.org/10.1177/1745691620921521

Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological review*, *103*(2), 336.

Hart, T. R. (2002). Ephilanthropy: Using the internet to build support. *International Journal of Nonprofit and Voluntary Sector Marketing*, *7* (4), 353–360. https://doi.org/10.1002/nvsm.192

Hörisch, J. (2015). Crowdfunding for environmental ventures: An empirical analysis of the influence of environmental orientation on the success of crowdfunding initiatives. *Journal of cleaner production*, *107*, 636–645. https://doi.org/10.1016/j.jclepro.2015.05.046

Hsee, C. K. (1996). The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organizational behavior and human decision processes*, *67*(3), 247–257. https://doi.org/10.1006/obhd.1996.0077

Hsee, C. K., Zhang, J., Lu, Z. Y., & Xu, F. (2013). Unit asking: A method to boost donations and beyond. *Psychological science*, *24*(9), 1801–1808. https://doi.org/10.1177/0956797613482947

Hsee, C. K., Zhang, J., Wang, L., & Zhang, S. (2013). Magnitude, time, and risk differ similarly between joint and single evaluations. *Journal of Consumer Research*, *40*(1), 172–184. https://doi.org/10.1086/669484

Hünermund, P., & Louw, B. (2020). On the nuisance of control variables in regression analysis. *arXiv preprint arXiv:2005.10314*. https://doi.org/10.48550/arXiv.2005.10314

Hünermund, P., Louw, B., & Caspi, I. (2021). Double machine learning and automated confounder selection – a cautionary tale. *arXiv preprint arXiv:2108.11294*. https://doi.org/10.48550/ARXIV.2108.11294

Jenni, K., & Loewenstein, G. (1997). Explaining the identifiable victim effect. *Journal of Risk and uncertainty*, *14*(3), 235–257. https://doi.org/10.1023/A:1007740225484

Josefy, M., Dean, T. J., Albert, L. S., & Fitza, M. A. (2017). The role of community in crowdfunding success: Evidence on cultural attributes in funding campaigns to "save the local theater". *Entrepreneurship Theory and Practice*, *41*(2), 161–182. https://doi.org/10.1111/etap.12263

Keith, K. A., Jensen, D., & O'Connor, B. (2020). Text and causal inference: A review of using text to remove confounding from causal estimates. *arXiv preprint arXiv:2005.00649*. https://doi.org/10.48550/arXiv.2005.00649

Kenworthy, N., & Igra, M. (2022). Medical crowdfunding and disparities in health care access in the United States, 2016–2020. *American Journal of Public Health*, *112* (3), 491–498. https://doi.org/10.2105/AJPH.2021.306617

Kogut, T., & Ritov, I. (2005a). The "identified victim" effect: An identified group, or just a single individual? *Journal of Behavioral Decision Making*, *18* (3), 157–167. https://doi.org/10.1002/bdm.492

Kogut, T., & Ritov, I. (2005b). The singularity effect of identified victims in separate and joint evaluations. *Organizational behavior and human decision processes*, *97* (2), 106–116. https://doi.org/10.1016/j.obhdp.2005.02.003

Kubo, T., Verıssimo, D., Uryu, S., Mieno, T., & MacMillan, D. (2021). What determines the success and failure of environmental crowdfunding? *Ambio*, *50*, 1659–1669. https://doi.org/10.1007/s13280-021-01522-0

Kuo, Y.-F., Lin, C. S., & Liu, L.-T. (2022). The effects of framing messages and cause-related marketing on backing intentions in reward-based crowdfunding. *Journal of Retailing and Consumer Services*, *64*, 102799. https://doi.org/10.1016/j.jretconser.2021.102799

Lau, J. H., & Baldwin, T. (2016). An empirical evaluation of doc2vec with practical insights into document embedding generation. *arXiv preprint arXiv:1607.05368*. https://doi.org/10.48550/arXiv.1607.05368

Le, Q., & Mikolov, T. (2014). Distributed representations of sentences and documents. *International conference on machine learning*, 1188–1196.

Lee, S., & Feeley, T. H. (2016). The identifiable victim effect: A meta-analytic review. *Social Influence*, *11*(3), 199–215. https://doi.org/10.1080/15534510.2016.1216891

Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic perspectives*, *21*(2), 153–174. https://doi.org/10.1257/jep.21.2.153

Li, M.-R., & Yin, C.-Y. (2022). Facial expressions of beneficiaries and donation intentions of potential donors: Effects of the number of beneficiaries in charity advertising. *Journal of Retailing and Consumer Services*, *66*, 102915. https://doi.org/10.1016/j.jretconser.2022.102915

Li, X., & Hsee, C. K. (2019). Beyond preference reversal: Distinguishing justifiability from evaluability in joint versus single evaluations. *Organizational Behavior and Human Decision Processes*, *153*, 63–74. https://doi.org/10.1016/j.obhdp.2019.04.007

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft coco: Common objects in context. *European conference on computer vision*, 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

Liu, Q., Kusner, M. J., & Blunsom, P. (2020). A survey on contextual embeddings. *arXiv preprint arXiv:2003.07278*.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Mimno, D., & Lee, M. (2014). Low-dimensional embeddings for interpretable anchor-based topic inference. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1319–1328.

