Screen Set Size and Charitable Giving: A Field Experiment on a Peer-to-Peer Online Donation Platform*

Masyhur Hilmy
University of New South Wales

Gedeon Lim
The University of Hong Kong

Yohanes E. Riyanto
Nanyang Technological University

March 16, 2026

Abstract

This paper examines how screen set size shapes charitable giving in a field experiment on Bagirata, an Indonesian peer-to-peer donation platform during COVID-19. We randomize 2,405 donor sessions to view 3, 8, or 10 beneficiaries per screen, with the 10-beneficiary screen as the control condition, and study both between-donor effects (how screen set size affects donation outcomes) and within-donor effects (how beneficiary attributes shape allocation choices). We find an inverted-U pattern across the three menus. Donation frequency is highest for the intermediate menu. Donors in the 8-beneficiary arm make 0.26 more donations than those in the 10-beneficiary control. This pattern is consistent with a trade-off between variety and overload. Relative to the 10-beneficiary control, donors in the 3-beneficiary arm refresh much more often but see fewer total beneficiaries, whereas donors in the 10-beneficiary arm spend less time per beneficiary. In within-donor analyses using random beneficiary draws, beneficiaries perceived as more deserving, in particular, breadwinners with dependent children, receive larger gifts. Using machine-learning-based text analysis, we construct a deservingness index that is strongly associated with donation outcomes. We also find suggestive evidence of female in-group bias. Female donors are more likely to donate to beneficiaries with female-sounding names. Finally, donation likelihood varies with display order, with intermediate positions receiving fewer donations on average. Overall, the results suggest that intermediate screen set sizes can improve donation outcomes by balancing variety against cognitive constraints, offering practical guidance for the choice architecture of online donation platforms.

JEL Classification: D64, O10, C93, D91

Keywords: Field Experiment, Charitable Giving, Online Donations, Choice Architecture

*Previously circulated as "Less is More? Maximizing Charitable Donations during Crises: An Online Field Experiment" and "What Determines Online Charitable Giving to Pandemic Victims? Evidence from a Field Experiment on Choice Overload & the Deservingness of Beneficiaries". We thank Sam Bazzi, Ray Fisman, Siddharth George, Josh Goodman, Ben Marx, Marie Claire Villeval and participants at NEUDC 2021 conference at Boston University, the Asia Meeting of the Econometric Society (AMES) South, Central and Western Asia 2023 in New Delhi, and the 2023 GATE Lyon - NTU Joint Workshop for useful comments. We thank the team members at Bagirata for graciously sharing their data with us. We also thank Jonathan Alfrendi, Amangku Brahmanaputra, Aisy Ilfiyah, and I Putu Yoga Tunas Sugitha for excellent research assistance. This project received ethics approval from the University of Hong Kong (HREC Reference Number EA200065), and is pre-registered at the Open Science Framework (OSF) (https://osf.io/c4xgd). Hilmy acknowledges support from the Institute for Economic Development at Boston University. Lim acknowledges research grant support from The University of Hong Kong. Riyanto acknowledges research grant support from Nanyang Technological University. All errors are our own.

Hilmy: m.hilmy@unsw.edu.au. Lim: Faculty of Business and Economics. Pokfulam Rd., Hong Kong SAR. Email: gedeonl@hku.hk. Riyanto: yeriyanto@ntu.edu.sg.

1. Introduction

Charities raise hundreds of billions of dollars annually from individual giving. Globally, the proportion of donors who give through online channels has been growing rapidly (Paxton, 2020; Clark et al., 2019). In developing countries, peer-to-peer giving through these channels can be an important source of aid. The lack of government infrastructure for making timely transfers (Hanna and Olken, 2018); low setup costs for digital platforms (Bhargava, 2022); and the widespread use of mobile e-payment systems have facilitated the delivery of direct aid to both charitable organizations and individual beneficiaries through mobile transfers (Suri et al., 2023). Nevertheless, the existing literature has predominantly focused on donations to charitable organizations, with comparatively limited attention given to peer-to-peer transfers from donors to individual recipients (Kessler et al., 2019; Schmitz, 2021).

Individual donors typically face numerous options to donate during times of crises, but the effect of set size on donation decisions is still not well-understood. Under the assumption of limited attention, laboratory experiments in the consumer choice literature have found that a smaller choice set leads to decreased choice overload.[1] However, both private consumption decisions (where other-regarding preferences are absent) and the laboratory setting might not be reflective of typical giving. Outside of the laboratory, credible causal estimates of screen set size on donations are rare, as well as examinations of how screen set size influences behavior.

In this paper, we run a field experiment to study the effect of screen set size on donor behavior in peer-to-peer giving. Our experiment was run in partnership with an online donation platform in Indonesia during the country's largest modern public health crisis, the COVID-19 pandemic. The experiment took place between October 4th, 2020 to June 9th, 2021. We randomize, at the donor-level, the number of alternative beneficiaries per screen that a donor views (henceforth referred to as *screen-size*). Donors make donation decisions based on the menu of displayed cards (see Figure 1), which is randomly selected by the platform's algorithm. Each beneficiary card contains a self-written narrative that details why he/she is asking for a donation (typically due to COVID-19-related job losses). Based on this information, potential donors are free to choose which beneficiary (or beneficiaries) to support and the amount that they wish to donate. Donations are made directly through a digital payment system.

[INSERT FIGURE 1 HERE]

Our experimental setup involves two levels of randomization. One at the donor-session level, and another at the beneficiary-level. We leverage this to study (i) the impact of screen size on donor behavior (*between-donor analysis*), and (ii) the determinants of donations within a single donor-session (*within-donor analysis*). At the donor-session level, upon entering the

[1] See Chernev et al. (2015) for a meta-analysis. There are also studies on the effects of the number of recipients on donation, but nearly all study giving through intermediaries, which abstracts from the very salient identification with victims in direct peer giving (Corazzini et al., 2015; Schmitz, 2021; Soyer and Hogarth, 2011). This literature has also only examined a limited number of choices, between one and three options.

platform, potential donors are randomized to view either 3, 8, or 10 beneficiaries per screen. At the beneficiary-level, within each donor session the platform displays a random selection of beneficiaries from its database. This guarantees that the array of beneficiary characteristics displayed to donors both within and across screens is as good as random.

In our *between-donor* analysis, we analyze the effect of screen-sizes on donor behavior. Leveraging our randomized design, we analyze differences in the number and value of donations given at the donor-session level. To characterize mechanisms, in exploratory analyses, we analyze the platform's back-end database to explore differences in donor behavior.

We find that the total number of donations is highest in the 8-screen size treatment. Specifically, donors assigned to an 8-screen make 0.26 more donations compared to 10-screens. Meanwhile, a 3-option screen did not induce donors to make more donations compared to the 10-option screens during the entirety of our experimental period. We thus observe an inverted-U pattern in the effect of the number of alternatives on donation counts. This inverted-U pattern is also observable in the amount of donation received by beneficiaries during normal traffic period, i.e., not during mobility lockdown or intensive social media campaign period, with donors in the 8-screen donating 3.7 USD more than the control group. These estimates incorporated COVID-19-specific variables in our regressions, including weekly mean positivity rate and daily death rate. We did not see a significant difference in the probability of making any donation across treatment arms.

In our framework, screen set size can influence donor behavior through the trade-off between variety (match quality) and cognitive overload. A bigger screen set makes it more likely for donors to encounter a more deserving individual, but information from each additional beneficiary may increase the cognitive cost. We find suggestive evidence for the former—donors in 3-option sets see fewer potential beneficiaries above a certain threshold of deservingness but there is no difference in 8 and 10-option sets. We analyze donors' refresh rates: after viewing the first screen, donors can decide to (i) make a donation to zero, one, or more than one beneficiary and (ii) to hit the refresh button at the bottom of the screen to obtain a fresh draw of cards. Therefore, the refresh rates proxy for donor effort to increase the variety of beneficiaries that they are exposed to. We find that donors in the 3-option arm refresh about 1.8 *more times on average* than donors in the 10-option (control) arm. In contrast, we find no differences in refresh rates between 8- and 10-screens. By the end of their search, donors in 3-screen sizes are still exposed to fewer beneficiaries despite their higher rate of refresh, as they continue to see fewer beneficiaries in all subsequent screens, suggesting a lower overall cognitive load for this group. At the same time, our data shows that donors are spending less time per beneficiary when they are exposed to more options in a set, possibly indicative of donors *giving up*, when faced with too many options. We further provide a suggestive piece of evidence for the convexity of cognitive costs: past the 8-option set (threshold), donors see a relatively lower dispersion in the lengths of beneficiary appeals vis-a-vis 10 option-sets despite just a difference of 2 beneficiaries.

Taken together, these results suggest that the 8-option screen set size balances the variety-cognitive overload trade-off. These donors face lower cognitive loads than those in 10-option sets, but, unlike in 3-option sets, this choice architecture did not necessitate a higher effort to search for donation targets yet enabled donors to encounter a beneficiary mix of sufficiently high match quality, leading to improved donation outcomes. In another words, the 3-option screens induce greater but less optimal donor (search) effort: donors refresh more often to search for more optimal targets, but, due to search fatigue, end up being exposed to fewer total beneficiaries, with no difference in donation rates compared to the 10-option beneficiaries.

On the other hand, we find a lack of evidence to support the role of decision complexity (Enke and Shubatt, 2024) and regret in driving the observed results. Analyses incorporating a narrative-based measure of decision complexity (proxied by unique words) produce patterns similar to our main analysis. Separately, analyses in sub-samples where regret is unlikely to play a role yield results consistent with the main findings.

In our *within-donor* analysis, we study the effects of deservingness, in-group bias, and beneficiary display order on donation decisions. These are factors that could also influence donation. Importantly, due to the platform's algorithm, a beneficiary is drawn randomly within and across screens. Hence, we continue to use OLS regressions at the beneficiary-donor dyad level with donor session fixed effects. To study deservingness, we leverage comprehensive beneficiary information displayed to donors, including detailed beneficiary narratives.[2] From these narratives, we code, both by hand and natural language processing (NLP) based textual analysis, an exhaustive set of characteristics that donors might perceive as signaling different dimensions of deservingness and capturing donor-beneficiary in-group biases. To study display order, we regress donation outcomes on indicators for display order.

We document three main findings in our within-donor analysis. First, we find that beneficiaries perceived as more deserving are more likely to receive donations. Specifically, those perceived as breadwinners with a dependent child (0.7 pp), those in the education sector (1.3 pp), and those who provided longer narratives (0.5 pp for every 50 words of the appeal) receive more donations. We corroborate this with a natural language processing (NLP) based textual analysis. Second, we find some evidence of in-group bias: female donors are more likely to donate to beneficiaries with female-sounding names. In contrast, donors are not more likely to donate to co-religionists or co-ethnics. Third, we find that beneficiaries whose information cards appear centrally within a screen are less likely to receive donations.

In summary, our between-donor results suggest that fewer alternatives are not always better. One possible interpretation is that the benefits of alleviating cognitive overload from multiple alternatives may be insufficient to outweigh the benefit from increasing the likelihood of finding a good match. In addition, our within-donor results suggest that, policy-wise it might be attractive to (i) pre-select and highlight key beneficiaries' characteristics that are expected to attract donations; (ii) leverage context-dependent, in-group biases by matching beneficiaries with similar identities to those of donors; and (iii) position beneficiaries with the highest marginal benefit of receiving donations at the start or end of the screen. Our findings thus provide two insights for policy-makers and online platforms seeking to leverage behavioral heuristics and optimize

[2]See Appendix Table A.1 and A.2 for examples.

donor behavior. First, our findings highlight the potential for online donations in developing countries, where low state capacity makes online donations a practical solution to redirect funds during large-scale disasters efficiently. Second, differences in the amount, type, and presentation of information to donors can make a large difference in donation outcomes.

To guide our empirical investigation, we also present a theoretical framework that models donor behavior as a dynamic search process with cognitive costs. The framework integrates three key elements, namely, (1) the trade-off between variety and cognitive overload in screen-set-size, (2) donor-beneficiary match quality based on perceived deservingness, and (3) session-level giving constraints. We use the framework primarily as an organizing tool to guide what mechanisms to investigate, such as refresh behavior, total exposure, and within-screen attention, and to clarify the empirical predictions we test, including the relationship between screen size and donation outcomes, as well as systematic links between narrative-based match quality and donation allocation within a session.

In all, our paper makes four novel contributions. First, we contribute to the literature about screen set size on individual decision-making behavior. While existing work has mostly focused on screen set size on consumer choice (Iyengar and Lepper, 2000; Iyengar and Kamenica, 2010; Reutskaja et al., 2011), we focus instead on the effects of the screen set size on charitable giving to individuals (peer-to-peer giving).[3] Our paper is among the first large-scale field experiments to study the effects of screen set size in charitable giving to individuals. Relative to laboratory experiments, donor motivations in our setting are likely to be more similar to those in other online, peer-to-peer giving platforms. This enhances the external validity of our findings and offers insights for broader applications to altruistic decision-making processes.

A related work to our paper is Corazzini et al. (2015), who uses a public goods game in a lab to study the effects of the number of alternatives on donor behaviors. In that setting, donors have a fixed budget constraint and clear strategic considerations due to matching benefits for making donations. Our paper is also related to Eckel et al. (2020), whose lab experiment tests sequential vs simultaneous donation requests from eight charities to university student subjects. Schmitz and Grieder (2024) test coordinated fundraising where donors are exposed to two coordinated causes or one cause in individual campaigns. Another related work is Iyengar and Kamenica (2010), who find that less is more: a smaller screen set size leads to more optimal decision-making behavior over asset allocations. In contrast, our paper suggests that less is not necessarily more. We show that, in charitable peer-to-peer giving, displaying a screen set of intermediate size (8-screens) boosts donations by possibly providing a larger number of beneficiaries that are potentially better matches with the donor's own preferences. Furthermore, our results demonstrate that the number of donations is possibly a concave function of screen set size: donors in 10-screens make fewer donations than those in 8-screens.

Second, this paper extends work that uses field experiments to study the determinants of charitable giving (Adena et al., 2024; Adena and Hager, 2024). This literature has examined,

[3]Kiva and GoFundMe are close analogues, but both are instances of *conditional* giving. GiveDirectly is an example of unconditional giving, but again, the platform functions as an intermediary that channels donations to beneficiaries only after the donor has completed the donation process.

e.g., the effect of competition and coordination among charities (Scharf et al., 2022; Schmitz and Grieder, 2024; Gallier et al., 2023) and the effect of natural disasters on charitable giving (Deryugina and Marx, 2021; Schwirplies, 2023). Most prior work studies donations made through charitable organizations as intermediaries. This paper, in contrast, studies peer-to-peer giving, which is (i) more accurately viewed as a two-sided market, with donors on one side and beneficiaries on the other; and (ii) where behavioral biases are likely to be more salient given individuals being confronted with real-life beneficiaries. This paper provides the first large-scale evidence of the determinants of peer-to-peer giving: a setting that might become increasingly policy-relevant.

Third, we contribute to the literature on how perceptions of deservingness affect altruistic behavior. To the best of our knowledge, our study is the first to examine fairness principles (Konow, 2000; Cappelen et al., 2007) in a real-world field experiment setting. We innovate by presenting donors with a full menu of beneficiaries in a real-world setting, implicitly forcing donors to compare beneficiaries against each other when making donation decisions. This allows us to provide a direct test of the specific dimensions of perceived deservingness that are *comparatively* more important in altruistic decisions.

Fourth, we contribute to a small but growing literature on identity, social distance, and in-group biases in charitable giving (Charness and Holder, 2019; Kessler and Milkman, 2018; Adena et al., 2024). Kessler and Milkman (2018) investigates the positive effects of identity priming of *donors* on generating more charitable donations to the Red Cross. This paper, in contrast, focuses on (i) directed peer-to-peer giving (rather than to organizations) and hence, demonstrates the benefits from ensuring identity concordance between both donors *and* recipients; (ii) shows the potential importance of taking into account the context in which giving takes place, to enhance the efficacy of identity primes on giving behavior. Specifically, Indonesia is one of the most ethnically and religiously diverse countries in the world. Yet, in contrast to much of the literature that would suggest greater benefits from priming ethnic or religious identity, our results on female gender in-group bias show that the overwhelmingly dominant narrative of COVID-19 and predominantly male breadwinners (in Indonesia) increased the relative saliency of female vulnerability to female donors. Hence, our results suggest that, in direct giving, there are potentially greater benefits from adjusting identity primes depending on *both* country and crises-specific contexts.[4]

This paper is organized as follows. Section 2 describes the study context, and Section 3 describes our theoretical framework. Section 4 presents our experimental design, and Section 5 presents the order of analysis. Section 6 discusses our results. Section 7 concludes.

[4]A caveat is in order. Our experiment varies screen size at three discrete points, $K \in \{3, 8, 10\}$. This enables clean comparisons across the interface designs we test, but it does not identify a smooth "inverted-U" relationship in K or a global maximization rule that necessarily generalizes beyond these specific screen-size menus. We therefore emphasize the relative performance of the designs we study (notably the intermediate 8-card screen versus 10), and we present the within-screen evidence on identity concordance and display position as complementary, suggestive design insights rather than primary causal estimates.

2. Online Charitable Giving Platform *Bagirata*

Globally, Indonesians rank among the top 10 most prolific givers, with much of this giving taking place through informal organizations (Charities Aid Foundation, 2019; Noor and Pickup, 2017). According to the Gallup World Poll, 78% of respondents in Indonesia donated money, 53% volunteered their time, and 40% helped a stranger (Charities Aid Foundation, 2018).[5] The ubiquity of such giving behavior played an important role in Indonesian society's largely grassroots-driven response to COVID-19.

When the government started imposing lockdowns to contain COVID-19, the resulting mobility restrictions led to 2.6 million workers being laid off and significantly reduced income among millions of others (Aria, 2021). The widespread economic stress led to the emergence of various bottom-up initiatives to help those in need. For example, COVID-19–related fundraisers on *Kitabisa*, a popular Indonesian crowdfunding platform, collectively raised USD 3.5 million in the first week of lockdown. These fundraisers leveraged the increasing adoption of digital financial services to facilitate micro-donations.[6] Our study focuses on one such bottom-up fundraising platform: *Bagirata*.

At the heart of the *Bagirata* platform is an online, centralized beneficiary database. The beneficiaries are individuals suffering from COVID-19–related income and job losses. To be registered as a beneficiary, individuals submit details such as their employment status, economic situation, social media handles, mobile payment QR codes, and contact information to *Bagirata*. This information is then verified by *Bagirata*, and only successfully validated applicants are included in the beneficiary database (a group henceforth referred to as potential beneficiaries).[7]

When a user enters the platform as a potential donor from the landing page, the platform's algorithm randomly draws and presents a set of beneficiary cards (Figure 1, see also Figure A.1 for the landing page).

[INSERT FIGURE 1 HERE]

These cards are based on the information provided by registered beneficiaries. In Section 4, we discuss how our experimental manipulation leverages this algorithm and how the experience of potential donors differs based on the treatment group to which they are assigned. Potential donors then decide if and how they want to donate. Specifically, after viewing the first set of beneficiaries, donors can make the following decisions that are non-mutually exclusive. First,

[5] This high level of giving is often linked to *zakat*/ alms, a pillar of Islam, Indonesia's majority religion. The National Board of Zakat reported collecting USD 434 million of alms in 2017 (Baznas, 2019).

[6] A J-PAL Southeast Asia survey found that 21% of men and 22% of women used digital financial services for the first time during the COVID-19 outbreak (J-PAL SEA, 2020). Combined with existing users, this influx of users raised the proportion of active users to 75% of men and 70% of women.

[7] All recipients were soliciting donations specifically for COVID-19 relief. *Bagirata* received coverage from various media outlets; e.g., see https://youtu.be/wrhxL5vfMQQ. See Table A.1-A.2 for a selection of appeal narratives written by beneficiaries.

they can donate to zero, one, or more than one beneficiary from the list on display. Second, they can obtain a new set of beneficiary cards by clicking the refresh button at the bottom of the screen. The number of beneficiaries shown in the new draw will be identical to the first screen. In the second screen, they are faced with the same decision-making problem, which is repeatable ad infinitum screen-by-screen until they leave the website. Donations are transferred directly from potential donors to their chosen beneficiaries through one of the three popular digital payment systems in Indonesia. After donating, donors are prompted to confirm their donation by reporting the donation amount and donation status on the *Bagirata* platform. Our analysis includes all donations verified in this manner.

To facilitate the donation process, the platform allows potential donors to donate anonymously. The only identifiable information that donors voluntarily provide is their email address. This design has two implications. First, we do not observe all donor characteristics. We address this by conducting a follow-up user survey where we collect email handles, thereby enabling us to match a subset of donation data from *Bagirata*'s back-end database to donor characteristics. Throughout the paper, however, our analysis focuses on the full set of donation data. In cases where our analysis uses the subset of matched data, we explicitly state so. Second, we cannot identify a donor that initiates multiple sessions if he/she does not provide an email address.[8]

The beneficiary side of the platform can be described as follows. Each beneficiary is displayed as a compact card (Figure 1), which provides a set of standardized information. This includes the beneficiary's name, occupation, area of residence, and whether he possesses any social media accounts (Instagram, Facebook, or Twitter). Furthermore, it provides a brief narrative on the impact of COVID-19 on the beneficiary's life and the reasons why monetary assistance is needed, outlines the minimum amount of monetary assistance required, and details the duration for which the assistance would be needed. The card also displays the total amount of donations collected thus far as a share of the ask amount and indicates the e-payment channels through which donations can be transferred.

3. Theoretical Framework

We develop a theoretical framework to understand how screen set size affects donor behavior in online charitable giving platforms. Our model integrates three key elements: (i) dynamic search with cognitive costs, (ii) donor-beneficiary match quality, and (iii) session-level giving constraints. The framework yields testable predictions about the non-monotonic relationship between screen size and donation outcomes, as well as predictions about within-screen donation patterns.

[8] Such a donor will appear in our dataset as multiple sessions. From our user survey, 13% of donor-sessions (N=312) have a nonunique email associated with them. And out of these tagged donor-sessions, N=190 have a unique email tag.

3.1. Basic Setup and Dynamic Search

We model a donor's decision on an online charitable giving platform, such as Bagirata, as a dynamic search and optimal stopping problem. As in Bagirata, the donor sees a screen displaying a set of K potential beneficiaries. After viewing one screen, the donor must decide whether to donate now (select one or more beneficiaries and give) or continue searching. Importantly, once a screen is passed, those particular beneficiaries cannot be revisited (no recall). If the donor chooses to continue browsing by clicking refresh, a new set of K beneficiaries (not seen before) is shown at the cost of expending additional time and cognitive effort.

We denote by $C(K)$ the cognitive cost of processing a screen of size K, which we assume is increasing and convex in K:

$$C'(K) > 0, \quad C''(K) \geq 0. \tag{1}$$

Intuitively, larger screen sets provide more variety but impose higher cognitive load (attention and comparison costs), whereas smaller sets are easier to process but may require more rounds of search to find a good match. Each time the donor clicks "refresh" to see K new options on a new screen, they incur the cost $C(K)$.