Mollick, E. (2014). The dynamics of crowdfunding: An exploratory study. *Journal of business venturing*, *29*(1), 1–16. https://doi.org/10.1016/j.jbusvent.2013.06.005

Nie, L., Chan, K. N., & Lam, W.-F. (2022). Social expectations for charitable giving in china. *Nonprofit and Voluntary Sector Quarterly*, 08997640221100721. https://doi.org/10.1177/08997640221100721

Priante, A., Ehrenhard, M. L., van den Broek, T., Need, A., & Hiemstra, D. (2022). "mo" together or alone? investigating the role of fundraisers' networks in online peer-to-peer fundraising. *Nonprofit and Voluntary Sector Quarterly*, *51*(5), 986–1009. https://doi.org/10.1177/08997640211057456

Reimers, N., & Gurevych, I. (2019). Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084*.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, *28*.

Rhue, L., & Robert, L. P. (2018). Emotional delivery in pro-social crowdfunding success. *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3170427.3188534

Ritov, I., & Baron, J. (2011). Joint presentation reduces the effect of emotion on evaluation of public actions. *Cognition and Emotion*, *25* (4), 657–675. https://doi.org/10.1080/02699931.2010.512512

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American journal of political science*, *58*(4), 1064–1082. https://doi.org/10.1111/ajps.12103

Rogers, A., Kovaleva, O., & Rumshisky, A. (2020). A primer in bertology: What we know about how bert works. *Transactions of the Association for Computational Linguistics*, *8*, 842–866. https://doi.org/10.1162/tacl_a_00349

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in methods and practices in psychological science*, *1*(1), 27–42. https://doi.org/10.1177/2515245917745629

Ruehle, R. C., Engelen, B., & Archer, A. (2021). Nudging charitable giving: What (if anything) is wrong with it? *Nonprofit and Voluntary Sector Quarterly*, *50* (2), 353–371. https://doi.org/10.1177/0899764020954266

Savchenko, A. V. (2021). Facial expression and attributes recognition based on multi-task learning of lightweight neural networks. *2021 IEEE 19th International Symposium on Intelligent Systems and Informatics (SISY)*, 119–124. https://doi.org/10.1109/SISY52375.2021.9582508

Shaw, L. L., Batson, C. D., & Todd, R. M. (1994). Empathy avoidance: Forestalling feeling for another in order to escape the motivational consequences. *Journal of Personality and Social Psychology*, *67*(5), 879. https://doi.org/10.1037/0022-3514.67.5.879

Sisco, M. R., & Weber, E. U. (2019). Examining charitable giving in real-world online donations. *Nature communications*, *10*(1), 1–8. https://doi.org/10.1038/s41467-019-11852-z

Slovic, P. (2007). If I look at the mass I will never act: Psychic numbing and genocide. *Judgment and Decision Making*, *2*(2), 79–95. https://doi.org/10.1017/S1930297500000061

Small, D. A., & Verrochi, N. M. (2009). The face of need: Facial emotion expression on charity advertisements. *Journal of marketing research*, *46* (6), 777–787.

Therneau, T., Atkinson, B., & Ripley, B. (2015). Rpart: Recursive partitioning and regression trees. *R package version*, *4*, 1–9.

van Teunenbroek, C., & Dalla Chiesa, C. (2022). Motives for giving via crowdfunding in eight categories: A review and theoretical framework of donation- and reward-based crowdfunding motives. *International Crowdfunding Research Conference*, 1.

van Teunenbroek, C., Bekkers, R., & Beersma, B. (2020). Look to others before you leap: A systematic literature review of social information effects on donation amounts. *Nonprofit and Voluntary Sector Quarterly*, *49* (1), 53–73. https://doi.org/10.1177/0899764019869537

van Teunenbroek, C., Chiesa, C. D., & Hesse, L. (2023). The contribution of crowdfunding for philanthropy: A systematic review and framework of donation and reward crowdfunding. *Journal of Philanthropy and Marketing*, e1791. https://doi.org/10.1002/nvsm.1791

van Teunenbroek, C., & Hasanefendic, S. (2023). Researching the crowd: Implications on philanthropic crowdfunding and donor characteristics during a pandemic. *Journal of Philanthropy and Marketing*, *28*(1), e1773. https://doi.org/10.1002/nvsm.1773

Västfjäll, D., Slovic, P., Mayorga, M., & Peters, E. (2014). Compassion fade: Affect and charity are greatest for a single child in need. *PloS one*, *9* (6), e100115. https://doi.org/10.1371/journal.pone.0100115

Wang, X., Zhang, X., & Tong, A. (2022). The impact of linguistic style of medical crowdfunding philanthropic appeals on individual giving. *Procedia Computer Science*, *199*, 293–300. https://doi.org/10.1016/j.procs.2022.01.036

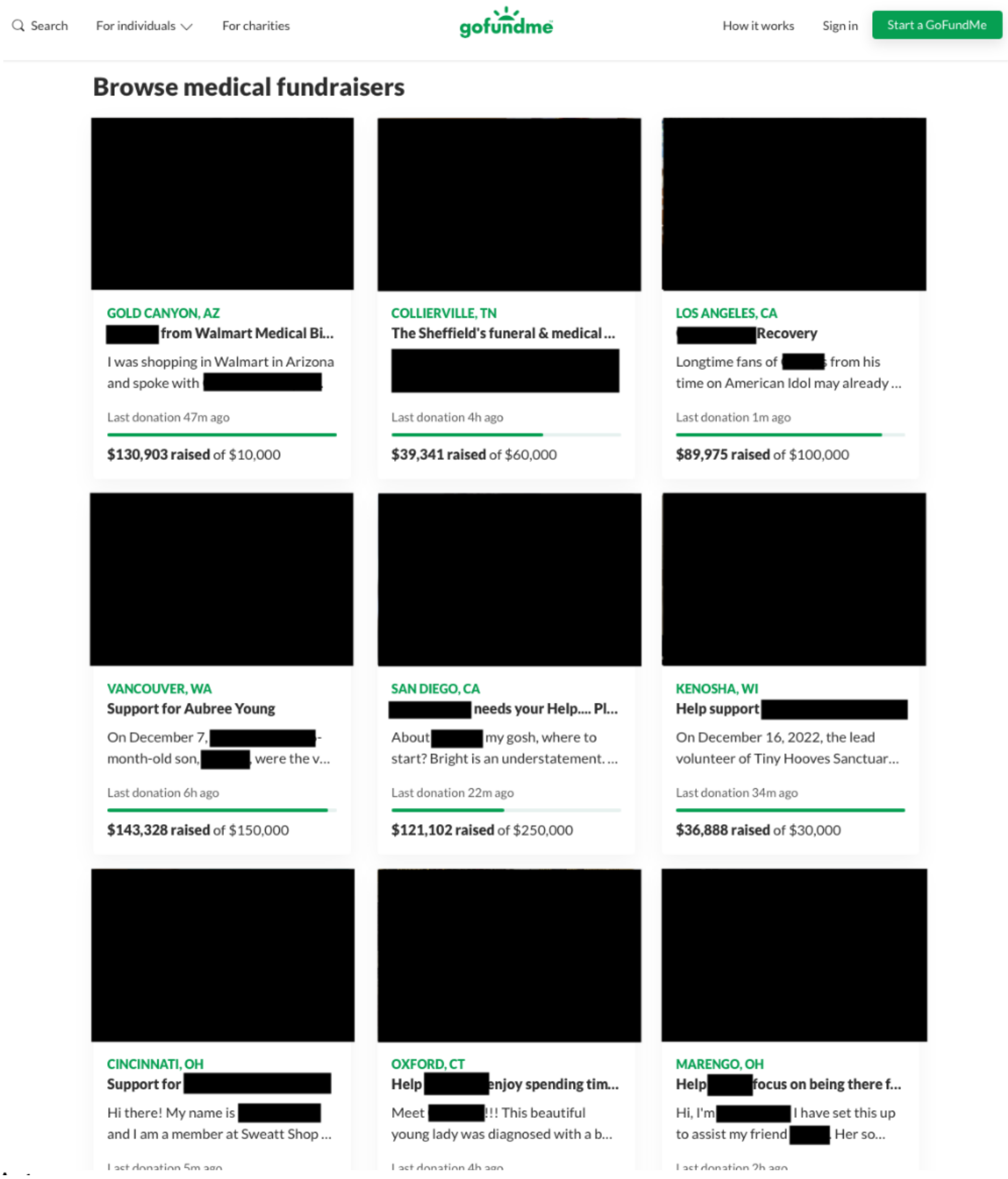
Westreich, D., & Greenland, S. (2013). The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. *American journal of epidemiology*, *177*(4), 292–298. https://doi.org/10.1093/aje/kws412

Woods, A. L., Wu, F. Y., & Hebl, M. R. (2023). Giving to Matthew, Emily, Jose, or Maria: A field study examining the impact of race and gender on donation requests. *Nonprofit and Voluntary Sector Quarterly*, *0* (0), 08997640221140314. https://doi.org/10.1177/08997640221140314

Wu, Y., Zhang, X., & Xiao, Q. (2022). Appeal to the head and heart: The persuasive effects of medical crowdfunding charitable appeals on willingness to donate. *Information Processing & Management*, *59*(1), 102792. https://doi.org/10.1016/j.ipm.2021.102792

7 Appendix

Figure A1 Project overview page for medical projects on GoFundMe.com. Identifying information blacked out.



7.1 Detailed information regarding machine learning algorithms

To detect the number of persons on a project's profile picture, I used the "faster_rcnn_r50_caffe_fpn_mstrain_3x_coco" model provided by the mmdetection Python library. This is a Faster R-CNN model (Ren et al., 2015) that was pretrained to detect persons in the COCO dataset (Lin et al., 2014) The model can be downloaded with this link. To detect the facial emotions expressed by the people on the project profile picture, I used the hsemotion Python library (Savchenko, 2021). I used the "enet_b2_7" model provided by this library.