3.2. Utility and Donation Choice

When faced with a given beneficiary i, the donor derives utility from giving that depends on the match quality q_i of the beneficiary and the chosen donation amount d. The donor's instantaneous utility from donating amount d to beneficiary i is:

$$U_i(d) = B(q_i, d) + W(d) - d, \tag{2}$$

where $B(q_i, d) = q_i \cdot b(d)$ denotes altruistic benefit, with $b(\cdot)$ increasing and concave in d; $W(d)$ denotes warm-glow benefit, also increasing and concave; and donating d units entails a linear cost of d.

The donor chooses the optimal donation d_i^* by solving the first-order condition:

$$\frac{\partial U_i(d)}{\partial d} = \frac{\partial B(q_i, d)}{\partial d} + \frac{\partial W(d)}{\partial d} - 1 = 0. \tag{3}$$

Let $B_d(q_i, d)$ and $W'(d)$ denote the marginal benefit to the donor from altruism and warm-glow, respectively. The condition can be expressed as:

$$B_d(q_i, d_i^*) + W'(d_i^*) = 1. \tag{4}$$

This ensures that the marginal altruistic plus warm-glow benefit equals the marginal cost. Because B and W are concave, this yields a unique optimum. Importantly, since B_d increases with q_i, we have:

$$\frac{\partial d_i^*}{\partial q_i} > 0, \tag{5}$$

implying that more "deserving" recipients receive larger donations. This means that donors might give more to beneficiaries with compelling stories or shared identity traits (e.g., a female donor giving to a female recipient).

3.3. Stopping Rule and Reservation Utility Threshold

Let $U_{\max} = \max_{i=1,\ldots,K} U_i(d_i^*)$ denote the best utility from donating on a given screen. If the donor chooses not to donate, she may refresh and see K new options by paying $C(K)$. We model this as a dynamic optimal stopping problem with uncertainty about future beneficiaries.

Let $F_K(u)$ denote the cumulative distribution function of $U_{\max}$ over screens of size K. The donor's optimal policy is to stop and donate if $U_{\max} \geq R(K)$, where $R(K)$ is the reservation utility threshold, and to continue searching otherwise.[9] This threshold satisfies the standard optimal stopping condition (Rogerson et al., 2005):

$$\int_{R(K)}^{\infty} (u - R(K))\, dF_K(u) = C(K). \tag{6}$$

The left-hand side represents the expected benefit of another search round (the option value of continued search), while the right-hand side is the search cost. Thus, the donor is indifferent between stopping and continuing when $U_{\max} = R(K)$.

Lemma 1 (**Reservation Threshold and Screen Size**). *If the distribution of match qualities has full support and $C(K)$ is increasing in K, then the reservation threshold $R(K)$ is non-increasing in K for sufficiently large K.*

Proof. As K increases, the distribution F_K first-order stochastically dominates $F_{K'}$ for $K > K'$ (more options increase the maximum). However, with convex $C(K)$, the cost of continued search rises faster than the option value. When $C'(K)$ is sufficiently large, the right-hand side of Equation (6) dominates, requiring $R(K)$ to fall to maintain equality. □

[9] On *Bagirata*, donating and refreshing are not mutually exclusive: a donor may make a gift on a given screen and still continue browsing, potentially donating again on later screens. Our stopping rule should therefore be read as follows: at each screen the donor decides whether to continue searching (request another draw) or to pause searching and allocate one or multiple gifts among the beneficiaries currently displayed (the "terminal" screen for that search step). After any such allocation, the donor may either exit or resume browsing, which is naturally represented as the same screen-level decision being repeated until the session ends (or a session-level giving goal is met). This interpretation keeps the optimal-stopping intuition, trading off expected gains from further search against its cost, while remaining consistent with multiple donations within a session.

3.4. Multiple Donations and Session-Level Giving

In our setting, donors may make multiple donations within a session (e.g., donating to more than one beneficiary). We extend the model to capture this feature in two ways.

First, making each additional donation requires another beneficiary selection and amount choice. We capture this by allowing a per-donation selection cost $\tau(K)$ that is weakly increasing in K:

$$\tau'(K) \geq 0. \tag{7}$$

Intuitively, when the screen is larger, making repeated selections entails higher comparison and decision costs, which can discourage multiple donations even conditional on donating at least once.

Second, donors may have a session-level giving target (or soft budget) that is largely set before they begin browsing. For example, a donor may come to the platform intending to give "about \$X today" or to allocate "one to a few small gifts" based on internal norms, limited liquidity, or a desire to be consistent with their self-image as a giver. In reduced form, this can be captured by assuming that the donor's utility is strongly concave in total giving D over the session:

$$\frac{\partial^2 W(D)}{\partial D^2} \ll 0 \quad \text{for } D \approx \bar{D}, \tag{8}$$

where $\bar{D}$ is the intended giving target. This implies that once the donor has given an amount close to her intended total, the marginal warm-glow (or moral satisfaction) from giving an additional dollar declines rapidly, making further increases in the total amount less attractive.

3.5. Set Size and Donation Behavior: Main Predictions

Our framework yields several testable predictions about how screen size K affects donation behavior. The key insight is that the relationship between K and donation outcomes is *non-monotonic* due to two opposing forces:

1. **The Variety Effect:** Larger K increases the likelihood of encountering a high q_i, thus increasing $U_{\max}$ and the probability of finding a good match.
2. **The Cognitive Overload Effect:** As K grows, $C(K)$ increases at an increasing rate, discouraging extensive search and potentially pushing $R(K)$ downward (Lemma 1). In addition, when donors make multiple donations, the per-donation selection cost $\tau(K)$ increases with K, further discouraging repeated giving on larger screens.

Proposition 1 (Non-Monotonic Relationship: Number of Donations). *There exists an interior optimum $K^* \in (K_{\min}, K_{\max})$ that maximizes the expected number of donations per* *session. For $K < K^*$, the variety effect dominates; for $K > K^*$, the cognitive overload effect dominates.*

When K is very small, each screen contains a limited variety, so donors must refresh more often to find additional satisfactory matches. Because each refresh reveals only a few new beneficiaries, cumulative exposure grows slowly while search effort accumulates, making it less likely that donors encounter enough "good" matches to support multiple donations before they disengage.

At the other extreme, when K is very large, evaluating and comparing many profiles becomes cognitively taxing, and the per-donation selection cost $\tau(K)$ is higher, which discourages making repeated gifts even conditional on donating at least once.

A moderate screen set K^* balances these forces: it provides enough variety per screen to generate additional acceptable matches, while keeping processing and selection costs manageable. As a result, donors are most likely to split their giving across multiple recipients at intermediate levels of K, producing an inverted-U pattern in donation counts.

Proposition 2 (**Stability of Total Donation Amount**). *If donors enter the platform with a rough session-level giving target (Equation (8)), the total donation amount may be relatively stable across different screen-set sizes, even as the number of donations varies.*

If donors have a soft budget $\bar{D}$, then modest interface changes mainly affect how donors allocate giving rather than how much they give overall. Varying K changes how easily donors can identify multiple acceptable recipients and how costly it feels to make repeated decisions, which in turn affects the number of donations and the average amount per donation. However, the donor's intended total (or the point at which marginal satisfaction from additional giving becomes small) remains largely unchanged, so total donations need not vary much on average.

Proposition 3 (**Refresh Behavior and Total Exposure**). *Donors in smaller screen sizes (K small) will refresh more frequently but accumulate lower total beneficiary exposure. Donors in larger screen sizes (K large) will refresh less frequently, with total exposure $K \cdot \mathbb{E}[N]$ that may initially increase with K but eventually flatten or decline.*

When K is small, any given screen is less likely to contain a beneficiary whose utility exceeds the donor's reservation threshold, so the donor refreshes more frequently. However, each refresh reveals only a few new beneficiaries, so total exposure accumulates slowly. As K increases, each screen contains a richer set of options, making it more likely that at least one beneficiary meets the donor's threshold and the donor stops searching sooner. Hence, the expected number of screens viewed, $\mathbb{E}[N]$, tends to fall with K. Total exposure is the product $K \cdot \mathbb{E}[N]$, so increasing K can raise exposure initially (because the larger number of beneficiaries per screen outweighs the modest reduction in screens viewed). But further increases in K yield diminishing gains in exposure if $\mathbb{E}[N]$ falls more sharply due to higher processing costs.

3.6. Match Quality and Within-Screen Predictions

The model also yields predictions about which beneficiaries receive donations and how large those donations are. Since d_i^* increases in q_i (Equation (5)), beneficiaries who are perceived as more deserving, or who generate a stronger sense of connection for the donor, tend to attract larger gifts conditional on being selected. Moreover, a higher q_i increases the likelihood that a beneficiary clears the donor's reservation threshold, making it more likely that the donor donates to them in the first place.

Proposition 4 (Deservingness and Donation Outcomes)**.** *Beneficiaries with higher match quality q_i (e.g., those perceived as more deserving, with dependents, or from preferred occupational sectors) are:*

(i) More likely to receive a donation (higher $\Pr(U_i(d_i^*) \geq R(K))$);

(ii) Receive larger donation amounts (higher d_i^*).

Proposition 5 (In-Group Preferences)**.** *If sharing an identity trait with the donor (e.g., gender, nationality, or other salient group membership) raises perceived match quality q_i, then in-group beneficiaries are more likely to be chosen and to receive larger donations than otherwise similar out-group beneficiaries.*

Proposition 6 (Display Order Effects)**.** *When cognitive load is high (large K), donors may exhibit position-dependent attention patterns within screens, with beneficiaries at the beginning and end of screens receiving more attention than those in the middle (serial position effects).*

When K is large, donors face a higher cognitive load in processing all beneficiaries. This can lead to heuristic-based decision making where donors rely on positional cues: beneficiaries at the top of the screen are processed first (primacy effects), while those at the bottom benefit from recency effects as donors prepare to make a decision or scroll. Beneficiaries in the middle receive less attention, creating a "dipping" pattern in donation probability across display positions.

4. Experimental Design, Procedures, and Hypotheses

Section 3 presented a framework explaining why screen size can matter on a platform like *Bagirata*, where donors search over beneficiaries, face some cognitive frictions, and differ in how well they feel matched to the profiles they see, all under a session-level giving constraint. Section 4 describes our platform environment, field experiment design, and procedures to further evaluate the theoretical insights. We organize the analysis in two layers—*between sessions*, where we describe how donation and browsing outcomes vary with screen size, and *within sessions*, where we examine how donors allocate giving across beneficiaries, including the roles of perceived deservingness/match quality, in-group concordance, and display position.

4.1. Experimental Design and Procedures

We administered our experiment to all potential donors who visited the Bagirata website during our study period (October 4th, 2020 to June 9th, 2021).[10] We manipulate screen-size by randomly assigning potential donors to one of the following three between-subject experimental treatments, featuring a 3-, 8-, or 10-set of beneficiaries.

[INSERT FIGURE 2 HERE]

Upon accessing the platform, each donor has an equal chance of being assigned to one of the three treatments. Figure 2 illustrates the treatment assignment. The donor assigned to the 3-beneficiary treatment would see three beneficiaries on their screen. Similarly, those assigned to the 8- and 10-beneficiary treatments would see eight and ten beneficiaries on their screens, respectively. The treatment assignment remains effective for a three-hour session.[11] We do not distinguish between donors and sessions, and hence refer to our unit of analysis as being at the *donor-session* level.

The 10-beneficiary set reflects the default configuration of the platform. As we outlined in the theoretical framework, the set size might influence our outcomes. Compared to a big set, a smaller set may make donors compare options more comfortably, but it also limits the alternatives that they see. Conversely, a big set presents more options for the donors, but beyond a certain threshold, we expect diminishing or even negative returns. In other words, the cognitive burden of evaluating too many options might outweigh the benefits. We selected the 3-beneficiary set to clearly contrast with the larger, 10-beneficiary set. Additionally, as the relationship between set size and outcomes might not be linear, we selected an additional comparison point (8) that is offset from the median value between 3–10.

These set sizes are also informed by the literature on optimal decision-making under choice variations. First, a seminal psychology paper by Miller (1956) suggests that humans can effectively process around seven items, plus or minus two, simultaneously. When this limit is exceeded, our ability to process information decreases significantly. In other words, the cognitive load does not increase steadily with each additional (beneficiary) option. Instead, donors experience a cognitive threshold that is roughly five to nine choices, after which they become more likely to skim the information rather than carefully consider each option. Miller's "7 ± 2"

[10] *Bagirata* connects potential donors and beneficiaries as a two-sided platform. Figure A.1 provides a screen capture of the landing page. The button "*mulai mendistribusikan dana*" is for donors to browse beneficiary cards, while the button "*masuk sebagai penerima dana*" is for beneficiaries to click to initiate the process of asking for aid.

[11] This duration was selected by the platform prior to our research collaboration. During our discussions, the platform founders highlighted their rationale as that of seeking a balance between the minimum length necessary to record multiple donation activities from a single donor (visit) and limiting the length for practical reasons related to data storage. Back-end data suggests that this time window was adequate as the majority of donor(-sessions) completed their donation activities (as recorded by the timestamp of their last donation) within two hours—well within the 3-hour time window. Importantly, for our research design, a single donor who refreshed the donation page and/or re-accessed the platform in this window with the same device would remain in the same screen-size treatment.

rule suggests that the 3-beneficiary sets maps to a low cognitive load, 8-beneficiary sets to an upper range of a manageable cognitive load, and 10-beneficiary sets to a set that exceeds the threshold, triggering potential cognitive overload. Second, Andreoni (2007) provides a benchmark for altruism parameters in the presence of varying group sizes. His experiment led to estimates that for the majority of givers, a \$1 donation to one person is equivalent to one in which n people receive $\$1/n^{0.68}$ each.[12] Additionally, the specific set sizes of 3 and 8 were also implemented in Soyer and Hogarth (2011).

After viewing the first screen, donors have the option of clicking a button at the bottom of each page to trigger a fresh draw of beneficiaries (*refresh*).[13] There is no limit to the number of times potential donors can click refresh. It is important to note that our treatment assignment can lead to differences in the frequency of refresh and variations in the actual number of beneficiaries seen. This variation occurs both within and across treatments. We interpret refresh as a key measure of donors' search effort, a possible mechanism for any differences in donor behavior that we may observe.

A related design concern is that the refresh control ("acak") is located at the bottom of the screen, so reaching it requires more scrolling in the 10-card arm than in the 3-card arm. Two points clarify why this does not overturn our interpretation of refresh as a search-effort proxy. First, placing the refresh option at the end of the list is intentional and common across treatments: it ensures that a "new draw" is requested only after the donor has had the opportunity to view the currently offered set. The extra scrolling in larger-K arms is therefore mechanically bundled with the treatment itself (showing more cards), rather than an orthogonal friction unrelated to screen-set size. Second, donors are not restricted to the bottom button to obtain a new draw: as noted in footnote 11, our refresh measure also captures standard browser refresh/back actions, which have comparable physical effort across arms. Taken together, the treatment differences in refresh are best read as reflecting donors' willingness to continue searching after being exposed to larger menus, not merely a scrolling artifact.

Throughout the paper, we focus on two levels of analysis: (i) the impact of screen size on donor behavior, where the unit of analysis is at the donor-session level (*between-donor analysis*), and (ii) the determinants of donations within a single donor-session, where the unit of analysis is at the beneficiary display level (*within-donor analysis*).

For the between-donor analysis, we examine whether there is any significant difference in the donation behaviors of the donors between different treatment arms. If donations differ, we explore the possible reasons and the mechanisms underlying the differences across various screen sizes. Specifically, we ask whether a smaller screen size prompts potential donors to find a larger sample of beneficiaries by clicking the refresh button more frequently.

[12] In our set-up, the equivalent donation for n=8 (0.24) is half the size of that of n=3 (0.47). The platform's existing configurations prevented us from proposing a treatment with n = 24 (equiv. donation 0.12), but our 10-screen set size would still be associated with a lower equivalent donation.

[13] As the screenshot in Figure 1 shows, this button was labeled "*acak*" at the bottom left corner, which has the literal translation "to randomize." Hereafter, we refer to this action, and also the action of pressing the back button, as a "refresh" action to combine it with a browser refresh.

For the within-donor analysis, a potential donor would encounter multiple beneficiaries across sets/screens in a session. Here, we consider each dyadic pair of a potential donor and a beneficiary within a donor-session as a single unit of observation. Crucially, we leverage the platform's algorithm to study the effect of beneficiary characteristics on donation behaviors. In particular, the platform's algorithm selects a random card from the database of all potential beneficiaries for each screen that the donors see, allowing us to leverage the as-good-as-random display of beneficiary characteristics to study the effects of deservingness on donor behavior.

In both the desktop and mobile versions of the website, the beneficiary cards are displayed to donors in vertical succession. The random draw from the beneficiary database that the platform performs for each card also means that the order in which beneficiary cards are displayed is random. This allows us to estimate the effect of sequential order on donations, i.e., whether there are differences in donation outcomes between beneficiaries displayed closer to the top vis-à-vis those displayed closer to the bottom of each draw.

4.2. Hypotheses

Building on the theoretical framework developed in Section 3, we derive testable hypotheses about donor behavior. Our hypotheses distinguish between *between-donor effects* (how screen size affects donation outcomes) and *within-donor effects* (what affects donation decisions within a given donor session).

4.2.1. Between-Donor Hypotheses

Our proposition 1 predicts a non-monotonic, inverted-U relationship between screen size and the number of donations. When the screen size is too small, donors face high search costs and limited variety, leading to fewer donations. When the screen size is too large, cognitive overload and high per-donation selection costs discourage multiple donations. An intermediate screen size balances these forces.

Hypothesis 1 (Inverted-U Relationship: Number of Donations)**.** *The number of donations per donor session follows an inverted-U pattern with respect to screen size K. Specifically, donors in the 8-beneficiary treatment will make more donations than donors in both the 3-beneficiary and 10-beneficiary treatments.*

Proposition 2 suggests that total donation amounts may be relatively stable across screen sizes due to session-level giving targets. However, if cognitive overload significantly reduces engagement, we may observe some variation.

Hypothesis 2 (Total Donation Amount)**.** *The total donation amount per donor session may be relatively stable across treatments due to session-level giving targets.*

The probability of making any donation depends on whether donors encounter at least one beneficiary whose match quality exceeds the reservation threshold. The variety effect (more options increase the chance of a good match) and the cognitive overload effect (higher costs may reduce engagement) may roughly offset each other.

Hypothesis 3 (Probability of Donating)**.** *The probability of making any donation does not differ significantly across screen size treatments, as the variety effect and cognitive overload effect approximately cancel each other out.*

Proposition 3 yields predictions about search behavior.

Hypothesis 4 (Refresh Behavior and Exposure)**.** *(i) Donors in the 3-beneficiary treatment will refresh more frequently than donors in the 8- and 10-beneficiary treatments. (ii) Despite higher refresh rates, donors in the 3-beneficiary treatment will have lower total beneficiary exposure compared to donors in the 8- and 10-beneficiary treatments.*

The above hypotheses provide us with a unified framework to test the impact of our experimental manipulation on donation outcomes and search behavior.[14] Notably, on the effects of changes in the number of donation options on donor behavior, our proposed, inverted-U pattern in H1 could potentially reconcile a number of diverging results in the extant literature. Specifically, Bernasconi et al. (2009) provides evidence from a public goods game that splitting a single public goods project into two could lead to an increase in total contributions through, perhaps, providing more detailed information on each project (the "unpacking effect"). On the other hand, two strands of literature suggest the opposite effect. First, lab experiments suggest that competition among charities (an increase in options) might lead to reductions in giving (Filiz-Ozbay and Uler, 2019). Second, a rich literature in psychology demonstrates that increasing the salience of individual beneficiaries (i.e. with fewer options) can garner more donations (Sudhir et al., 2016). Our theoretical framework suggests that a non-monotonic effect of set size could be a potential reason for these discrepancies.

4.2.2. Within-Donor Hypotheses

Within a given donor session, we exploit the platform's algorithm to test the following additional hypotheses. We note that while these tests do not follow from explicit experimental manipulation, the algorithms' random draws allow us to produce exploratory evidence and disentangle possible mechanisms that drive our main results. These analyses remain connected to our theoretical framework in the preceding section. Proposition 4 from the framework predicts that beneficiaries with higher perceived match quality (deservingness) will receive more donations and larger amounts.

Hypothesis 5 (Deservingness)**.** *(i) Beneficiaries perceived as more deserving (e.g., breadwinners with dependents, those in certain occupational sectors) are more likely to receive donations. (ii) Beneficiaries perceived as more deserving receive larger donation amounts.*

Proposition 5 predicts in-group bias based on shared identity traits.

[14]Our study is preregistered at the Open Science Framework (OSF). We discuss our pre-registration in more detail in Appendix B. We state which analyses were pre-registered and which are exploratory in our results sections.

Hypothesis 6 (In-Group Bias)**.** *Donors are more likely to donate to beneficiaries who share their identity characteristics (gender, religion, ethnicity, or location), and these donations are larger in amount.*

The conception of in-group biases in our analysis relates to the concept of social distance (Chen and Li, 2009; Meer and Rigbi, 2013; Sudhir et al., 2016) and here we investigate four dimensions of identity: gender, religion, ethnicity, and location. We note that as the different identities are not uniformly distributed in our beneficiary database, we also investigate if the distribution of characteristics influences our donors' decision to donate.

Proposition 6 predicts position-dependent attention patterns, especially for larger screen sizes.

Hypothesis 7 (Display Order Effects)**.** *(i) Beneficiaries displayed at the top and bottom of screens are more likely to receive donations than those in the middle (serial position effects). (ii) This dipping pattern is more pronounced in larger screen sizes (8- and 10-beneficiary treatments) compared to the 3-beneficiary treatment.*

Additional Analysis. We also explore additional analysis to consider alternative mechanisms that could also influence decision donation. We preview one of these analyses here: we explored the influence of existing donations. Earlier donations by other donors may influence donors to either crowd in the existing donations or pull back and give to other beneficiaries. On one hand, donors may see donations from other donors as a positive signal (a "social proof" as highlighted by psychologists (Cialdini, 2007)) that might lead to a further crowding-in of donations (Huck and Rasul, 2011; Meer, 2014). On the other, the institutional context of our setting might lead to an opposite effect. The donation platform's name, *Bagirata*, might nudge donors to divide their donations more equally across multiple donors (the literal translation of "*Bagirata*" means to "give equally").

5. Empirical Strategy

Our empirical analysis proceeds at two levels: (i) *between-donor analysis*, examining how screen size affects donation outcomes at the donor-session level, and (ii) *within-donor analysis*, examining how beneficiary characteristics affect donation decisions at the donor-beneficiary dyad level.

5.1. Between-Donor Analysis

For the between-donor analysis, we estimate the effect of screen size on donation outcomes using the following specification:

$$Outcome_i = \alpha_1 + \beta_1\, ScreenSize3_i + \beta_2\, ScreenSize8_i + \gamma' X_t + \varepsilon_{1i}. \tag{9}$$

where $Outcome_i$ is a donation outcome measure for donor session i; $ScreenSize3_i$ and $ScreenSize8_i$ are indicator variables equal to 1 if donor i was assigned to the 3- or 8-beneficiary screen size, respectively (with 10-beneficiary as the omitted category); X_t is a vector of control variables including COVID-19 related measures (weekly average positivity rate and fatality counts) and time fixed effects (month, biweekly period, and day-of-week indicators); and ε_{1i} is the idiosyncratic error term.