7.2 Main results with outliers

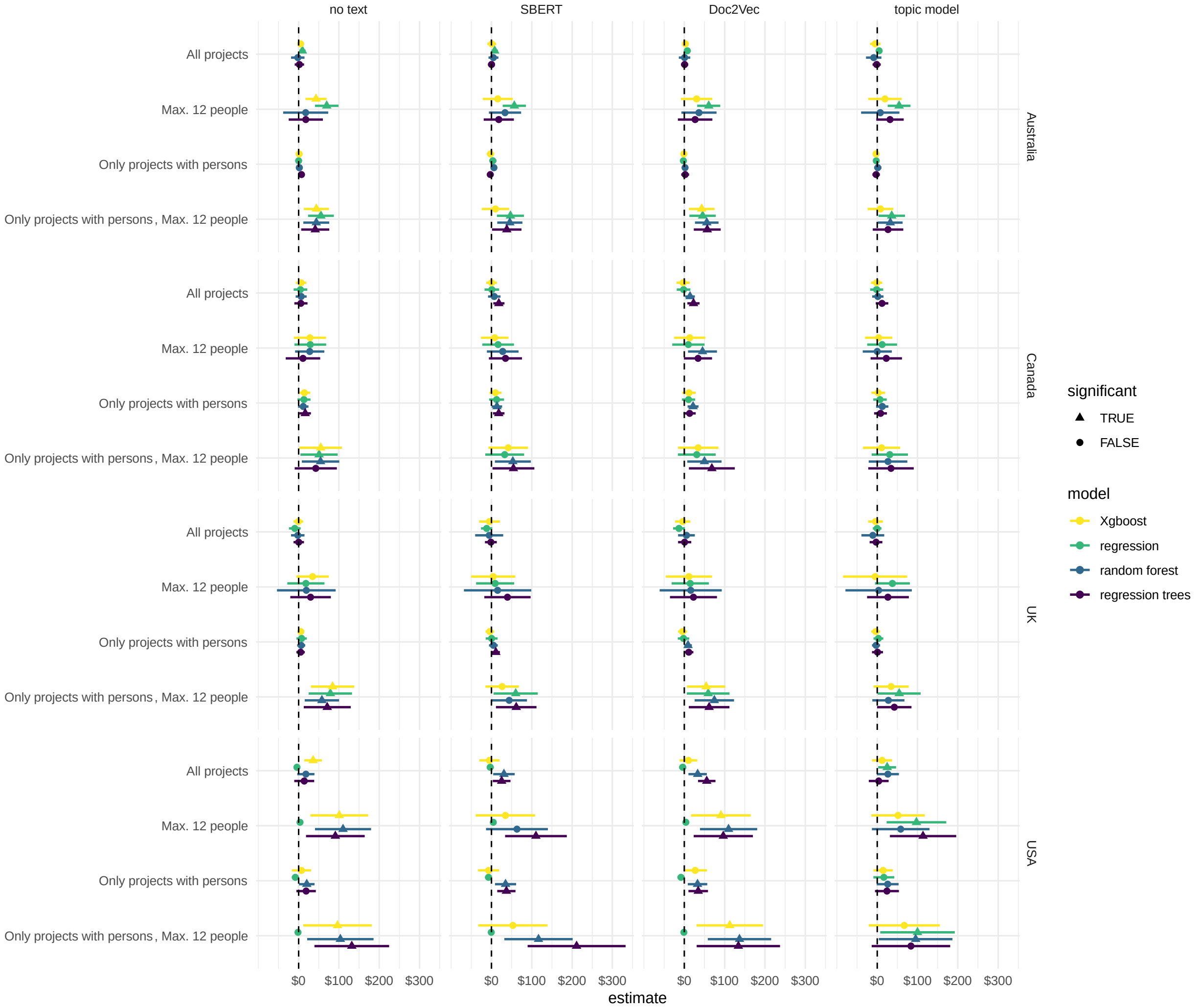
Table A1 Summary Statistics of all projects.

Variable	NotNA	Mean	Median	Sd	Min	Max
Number of persons on campaign picture	64848	2.597	1	5.445	0	84
At least one person on campaign picture	64848					
... No	26152	40.3%				
... Yes	38696	59.7%				
Max. 12 people on campaign picture	64848					
... No	3088	4.8%				
... Yes	61760	95.2%				
Amount raised	64848	6168.056	800	25684.513	0	1157925.116
Target amount	64848	13094.568	3475.118	41662.061	0.695	1108890
Created x days ago	64848	63.823	45	79.571	1	3089
Length of description (words)	64848	207.501	153	210.819	0	8256
Total updates of campaign	64848	0.657	0	2.382	0	117
Total photos of campaign	64848	1.877	1	3.394	0	218
Number of social media shares	64848	112.604	12	881.456	0	182965
Number of campaign hearts	64848	55.907	17	268.53	0	23577
Number of comments	64848	2.273	0	12.422	0	1074
Organized by	64848					
... organized by an organization	3660	5.6%				
... organized by a person	61188	94.4%				
Number of people organizing	64848	1.122	1	1.475	0	138
Team fundraiser	64848					
... No	59571	91.9%				
... Yes	5277	8.1%				
Organized for	64848					
... not organized for anyone	51221	79%				
... organized for an organization	4864	7.5%				
... organized for a person	8763	13.5%				
Country	64848					
... Australia	16325	25.2%				
... Canada	14757	22.8%				
... UK	17098	26.4%				
... USA	16668	25.7%				

Table A2 Summary Statistics of projects with at least one detected face.

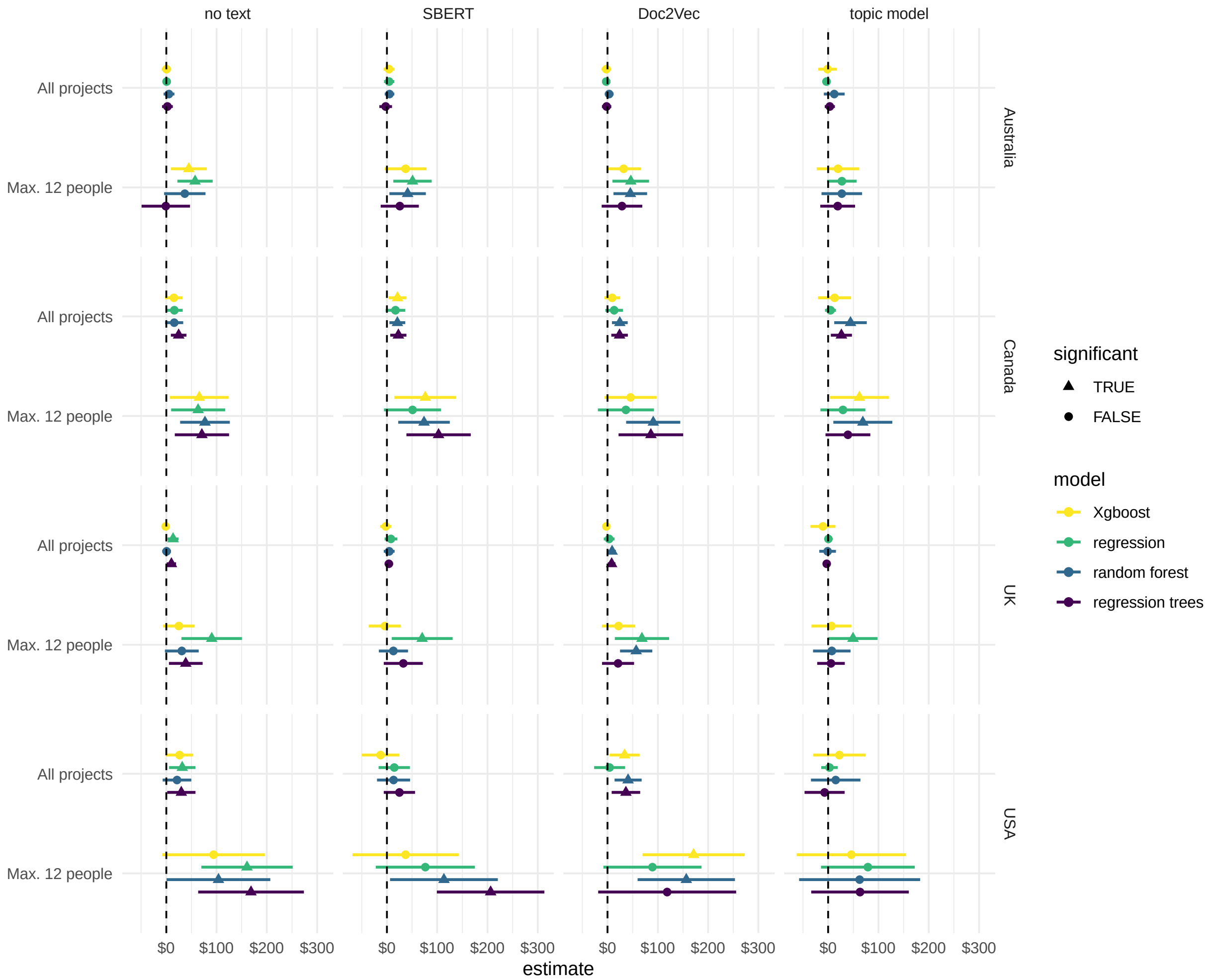
Variable	NotNA	Mean	Median	Sd	Min	Max
Number of persons on campaign picture	29910	4.516	2	6.716	0	80
At least one person on campaign picture	29910					
... No	819	2.7%				
... Yes	29091	97.3%				
Max. 12 people on campaign picture	29910					
... No	2550	8.5%				
... Yes	27360	91.5%				
Mean facial emotion: angry	29910	0.062	0.019	0.113	0	0.99
Mean facial emotion: disgust	29910	0.048	0.015	0.096	0	0.981
Mean facial emotion: fear	29910	0.025	0.003	0.064	0	0.964
Mean facial emotion: happy	29910	0.56	0.611	0.371	0	1
Mean facial emotion: sad	29910	0.063	0.021	0.108	0	0.98
Mean facial emotion: surprise	29910	0.047	0.011	0.09	0	0.965
Mean facial emotion: neutral	29910	0.194	0.118	0.211	0	0.972
Amount raised	29910	4714.578	1034.9	15801.836	0	787831
Target amount	29910	12871.291	3890.6	36190.548	0.695	800865
Created x days ago	29910	60.59	42	80.805	1	3089
Length of description (words)	29910	225.428	169	221.303	1	8256
Total updates of campaign	29910	0.723	0	2.495	0	117
Total photos of campaign	29910	1.916	1	3.55	0	218
Number of social media shares	29910	165.059	38	1238.569	0	182965
Number of campaign hearts	29910	76.528	24	294.934	0	23577
Number of comments	29910	3.117	1	12.7	0	720
Organized by	29910					
... organized by an organization	1240	4.1%				
... organized by a person	28670	95.9%				
Number of people organizing	29910	1.14	1	1.302	0	105
Team fundraiser	29910					
... No	27230	91%				
... Yes	2680	9%				
Organized for	29910					
... not organized for anyone	22901	76.6%				
... organized for an organization	1642	5.5%				
... organized for a person	5367	17.9%				
Country	29910					
... Australia	6965	23.3%				
... Canada	6486	21.7%				
... UK	7065	23.6%				
... USA	9394	31.4%				