Because *Bagirata* is a live platform, the "% of ask already funded" shown on a beneficiary card can change over time as earlier donors contribute. This does not undermine the identification of the screen size effect because the screen size is randomized at session entry, and the platform does not operate separate beneficiary pools by treatment arm, so any aggregate evolution in funding progress applies to all conditions. We therefore interpret our coefficients as *intent-to-treat* effects of screen size on donor behavior, conditional on the platform conditions at entry. This is the policy-relevant effect of displaying more versus fewer options on a screen, holding fixed the platform's matching and updating process as it operates in practice.

The coefficients of interest are β_1 and β_2, which measure the difference in donor behavior between donors assigned to the 3- or 8-beneficiary screen arms relative to donors assigned to the 10-beneficiary screen arm. Our between-donor randomized assignment in the number of beneficiaries per screen underpins our use of the OLS specification. Under random assignment, $\mathbb{E}[\text{ScreenSize}_i \cdot \varepsilon_{1i}] = 0$, ensuring consistent estimation of the treatment effects.

We examine three primary outcome variables. First, *number of donations,* which is the total count of donations made during the session. Second, *total donation amount,* which is the sum of all donations (in USD) made during the session. Third, the *donation indicator*, a binary variable equal to 1 if the donor made at least one donation during the session.

To explore mechanisms, we also estimate Equation (9) with the following additional outcome variables. *Refresh button clicks,* which is the number of times the donor clicked the refresh button. *Total beneficiary exposure,* which is the total number of unique beneficiaries viewed ($K\times$ number of screens viewed). *Time per beneficiary,* which is the average time spent per beneficiary (in minutes). We use robust standard errors throughout to account for heteroskedasticity. In some specifications, we restrict the sample to periods of "normal traffic" (excluding the mobility lockdown period and social media campaign periods) to ensure our results are not driven by unusual platform traffic.[15]

5.2. Within-Donor Analysis

For the within-donor analysis, we exploit the random draw of beneficiaries within each donor session to estimate the effect of beneficiary characteristics on donation outcomes. The unit of observation is the donor-beneficiary dyad. Because the platform's algorithm randomly draws ben-

[15]The relevant mobility lockdown in our study period is the first PPKM policy, enacted between 10-20 January 2021. The social media campaign referred to a May Day publicity campaign, which was active between April 23rd and May 13th, 2021. See Figure A.2 for a plot of visits and COVID statistics over time.

eficiaries from the database for each screen, beneficiary characteristics are as-good-as-randomly assigned, conditional on the donor session. This provides within-session variation that helps isolate the contribution of beneficiary attributes to donor behavior.

Our within-donor analysis examines three dimensions of beneficiary characteristics that may affect donation decisions: (i) *deservingness*—whether beneficiaries possess characteristics donors view as meriting support; (ii) *in-group identity*—whether beneficiaries share identity traits with donors; and (iii) *display order*—whether beneficiaries' position within a screen affects their likelihood of receiving donations.

Baseline Within-Donor Specification We estimate the following baseline specification at the donor-beneficiary dyad level:

$$Y_{ijkl} = \alpha_2 + \beta_3' \mathbf{X}_j + \phi_i + \varepsilon_{2ijkl}, \tag{10}$$

where Y_{ijkl} is the donation outcome (either an indicator for whether a donation was made or the donation amount) for donor session i seeing beneficiary j in the k-th screen at position l within that screen; $\mathbf{X}_j$ is a vector of beneficiary characteristics; ϕ_i are donor-session fixed effects; and ε_{2ijkl} is the error term.

The donor-session fixed effects ϕ_i absorb all time-invariant characteristics of the donor session, including the screen size treatment assignment. This ensures that identification of β_3 comes from variation in beneficiary characteristics *within* a donor session, not across donor sessions. Standard errors are clustered at both the donor-session and beneficiary levels to account for potential correlation in errors within donor sessions and within beneficiaries.

We measure narrative-based beneficiary characteristics through three complementary approaches. First, we hand-code salient attributes from beneficiaries' self-written appeals, guided by donors' stated markers of deservingness. Second, we construct a composite text-based deservingness index using Keyness Statistics (KS) and Latent Semantic Scaling (LSS) (Zollinger, 2022), which summarize how closely each appeal resembles narratives that receive donations versus those that do not. Third, we complement these measures with exploratory NLP summaries, emotion ratings QWEN3 (Yang et al., 2025) and topic grouping IndoBERT to cluster narratives into small number (Koto et al., 2020; Aji et al., 2022), to capture affective tone and thematic content in appeals that may matter for donors but are not well captured by specific hand-coded indicators or frequency-driven key terms. We describe each approach below.

Hand-Coded Beneficiary Characteristics To obtain a complete set of beneficiary characteristics that donors might consider, we hand-coded thousands of self-written beneficiary narratives. Our coding scheme was informed by donor survey responses about what makes beneficiaries deserving. Table A.3 summarizes donor responses: beneficiaries who are breadwinners (86%); individuals in persistent poverty (85%); those hit by unforeseen circumstances such as disasters, illnesses, or job loss (82%); and female beneficiaries (69%) were thought of as the *most* deserving of donations.

To ensure accuracy and mirror donors' reading of narratives, we tasked two Indonesian research assistants with manually reading each narrative.[16] Following donor survey responses, we coded characteristics including:

- **Household status**: indicators for breadwinner, having dependents, and mentioning children;
- **Occupational sector**: hospitality/retail/food service, education/government/health, art and creatives, transportation, finance/IT, and other;
- **Demographics**: gender and religious identity (imputed from names), regional location;
- **Narrative features**: narrative length (word count).

We illustrate this coding process in Table A.1. For example, a beneficiary who writes "*The restaurant where I worked closed .... I am a mother with two children.*" is coded as a breadwinner with dependent children in the hospitality sector. In contrast, a beneficiary with an occupation as a graphic designer and writes "*My workplace is closed indefinitely, and my salary is cut by 50% while working from home.*" is coded as having no dependents in the art and creatives sector.

To test Hypothesis 5 on deservingness, we estimate:

$$Y_{ijkl} = \alpha_3 + \sum_{m=1}^{M} \beta_{4m}\, Deservingness_{mj} + \delta'\, Controls_j + \phi_i + \varepsilon_{3ijkl}. \tag{11}$$

where $Deservingness_{mj}$ is the m-th deservingness characteristic of beneficiary j (e.g., breadwinner status, having dependent children, occupational sector); $Controls_j$ includes beneficiary-level controls (ask amount, ask duration, narrative length, display order); and ε_{3ijkl} is the error term clustered at donor-session and beneficiary levels. The coefficients β_{4m} capture the effect of each deservingness characteristic on donation outcomes, holding constant other beneficiary attributes.

Textual Analysis: Deservingness Index To corroborate our hand-coded measures and capture dimensions of deservingness that may not be explicitly coded, we constructed a composite deservingness index using textual analysis. We employ Keyness Statistics (KS) and Latent Semantic Scaling (LSS) (Zollinger, 2022) to classify narratives based on their similarity to those that received donations versus those that did not.

The KS method computes a χ^2 statistic for each term in a narrative, ranking the most frequently mentioned terms for beneficiaries who received at least one donation versus those who did not. We used KS to identify seed words for LSS processing, which combines word-based supervision with an unsupervised model for synonymous words to process our noisy

[16] To minimize biases in coding, the two research assistants have complementary backgrounds: one is female, and the other is male; their ethnic backgrounds include Javanese and Batak from Sumatra; and their religious affiliations encompass Muslim and Protestant Christian. Disagreements in coding are resolved through detailed manual review by one of the authors.

corpus, where beneficiaries wrote in a mix of formal and informal language.[17] We rescale the LSS statistic for each narrative to range from 0 (least similar to successful narratives) to 1 (most similar to successful narratives).

To test Hypothesis 5 using the textual measure, we estimate:

$$Y_{ijkl} = \alpha_4 + \beta_5 \, DeservingnessIndex_j + \lambda' \, Controls_j + \phi_i + \varepsilon_{4ijkl}. \tag{12}$$

where $DeservingnessIndex_j \in [0, 1]$ is the LSS-based deservingness index for beneficiary j. The coefficient β_5 captures the effect of a one-unit increase in the deservingness index on donation outcomes.

Exploratory NLP: Emotion and Topic Features. In addition to the KS/LSS-based index, we implement two complementary tools for narrative summary. Using two (pre-trained) language models, we (i) apply QWEN3 (Yang et al., 2025) to classify each appeal into an emotion category and relate these ratings to donation outcomes, and (ii) use IndoBERT-based representations (Koto et al., 2020; Aji et al., 2022) to cluster narratives into a small number of recurring themes (topics) and estimate how these themes predict giving. These NLP summaries differ from KS and LSS, which are corpus-anchored measures. KS identifies terms disproportionately associated with narratives that receive (versus do not receive) donations, and LSS uses those terms as supervision to produce a continuous narrative score along a donation-associated semantic dimension. In contrast, emotion ratings and topic groupings summarize what is expressed in the narrative, its emotional register and thematic focus. Hence, we treat these analyses as complementary, exploratory features that descriptively augment (rather than replace) the KS/LSS deservingness index and our hand-coded measures.

In-Group Bias To test Hypothesis 6, we use a matched sample of donors and beneficiaries where we observe both donor and beneficiary characteristics. We match donors from our user survey to beneficiaries using email addresses, yielding 78 donor sessions matched with 1,283 beneficiaries (2,396 observations).

We estimate:

$$Y_{ij} = \alpha_5 + \beta_6' \, Concordance_{ij} + \omega' \, Controls_j + \phi_i + \varepsilon_{5ij}. \tag{13}$$

where $Concordance_{ij}$ is a vector of indicators for whether donor i and beneficiary j share the same gender, religion, ethnicity, or location. We code donor characteristics from survey responses and beneficiary characteristics from names (gender, religion) and residence (ethnicity).[18] [19]

[17] In political science, this method has been used to identify right- versus left-leaning voters from self-written voter descriptions (Zollinger, 2022). We describe this methodology in detail in Appendix B.

[18] We illustrate our coding process as follows. For example, a beneficiary has a first name that is a masculine Javanese word and a surname that is an Arabic word; our assistant coded his name as both masculine and Muslim. Furthermore, as this beneficiary resides in Central Java, an area with a predominantly ethnic Javanese population, we coded his ethnicity as Javanese, which is concordant with information from his name. For another beneficiary whose name resembles an Arabic word related to the popular male Muslim name Muhammad, our assistants inferred his name to be masculine and Muslim.

[19] In a small number of cases, gender imputed from first names may conflict with gender cues in the narrative

Display Order Effects To test Hypothesis 7 regarding display order effects, we estimate:

$$Y_{ijkl} = \alpha_6 + \sum_{p=1}^{K} \beta_{7p}\, \mathbb{1}[Position_{ijkl} = p] + \phi_i + \varepsilon_{6ijkl}. \tag{14}$$

where $\mathbb{1}[Position_{ijkl} = p]$ is an indicator for whether beneficiary j is displayed at position p within the screen, and K is the screen size (3, 8, or 10). We also estimate specifications with indicators for top positions (first 3–4), middle positions, and bottom positions (last 3–5) to test for serial position effects.

6. Results

6.1. Donations Statistics, Beneficiary Characteristics, and Balance Tests

Before turning to the main results, we summarize key features of the data and document the balance of our experimental assignment. Our main data set comprises 2,405 unique donor–sessions and 2,054 unique beneficiaries. Table 1, Panel A presents summary statistics of donation activities from the *beneficiary's* perspective (N=2,054). Each beneficiary is randomly drawn to be displayed to donors 26.2 times on average.[20] Across all beneficiaries, 81% received at least one donation, with the average beneficiary receiving 2.09 donations for a cumulative sum of USD 17.84 (Panel A).[21] Compared to the average annual beneficiary earnings in our matched user survey (USD 1,882), the total donations received by the average beneficiary correspond to roughly 11% of average monthly earnings.

Donation activity and browsing patterns. Appendix Table A.4 presents selected summary statistics on session-level donation outcomes by treatment assignment. The probability of sending any donation is similar across arms (0.17–0.19), while the distribution of donation outcomes is highly right-skewed and features a large mass at zero (the median is zero for all three outcomes across arms). In the 3-option, 8-option, and 10-option arms, the mean number of donations per donor–session is 0.46, 0.59, and 0.42, respectively, and the mean total donations per session are USD 4.39, 6.67, and 5.08.

Appendix Table A.5 summarizes session-level browsing behavior. On average, visitors assigned to smaller screen sets view more sets before exiting (mean sets seen: 4.6 in the 3-option arm versus 3.2 and 2.6 in the 8- and 10-option arms). Donors who eventually give tend to browse more than non-donors (overall means: 5.3 versus 3.0 sets). Conditional on making a donation, giving typically occurs within the first few sets: the average "earliest set" in which a donation is made is 2.6 overall (3.9, 1.9, and 2.1 in the 3-, 8-, and 10-option arms, respectively).

text (e.g., a beneficiary with a feminine-coded name whose narrative states he is "providing for my pregnant wife"). We treat this as measurement noise in the concordance indicator

[20] Each beneficiary is limited to only one appearance per session. Hence, on average, beneficiaries are displayed about 26 times across unique sets.

[21] While donations are made in Indonesian rupiah (IDR), throughout the analysis, we express donation amounts in US dollars. We use a conversion rate of USD 1 = IDR 14,000.

These descriptive patterns motivate our focus on the trade-off between variety and the frictions associated with processing larger screen sets. Note that donors are not constrained to making at most one donation. 65% of donors who donated actually sent multiple donations.

Beneficiary characteristics and appeals. On average, a beneficiary asked for USD 139 per month over a duration of 2.2 months (Table 1, Panel A). Our systematic coding from beneficiary narratives and names allows us to classify beneficiaries along several salient dimensions, including gender, religion, household status (e.g., whether the beneficiary describes being a breadwinner or having dependents), region, and employment sector. From names, we impute that most beneficiaries have masculine names (63%) and Muslim names (82%) (Table 1, Panel B). With respect to household structure, 22% of beneficiaries mention being the family breadwinner or having dependents, and 12% explicitly mention dependent children. In terms of employment, most beneficiaries are in hospitality/retail/food service (61%), followed by creatives (16%) and transportation (6%, largely ride-share drivers). Geographically, most beneficiaries are located in the Jakarta metro area (67%), with 9% based outside of Java.

Table A.6 summarizes beneficiaries' total appeal amounts (computed as monthly request times requested duration) by subgroup. Total appeals vary across sectors and household categories, providing meaningful heterogeneity in needs within the platform (e.g., average total appeals range from roughly USD 196 in healthcare to about USD 426 among creatives). Tables A.7–A.8 complement this picture by describing exposure and realized giving by beneficiary subgroup. For example, beneficiaries who mention being a breadwinner or having dependents are more likely to receive at least one donation (90% versus 78%) and receive larger cumulative donations on average (USD 27.80 versus 14.95). Similarly, beneficiaries who mention dependent children exhibit higher receipt rates (93%) and larger cumulative donations (USD 31.74) than those who do not. These are descriptive differences, but they foreshadow the importance of beneficiary attributes for within-session allocation decisions that we examine later in the paper.

Donor–beneficiary comparison in the matched survey subsample. Table 1, Panel B summarizes characteristics of beneficiaries and donors in our user survey subsample. Beneficiaries report lower levels of education and substantially lower earnings than donors: the average donor earns USD 8,626 per year, almost five times the average beneficiary earnings (USD 1,882). Despite this disparity, donors and beneficiaries report allocating a similar share of income to charity (about 5–6%). For context, the millennial age group in the US reports giving on average only 0.9% of income (Clark et al., 2019), suggesting that giving norms in our setting may differ from those documented in developed-country samples.

Balance tests. We next verify that the platform's randomization of visitors into treatment arms yields comparable groups. Table A.9 shows balance in donor entry timing across treatment arms (across time periods, day-of-week, time-of-day, and levels of average COVID positivity rates). Because the platform collects limited information on visitors, we additionally examine the balance using the matched survey subsample. In Table A.10, donor characteristics do not differ statistically significantly across treatment groups along key observables (including gender, marital status, religion, education, location, earnings, workplace type, and COVID-related experiences).

With respect to beneficiaries, Table A.11 confirms that the average mix of beneficiary characteristics displayed is comparable across the 3-, 8-, and 10-option arms. Together, these checks suggest that any differences in outcomes across arms in the next section are unlikely to be driven by systematic differences in who enters the platform or which beneficiaries are shown. Within a donor session, the characteristics of the beneficiaries are also not correlated with their position within a set (Table A.12). In other words, there are no systematic differences between beneficiaries shown at the top, middle, and bottom of each set.

6.2. Between–Donor Analysis

This section presents the results of the tests of our predictions. We analyze three donation outcomes in Section 6.2.1 and the trade-off between variety and cognitive overload in Section 6.2.2. We discuss alternative explanations such as decision complexity and regret in Section 6.2.3.

6.2.1. Main Results: Donation Outcomes

We first investigate whether differences in screen set size lead to differences in the number of donations among donors across treatment arms. Table 2 presents OLS results. Throughout, we control for two measures of COVID-19 severity of weekly mean positivity rate and daily deaths. In Column 1, the estimated coefficient for the 8-option set group is 0.185 higher than the 10-option set control groups, and it is significant at the 5% level. The coefficient for the 3-option set group, on the other hand, is not statistically significantly different from zero. The COVID variables we included are the weekly average positivity rate and the fatality counts. Restricting the regression to the subsample outside the first mobility restriction period in 2021 and outside the May Day social media campaign increases the estimated coefficient for the 3-option sets. We see that during normal traffic, the number of donations made by donors in the 3-option sets is higher than the 10-option sets control group, but this coefficient is still smaller than the coefficient for the 8-option sets. This sample restriction to periods during normal traffic to the platform also increased the estimated coefficients for the 8-option sets. The raw statistics in Table A.4 display a consistent pattern across the mean and 90th percentile of donations per session, with the 8-option group showing the highest values. This is our first result.

Result 1. *The average number of donations that donors make increases with a reduction in the number of potential beneficiaries, but further reduction in the screen set size also reduces the number of donations. Thus, the effect of the number of potential beneficiaries on the number of donations is non-linear and follows an inverted-U pattern.*

The above result should be read as a statement about the *pattern of point estimates* across our three screen designs, rather than as a statistically established difference between the 3-

card and 8-card arms. Table 2 reports effects relative to the 10-card benchmark. The 8-card condition is significantly higher than 10 on several donation outcomes, while the 3-card condition is typically not distinguishable from 10.[22] The wording "further reduction" is therefore intended to summarize the ordering $\widehat{\beta}_8 > \widehat{\beta}_{10}$ and $\widehat{\beta}_3 \lesssim \widehat{\beta}_{10}$ (in point estimates), which is consistent with diminishing returns to variety at very small menus, but it does not by itself imply that the 8-card and 3-card arms differ significantly from each other.

We next investigate the impact of screen set size on the donation amount. In Table 2, Column 3 shows that the differences in donation amounts across groups are not statistically significantly different from zero. However, in the subsample of normal traffic period to the platform, the total donation from the 8-option group is statistically significantly higher than the 10-option set control group (Column 4, Table 2).[23] The total donation for the 3-option group in this subsample still does not differ from the control group.

Result 2. *The total donation amount does not differ significantly between donors to whom different numbers of potential beneficiaries are being displayed. However, an inverted-U pattern in total donation amounts can be recovered in some situations.*

Our analysis of whether differences in screen set size lead to differences in the probability of making a donation among donors in different treatment arms is presented in Column 5-6 of Table 2. The result shows that, compared to the 10-option sets group, the probability of sending any donation does not differ significantly from those in the 3- and 8-option sets groups. Restricting the regression to periods of normal traffic did not significantly change the estimated coefficients for the screen-set-size indicators. Regressions using probit and logit also fail to reject the null hypothesis for this outcome (not shown). These regression results are broadly consistent with the raw differences in Table A.4. The 8-option arm shows an increase of about 2 p.p. in the share of donors who give, which is close to the corresponding regression estimate in Table 2 (Col. 5). The key difference is statistical precision. Once we account for sampling variation (and controls, where applicable), we cannot reject zero at conventional levels of statistical significance. In sum, the evidence does not support a statistically significant increase in the extensive margin of a donation.

Result 3. *The probability of making any donation does not differ significantly between donors to whom different numbers of potential beneficiaries are being displayed.*

[INSERT TABLE 2 HERE]

The key results above are naturally situated in the context of beneficiaries seeking COVID-19 relief, although the long duration of our field experiment also allows us to confirm the patterns of donation outcomes outside specific times of intense COVID crisis. These findings also allow us

[22] Except for Column (2)

[23] One possible explanation is that our "soft" intervention of varying screen-set size might have been more effective during weeks of normal traffic. In another words, weeks of severe COVID movement restrictions and social media campaigns might have attracted donors with an innately higher propensity to make donations, regardless of screen-set size (*selection*).

to connect with existing studies that investigate charitable actions during disasters. Deryugina and Marx (2021) find that a tornado disaster increased total donations by individuals close to the tornado's path, leading to the conclusion that the supply of donations is not fixed. In the temporal dimension, Scharf et al. (2022) documented that donors respond to charity appeals during a crisis with increased donations that were not offset by lower donations later in time. In contrast, Jayaraman et al. (2023) argued that donations on a German online giving platform declined over a 2-month period following other disasters after Typhoon Haiyan in 2013 and the 2015 Nepal earthquake, large natural disasters that already attracted massive media attention and funding. Our study builds on these studies by showing how donation behavior responds to variation in choice set sizes. Importantly, we further provide specific tests of mechanisms leveraging on rich, backend data. We discuss this in the next section.

6.2.2. Mechanisms: Refresh Behavior and Total Exposure (Cognitive Load vs Match Quality)

Why does screen set size influence donor behavior? In our framework, variety competes with cognitive overload. Variety effect due to a larger screen set size increases the likelihood of encountering a high-quality match, i.e., a more deserving individual. At the same time, beneficiary information is noisy, and a larger screen set potentially increases cognitive cost.

We test for this trade-off using rich back-end data. Specifically, as proxies for search behavior and cognitive load, we examine the effect of screen size on donor refresh rates; total beneficiary exposure; and time spent per beneficiary. As a measure of match quality, we use our beneficiary-specific deservigness index (described in Section 5.2) to construct aggregate measures of deservingness per screen-set. Specifically, we interpret a higher level of *deservingness per screen* as reflecting greater match quality.

The platform is designed so that donors can tap the *refresh* button at the bottom of each screen to obtain a new random draw of beneficiaries. Furthermore, any differences in refresh rates would also affect the total number of beneficiaries that each donor is exposed to. Hence, as a proxy for search behavior, we regress refresh rates on our treatment indicators. To proxy for cognitive load, we use the total number of beneficiaries and the time spent per beneficiary. We interpret a higher number of beneficiaries a donor is exposed to, as suggestive of a *higher* cognitive load. [24]

Table 3 displays regression results. Columns (1)-(2) estimates that, on average, donors assigned to 3-option screen sizes click *refresh* 1.6-1.9 additional times relative to donors in 10-option screen sizes. In contrast, there is no statistically significant difference between refresh rates in 8- and 10-option screen sizes. Columns (3)-(4) shows that donors in 3-option screen sizes are exposed to 13-15 *fewer* beneficiaries relative to 10-option screens. This suggests that 3-option screen donors feel that they have viewed too few beneficiaries and hence continue to defer the donation decision and click refresh. Yet, 3-option screen donors continue to view fewer

[24] Our within–donor analysis on the influence of display order provides another indication of the role of cognitive overload that donors face in processing beneficiaries' information. We defer this analysis to Section 6.3.4.

potential beneficiaries in all subsequent screens. Again, there is little difference between 8- and 10-option screen sizes.