Figure A2 Effect of the number of persons on a project's profile picture on the amount of funds raised in the full sample (i.e., with outliers).



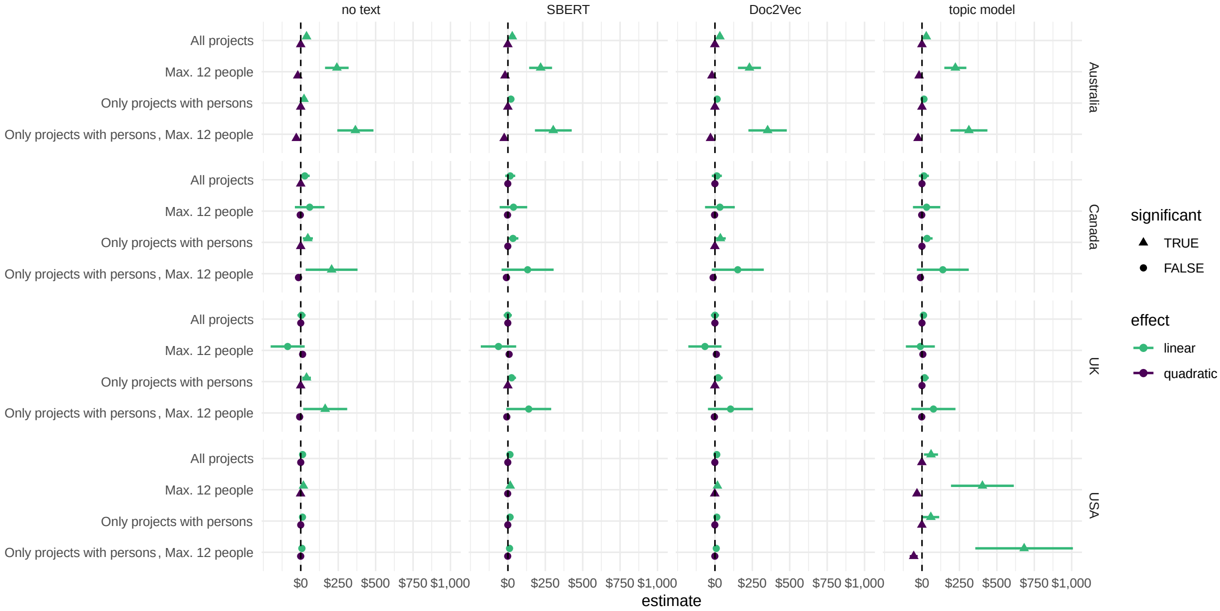
Notice: Error bars denote 95% confidence intervals.

Figure A3 Effect of the number of persons on a project's profile picture on the amount of funds raised for models that do control for the facial emotions expressed by the persons in the full sample (i.e., with outliers).



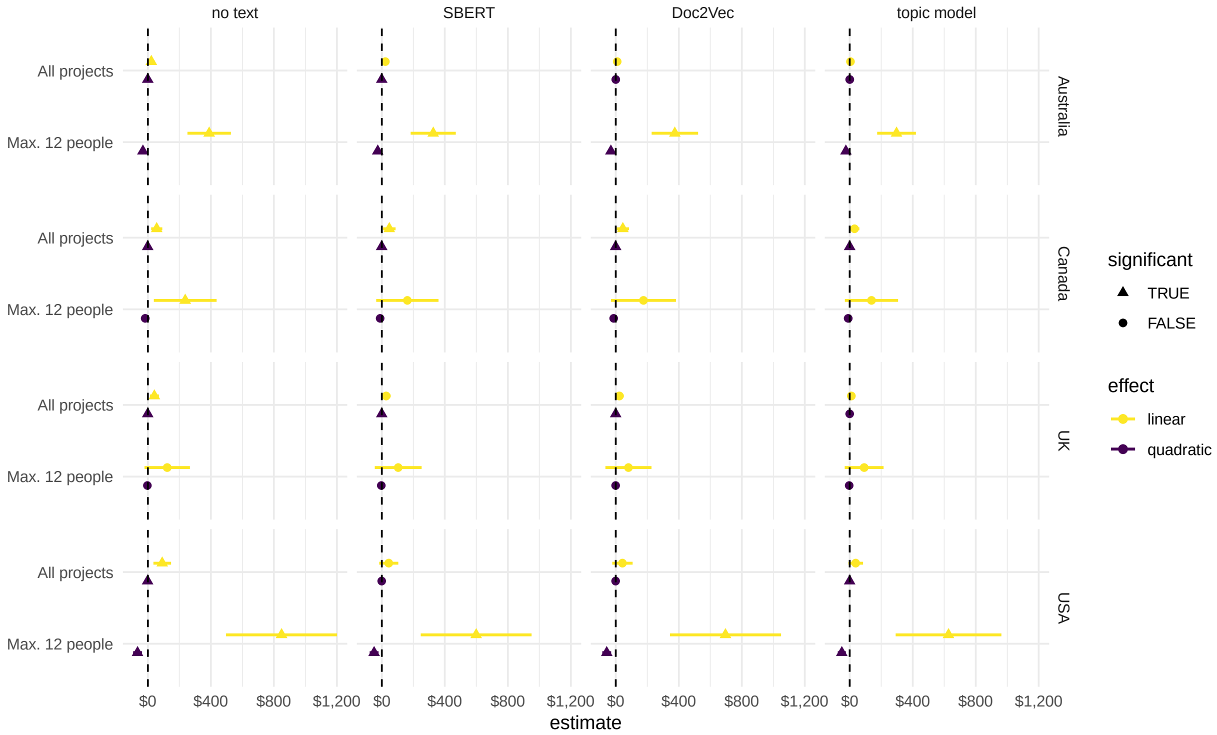
Notice: Error bars denote 95% confidence intervals.

Figure A4 Linear and nonlinear effect of the number of persons on a project's profile picture on the amount of funds raised in the full sample (i.e., with outliers).



Notice: Error bars denote 95% confidence intervals.

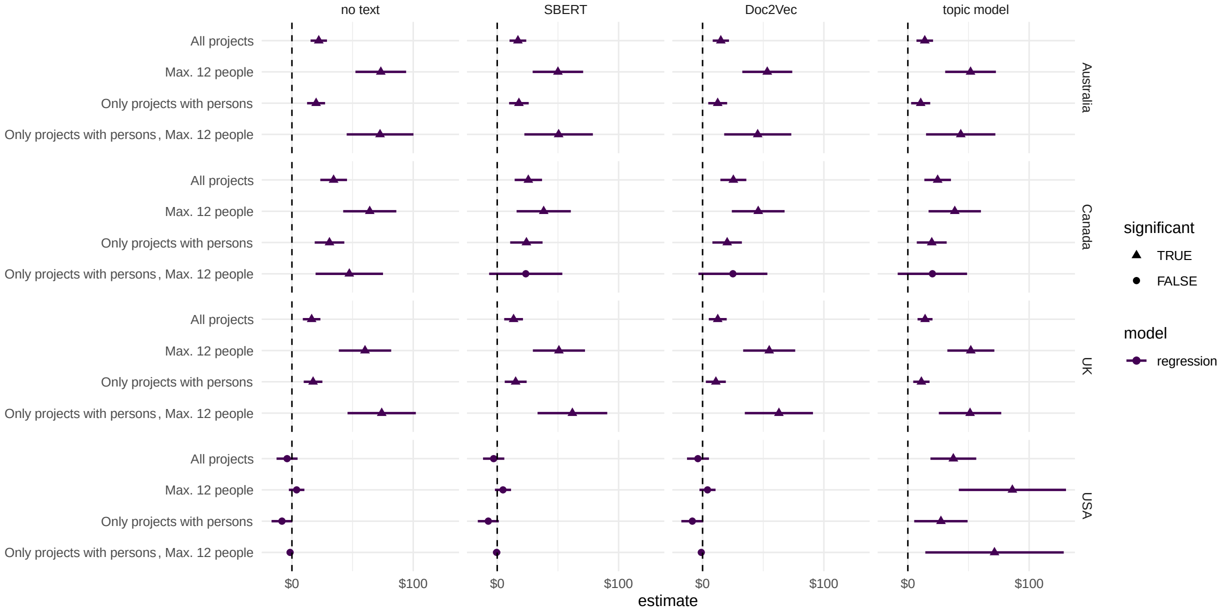
Figure A57 Linear and nonlinear effect of the number of persons on a project's profile picture on the amount of funds raised for models that do control for the facial emotions expressed by the persons in the full sample (i.e., with outliers).



Notice: Error bars denote 95% confidence intervals.