Result 4a. *The number of refresh count is highest among donors in small screen size compared to donors in large screen sizes.*

Result 4b. *Total beneficiary exposure to beneficiaries is lower for donors in the small screen size compared to donors in the largest screen sizes, but the difference for the two large screen sizes is not significant.*

We also find that donors spend *less* time per beneficiary when they are exposed to a large screen-set size. In the subsample of sessions where donations occurred (Table 3, Column 5), a donor spends an average of 0.92 minutes (55 seconds) longer on each beneficiary in a 3-beneficiary screen size donor session. Imputing the median value of deliberation time among donors to all non-donating sessions, donors in the 10-option sets spend the least amount of time per beneficiary compared to the 8- and 3-option sets (Columns 6 and 7). Why might this be the case? We argue that a *lower* number of minutes spent per beneficiary in the 10-opt sets vis-a-vis the 3-opt sets might actually reflect an excessively high cognitive load. Beyond a threshold, donors in larger (10-)screen sizes encounter too large a number of beneficiaries (possibly both within and across screens), leading them to engage less with beneficiaries and make fewer donations.

In Table A.13, we further explore the role of beneficiary text appeals as a primary source of cognitive load. We interpret a *greater dispersion* in beneficiary appeal length (*Maximum - minimum appeal length (mean/set)*) as potentially reflecting the imposition of a *greater cognitive load* on donors. In Columns (1)-(2), we find that appeals for 3- and 8-option set donors are more uniform in length than those in the 10-option set, with smaller differences in appeal length between the shortest and longest appeals. These results suggest that greater dispersion in beneficiary appeals might significantly increase the cognitive load of donors in the 10-option sets and affects the time they spend deliberating on donation choices. Interestingly, Columns (1)-(2) suggest that donors in the 8-option sets continue to face a comparatively lower cognitive load than 10-option sets despite just a difference of 2-options per set. We interpret this as being in line with a *Cognitive Overload Effect*. 8-option sets might constitute a cognitive threshold where, due to $C(K)$ (the cognitive cost of processing a screen of size K) being increasing and convex in K, any further increase in the number of options leads to excessively high cognitive costs, and a disproportionately sharp decrease in donation behavior.

In terms of match quality, Table A.18, shows that donors in 3-option sets see fewer potential beneficiaries above a certain threshold of deservingness. If, as posited in our theoretical framework, donors were to donate only to those with a median value of the deservingness index or above, donors in 3-option sets see 6.4–7.8 fewer beneficiaries meeting this criterion compared to donors in 10-option sets. For higher thresholds (p75, p90), 3-option sets still see 1.3–3.9 fewer beneficiaries meeting the criteria. On the other hand, the total number of 'suitable' beneficiaries that donors in 8-option sets see, is not statistically significantly different than the 10-option control group.

Taken together, these results suggest that the 8-option screen set size balances the variety-cognitive overload trade-off. These donors face lower cognitive loads than those in 10-option sets (in terms of dispersion in beneficiary appeal length), but, unlike in 3-option sets, this choice architecture did not necessitate a higher effort to search for donation targets yet enabled donors to encounter a beneficiary mix of sufficiently high match quality, leading to improved donation outcomes. In another words, the 3-option screens induce greater but less optimal donor (search) effort: donors refresh more often to search for more optimal targets, but, due to search fatigue, end up being exposed to fewer total beneficiaries, with no difference in donation rates compared to the 10-option beneficiaries.

6.2.3. Robustness Checks and Additional Mechanism Tests

Concurrent campaigns and the saliency of donation during COVID-19. We investigate whether social media campaigns run by the *Bagirata* platform increased the salience of (making) donations during the emergency period of the pandemic. We use Google Trends data to track the weekly popularity of three donation-related keywords: *donasi*, *sedekah* (Islamic charity), and *bansos* (social assistance). We find that the overall popularity of these terms as recorded by Google did not change during the weeks of the first mobility restriction in 2021 (*PPKM*), the periods immediately following this mobility restriction, or around May Day when the social media campaign was underway (Table A.14). These two periods were marked by spikes in platform visit statistics. Instead, there are some associations with fatality or positivity rates, although, in aggregate across contemporary and lagged values, effects appear muted.

Decision complexity. Many decision problems are complex processes, and it is plausible that the complexity plays an important role in observed decisions agents are making (Enke and Shubatt, 2024). We examine the role of complexity in our context through additional analysis. In our experiment, individual donors face identical decision structures in each set (except for the number of alternatives and the target beneficiaries). This structure suggests that the complexity for our donors lies in assessing the appeal of the text that they see. We consider key elements of text complexity (unique words and hard words) and their interaction with our experimental manipulation. In Table A.15, we consider text to be more complex if it has more unique words within the text or harder words. We define hard words as words infrequently used in the corpus derived from the newspaper *Kompas* and the Indonesian Wikipedia. We then calculate, for each donor, whether they are exposed to a set with a complexity measure above the median, and include this indicator in the regression, interacted with the set size indicator, and on its own.

Table A.15, Columns 1-4 show that two donation outcomes for the 8-option set group remain higher than the control group. We also see improved precision for the 3-option set for the number of donations. Estimates for the complexity measure show that seeing texts with a share of unique words above the median is positively associated with making more donations, but the pattern of its interaction with the set-size indicators across various donation outcomes is largely not statistically significantly different from zero. For our measure of hard words (Columns 5-8), the coefficient sizes remain similar in magnitude for the 8-option set indicator, although they are no longer statistically significant. By and large, the positive relationship between the set of 8 and donation outcomes persists regardless of the text's complexity. This suggests that the role of decision complexity may be limited compared to the trade-off between variety and cognitive overload.

Regret. Donors can refresh the selection of beneficiaries at the end of each screen, but once they do, they cannot move back to a previous screen. Could donors' charity decisions be influenced by regret? With a new draw, they could lose access to beneficiaries they perceived as more deserving than those newly displayed. We explore the role of regret by analyzing sub-samples where regret motives are likely to be minimized. We present results in Table A.16.

First, we replicate our baseline, donor-level analysis by restricting our sample to the first three beneficiaries, in the first screen, seen by each donor (Table A.16, Columns 1-3) where "regret" is less likely to have been activated. The intuition here, is that, in the first screen, donors across all treatment arms have yet to hit refresh and hence are unlikely to have experienced any form of "moral" risk or "regret" in this initial draw of beneficiaries. Furthermore, we restrict our analysis to the first three beneficiaries in an attempt to disentangle our estimates from that of display order effects. Reassuringly, results remain largely similar to our full baseline sample, suggesting that "regret" is unlikely to be a main driver of our observed results.

Second, we turn to beneficiary-donor dyad-level regressions using a specification that includes *beneficiary* fixed effects and fixed effects for both the *screen* in which a beneficiary appears in, and a beneficiary's *cumulative, sequential position across screens*. This specification approximates a comparison of donation outcomes for the same beneficiary when they appear in the same ordered screen (which holds constant the number of refreshes made by the donor) and same sequential position across screens, but in different screen set sizes. Again, consistent with our baseline results, Columns 4 -5 of Table A.16 shows that the same beneficiary is both more likely to receive a donation and receive a marginally higher donation amount, when they are displayed to donors in 8-option sets.

6.3. Results: Within–Donor Analysis

We have documented the effect of screen set sizes on the donation outcomes of donors, demonstrating a key trade-off between variety/match quality and cognitive overload. In this section, we analyze dimensions of quality and information that beneficiary narratives might have provided to donors. Section 6.3.1 uses hand-coded characteristics and language models (QWEN3 and IndoBERT) to explore the demographics and stories that beneficiaries wrote to appeal to donors, to further pin down the characteristics of appeals which were more successful in eliciting donations. Section 6.3.2 uses Keyness Statistics (KS) and Latent Semantic Scoring (LSS) to provide a systematic analysis of a single, key dimension of deservingness. Importantly, emotional scoring reveals *how* beneficiaries wrote their appeals (emotional tone). In contrast, KS and LSS identifies *what* beneficiaries wrote (i.e. the lexical content associated with successfully obtaining donations). Both provide complementary yet potentially distinct explanations for understanding donor behavior. Last, we look at possible biases resulting from in-group bias (Section 6.3.3) and within-screen, display order (Section 6.3.4).

6.3.1. Demographics and Narratives

What kinds of beneficiaries did donors make donations to? Donors did not give to all potential beneficiaries to whom they were exposed to; neither did all beneficiaries in the platform's database receive a donation. Even among beneficiaries who received donations, the amount of each donation and the number of donors that made a donation varied. A key piece of information for donors to consider were the rich, textual narratives that beneficiaries wrote to appeal for donations. We begin by providing a descriptive look of the top ten narratives that were the most successful in eliciting donations. These descriptive comparisons motivate and complement our later regressions with hand-coded indicators of deservingness; NLP analyses (emotion ratings QWEN3 and topic groupings IndoBERT); and our KS/LSS-based deservingness index. Table A.1 and A.2 presents the top ten appeals by donation amount and frequency of donations.

Table A.1 lists the top ten appeals of beneficiaries that received the highest donation amounts. Half of these beneficiaries explicitly mention being the primary breadwinner in their family and having child dependents. Six beneficiaries mention COVID or the mobility restriction instituted to contain it, and the rest mention a closed workplace. In terms of occupational sector, beneficiaries appear to have a wide range of occupations, ranging from catering workers, to teachers, and even entrepreneurs, graphic designers, and freelancers. Similarly, in terms of geographical location, beneficiaries are located throughout all provinces of Java (Indonesia's most populous island), as well as Bali and Sumatra (Outer Islands). Last, the frequency of beneficiaries with Muslim and feminine names roughly tracks the overall distribution. Linguistically, however, none of these appeals are free from grammatical errors or non-standard spelling in the Indonesian language. This is typical of texts in online corpora outside of journalistic outlets and encyclopedias and motivates our choice of NLP tools.

Surprisingly, Table A.2 shows little overlap between the top ten appeals by donation amount and donation frequency. Specifically, this list shares a mere two beneficiaries with the former, pointing to the potential limitations of hand-coded analyses. Reassuringly, however, in terms of specific characteristics, the share of beneficiaries who write they are the primary breadwinner in their family (with young children) and who directly reference COVID-imposed restrictions in their appeals, is similarly high across both lists (eight and seven appeals, respectively). Similarly, we see a wide variety of occupational sectors and geographical locations. Interestingly, a key pattern that emerges, is that four of these beneficiaries appear to have been fairly successful in telling a coherent narrative of persistently heavy financial burdens despite a gradually improving work situation and their best efforts (e.g. Donor #2,#3,#4, #10). This might have successfully conveyed an overall narrative of *hopelessness* that appealed more strongly to donors.

Hand-Coded Characteristics We estimate Equation (10), regressing our donation measures on hand-coded characteristics of each narrative. Figures 3 and 4 display selected coefficient estimates and provide suggestive evidence that breadwinners in particular, are more likely to receive a donation. Next, we consider occupational sectors. Education workers appear to be more deserving than workers in the hospitality industry (the omitted group). Beneficiaries in the education sector are 1.3 pp more likely to attract donations and receive USD 0.18 more in donation value. In contrast, we find little evidence that individuals reporting economic shocks are more likely to receive donations. Turning to our full regression results (Table A.17), the coefficient estimate for being laid off (retrenched) is statistically insignificant. Nevertheless, with more than 2,000 beneficiary narratives; many of these hand-coded characteristics being potentially correlated; and that donors themselves might have overlooked certain characteristics in their survey responses. We now turn to richer, NLP-based methods of textual analysis.

[INSERT FIGURE 3 AND 4 HERE]

Sentiment Analysis and Topic Grouping. We conduct an exploratory NLP analysis of affective tone and themes in appeal narratives. Sentiment analysis using natural language processing (NLP) and other computational methods has emerged as an important tool for researchers working with large-scale unstructured text wanting to detect positive or negative sentiments in a given piece of text. Nevertheless, in our context, we are mindful of two limitations that might potentially limit the efficacy of these tools. First, while classic NLP resources to process English text corpora are readily available, equivalent resources for Indonesian texts are much more sparse (Aji et al., 2022). Hence, simply translating texts from Indonesian to English may flatten much of the variation in sentiment and tone stemming from, e.g. the informality of the original text. Second, basic sentiment analysis, (which maps a text to positive/negative sentiment) is likely to produce little variation from our text corpus, given that beneficiary narratives overwhemingly cited layoffs, income loss, and general uncertainties from the COVID pandemic.

To that end, we run two forms of textual analysis. First, we use QWEN3, a Llama-based LLM to rate the emotion in each beneficiary's text appeals (Yang et al., 2025). QWEN rates the appeal in one of six difficult emotions such as anxiety or frustration and we regress our donation outcomes on these ratings. Second, we use IndoBERT, a self-supervised neural network that encodes Indonesian texts into topics (Koto et al., 2020; Aji et al., 2022). IndoBERT's encoding uncovers themes from sentences in narrative texts and hierarchically groups them into 9 major themes. We again regress indicators of each theme on donation outcomes.

Figure A.3 displays results. The left panel displays results from regressing donation indicator and amount on emotional scores. These results suggest that hopelessness (anxiety) is positively (somewhat negatively) associated with donation amounts. The right panel displays results from regressing the same outcome variables on 9 major themes. Household financial burdens appear to have the strongest positive association with both the probability of receiving a donation and the donation amount. We interpret this theme as potentially being related to that of being a breadwinner (with dependent children), which, as we saw, was overrepresented in both the list of appeals with the highest donation amounts and counts. Taken together, these results naturally prompt the question of whether it might be possible to combine our earlier analysis of hand-coded characteristics of beneficiaries with a more systematic NLP approach. We turn to this in the next subsection.

6.3.2. Deservingness and Match Quality

As the appeals and characteristics of potential beneficiaries exhibit a high degree of diversity, donors need to systematically evaluate each appeal to determine whether they would derive utility from donating. Thus, some of the potential beneficiaries could be deemed more deserving of a donation than others through an evaluation of their appeals. From the standpoint of regression analysis, however, it is unclear what the optimal way to (dis-)aggregate individual characteristics from each beneficiary appeal might be. To that end, we turn to a computational approach to systematically score and group beneficiaries into how deserving they are of charity.

We employ textual analysis to construct a *deservingness index* for each narrative using *Keyness Statistics* (KS) and *Latent Semantic Scaling* (LSS) (Zollinger, 2022). The KS method approximates asking donors for the motivations behind their decision to donate to a beneficiary by computing a χ^2 statistic for each term that appears in a narrative and ranks the most frequent terms among beneficiaries who received at least one donation versus those who did not receive any donations. The LSS statistic provides a composite score index for each beneficiary narrative (i.e., a *deservingness index*). The index takes the value of 0 for narratives scored as having the highest similarity to narratives that were the *least* likely to receive donations. The index takes the value 1 for narratives scored as having the highest similarity to narratives that were the *most* likely to receive donations.

There are two potential advantages to using KS and LSS over classical NLP methods. First, classical approaches like basic sentiment analysis—which maps a text to positive versus negative sentiments—might yield little variation from our text corpus of beneficiary narratives which comprise mainly of COVID-19 driven layoffs, losses, and uncertainties (i.e. overwhemingly *negative* sentiments) Second, we were concerned that traditional NLP methods might perform less reliably on Indonesian-language text given a much thinner resource base and the prevalence of code-mixing and colloquialisms in our narratives (Aji et al., 2022). In contrast, KS only requires us to provide a reference corpus to map a target corpus into output measures, a process that we judged to be largely language-agnostic. Similarly, LSS combines word-based supervision and an unsupervised model for synonymous words to generate a measure that might, again, be less overtly influenced by unrelated variation in other linguistic features of our corpus.

Figure 5 displays our *keyness* statistics. The black bars depicted in the upper part of the figure show the terms mentioned with the greatest relative frequency, and results align closely with our hand-coded results. Keywords positively associated with donations are those related to beneficiaries with child dependents or affiliations with the education sector. We used this result to identify the appropriate seed words for LSS processing to provide a composite score for each narrative.

[INSERT FIGURE 5 HERE]

Table 4 presents regression results of the probability of receiving a donation and the donation amount on our deservingness index (together with a parsimonious set of controls). Column 1 shows that a beneficiary narrative with a higher deservingness index is more likely to receive a donation, and this significance remains (Columns 2–3) when we add additional control variables. We find similar results for donation amounts in Columns 4–6.

[INSERT TABLE 4 HERE]

Taken together, we interpret that perceptions of deservingness are a core component of the match quality between donors and beneficiaries, as outlined in our theoretical framework. The analysis using hand-coded variables and textual approaches suggests that perceptions of deservingness matter, above and beyond observable demographic characteristics, and that a primary driver of deservingness perceptions is the presence of dependents. More broadly, our results provide novel empirical evidence for the accountability principle (Konow, 1996, 2000): donors are more likely to donate to beneficiaries whose neediness corresponds to factors they cannot reasonably influence or change in the short run through their own effort.

Result 5. *Beneficiaries with characteristics associated with the perception of being more deserving are more likely to receive donations and receive larger donation amounts.*

6.3.3. In-group Bias

An alternative mechanism driving differences in donor behavior is in-group biases. Donors may be more generous to individuals who share a similar identity. For example, individuals might have a higher level of trust and sympathy towards in-group members, leading to higher donations when the donor and recipient share an identity.[25]

We test for in-group biases by matching beneficiary demographics from the beneficiary database and donor demographics from our primary donor survey. We caveat that this analysis relies on a smaller sample, as only a subset of donors left their email addresses on both the platform and our donor survey.[26] We estimate Equation (13) and Table 5 presents the full results.

[25] Altruistic decision-making shares similarities with the decision process about a public benefit that will accrue to someone other than the donor. In this vein, researchers have argued that heterogeneous communities contribute less to social organizations and activities (Alesina and La Ferrara, 2000; Miguel and Gugerty, 2005; Okten and Osili, 2004). Individuals might be less willing to contribute to a public good if it benefits other groups, due to mistrust across groups or an inability to enforce within-group reciprocity (Alesina and La Ferrara, 2002; Habyarimana et al., 2007). This idea relates to the concept of social distance (Chen and Li, 2009; Meer and Rigbi, 2013; Sudhir et al., 2016). For this analysis, we investigate four dimensions of identity: gender, religion, ethnicity, and location.

[26] As noted, the donor survey was decoupled from the donation process to minimize the possibility that any perceived reduction in anonymity would discourage donations.

Across regressions, we find evidence of in-group bias for female gender identity and, to a lesser extent, ethnic identity. For the donation indicators, the coefficient estimates for the concordance indicators for female gender identity are statistically different from zero. We find similar results for the regression with donation amounts as the outcome variable. In comparison, the coefficients for all other concordance variables are statistically indistinguishable from zero except for ethnicity alignment, which is marginally significant at the 10% level in donation indicator regression.[27]

Result 6. *There is suggestive evidence that female donors are more likely to donate to feminine beneficiaries and send a larger donation. However, there is insufficient evidence of donor-beneficiary concordance across other dimensions of identity.*

[INSERT TABLE 5]

These results provide novel evidence complementary to that of existing studies on in-group biases in non-charitable giving contexts. We introduce a novel finding: the activation of in-group bias in altruistic settings is highly context-dependent and can be attenuated by the *nature* of the disaster. First, our weak results on ethnic identity concordance could potentially be explained by the fact that COVID-19 was a global public health disaster that, arguably, affected all individuals regardless of ethnic identity. Hence, ethnic biases might have been less central in donors' decision-making processes. Second, in contrast, our strong results on female identity concordance suggest an interactive effect between gender in-group bias and deservingness. Female donors were more likely to consider female beneficiaries to be more deserving of donations.

6.3.4. Display Order

Additional evidence for the effect of cognitive load on donation outcomes comes from the random draw and random assignment of beneficiaries to sets. We investigate the effect of the display order within the set on donation outcomes. Without cognitive overload, donors would consider the alternatives equally, independent of their position within the consideration set. On the other hand, donors in 8- and 10-screen sizes are likely to be susceptible to cognitive overload. Figure A.4 shows the raw donation rate by position in the set across sessions and across donors. We see that a higher share of beneficiaries at the beginning and toward the end of a set receive a donation than those in the middle. Formally, we test whether donors choose the beneficiary to donate to regardless of where they are placed in the set, following the specification in equation 14.

Table 6 displays the estimated coefficients. We hold set size constant with donor–session

[27] Given the overall imbalance in the religious (Muslim) identity and gender of beneficiaries in our sample, Table A.19 and A.20 presents robustness checks where we restrict our regression sample to screens where either (i) 50% of beneficiaries had female-coded names and 50% had male ones or (ii) 50% of beneficiaries had Muslim names and 50% had non-Muslim ones. Our sample size is greatly reduced, resulting in greater imprecision of our coefficient estimates but, reassuringly, results are similar to that of Table 5.

fixed effects ϕ_i and obtain the estimates for the effect of display order on donation outcomes of each beneficiary relative to all other beneficiaries within the same screen. Pooling our sample across all treatment arms, we observe that individuals placed higher are more likely to receive donations (Column 1). We also find that donation propensity varies nonlinearly with display order within a screen. In the pooled sample, beneficiaries shown at both the top and bottom of the screen are more likely to receive donations (Column 2). We further split our sample by treatment arm, where we restrict our regression sample to 3-, 8-, and 10-set in Columns 3, 4, and 5 respectively. Interestingly, this nonlinearity is most pronounced in the 8-option screen. Estimates in the 3-option screen are comparatively imprecise, and the 10-option screen appears largely driven by a "top-of-screen" advantage rather than a symmetric top-and-bottom pattern.

Result 7. *Beneficiaries displayed at the top and bottom of screens are more likely to receive donations than those in the middle.*

6.3.5. Additional Within–Donor Analysis

Optimal Number of Beneficiaries. Our three-arm experiment with 3-, 8-, or 10-beneficiaries in a screen has allowed us to analyze the effect of screen set size on charitable donations, but we can also conceptualize an ideal experiment with an incremental increase of 1 beneficiary per treatment arm. Such fine-grained variation, however, would require a much larger sample size to maintain statistical power. In our setting, our matched survey sample can be leveraged to obtain insights into possible tipping points between screen set size. In particular, we can exploit the variation in e-payment compatibility between the donor and beneficiary in conjunction with the experimental variation.

On the platform, beneficiaries can list different e-payment channels (*Gopay*, *Jenius*, *Dana*), but our survey indicates that donors may use only some of them. Beneficiaries with, e.g., only a *Dana* e-payment will thus be effectively out of the donor's consideration set if the donor doesn't also use this channel, and therefore their effective options could be fewer. Using this restricted sample, we examine the effect of the number of effective beneficiaries in a set on donors' donation outcomes and use its estimates to plot in Figure A.5 the predicted outcomes for different numbers of beneficiaries. We find inverted-U patterns in donation outcomes, with a maximum predicted outcome at the 6-7 option range, consistent with results from coarser treatment assignments.