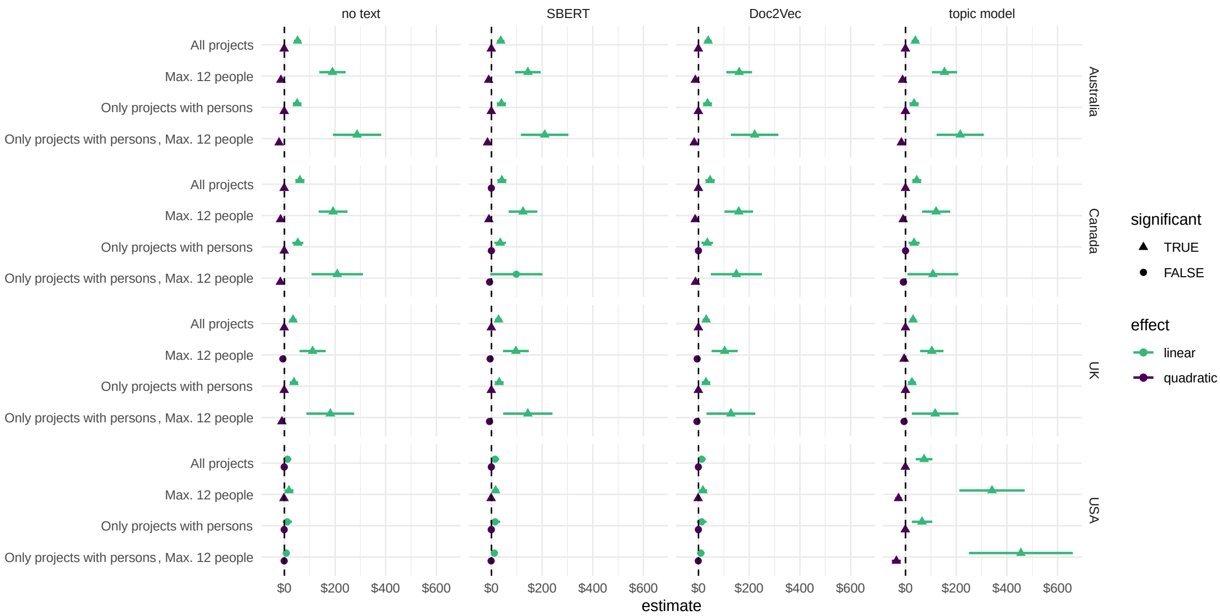
7.3 Results of models that do not control for the number of social media shares, number of comments and number of campaign hearts

Figure A6 Effect of the number of persons on a project's profile picture on the amount of funds raised for models that do not control for social media shares, number of comments and number of campaign hearts.



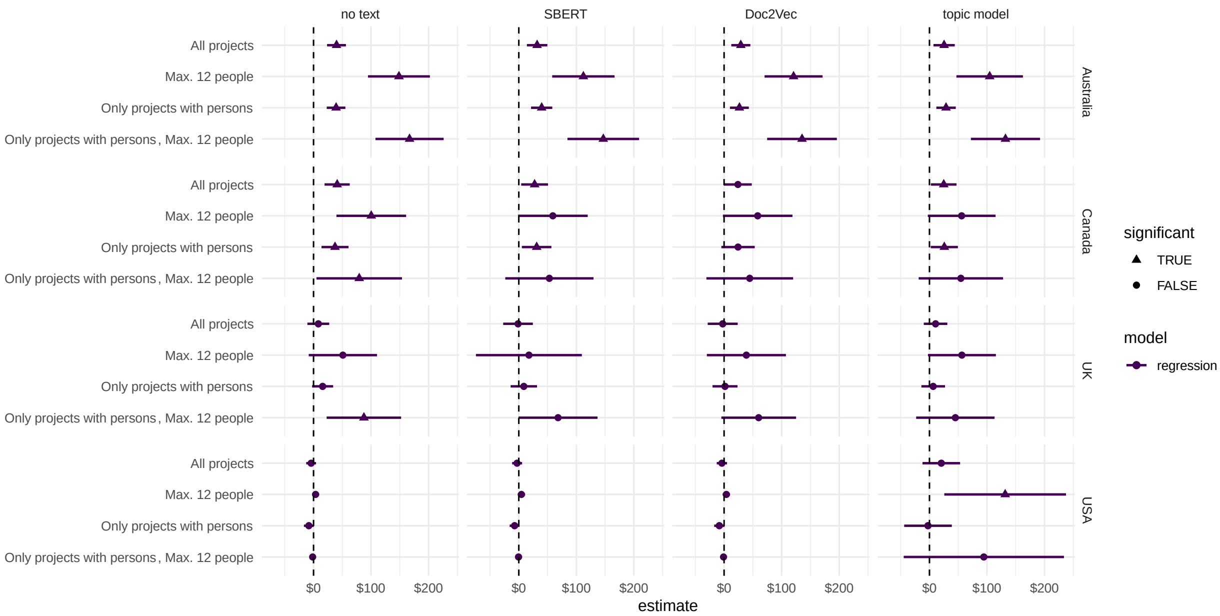
Notice: Error bars denote 95% confidence intervals.

Figure A7 Linear and nonlinear effect of the number of persons on a project's profile picture on the amount of funds raised for models that do not control for social media shares, number of comments and number of campaign hearts.



Notice: Error bars denote 95% confidence intervals.

Figure A8 Effect of the number of persons on a project's profile picture on the amount of funds raised in the full sample (i.e., with outliers) for models that do not control for social media shares, number of comments and number of campaign hearts.



Notice: Error bars denote 95% confidence intervals.

Figure A9 Linear and nonlinear effect of the number of persons on a project's profile picture on the amount of funds raised in the full sample (i.e., with outliers) for models that do not control for social media shares, number of comments and number of campaign hearts.

Notice: Error bars denote 95% confidence intervals.