The Influence of Existing Donations. We also test whether donors are influenced by earlier donations by other donors. A donor's desire to donate to a particular beneficiary may be lower if they already received some donations, as the platform's name sends a message to prioritize spreading their charity. On the other hand, donors overwhelmed by options could see existing donations as a social proof to rally behind (Cialdini, 2007). On the platform, donors see the share of the ask amount that a beneficiary has collected in donations (Figure 1). Table A.21 presents regressions from a sparse specification with just the share of donation ask, and we find that donors crowd in existing donations. They are more likely to donate to beneficiaries who have received donations and to donate a higher amount.

7. Conclusion

This paper documents that donors are susceptible to choice overload in the context of online charitable giving in a developing country. Beneficiaries displayed in an 8-beneficiary screen size receive the most donations relative to 3- and 10-screen sizes (pre-registered). In an exploratory analysis, we show that higher donations in 8-screens possibly arise from 8-screen sizes, both reducing choice overload and improving search behavior through (i) presenting a relatively smaller number of alternatives vis-a-vis 10-screens; and (ii) enabling greater donor-beneficiary preference alignment by presenting a sufficiently large number of alternatives vis-a-vis 3-screens.

Within-screens, we find strong evidence of deservingness and female gender in-group biases leading to higher donations. Furthermore, we document novel evidence of dipping behavior: beneficiaries placed in the middle of screens are the least likely to receive donations.

Our results provide novel, policy-relevant evidence of a low-cost way to possibly attenuate suboptimal heuristics in online charitable giving platforms: presenting just the right number of alternatives. This could reduce informational overload by allowing donors to pay more attention to each beneficiary choice and the attendant characteristics that platforms deem most correlated with the highest marginal value of donations.

In addition, our findings have implications for thinking about the optimal way to optimize charitable giving above and beyond those associated with crisis response. In particular, given the near-zero setup and transaction costs of online donation platforms, our findings suggest that policymakers could play a valuable role in overseeing and guiding these platforms to help minimize donor fatigue and reduce the potential de-personalization of donation experiences (Andreoni and Payne, 2013). Our findings also offer the tantalizing possibility that small adjustments to choice architecture could increase giving to reparations and redistributive causes (e.g., climate change).

References

Adena, M. and A. Hager (2024): "Does online fundraising increase charitable giving? A nationwide field experiment on Facebook," *Management Science*.

Adena, M., R. Hakimov, and S. Huck (2024): "Charitable giving by the poor: A field experiment in Kyrgyzstan," *Management Science*, 70, 633–646.

Aji, A. F., G. I. Winata, F. Koto, S. Cahyawijaya, A. Romadhony, R. Mahendra, K. Kurniawan, D. Moeljadi, R. E. Prasojo, T. Baldwin, et al. (2022): "One country, 700+ languages: NLP challenges for underrepresented languages and dialects in Indonesia," in *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 7226–7249.

Alesina, A. and E. La Ferrara (2000): "Participation in heterogeneous communities," *The quarterly journal of economics*, 115, 847–904.

——— (2002): "Who trusts others?" *Journal of public economics*, 85, 207–234.

Andreoni, J. (2007): "Giving gifts to groups: How altruism depends on the number of recipients," *Journal of Public Economics*, 91, 1731–1749.

Andreoni, J. and A. A. Payne (2013): "Charitable giving," in *Handbook of public economics*, Elsevier, vol. 5, 1–50.

Aria, P. (2021): "2,56 Juta orang menganggur akibat pandemi, 24 Juta pekerja potong gaji," *Katadata.co.id*.

Baznas (2019): *Zakat Outlook 2019*, 2019.

Bernasconi, M., L. Corazzini, S. Kube, and M. A. Maréchal (2009): "Two are better than one!: individuals' contributions to "unpacked" public goods," *Economics letters*, 104, 31–33.

Bhargava, H. K. (2022): "The creator economy: Managing ecosystem supply, revenue sharing, and platform design," *Management Science*, 68, 5233–5251.

Cappelen, A. W., A. D. Hole, E. Ø. Sørensen, and B. Tungodden (2007): "The pluralism of fairness ideals: An experimental approach," *American Economic Review*, 97, 818–827.

Charities Aid Foundation (2018): "World Giving," Tech. rep.

——— (2019): "CAF World Giving Index 10th Edition," Tech. rep.

Charness, G. and P. Holder (2019): "Charity in the laboratory: Matching, competition, and group identity," *Management Science*, 65, 1398–1407.

Chen, Y. and S. X. Li (2009): "Group identity and social preferences," *American Economic Review*, 99, 431–457.

Chernev, A., U. Böckenholt, and J. Goodman (2015): "Choice overload: A conceptual review and meta-analysis," *Journal of Consumer Psychology*, 25, 333–358.

Cialdini, R. B. (2007): *Influence : the psychology of persuasion.*, HarperCollins Publishers.

Clark, C. J., X. Han, and U. O. Osili (2019): "Changes to the Giving Landscape," Tech. rep., Indiana University Lilly Family School of Philantrophy.

Corazzini, L., C. Cotton, and P. Valbonesi (2015): "Donor coordination in project funding: Evidence from a threshold public goods experiment," *Journal of Public Economics*, 128, 16–29.

Deryugina, T. and B. M. Marx (2021): "Is the supply of charitable donations fixed? Evidence from deadly tornadoes," *American Economic Review: Insights*, 3, 383–398.

Eckel, C., B. Guney, and N. Uler (2020): "Independent vs. coordinated fundraising: Understanding the role of information," *European Economic Review*, 127, 103476.

Enke, B. and C. Shubatt (2024): "Quantifying lottery choice complexity," Tech. rep., National Bureau of Economic Research.

Filiz-Ozbay, E. and N. Uler (2019): "Demand for giving to multiple charities: An experimental study," *Journal of the European Economic Association*, 17, 725–753.

Gallier, C., T. Goeschl, M. Kesternich, J. Lohse, C. Reif, and D. Römer (2023): "Inter-charity competition under spatial differentiation: Sorting, crowding, and spillovers," *Journal of Economic Behavior & Organization*, 216, 457–468.

Habyarimana, J., M. Humphreys, D. N. Posner, and J. M. Weinstein (2007): "Why does ethnic diversity undermine public goods provision?" *American political science review*, 101, 709–725.

Hanna, R. and B. A. Olken (2018): "Universal basic incomes versus targeted transfers: Anti-poverty programs in developing countries," *Journal of Economic Perspectives*, 32, 201–226.

Huck, S. and I. Rasul (2011): "Matched fundraising: Evidence from a natural field experiment," *Journal of Public Economics*, 95, 351–362.

Iyengar, S. S. and E. Kamenica (2010): "Choice proliferation, simplicity seeking, and asset allocation," *Journal of Public Economics*, 94, 530–539.

Iyengar, S. S. and M. R. Lepper (2000): "When choice is demotivating: Can one desire too much of a good thing?" *Journal of personality and social psychology*, 79, 995.

J-PAL SEA (2020): "Online Survey on Digital Financial Service Use during COVID-19 in Indonesia," Tech. rep.

Jayaraman, R., M. Kaiser, and M. Teirlinck (2023): "Charitable donations to natural disasters: evidence from an online platform," *Oxford Economic Papers*, 75, 902–922.

Kessler, J. B. and K. L. Milkman (2018): "Identity in charitable giving," *Management Science*, 64, 845–859.

Kessler, J. B., K. L. Milkman, and C. Y. Zhang (2019): "Getting the rich and powerful to give," *Management Science*, 65, 4049–4062.

Konow, J. (1996): "A positive theory of economic fairness," *Journal of Economic Behavior & Organization*, 31, 13–35.

——— (2000): "Fair shares: Accountability and cognitive dissonance in allocation decisions," *American economic review*, 90, 1072–1092.

Koto, F., A. Rahimi, J. H. Lau, and T. Baldwin (2020): "IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP," in *Proceedings of the 28th International Conference on Computational Linguistics*, 757–770.

Meer, J. (2014): "Effects of the price of charitable giving: Evidence from an online crowdfunding platform," *Journal of Economic Behavior and Organization*, 103, 113–124.

Meer, J. and O. Rigbi (2013): "The effects of transactions costs and social distance: Evidence from a field experiment," *B.E. Journal of Economic Analysis and Policy*, 13, 271–296.

Miguel, E. and M. K. Gugerty (2005): "Ethnic diversity, social sanctions, and public goods in Kenya," *Journal of public Economics*, 89, 2325–2368.

Miller, G. A. (1956): "Information and memory," *Scientific American*, 195, 42–47.

NapoleonCat (2023): "Social media users in Indonesia at the end of 2022," https://napoleoncat.com/stats/social-media-users-in-indonesia/2022/, accessed: 23 July 2023.

Noor, Z. and F. Pickup (2017): "The Role of Zakat in Supporting the Sustainable Development Goals," *BAZNAS-UNDP Brief*.

Okten, C. and U. O. Osili (2004): "Social networks and credit access in Indonesia," *World Development*, 32, 1225–1246.

Paxton, P. (2020): "What Influences Charitable Giving?" in *The Nonprofit Sector: A Research Handbook*, ed. by W. W. Powell and P. Bromley, 543–557.

Reutskaja, E., S. Iyengar, B. Fasolo, and R. Misuraca (2011): "Cognitive and affective consequences of information and choice overload," *Journal of Consumer Research*, 37, 425–442.

Rogerson, R., R. Shimer, and R. Wright (2005): "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, 43, 959–988.

Scharf, K., S. Smith, and M. Ottoni-Wilhelm (2022): "Lift and shift: the effect of fundraising interventions in charity space and time," *American Economic Journal: Economic Policy*, 14, 296–321.

Schmitz, J. (2021): "Is charitable giving a zero-sum game? The effect of competition between charities on giving behavior," *Management Science*, 67, 6333–6349.

Schmitz, J. and M. Grieder (2024): "Coordinated fundraising: Evidence from a randomized field experiment," *Available at SSRN 4849150*.

Schwirplies, C. (2023): "Does additional demand for charitable aid increase giving? Evidence from Hurricane Sandy," *Journal of Economic Behavior & Organization*, 209, 53–73.

Soyer, E. and R. M. Hogarth (2011): "The size and distribution of donations: Effects of number of recipients," *Judgment and Decision making*, 6, 616–628.

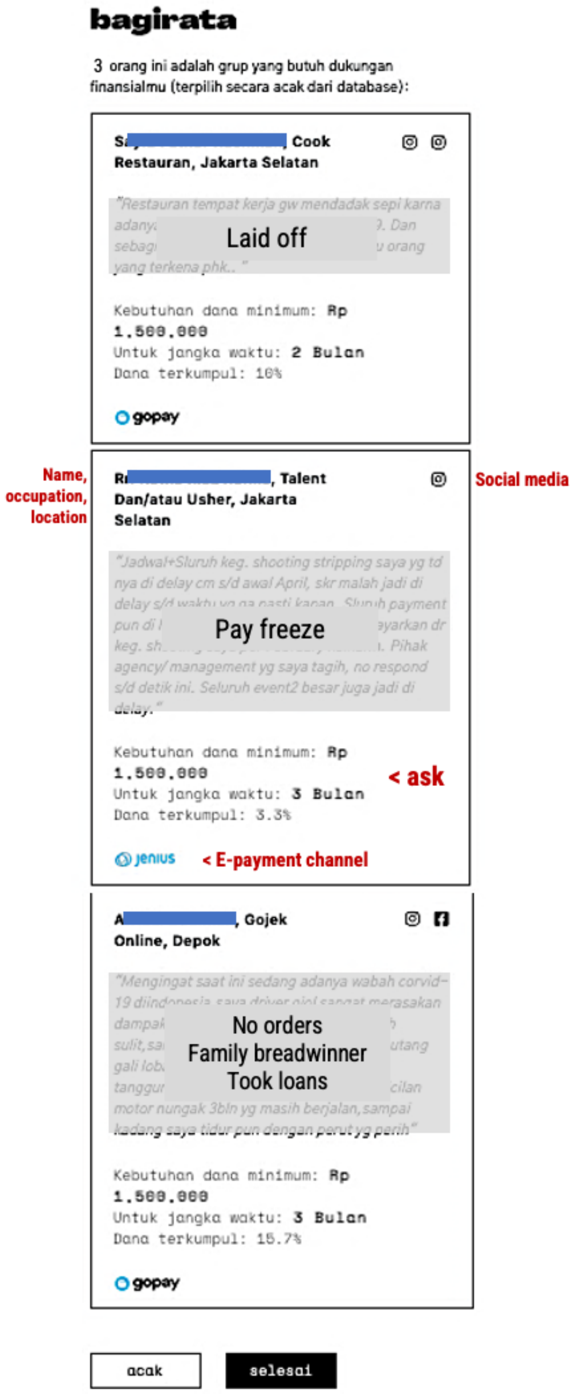
Sudhir, K., S. Roy, and M. Cherian (2016): "Do sympathy biases induce charitable giving? The effects of advertising content," *Marketing Science*, 35, 849–869.

Suri, T., J. Aker, C. Batista, M. Callen, T. Ghani, W. Jack, L. Klapper, E. Riley, S. Schaner, and S. Sukhtankar (2023): "Mobile Money," *VoxDevLit*, 2, –.

Yang, A., A. Li, B. Yang, B. Zhang, B. Hui, B. Zheng, B. Yu, C. Gao, C. Huang, C. Lv, et al. (2025): "Qwen3 technical report," *arXiv preprint arXiv:2505.09388*.

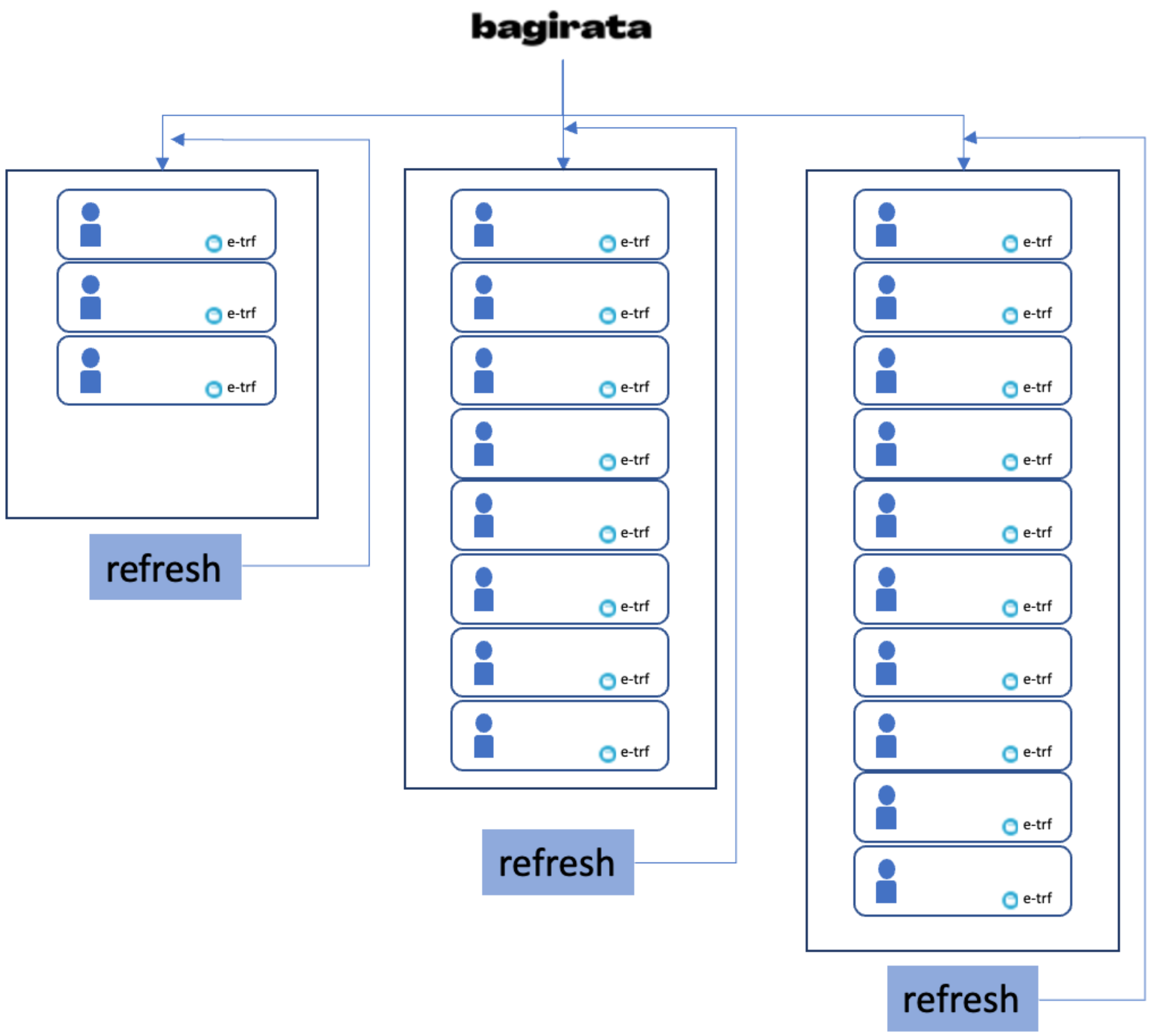
Zollinger, D. (2022): "Cleavage Identities in Voters' Own Words: Harnessing Open-Ended Survey Responses," *American Journal of Political Science*.

Figure 1: Set of Beneficiary Cards Presented to Donors on the Platform



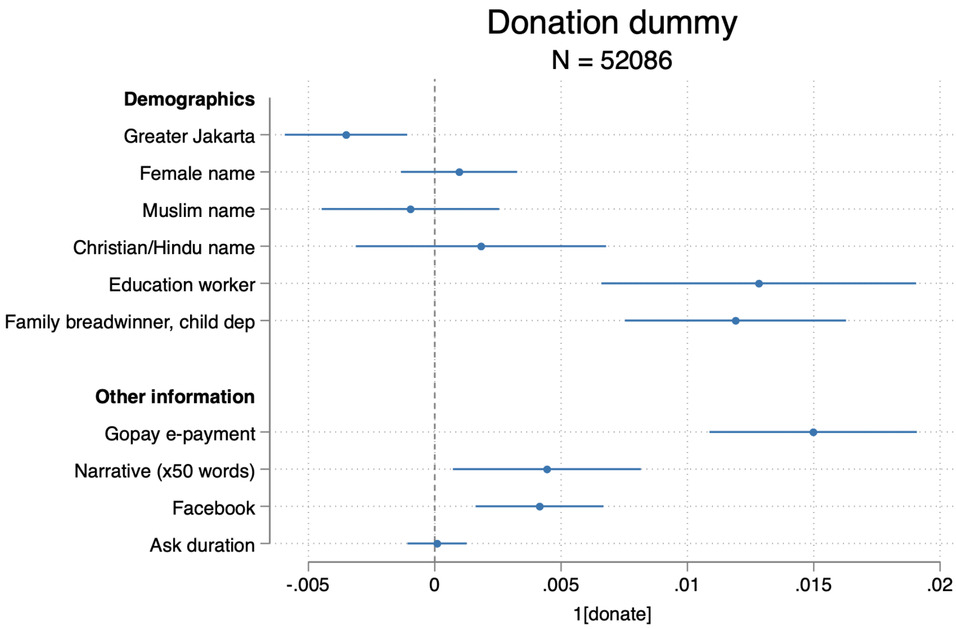
Note: An example of a set of beneficiary cards that potential donors encounter on the *Bagirata* platform. In this example, the donor was randomly assigned to view sets of three beneficiaries at a time. The randomization of the screen set size and the random selection of beneficiaries from the database to be displayed took place after the visitor clicked the button on the landing page expressing her wish to donate. Donors are informed that beneficiaries are randomly selected (as indicated by the top text below the *Bagirata* logo). Each beneficiary card includes the beneficiary's name, occupation, and location (top left), social media links (top right), a free text narrative appeal from the beneficiary (center), nominal ask, duration of ask, overall donation progress, and a link to e-payment channels (bottom). In this example, key aspects of the appeal in English have been superimposed onto the original Indonesian text in the center. For detailed sample appeals and their English translations, see Tables A.1 and A.2. Cards are arranged in a vertical sequence on the website, requiring users to scroll to subsequent cards in the set. Donors have the option to click the "*acak*" button to generate a fresh random selection of beneficiaries or to directly donate through the e-payment link provided.

Figure 2: Schematic of Randomization Procedures for Platform Visitors



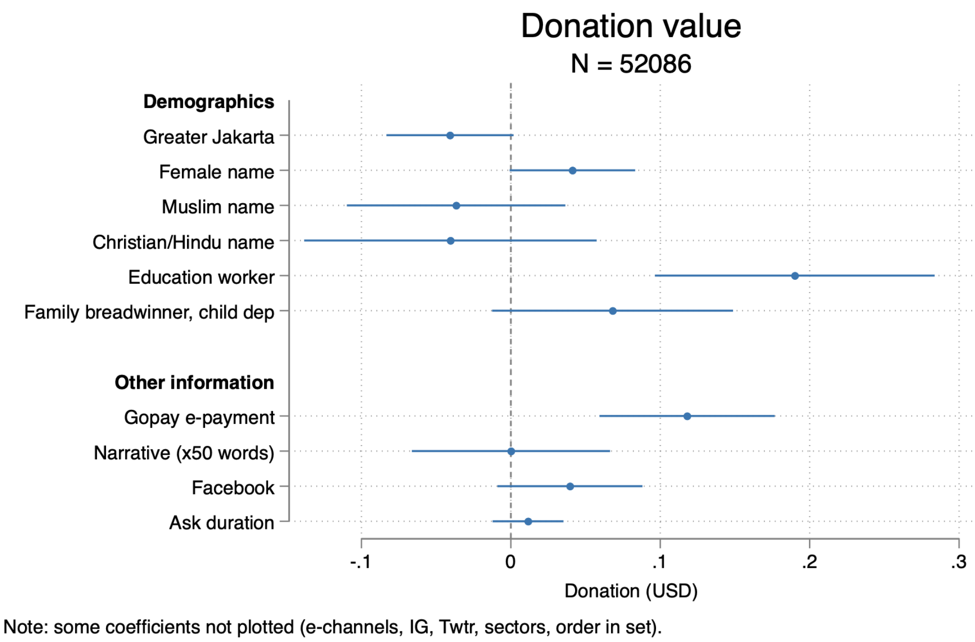
Note: Schematic of randomization procedures for platform visitors. Visitors are randomly assigned with equal probability to one of our three treatment groups, which present sets of 3, 8, or 10 beneficiaries. This randomization scheme is maintained throughout the duration of a web session, which typically lasts three hours. Within a web session, every time a donor refreshes the webpage or clicks the "*acak*" button (see Figure 1), she would encounter a new display set of the same number beneficiaries within her assigned treatment group.

Figure 3: Effects of Beneficiary Characteristics on Donation Indicator



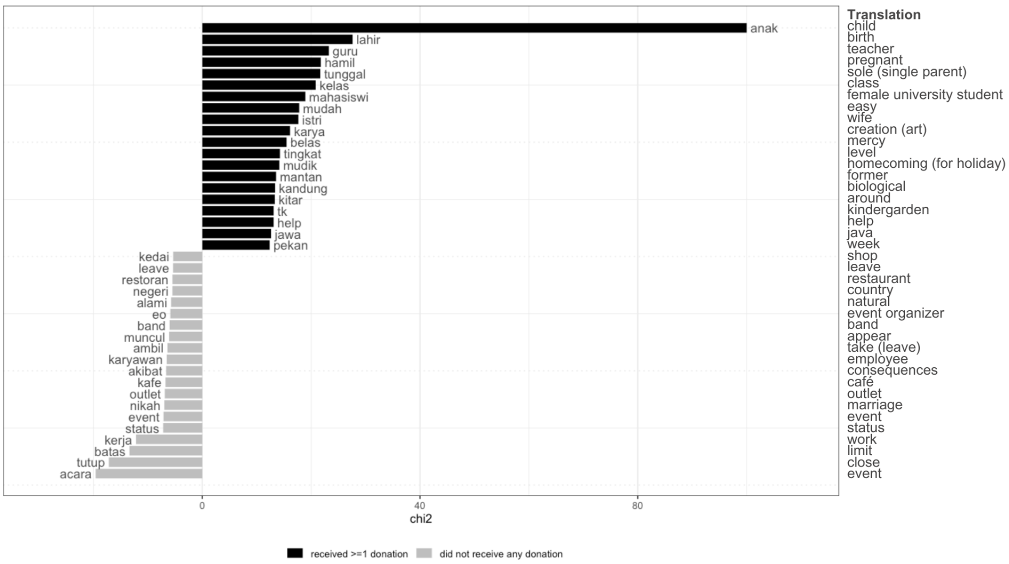
Note: Chart plots coefficients from $Y_{ijkl} = \alpha_2 + \beta_2 Characteristics_j + DonorFE_i + \varepsilon_{1,ijkl}$. Range for each coefficient indicates the 90% confidence interval. Full regression results can be found in Table A.17

Figure 4: Effects of Beneficiary Characteristics on Donation Values



Note: Chart plots coefficients from $Y_{ijkl} = \alpha_2 + \beta_2 Characteristics_j + DonorFE_i + \varepsilon_{1,ijkl}$. Range for each coefficient indicates the 90% confidence interval. Full regression results can be found in Table A.17

Figure 5: Keyness Statistics on Donor Behavior: Characteristics of Beneficiaries Who Received a Donation versus Those Who Did Not



Note: Black (gray) bars show terms mentioned with greatest relative frequency in beneficiary narratives that received at least one donation relative to those that did not receive any donations (and vice versa).

Table 1: Summary Statistics of Platform Users

A. Donations and Appeals on Platform	(1) Mean	(2) SD	(3) Max	(4) Count
Received any donations	0.81	0.40	1	2054
Number of donations	2.09	2.14	27	2054
Total received donations (USD)	17.84	26.01	646	2054
Frequency being displayed to donors	26.21	18.52	68	2054
Narrative length (words)	30.13	14.87	70	2054
Appeal duration (month)	2.19	0.87	3	2054
Appeal (USD, winsorized)	139.09	91.06	643	2054

B. Characteristics	Platform Database		Platform User Survey (Averages/Shares)	
	Count	% of Benef.	Recipients	Donors
	(5)	(6)	(7)	(8)
Gender				
Masculine name	1302	63%		
Male			57%	30%
Household status				
Breadwinner/mentions dependent(s)	462	22%		
Mentions child(ren) as dependents	254	12%		
Married			43%	34%
Household size			3.7	3.2
Religion				
Muslim name	1678	82%		
Islam			87%	68%
Region/Ethnicity				
Jakarta metro area	1385	67%		
Java, non-Jakarta Metro	491	24%		
Outside Java	178	9%		
Javanese			48%	56%
Employment sector				
Hospitality, retail, food service	1243	61%	35%	7%
Government, education, or health	111	5%	3%	17%
Art and creatives	326	16%		
Transportation	131	6%		
Finance or IT			3%	21%
Other	243	12%	38%	36%
Other characteristics				
Age			30	29
Years of education			13	15
Earning (USD)			$1,882	$8,626
Earning for charity			5%	6%
Mobile money platform in use			1.4	2.3
Employer is corporation/international			17%	49%
Employer is small			38%	13%
Amount received from platform (USD)			$25.68	$0
Amount donated via platform (USD)			$0	$26.33
Obs	2054		60	216

Note: Columns 1-6 display statistics from the platform database. Columns 7 and 8 display statistics from responses to a user survey fielded between October 2020 to July 2021. Survey is voluntary and decoupled from the donation process. Some characteristics are not exactly identical as they were either generated from imputation (gender from masculine name) or from direct survey questions. The appeals statistics in Panel A were generated over the period of the experiment. The donation statistics were generated over the lifetime of the platform, from April 2020-June 2021, i.e., including a six-month period prior to the start of the experiment in 4 October 2020. See text for details.

Table 2: Impact of Set Size on Donation Outcomes

	(1) Number of Donation	(2) Number of Donation	(3) Donations (USD)	(4) Donations (USD)	(5) 1(Donation)	(6) 1(Donation)
3-opt sets	0.042	0.163*	-0.658	0.420	-0.005	0.018
	(0.070)	(0.091)	(1.043)	(1.066)	(0.019)	(0.023)
8-opt sets	0.185**	0.256**	1.880	3.704**	0.017	0.023
	(0.092)	(0.124)	(1.500)	(1.888)	(0.019)	(0.023)
Covid positivity rate	-0.108***	-0.116***	-1.585***	-1.421**	-0.020**	-0.019**
	(0.036)	(0.041)	(0.589)	(0.686)	(0.008)	(0.009)
Covid deaths	-0.001	-0.002*	-0.010	0.005	0.000	-0.000
	(0.001)	(0.001)	(0.011)	(0.019)	(0.000)	(0.000)
Constant	1.565***	1.860***	23.397***	20.424**	0.321***	0.338***
	(0.443)	(0.593)	(6.882)	(8.090)	(0.098)	(0.112)
Sample	All	Normal Traffic	All	Normal Traffic	All	Normal Traffic
Observations	2405	1544	2405	1544	2405	1544

Notes: Regression of donation outcomes on screen set size, adjusting for two COVID variables: positivity rate and death counts. Regressions in Columns (2), (4), and (6) are restricted to observations that did not occur during the mobility lockdown period or social media campaigns. Covid positivity rate is the weekly mean (last 7 days) of share tested people who are positive, as recorded in the Jakarta government dashboard. Covid deaths are the number of fatalities on the day, as recorded on Wikipedia. The mobility lockdown period was defined following the first PPKM policy enacted in Java and Bali between 10-30 January, 2021. The social media campaigns was launched for May day 2021, between 23 April and 13 May. Observation unit is a donor-session. Robust standard errors are displayed in parentheses. Regressions include time FEs that consist of indicators for months, fortnights, and day of week. Sample is from 4 October 2020 to 9 June 2021. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Impact of Set Size on Donor Behavior from Back-End Data (Mechanisms)

	(1) Refresh Button Action	(2) Refresh Button Action	(3) Total Beneficiary Exposure	(4) Total Beneficiary Exposure	(5) Minutes Spent per Beneficiary	(6) Imputed Minutes per Beneficiary	(7) Imputed Minutes per Beneficiary
3-opt sets	1.852***	1.596**	-13.082***	-15.792***	0.920**	0.460***	0.536***
	(0.509)	(0.736)	(2.696)	(4.018)	(0.394)	(0.066)	(0.101)
8-opt sets	0.513	0.306	-0.985	-2.989	0.137	0.175***	0.187***
	(0.427)	(0.630)	(3.636)	(5.384)	(0.164)	(0.031)	(0.046)
Covid positivity rate	-0.042	-0.027	-0.438	-0.574	-0.235**	-0.049***	-0.041**
	(0.212)	(0.278)	(1.525)	(1.976)	(0.096)	(0.018)	(0.020)
Covid deaths	0.004	0.014	0.026	0.105	0.000	-0.000	-0.000
	(0.005)	(0.014)	(0.037)	(0.114)	(0.001)	(0.000)	(0.001)
Constant	4.780	3.941	35.507**	31.318	3.329***	0.981***	0.933***
	(2.913)	(3.239)	(17.802)	(19.992)	(1.198)	(0.224)	(0.282)
Sample	All	Normal Traffic	All	Normal Traffic	Donors	All	Normal Traffic
Observations	2405	1544	2405	1544	426	2405	1544

Notes: Regression of detailed measures of donor behavior on screen set size, adjusting for two COVID variables: positivity rate and death counts. Regressions in Columns (2), (4), and (7) are restricted to observations that did not occur during the mobility lockdown period or social media campaigns. Covid positivity rate is the weekly mean (last 7 days) of share tested people who are positive, as recorded in the Jakarta government dashboard. Covid deaths are the number of fatalities on the day, as recorded on Wikipedia. The mobility lockdown period was defined following the first PPKM policy enacted in Java and Bali between 10-30 January, 2021. The social media campaigns was launched for May day 2021, between 23 April and 13 May. Observation unit is a donor-session. Robust standard errors are displayed in parentheses. Regressions include time FEs that consist of indicators for months, fortnights, and day of week. Sample is from 4 October 2020 to 9 June 2021. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: *Deservingness* (Latent Semantic Scale) and Donation Outcomes

	(1) 1(Donate)	(2) 1(Donate)	(3) 1(Donate)	(4) Donation (USD)	(5) Donation (USD)	(6) Donation (USD)
Deservingness (index)	0.0822***	0.0687***	0.0685***	0.8871***	0.5076*	0.5022*
	(0.0168)	(0.0162)	(0.0163)	(0.2719)	(0.2627)	(0.2646)
% Ask fulfilled		0.0009***	0.0009***		0.0241***	0.0242***
		(0.0001)	(0.0001)		(0.0046)	(0.0046)
Set counter		-0.0000	-0.0000		-0.0001	-0.0001
		(0.0000)	(0.0000)		(0.0002)	(0.0002)
Ask amount (USD)		0.0000**	0.0000**		0.0011***	0.0011***
		(0.0000)	(0.0000)		(0.0003)	(0.0003)
Ask duration		0.0002	0.0001		0.0158	0.0155
		(0.0008)	(0.0008)		(0.0151)	(0.0151)
Greater Jakarta			-0.0015			-0.0357
			(0.0015)			(0.0250)
Order in set			-0.0006**			-0.0038
			(0.0002)			(0.0037)
Constant	-0.0183**	-0.0236***	-0.0201**	-0.1928	-0.3963**	-0.3572**
	(0.0084)	(0.0087)	(0.0088)	(0.1358)	(0.1567)	(0.1550)
FE	donor	donor	donor	donor	donor	donor
R2	0.244	0.253	0.253	0.193	0.213	0.213
Observations	52072	52072	52072	52072	52072	52072
Deservingness SD	0.177	0.177	0.177	0.177	0.177	0.177

Notes: Regression of donation outcomes on beneficiary characteristics with donor session fixed effects. Observation unit is donor–beneficiary dyad. Standard errors are clustered at donor and beneficiary levels and displayed in parentheses. Sample is from Oct 2020 to Jun 2021, excluding outliers. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: In-Group Bias: Regression of Donation Indicator and Value on Alignment of Donor–Beneficiary Characteristics

	(1)	(2)	(3)	(4)	(5)
	A. Outcome: 1(Donate)				
Female donor-feminine name beneficiary	0.0590***				0.0588***
	(0.0147)				(0.0148)
Muslim donor-muslim name beneficiary		-0.0047			-0.0001
		(0.0287)			(0.0284)
Ethnicity alignment donor-beneficiary			0.0151		0.0506*
			(0.0243)		(0.0292)
Donor-beneficiary in same district				-0.0009	0.0067
				(0.0175)	(0.0175)
Dep. Var. Mean	0.087	0.087	0.087	0.087	0.087
R2	0.138	0.132	0.132	0.132	0.140
Observations	2396	2396	2396	2396	2396
	B. Outcome: Donation (USD)				
Female donor-feminine name beneficiary	0.8680**				0.8846**
	(0.3729)				(0.3838)
Muslim donor-muslim name beneficiary		0.0139			0.0481
		(0.4266)			(0.4426)
Ethnicity alignment donor-beneficiary			0.4988		0.5253
			(0.3865)		(0.4018)
Donor-beneficiary in same district				0.2510	0.3419
				(0.3506)	(0.3688)
Dep. Var. Mean	0.957	0.957	0.957	0.957	0.957
R2	0.117	0.113	0.113	0.113	0.117
Observations	2396	2396	2396	2396	2396

Notes: Regression of the donation indicator and value on indicators for alignment between donor and beneficiary characteristics. The observation unit is a donor–beneficiary dyad. Standard errors are clustered at the donor email, session and beneficiary levels and displayed in parentheses. The sample is matched dyads between platform user survey and activity trace, with singletons omitted. The sample is comprised of 40 donor-emails in 78 sessions, presented with 1283 unique beneficiaries from the database. This is the only sample for which we can separately identify donors from sessions based on the email addresses that they entered in both the *Bagirata* database and the user–donor survey. All regressions include set counters and beneficiary order within set. All regressions include session FE (absorbing set size assignment), and donor email FE (absorbing donor-email-invariant indicators from survey indicating gender, religious affiliation, and ethnicity). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Impact of Display Order within Set on Donation

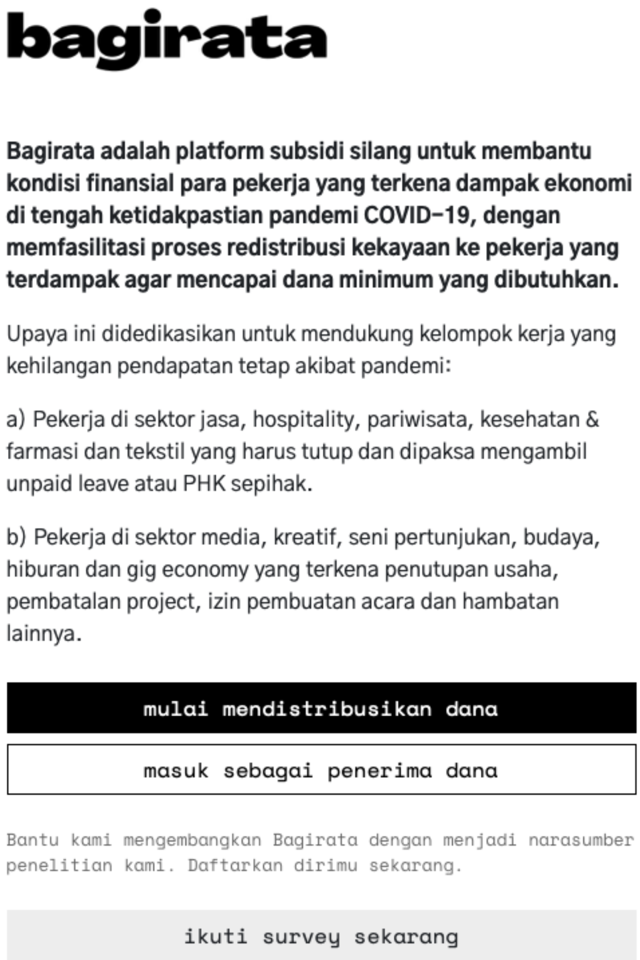
	(1)	(2)	(3)	(4)	(5)
	1(Donation)	1(Donation)	1(Donation)	1(Donation)	1(Donation)
Display order in set	-0.0006**				
	(0.0002)				
Top (4) in set		0.0070***	0.0065	0.0088***	0.0050*
		(0.0019)	(0.0043)	(0.0030)	(0.0026)
Bottom (3 or 5) in set (8 or 10)		0.0048***	0.0051	0.0063**	0.0028
		(0.0017)	(0.0038)	(0.0027)	(0.0024)
Constant	0.0253***	0.0178***	0.0299***	0.0166***	0.0130***
	(0.0011)	(0.0014)	(0.0023)	(0.0024)	(0.0021)
Sample	All	All	3-opt	8-opt	10-opt
FE	donor	donor	donor	donor	donor
Observations	52086	52086	10620	20776	20690

Notes: Regression of donation indicator on a continuous variable representing the position of the beneficiary's display position within a set, across all treatment groups (Column (1)), and two dummy variables representing the top and bottom (groups) in the set for each treatment group (all groups, then separately for set of 3, 8, and 10 in Columns (2)–(5)). Observation is a donor–beneficiary dyad. Standard errors are clustered at the donor and beneficiary levels and displayed in parentheses. Sample is from 4 Oct 2020 to 9 Jun 2021. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A. Appendix Tables and Figures

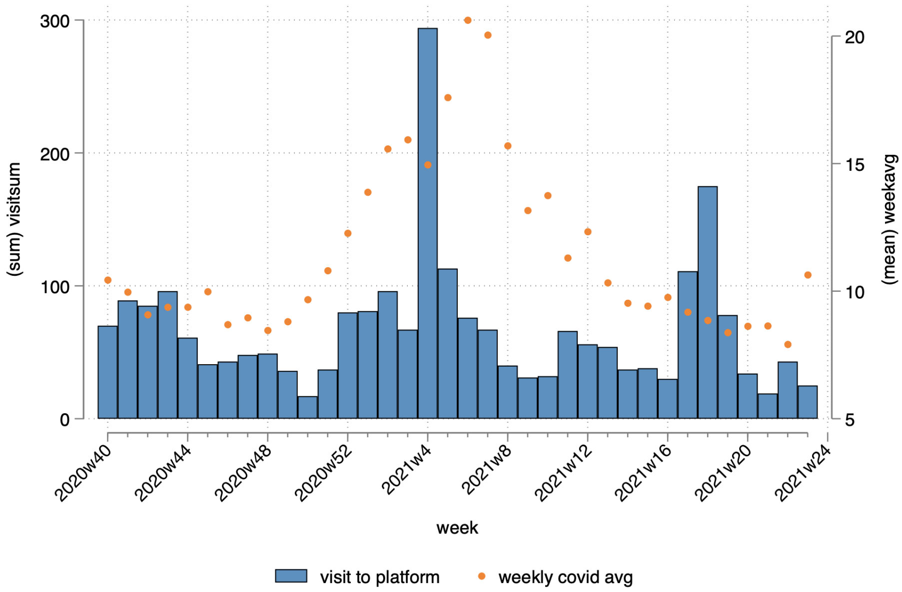
A.1. Appendix figures.

Figure A.1: Landing Page of the Platform



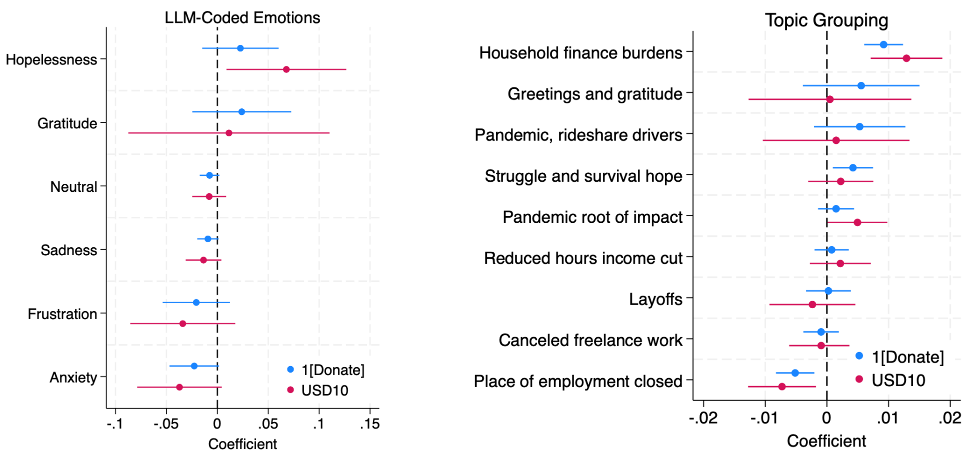
Note: This is the first page that potential donors see upon entering the *Bagirata* website.

Figure A.2: Unique Sessions on Platform over Time



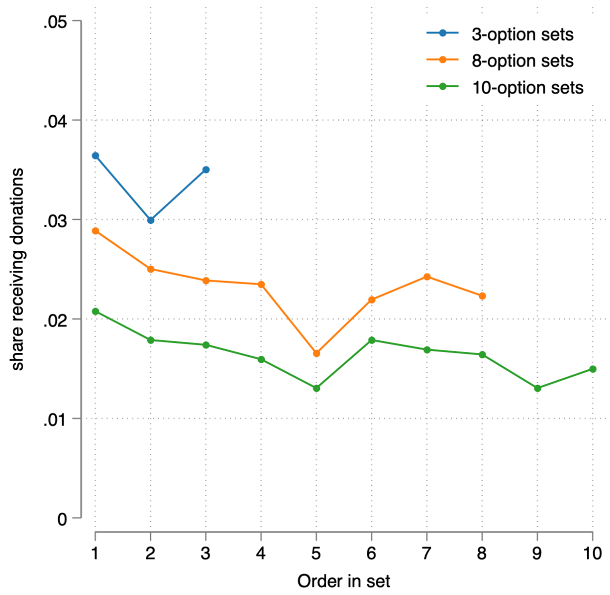
Note: The bar chart displays traffic to the platform over time, overlaid with Covid positivity rate in orange dots. The two spikes on the bar chart correspond to the mobility restriction (*Pemberlakuan Pembatasan Kegiatan Masyarakat*/PPKM) implemented in January 2021 (2021w4) and a Labor Day/May Day donation drive campaign (2021w17-19). Randomization remained ongoing during these two events.

Figure A.3: Topic Grouping and LLM-Coded Emotions



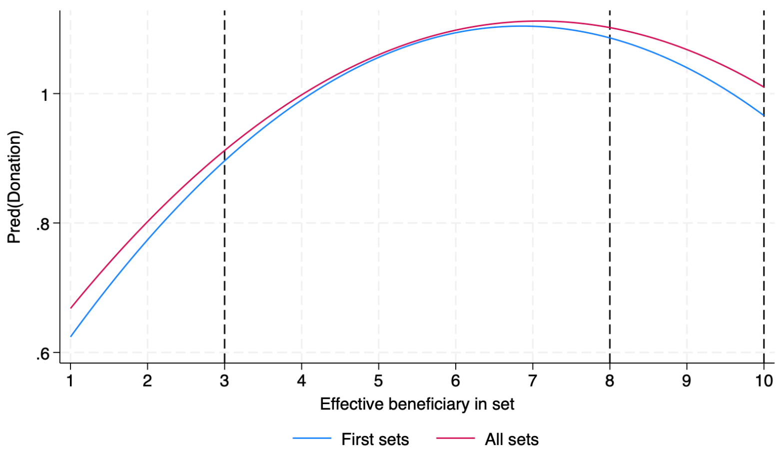
Note: Each dot represent an estimated coefficient for each topic/emotion. Blue dots are from a regression of the donation indicator on treatment indicators and indicators of emotion loading (left) or topic grouping (right). Emotion loading has a [0,1] scale of each emotion, using QWEN3 LLM to process narrative text. Red dots are from analogous regression with donations in 10 USD increments as the outcome variable, scale chosen for comparability with the donation indicator coefficients. Whiskers indicate 95% confidence interval. Standard errors are clustered at session level.

Figure A.4: Raw Donation Rate for Beneficiaries, by Position in Set



Note: Order in set is position within set, across screens and sessions. Vertical axis refer to the share of all beneficiaries who received donations when placed in the respective position.

Figure A.5: Inverted-U Pattern in Predicted Donation Outcomes using Effective Beneficiaries by Payment Channels



Note: Plot of predicted donation outcomes using parameters from regressions of effective beneficiaries, its square, and a constant in a matched survey-dyad sample on an indicator whether the donor makes any donation in the set. Blue line is parameterized with a restriction to the first set that the donor see in the matched sample, red line include sets after the first set. Regressions include indicators for 3- and 8-option sets, the experimentally induced variations.

A.2. Tables: Sample appeals with top donation amount and count

Table A.1: Appeals with Top Donation Amount

No.	Appeal (English translation/Indonesian text) and beneficiary characteristics
#1	I am a contract worker laid off due to low business since the Covid-19 pandemic. I have two children and my wife passed away two years ago. / "*Saya pekerja kontrak dan sudah dirumahkan karena sepinya orderan, semenjak adanya pandemi covid19, dengan anak 2. sedangkan istri meninggal dunia 2tahun yang lalu.*" A catering worker in East Java. Muslim name, not a feminine name, coded as a breadwinner. Asked US$71. Received a total of US$144 from five donations.
#2	The restaurant at the hotel where I worked is closed indefinitely. So I got furloughed indefinitely. I am a mother with two children. My children's need for milk, diapers, and other needs are quite a lot. / "*Restoran pada hotel di tempat saya bekerja tutup sampai waktu yang belum ditentukan. Jadi saya di rumahkan sampai batas waktu yang saya tidak tau sampai kapan.. Saya seorang ibu rumah tangga dengan 2 orang anak. Kebutuhan anak-anak saya untuk beli susu, popok bayi dan kebutuhan yang lainnya lumayan banyak.*" A service staff in Bali. Feminine name, not a Muslim name, coded as a breadwinner. Asked US$107. Received US$136 from four donations.
#3	Since Corona hit my income drops dramatically. I can't think of spending money, as I even need to borrow money to eat. / "*Semenjak adanya corona pemasukan saya berkurang dratis jangankan uang blanja, buat makan saja saya masih berhutang* " A freelancer in Greater Jakarta. Muslim name, not a feminine name, not a breadwinner. Asked US$107. Received US$126 from two donations.
#4	Because of the second mobility restriction I was made to take an unpaid leave. Until now I don't know when I can return to work. I have a family with a wife and two children (a 3-year-and-4-month old, and a 1-year-and-7-month old). / "*Karena PSBB ke dua ini saya terkena unpaid leave . Sampai saat ini saya belum dapat info kapan saya masuk kembali. Saat ini saya sudah berkeluarga dengan seorang istri, dua anak, 3th4bln & 1thn7bln.*" A bartender in Greater Jakarta. Muslim name, not a feminine name, coded as a breadwinner. Asked US$107. Received US$111 from two donations.
#5	I am a contract teacher, and my salary is only US$43 per month. Even just for meals this is not enough to live, and now we have Covid and my salary is cut. I had planned to marry in January, but what can I do and I don't even have savings. / "*Saya adalah seorang guru honorer, saat ini saya hanya berpenghasilan 600 rb dalam perbulan, untuk makan pun dalam kebutuhan perbulan pun jelas kurang, apalagi saat ini covid sedang melanda dan imbasnya gaji saya pun d potong, saya akan menikah januari tahun 2021 tapi apalah daya dana yang saya tabungpun nihil*" A teacher in Western Java. Muslim name, not a feminine name, not a breadwinner. Asked US$107. Received US$108 from six donations.
#6	I work in room service in a three-star hotel in Pekanbaru. I had only worked for six months as a contract worker. In mid-April I was furloughed without pay because guest occupancy at the hotel is low due to Covid-19 and now it is not clear when I can return to work. / "*Saya Bekerja di bidang Room Service di Hotel Bintang 3 Kota Pekanbaru. saya baru bekerja 6 bulan dengan status Karyawan Kontrak Pertengahan April saya di Rumahkan Tanpa di Gaji karna Menurunnya occupancy Tamu di hotel karna Covid-19 dan sekarang Blum jelas kapan mulai Bekerja kembali karna blum adanya pemanggilan dari Pihak perusahaan.*" A hospitality worker in Sumatera. Muslim name, feminine name, not a breadwinner. Asked US$93. Received US$108 from three donations.
#7	My coffeeshop closed, my online shop has few customers I am at home earning money for my mother, my wife, and two children. / "*Warung kopi usaha saya tutup dan online shop juga sepi dan dirumah menghidupi ibu . Istri dan dua anak.*" An entrepreneur in Eastern Java. Muslim name, not a feminine name, coded as a breadwinner. Asked US$107. Received US$107 from one donation.
#8	I migrated to Jakarta but since Covid and the mobility restrictions in March my workplace is closed and I do not have income. When it opens I only get shifts for three to five days and I am being paid daily. I have great difficulty to pay for my boarding and I am late in paying for three months of boarding while my work is closed. / "*Saya merantau di Jakarta semenjak covid dan psbb berlaku sejak Maret tempat kerja saya tutup dan saya tidak ada pemasukan, sekarang sudah buka tetapi dapat jatah masuk 3 sampai lima hari saja dan dibayar harian. Saya sangat kesulitan untuk membayar tempat kos karena ada tunggakan pembayaran kos selama 3 bulan tempat kerja saya tutup.*" A restaurant worker in Greater Jakarta. Muslim name, feminine name, not a breadwinner. Asked US$107. Received US$107 from one donation.

#9	My workplace is closed indefinitely, and my salary is cut by 50% while we work from home although the hours stay long. / "*tempat saya kerja tutup sampai waktu yang belum ditentutkan, dan gaji pun dipangkas 50% selama WFH yang ga kenal jam kerja* " A graphic designer in Greater Jakarta. Muslim name, not a feminine name, not a breadwinner. Asked US$107. Received US$107 from one donation.
#10	I got laid off because my workplace is closed. My husband was also laid off in March. He applied for a job again and got an offer, but it will only start in October. I need help to buy the needs of my infant child. / "*Saya terkena PHK karena outlet tempat kerja saya tutup. Suami juga terkena phk sejak bulan maret. Suami Sudah melamar pekerjaan, di terima, namun baru berangkat bulan oktober Saya butuh bantuan untuk beli keperluan anak saya yg msh balita.*" A worker in Central Java. Muslim name, feminine name, coded as a breadwinner. Asked US$107. Received US$107 in one donation.

Table A.2: Selection of Appeals with Top Donation Counts

Donors	Appeal (English translation/Indonesian text) and beneficiary characteristics
7 donors	The school where I work can only pay us quarterly if there's money to pay. Due to Covid-19, our enrollment went down, automatically reducing our hours and salary. / "*Sekolah/Madrasah tempat saya bekerja hanya bisa menggaji kami 3 bulan sekali itupun kalo ada gaji untuk di bagi. Dan akibat Covid 19 jumlah siswa kami menurun, itu otomatis akan mengurangi jumlah jam mengajar dan juga mengurangi gaji kamu.*" A teacher in Bali. Muslim name, not a feminine name, not a breadwinner. Asked US$107. Received US$61.
6 donors	I am a migrant from Jogja. My hotel fully closed between March and June. Thankfully it is back open but it is only 70% occupied on weekends. On weekdays the rate is below 20. Our work now pays a daily rate. I haven't gone home for six months because I don't have the money. Sadly my child is only 11 month old, infected by TBC and I am seeking medication. / "*Saya perantau dari jogja. Hotel saya tutup penuh dari maret sampai juni. Untuk sekarang sudah buka, namun hanya ramai 70% di weekend. Kalo weekday hanya belasan. Kerja sekarang hanya dihitung harian. Sudah 6 bulan belum mudik karena belum ada biaya. Sedihnya anak saya baru umur 11 bulan, tertular TBC dan sekarang sedang proses pengobatan.*" A hotel worker in West Java. Muslim name, not a feminine name, a breadwinner. Asked US$107. Received US$136.
6 donors	My contract from before 2020 is not extended because of workforce reduction from [company name]. While I don't work as a delivery worker for [store name] I sell snacks at a primary school. Thankfully I make ends meet to buy milk for my twin children and for eating day-by-day. Since Covid, schools are closed and my snacks aren't selling despite me hawking them around. I hope Bagirata can help me and my family. / "*Kontrak sblm thn baru 2020 tdk diperpanjang sy kena pengurangan dr [Nama Perusahaan], slama sy tdk kerja di delivery [Nama Toko] sy jualan cilor di skLhn sd. Untungnya lumayan bs utk beli susu anak sy yg kembar & buat mkn shari2, sejak covid19 semua skLhn diLiburkan & dagangan sy krg laku pdhl udh keliling. Mudah2an bagirata bs membatu sy dgn keluarga.*" A delivery driver in Greater Jakarta. Muslim name, not a feminine name, a breadwinner. Asked US$500. Received US$39.
6 donors	*Bismillah*, I have three children, a wife in a contract house. My work situation is improving, previously I work only 4 times a month now could be eight times depending on the schedule, with a wage of US$7 per shift. Two of my children are in school, one in grade 7 and one in grade 3. / "*Bismillah, saat ini sya mempunyai anak 3, isteri 1, rumah ngontrak. Kondisi pekerjaan sya saat ini membaik awalnya kerja hanya sebulan 4x sekarang bisa hingga sebulan 8x tergantung jadwal yg ada, dgn upah yg diberikan 1x kehadiran Rp 100.000 tanpa ada tambahan biaya yg lainnya. Ke2 anak sya sekolah setingkat smp kelas 1 dan sdi kelas 3.*" A hotel worker in Greater Jakarta. Muslim name, not a feminine name, a breadwinner. Asked US$107. Received US$41.
6 donors	Because of the pandemic I have not received wages from the school indefinitely. / "*Dikarenakan pandemi yang terjadi saat ini saya tidak menerima gaji dari sekolah untuk Waktu yang tidak bisa ditentukan.*" A teacher in East Java. Muslim name, feminine name, not a breadwinner. Asked US$107. Received US$68.

6 donors	I am a contract teacher, and my salary is only US$43 per month. Even just for meals this is not enough to live, and now we have Covid and my salary is cut. I had planned to marry in January, but what can I do and I don't even have savings. A teacher in Western Java. Muslim name, not a feminine name, not a breadwinner. Asked US$107. Received US$108. See #5 top donation for original text.
5 donors	I am a contract worker laid off due to low business since the Covid-19 pandemic. I have two children and my wife passed away two years ago. A catering worker in East Java. Asked US$71. Muslim name, not a feminine name, a breadwinner. Received a total of US$144. See #1 top donation for original text.
5 donors	I work on projects, and the business is low because of the pandemic. I barely make ends meet, no earning, need to provide for my nine-month pregnant wife, with an expected due date coming soon and I don't have money, and I need money for the birth, while work is quiet. / "*Pekerjaan saya proyek, proyek saya lagi sepi gara-gara ada pandemi ,sedangkan kebutuhan sehari-hari saya kurang ,tidak ada pemasukan ,saya harus menafkahi istri saya yang sedang mengandung 9 bulan dan bentar lagi mau lahiran saya blom ada uang ,dan saya butuh biaya buat lahiran anak saya ,apalagi kerja saya sedang sepi.* A laborer in East Java. Muslim name, feminine name, a breadwinner. Asked US$107. Received US$68.
5 donors	My workplace closed because many international flights are not running and many countries are in lockdown or paused by their government (prohibition to operate due to covid). I was forced to take unpaid leave indefinitely, while I have an infant child with many needs. / "*Perusaahan tempat kerja saya tutup karena penerbangan di dunia dalam keterpurukan, banyak negara2 yg lockdown dan penerbangan di berhentikan oleh pemerintah(larangan utk beroperasi terkait covid) saya di paksa ambil unpaid leave sampai batas yg belum di pastikan, sedangkan saya punya anak balita dan kebutuhan nya banyak..* A ticketing agent in Central Java. Muslim name, feminine name, a breadwinner. Asked US$105. Received US$57.
5 donors	After the end of my contract in May 2020 at the hotel where I worked, I adjunct for a semester, do freelance talks, freelance auditing, and sell breakfast food. In the beginning of 2021 I still have not found a job and for a month I need to make daily ends meet including for three children who are growing up. / "*Setelah Putus kontrak di May 2020 di hotelsaya bekerja sebagai dosen lepas selama satu semester saja , pembicara freelance, auditor freelance dan berjualan makanan sarapan , saya di awal thn 2021 belum dpt pekerjaan sama sekali sudah 1 bulan harus memenuhi biaya hidup sehari hari & tanggungan 3 anak yg beranjak dewasa.*" A freelancer in Sumatera. Muslim name, not a feminine name, a breadwinner. Asked US$107. Received US$43.
5 donors	My workplace is open again, but the daily rate is only US$3 per day or meal allowance. Base salary isn't being paid because the company cannot afford it now. Meanwhile, I have two children and I am a single parent. There are leftover loans to pay because for the last three months I was made to take unpaid leave. / "*Tempat kerja saya sudah buka kembali. Tapi hanya dibayarkan 45 rb sehari/uang makan saja. Tidak ada gaji pokok. Krn perusahaan tdk mampu utk membayar saat ini. Sedangkan saya punya anak 2 dan single parent. Byk tunggakan yg menumpuk krn 3 bulan kmrn di unpaid leave.*" A worker in Greater Jakarta. Muslim name, feminine name, a breadwinner. Asked US$107. Received US$37.

Table A.3: Platform Users' Self-Declared Reasons for Charitable Donations

Donors' responses to user survey on the platform	%
Donated to an organization/volunteered in the last year	92
Donated blood in the last year	18
Reasons to donate	
The beneficiary needs my donation	58
The organization is trustworthy	56
I support humanitarian causes	54
The organization uses donations effectively	50
Following religious teaching	43
I support education causes	41
I support health causes	41
I support a disaster relief program	40
I support the causes behind the fundraiser	38
I wished to not be bothered anymore by the fundraisers/beggars/buskers	3
Stated "very likely" to donate to beneficiaries with particular characteristics	
The beneficiary needs to take care of their family (children or elderly)	86
The beneficiary has been poor for a long time/came from a poor family	85
The beneficiary needs help because of an unexpected event (disaster, illness, layoff)	82
The beneficiary is a woman	69
The beneficiary lives in the same neighborhood as the donor	56
The beneficiary did not have a good education	53
The beneficiary has the the same religion as the donor	49
The beneficiary has the the same ethnicity as the donor	42
The beneficiary has also received donations from other donors	34
The beneficiary has a young age	32
Observations	216

Notes: Survey responses from Oct 2020 to July 2021.

Table A.4: Summary statistics

	3-opt sets		8-opt sets		10-opt sets	
	Mean	p90	Mean	p90	Mean	p90
	(1)	(2)	(3)	(4)	(5)	(6)
Number of donations	0.46	1	0.59	2	0.42	1
Total donations (USD)	4.39	14.29	6.67	10.71	5.08	10.71
Send any donation	0.17	1	0.19	1	0.17	1
N	774		824		807	

Notes: Summary statistics of main donation outcomes by treatment arm assignments. The median (p50) values for all three outcomes are zero across the treatment assignments.

Table A.5: Summary of Visits, Assignments by Donation Outcome

	Set = 3			Set = 8			Set = 10			Overall		
	Mean	Med	N	Mean	Med	N	Mean	Med	N	Mean	Med	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Donation (USD)												
Beneficiaries with donation	9.46	7.14	359	11.35	7.14	484	12.07	7.14	340	10.98	7.14	1,183
All displayed beneficiaries	0.32	0.00	10,620	0.26	0.00	20,776	0.2	0.00	20,690	0.25	0.00	52,086
Total set seen by visitors												
Nondonating	3.8	1	642	3.0	1	669	2.3	1	668	3.0	1	1979
Donating	8.5	5	132	3.8	2	155	4.0	2	139	5.3	3	426
All visitors	4.6	1	774	3.2	1	824	2.6	1	807	3.4	1	2405
When donation is made												
The earliest set	3.9	2	132	1.9	1	155	2.1	1	139	2.6	1	426

Notes: Table shows the mean set seen by visitors, disaggregated by eventual donation outcome (donating visitors versus nondonating visitors) and assignment to treatment arms (screen set size). Columns show the mean number of sets, median number of sets, and number of visitors in each category.

Table A.6: Summary statistics on beneficiaries' total appeal (in USD), by subsample

	Mean	SD	Count	% of Total
Employment sector				
Hospitality, retail, food service	344.40	1157.41	1,243	61%
Art and creatives	425.90	721.84	326	16%
Transportation	394.63	321.62	131	6%
Education	289.19	359.69	77	4%
Healthcare	195.59	100.56	34	2%
Other (incl. media, textile)	265.95	200.64	243	12%
Region				
Jakarta metro area	342.95	713.44	1,385	67%
Java, non–Jakarta metro	378.84	1534.68	491	24%
Outside Java	287.56	257.77	178	9%
Mobile money channels				
Go-pay	338.74	771.79	1,317	64%
Dana	353.63	1150.40	808	39%
Jenius	388.37	466.39	216	11%
Social media				
Instagram	355.21	1076.71	1,579	77%
Facebook	307.50	442.55	895	44%
Twitter	314.55	294.96	315	15%
Gender				
Masculine name	363.56	1140.01	1,302	63%
Feminine name	317.58	489.07	752	37%
Religion marker				
Muslim name	330.30	705.37	1,678	82%
Non-Muslim name	420.03	1660.99	376	18%
Household status				
Breadwinner/mentions dependent(s)	362.45	557.41	462	22%
No mention of dependents	342.17	1042.18	1,592	78%
Dependent children				
Mentions child(ren) as dependents	362.68	327.24	254	12%
No mention of a child	344.48	1012.50	1,800	88%

Note: % of total describes the proportion of each subgroup out of the 2,054 total beneficiaries. Total appeal is calculated from appeal per month times the number of months that the beneficiaries requested a donation.

Table A.7: Summary Statistics of the Display Counter and Donations among Platform Beneficiaries from Donor Perspective

	# times displayed		% receive	# donations		Donation (USD)	
Platform beneficiaries	Mean	SD	donations	Mean	SD	Mean	SD
Employment sector							
Hospitality, retail, food service	26.60	18.65	81%	2.06	2.05	17.94	29.05
Art and creatives	21.20	17.38	82%	2.09	1.85	17.86	19.27
Transportation	12.91	9.31	94%	3.82	3.62	26.31	24.64
Education	33.69	18.17	73%	2.26	2.32	23.14	23.97
Healthcare	40.38	16.54	91%	1.38	0.85	10.84	12.14
Other (incl. Media, Textile)	33.71	17.31	68%	1.34	1.34	12.04	17.20
Region							
Jakarta metro area	24.48	18.44	83%	2.24	2.20	18.64	27.45
Java, non–Jakarta metro	28.63	18.48	80%	1.91	2.14	17.29	23.41
Outside Java	32.96	17.02	65%	1.45	1.51	13.20	20.31
Mobile money channels							
Go-pay	25.10	18.16	88%	2.55	2.32	21.88	29.62
Dana	28.04	18.99	71%	1.58	1.90	13.15	20.14
Jenius	22.19	17.62	84%	2.26	1.97	20.54	23.85
Social media							
Instagram	25.19	18.43	81%	2.17	2.27	18.10	26.52
Facebook	27.75	18.53	80%	2.03	1.96	18.27	23.62
Twitter	23.09	17.94	80%	2.05	2.26	16.89	19.39
Gender codes							
Masculine name	26.37	18.48	78%	1.98	2.15	16.75	27.51
Feminine name	25.92	18.61	85%	2.27	2.12	19.74	23.09
Religion marker							
Muslim name	26.32	18.59	81%	2.13	2.22	18.17	27.00
Non-Muslim name	25.72	18.24	78%	1.91	1.78	16.39	20.99
Household status							
Breadwinner/has dependent(s)	25.61	18.22	90%	3.10	2.78	27.80	28.82
No mention of dependents	26.38	18.61	78%	1.80	1.82	14.95	24.40
Dependent children							
Mentions child(ren) as dependents	26.86	18.39	93%	3.39	2.71	31.74	31.81
No mention of a child	26.11	18.54	79%	1.91	1.99	15.88	24.47

Note: % receive donations describes the share of beneficiaries in the subgroup who receive donation out of the total beneficiaries in their respective subgroup.

Table A.8: Summary Statistics of Donations among Platform Beneficiaries with Respect to Frequency of Display to Donors

	N times displayed	Donation count	Share display receiving donation	Mean donation (USD)	Uncond. mean donation (USD)
Employment sector					
Hospitality, retail, food service	32,008	732	0.023	10.91	0.25
Art and creatives	6,632	126	0.019	10.89	0.21
Transportation	1,596	49	0.031	11.85	0.36
Education	2,523	87	0.034	12.98	0.45
Healthcare	1,353	28	0.021	7.95	0.16
Other (incl. Media, Textile)	7,974	161	0.020	10.53	0.21
Region					
Jakarta metro area	32,753	741	0.023	10.86	0.25
Outside Jakarta metro	19,333	442	0.023	11.18	0.26
Mobile money channels					
Go-pay	31,929	966	0.030	11.41	0.35
Dana	22,007	333	0.015	10.06	0.15
Jenius	4,607	110	0.024	9.14	0.22
Social media					
Instagram	38,442	836	0.022	10.60	0.23
Facebook	24,061	596	0.025	11.68	0.29
Twitter	7,018	155	0.022	9.95	0.22
Gender codes					
Masculine name	33,238	698	0.021	10.50	0.22
Feminine name	18,848	485	0.026	11.68	0.30
Religion marker					
Muslim name	42,737	957	0.022	11.10	0.25
Non-Muslim name	9,349	226	0.024	10.49	0.25
Household status					
Breadwinner/mentions dependent(s)	11,440	387	0.034	11.44	0.39
No mention of dependents	40,646	796	0.020	10.76	0.21
Children dependents					
Mentions child(ren) as dependents	6,597	260	0.039	12.01	0.47
No mention of a child	45,489	923	0.020	10.69	0.22

Table A.9: Randomization Balance Check Across Time for All Donors

	(1) Year 2021	(2) Xmas 2020 NY 2021	(3) 2021 May	(4) Weekdays	(5) Weekends	(6) Hour 5-8	(7) Covid weekly avg	(8) Covid deaths
3-opt sets	-0.026	-0.020	0.010	0.037	-0.037	0.013	-0.135	3.022
	(0.023)	(0.015)	(0.017)	(0.023)	(0.023)	(0.024)	(0.175)	(4.264)
8-opt sets	0.009	0.010	0.008	0.007	-0.007	-0.012	-0.015	-0.113
	(0.023)	(0.015)	(0.016)	(0.023)	(0.023)	(0.023)	(0.172)	(3.944)
Constant	0.696***	0.104***	0.123***	0.683***	0.317***	0.336***	12.038***	177.173***
	(0.016)	(0.011)	(0.012)	(0.016)	(0.016)	(0.017)	(0.123)	(2.896)
Observations	2405	2405	2405	2405	2405	2405	2405	2405

Notes: Regression of various time indicators and Covid variables on set size indicators. The time indicators are indicators for 2021, Christmas-New Year (25 December 2020-14 January 2021), May 2021, Weekdays (Monday-Friday), Weekends (Saturday-Sunday), and hours between 5-8 o'clock. The COVID weekly average is the weekly mean (last 7 days) of share tested people who are positive, as recorded in the Jakarta government dashboard. Covid deaths are the number of fatalities on the day, as recorded on Wikipedia. Observation unit is a donor-session. Robust standard errors are displayed in parentheses. Sample is from 4 October 2020 to 9 June 2021. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.10: Balance Check on Donor Characteristics, Sample of Matched Survey

	(1) Male	(2) Married	(3) Muslim	(4) Years of education	(5) Greater Jakarta	(6) Earning (xGDPpc)	(7) Workplace for profit	(8) Did blood donation	(9) Ever test Covid
3-opt sets	-0.025	-0.126	0.113	0.505	-0.088	0.145	0.148	-0.041	-0.082
	(0.131)	(0.111)	(0.138)	(0.439)	(0.128)	(0.641)	(0.137)	(0.064)	(0.134)
8-opt sets	-0.127	0.168	0.202	0.639	-0.075	0.577	-0.111	-0.014	0.002
	(0.121)	(0.126)	(0.132)	(0.393)	(0.123)	(0.607)	(0.129)	(0.069)	(0.128)
Constant	0.346***	0.269***	0.423***	15.923***	0.731***	2.756***	0.423***	0.077	0.654***
	(0.095)	(0.089)	(0.099)	(0.322)	(0.089)	(0.459)	(0.099)	(0.053)	(0.095)
Observations	86	86	86	86	86	86	86	86	86

Notes: Regression of donor characteristics on set size indicators for each donor-session that can be matched with platform user survey. Observation unit is a donor. Dependent variables on Columns (1), (2), (3), (5), (7), (8), (9) are indicators for gender, marital status, religion, location in Greater Jakarta/Jabodetabek, workplace is a for-profit enterprise, ever donated blood, and ever test for COVID. Robust standard errors are displayed in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11: Balance Check on Beneficiaries Average Characteristics Across Set Sizes

	(1) % set female name	(2) % set Muslim name	(3) % set is breadwinner	(4) % set mentions child	(5) % set works F&B or hospitality	(6) % set blue collar job	(7) % set in Greater Jakarta	(8) % set has facebook
3-opt sets	-0.003	-0.004	-0.002	-0.003	0.013	0.002	-0.009	-0.002
	(0.009)	(0.008)	(0.008)	(0.006)	(0.009)	(0.010)	(0.010)	(0.010)
8-opt sets	-0.005	-0.004	0.001	0.001	0.007	0.006	-0.002	0.002
	(0.007)	(0.005)	(0.006)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	0.368***	0.820***	0.220***	0.126***	0.613***	0.614***	0.643***	0.455***
	(0.005)	(0.004)	(0.004)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
Observations	2405	2405	2405	2405	2405	2405	2405	2405

Notes: Regression of set average characteristics on set size indicators for each donor-session. Observation unit is a donor-session. Robust standard errors are displayed in parentheses. Sample is from 4 October 2020 to 9 June 2021. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.12: Balance Check on Beneficiaries Characteristics In Different Positions in Set

	(1) Female name	(2) Muslim name	(3) Breadwinner	(4) Mentions child	(5) Works F&B or hospitality	(6) Blue collar job	(7) In Greater Jakarta	(8) Has facebook
Display order	-0.001	-0.001	-0.001	-0.000	0.001	-0.000	-0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.363***	0.825***	0.226***	0.131***	0.605***	0.618***	0.629***	0.460***
	(0.006)	(0.005)	(0.005)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)
Observations	52086	52086	52086	52086	52086	52086	52086	52086

Notes: Regression of beneficiary characteristics on display order within a set. Observation unit is a dyad. Regression controls for set size indicator. Robust standard errors are displayed in parentheses. Sample is from 4 October 2020 to 9 June 2021. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.13: Set-level Aggregate Appeal Characteristics by Screen Sets

	(1) Max-min appeal length mean/set	(2) Max-min appeal length mean/set
3-opt sets	-17.122***	-17.132***
	(0.417)	(0.520)
8-opt sets	-2.627***	-2.664***
	(0.312)	(0.387)
Constant	42.513***	41.677***
	(2.114)	(2.403)
Sample	All	Normal Traffic
Observations	2405	1544

Notes: Regression of differences between the longest and shortest beneficiary appeals per set. These variables are regressed on the set size indicators, time fixed effects indicators for different samples. Robust standard errors in parentheses. Sample is from 4 Oct 2020 to 9 Jun 2021. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.14: Weekly Popularity of Various Terms on Google Trends and COVID variables

	(1) *Donasi*	(2) *Donasi*	(3) Islamic charity	(4) Islamic charity	(5) Social assistance	(6) Social assistance
Jan 2021 mobility restriction	0.681		0.743		-0.435	
	(0.551)		(0.452)		(0.594)	
Post mobility restriction	-0.039		0.314		-0.137	
	(0.343)		(0.368)		(0.298)	
May day soc. media campaign	0.906		-0.430		-0.557	
	(0.627)		(0.644)		(0.437)	
Covid positivity rate		-0.000		-0.080		-0.073
		(0.106)		(0.060)		(0.086)
L.Covid positivity rate		-0.079		0.183**		0.171
		(0.112)		(0.070)		(0.130)
L2.Covid positivity rate		0.040		-0.042		-0.151
		(0.074)		(0.067)		(0.114)
Covid deaths		0.004***		0.000		0.001
		(0.001)		(0.002)		(0.003)
L.Covid deaths		-0.001		-0.004**		0.004*
		(0.002)		(0.002)		(0.002)
L2.Covid deaths		-0.000		0.004***		0.001
		(0.002)		(0.001)		(0.002)
L.term	0.561***	0.636***	0.702***	0.669***	0.607***	0.470**
	(0.200)	(0.211)	(0.129)	(0.118)	(0.211)	(0.191)
Constant	-0.118	-0.011	-0.066	-0.721	0.148	-0.243
	(0.163)	(0.556)	(0.165)	(0.445)	(0.217)	(0.440)
Observations	31	31	31	31	31	31

Notes: Regression of standardized popularity of various terms on COVID variables and social media campaign. Popularity of terms are as captured weekly by Google Trends: *donasi* (donation) in Columns (1-2), *sedekah* (Islamic charity) in Columns (3-4), *bansos* (social assistance) in Columns (5-6). Popularity of terms are standardized to have a mean of zero and a standard deviation of one. The Jan 2021 mobility restriction is an indicator for the second-fourth weeks of January 2021, when the first mobility restriction was imposed (*PPKM*). The post mobility restriction is an indicator for the next four weeks following the first restriction (Weeks 5-8 of 2021). The Mayday indicator is equal to one during the social media campaigns on Week 17-19 of 2021. "L." indicates lagged values of the relevant variable. Observation unit is a week. Robust standard errors are displayed in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.15: Interaction with Decision Complexity Using Narrative Measures

	Unique words				Corpus infrequent words			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	# donation	# donation	Donate (USD)	Donate (USD)	# donation	# donation	Donate (USD)	Donate (USD)
3-opt sets	0.162*	0.309***	0.707	1.292	0.103	0.213*	1.163	1.451
	(0.088)	(0.115)	(1.122)	(1.480)	(0.094)	(0.120)	(1.146)	(1.301)
... x Abv median complexity	-0.237*	-0.287	-2.701	-1.721	-0.120	-0.104	-3.622	-2.119
	(0.139)	(0.178)	(2.071)	(2.119)	(0.139)	(0.177)	(2.284)	(2.275)
8-opt sets	0.286***	0.342**	3.655*	5.186*	0.202	0.268	2.159	3.026*
	(0.110)	(0.142)	(1.873)	(2.665)	(0.135)	(0.178)	(1.359)	(1.772)
... x Abv median complexity	-0.202	-0.162	-3.534	-2.997	-0.033	-0.026	-0.551	1.347
	(0.181)	(0.248)	(2.905)	(3.720)	(0.173)	(0.218)	(3.070)	(3.845)
Abv median complexity	0.252***	0.261***	2.595	1.065	0.093	0.081	2.759	2.042
	(0.087)	(0.099)	(1.623)	(1.479)	(0.089)	(0.105)	(1.808)	(1.520)
Constant	1.689***	1.925***	24.311***	19.208**	1.775***	2.011***	24.110***	18.601**
	(0.466)	(0.614)	(7.598)	(8.546)	(0.475)	(0.632)	(7.608)	(8.656)
Sample	All	Normal Traffic	All	Normal Traffic	All	Normal Traffic	All	Normal Traffic
Observations	2405	1544	2405	1544	2405	1544	2405	1544

Notes: Regression at session level using various indicators of above median narrative complexity interacted with treatment indicators. Columns 1-4 use the share of unique words in appeal text, columns 5-8 use words that are not frequently used in the corpus derived from *Kompas* newspaper and Indonesian Wikipedia. Regressions are adjusted for Covid variables and a set of time indicator variables. Observation unit is a donor-session. Robust standard errors are displayed in parentheses. Sample is from 4 October 2020 to 9 June 2021. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.16: Impact of Screen Size on Donation Outcomes for Selected Beneficiary Groups

	(1)	(2)	(3)	(4)	(5)
	1[Donate]	# donation	Donate (USD)	1[Donate]	Donate (USD)
3-opt sets	0.020	0.033*	0.374	0.008	0.095
	(0.013)	(0.020)	(0.261)	(0.006)	(0.084)
8-opt sets	0.028**	0.040**	0.527*	0.009**	0.113*
	(0.013)	(0.019)	(0.294)	(0.004)	(0.059)
Constant	0.056***	0.077***	0.735***	0.018***	0.186***
	(0.008)	(0.012)	(0.154)	(0.002)	(0.034)
Observations	2405	2405	2405	51905	51905

Note: Regressions on selected beneficiary groups: cols. 1–3 use only the first three beneficiaries that donors see; cols. 4–5 are at the dyad level, with fixed effects for beneficiary, sequence, and screen. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.17: Beneficiary Characteristics and Donation Outcomes

	(1) 1(Donate)	(2) Donation (USD)
Breadwinner	0.007***	-0.009
	(0.002)	(0.041)
Transportation worker	-0.005	-0.092
	(0.005)	(0.128)
Laid off	0.001	-0.017
	(0.002)	(0.029)
Arts	-0.004**	-0.075**
	(0.002)	(0.034)
Education worker	0.013***	0.188***
	(0.004)	(0.056)
Narrative (x50 words)	0.005**	0.020
	(0.002)	(0.040)
Female name	0.001	0.042
	(0.001)	(0.026)
Muslim name	-0.002	-0.014
	(0.002)	(0.031)
Non-formal language	0.000	-0.005
	(0.001)	(0.023)
Facebook link	0.004***	0.042
	(0.002)	(0.030)
Instagram link	-0.001	-0.038
	(0.002)	(0.038)
Twitter link	-0.003*	-0.036
	(0.002)	(0.032)
Greater Jakarta	-0.004**	-0.042
	(0.001)	(0.026)
Order in set	-0.001**	-0.004
	(0.000)	(0.004)
Gopay e-channel	0.015***	0.118***
	(0.003)	(0.036)
Dana e-channel	0.001	-0.037
	(0.002)	(0.033)
Jenius e-channel	0.006	-0.003
	(0.004)	(0.048)
No donations yet	-0.019***	0.089
	(0.003)	(0.071)
% Ask fulfilled	0.001***	0.025***
	(0.000)	(0.006)
Set counter	-0.000	-0.000
	(0.000)	(0.000)
Ask amount (USD)	-0.000	0.001***
	(0.000)	(0.000)
Ask duration	0.000	0.012
	(0.001)	(0.014)
Constant	0.018***	-0.223*
	(0.005)	(0.130)
Dep. Var. Mean	0.023	0.249
R2	0.259	0.214
Observations	52086	52086

Notes: Regression of donation outcomes on beneficiary characteristics with donor session FE. Observation unit is donor–beneficiary dyad. Standard errors are clustered at the donor and beneficiary levels and displayed in parentheses. Sample is from Oct 2020 to Jun 2021, excluding outliers. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.18: Exposure to High-deservingness Score Beneficiaries by Screen Sets

	(1) # above p50 deserve index	(2) # above p50 deserve index	(3) # above p75 deserve index	(4) # above p75 deserve index	(5) # above p90 deserve index	(6) # above p90 deserve index
3-opt sets	-6.442***	-7.809***	-3.293***	-3.937***	-1.316***	-1.612***
	(1.364)	(2.026)	(0.671)	(0.991)	(0.285)	(0.421)
8-opt sets	-0.343	-1.393	-0.274	-0.789	-0.114	-0.383
	(1.837)	(2.709)	(0.915)	(1.342)	(0.378)	(0.552)
Constant	16.117*	13.955	7.500	6.294	2.214	1.559
	(9.011)	(10.182)	(4.642)	(5.308)	(1.860)	(2.118)
Sample	All	Normal Traffic	All	Normal Traffic	All	Normal Traffic
Observations	2405	1544	2405	1544	2405	1544

Notes: Regression of the total number of beneficiaries encountered with a within-screen-set deserving-ness index score higher than various thresholds (p50, p75, and p90). This variable is regressed on the set size indicators, time fixed effects indicators for different samples. Robust standard errors in parentheses. Sample is from 4 Oct 2020 to 9 Jun 2021. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.19: Robustness Check: In-group Bias and the Probability of Donation (Sample restricted to screens with balanced beneficiary characteristics)

	(1) 1[Donation]	(2) 1[Donation]	(3) 1[Donation]	(4) 1[Donation]	(5) 1[Donation]
Female donor-feminine name beneficiary	0.130**				
	(0.060)				
Male donor-masculine name beneficiary		-0.018			
		(0.086)			
Muslim donor-muslim name beneficiary			0.004		
			(0.112)		
Non-muslim donor-non muslim name beneficiary				-0.107	
				(0.086)	
Donor-beneficiary in same South Jakarta district					-0.071
					(0.087)
Observations	158	158	60	60	98

Note: Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.20: Robustness Check: In-group Bias and Donation Value
(Sample restricted to screens with balanced beneficiary characteristics)

	(1) Donation (USD)	(2) Donation (USD)	(3) Donation (USD)	(4) Donation (USD)	(5) Donation (USD)
Female donor-feminine name beneficiary	0.957*				
	(0.509)				
Male donor-masculine name beneficiary		-0.099			
		(0.961)			
Muslim donor-muslim name beneficiary			0.031		
			(0.797)		
Non-muslim donor-non muslim name beneficiary				-0.773	
				(0.617)	
Donor-beneficiary in same South Jakarta district					-0.587
					(0.639)
Observations	158	158	60	60	98

Note: Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.21: Regression of the Donation Outcomes on Beneficiary's Donation Progress

	(1) 1(Donate)	(2) 1(Donate)	(3) 1(Donate)	(4) 1(Donate)	(5) Donation (USD)	(6) Donation (USD)	(7) Donation (USD)	(8) Donation (USD)
% Ask fulfilled	0.001***	0.001***	0.004***	0.003***	0.027***	0.024***	0.109***	0.109***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.005)	(0.004)	(0.021)	(0.022)
FE	_cons	donor	beneficiary	donor beneficiary	_cons	donor	beneficiary	donor beneficiary
R2	0.015	0.252	0.076	0.303	0.029	0.212	0.147	0.312
Observations	52086	52086	52081	52081	52086	52086	52081	52081

Notes: Regression of the donation outcomes on beneficiary's donation progress (% of asked donation amount received so far from other donors). The observation unit is a donor–beneficiary dyad. Standard errors are clustered at the donor and session levels and displayed in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B. Further Supplementary Materials

Beneficiary Coding Guidelines. We coded gender and religion from the beneficiaries' names to create indicators for feminine names and Muslim names. We rely on beneficiary's location at the district level to approximate his neighborhood origin. We do not have explicit markers for education and age, but we use beneficiaries' writing style from their narrative appeals and use of social media to provide information. Assistants coded the use of nonformal written Indonesian with reliance on abbreviations, regional slang for pronouns, and (mis)use of punctuation marks, which are typically associated to individuals with lower education. We include indicators of social media links, which also provide hints about the beneficiary's age: a social media analytics tool company reports that Instagram is mostly used by younger age groups, while Facebook is more popular among older people in Indonesia. Specifically, slightly more than 50% of Instagram users in Indonesia are 13–24 years old, compared to 40% of Facebook users in the same age group. Facebook also has a larger share of users from the 35+ age group than Instagram at 28% versus 18%, respectively (NapoleonCat, 2023).

Note on Pre-registration. This project was preregistered at the Open Science Framework using an associated project that was created in Jan 29, 2021. This pre-registration was made public on May 21, 2021. The public pre-registration is accessible at https://osf.io/c4xgd and the original associated project is located at https://osf.io/wnz46. The pre-registration was created and made public prior to the full dataset made available to the researchers (see Olken JEP 2015 for discussion on timing). This pre-registration included hypotheses we tested, a sketch of the study design, and prospective sample sizes. We rewrote the hypotheses in the pre-registration follow the theoretical framework we outlined in Section 3. We were unable to test our original Hypothesis 2 as the features to do so were not implemented by our partner. Where the experiments are implemented, we pre-specified the outcomes of the number of donations, donation amount, and an indicator of donation. The original hypothesis 1, 3, and 4 are still reported in Tables 2, A.21, and 5 respectively. On the pre-specified study design, treatments 1-2-3 correspond to the use of 3-, 8-, and 10-option beneficiaries. We used treatment 3 (10-option beneficiary) as the control group because it was the arm that was originally implemented prior to the introduction of the experiment. Treatment 4, presenting donors with a menu of randomly selected beneficiaries of varying characteristics is implemented by randomly selecting beneficiaries from the beneficiary list. The implementation departs from the planned design by instead crossing this with treatments 1-2-3 (instead of only treatment 3). Treatment 5, providing information on the number of other donors who have donated to the same beneficiary, were scrapped from implementation. Treatment 6, information on the magnitude of donations that beneficiaries have received so far, were instead uniformly provided to all potential donors.

Keyness Statistics. This method computes a χ^2 statistic for each term that appears in a beneficiary narrative and ranks, across all narratives, the most frequently mentioned terms for beneficiaries who received at least one donation vis-à-vis those who did not receive any donations. In our context, this method approximates asking donors for the motivations behind their decision to donate to a specific beneficiary, based on various perceived measures of deservingness drawn from textual analysis of beneficiary narratives. In political science, this method has been used to identify right- versus left-leaning voters from self-written voter descriptions (Zollinger, 2022). The results for this statistic are displayed in Figure 5, although one should interpret the appearance and ranking of individual terms with caution (Zollinger, 2022).

Keywords positively associated with donation are those related to beneficiaries with dependent children or affiliations with the education sector. Narratives containing terms related to children, pregnancy and childbirth, or marriage are more likely to attract donations. Likewise, narratives containing the terms "teacher" or "college student" receive more favorable donation outcomes. In contrast, narratives that contain terms indicative of employment hardship, such as references to restaurant closures or cancelled events, are less likely to secure donations. The original Indonesian words for these translations are as follows: *anak, hamil/kandung, lahir, istri* for beneficiaries as family breadwinners; *guru, mahasiswi* for education-sector markers; and *kafe, restoran, tutup, acara, event, EO* for the hospitality industry and performing arts. We incorporate these individual seed words into a regression analysis by computing the deservingness index as a composite score for each beneficiary narrative using latent semantic scaling.

LSS: Latent Semantic Scaling. Latent semantic scaling (LSS) utilizes an initial set of user-

Table B.1: Top 10 Keywords: Keyness Statistics

donate = 1 (***deserving***)	anak	lahir	madrasah	guru	separuh	mother	goyang	hamil	tunggal	pantomim
donate = 0 (***undeserving***)	acara	tutup	batas	kerja	hibur	status	event	nikah	outlet	kafe

Notes: The first row lists 10 keywords among the narratives of beneficiary who received at least one donation.

defined "seed words" to assign scores to other words based on their contextual proximity to the seed words. In addition to these user-defined seed words, LSS requires a substantial corpus of documents, typically ranging from 5,000 to 10,000 documents. To calculate the semantic proximity between words in the corpus, LSS employs a word-embedding technique, generating word vectors that represent low-dimensional representations of word semantics. These produced word vectors are then used by LSS to calculate proximity scores for each word in relation to each seed word. The score of a given word to all predefined seed words is then weighted to calculate the proximity score of each word. Subsequently, LSS computes the proximity score of documents by weighting the proximity scores of individual words provided in the documents based on their frequency within the documents.

Table B.1 presents the seed words utilized in the computation, based on the keyness statistics. Words with closer contextual associations with the deservingness markers are assigned scores closer to 1, while words with closer contextual associations with undeservingness are assigned scores closer to -1. For example, the word "*mahasiswi*" (female student) receives the highest score, as it is contextually closer to the 10 deservingness seed words. Conversely, the word "*tutup*" (close(d)) receives the lowest score, as it is contextually closer to the 10 undeservingness seed words. This process is repeated for every single word that appears in a beneficiary's narrative. For each beneficiary narrative, *latent semantic scaling* maps *keyness statistics* to a composite score by computing and assigning a weighted proximity score for each word, in each narrative, to the seed words listed in Table B.1.

To illustrate this procedure, we discuss two beneficiary narratives, one with the lowest and one with the highest proximity score. Take the beneficiary narrative with the lowest proximity score, "*Saya bekerja sebagai Disk Jockey DJ paruh waktu untuk dua outlet [Group name] yaitu [Bar name] dan [Pizza name has the word party] dan minimal saya mendapat giliran 3 kali dalam sebulan. Itu adalah satu-satunya sumber pemasukan saya sebelum Covid 19 menyerang dan tempat itu tutup sampai waktu yang tidak ditentukan*". Collectively, every (stemmed) word in this narrative possesses minimal contextual similarities with any of the top 10 deservingness seed words. Instead, they demonstrate very close contextual meanings with the top 10 undeservingness seed words. For example, the word "party" shares a close contextual meaning with the seed word "event" and the word "bar" to the seed word "cafe."

In contrast, the document with the highest score, "*[School name] sebagai yayasan pengelola tenaga alih daya outsourcing yang menampung guru-guru praktikum di sekolah-sekolah swasta ditutup karena pendemi covid 19. Saya dan semua guru diberhentikan baik guru full time maupun part time Saya sebagai guru full time pun diberhentikan dan hanya menerima gaji terakhir saya bekerja tanpa pesangon*", contains several words that possess close, if not identical, contextual meanings with the deservingness seed words. For instance, the word "*guru*" appears multiple times in the document and is one of the top 10 seed words, contributing to the higher score assigned to this document.

Hence, we use the LSS statistic as our primary text-based measure of deservingness. We linearly rescale the LSS score to lie in $[0, 1]$, where 1 indicates maximal semantic similarity to our deservingness keywords (and minimal similarity to our undeservingness keywords), and 0 indicates the reverse. We refer to this rescaled measure as the *deservingness index*.