UCLouvain

UNIVERSITÉ DE NAMUR

Master's Thesis · Academic Year 2025–2026

Do Norm Sources Matter?

Experimental Evidence from Belgium

Théo Mauroy

Supervised by Giovanna d'Adda & Johannes Johnen

In partial fulfillment of the requirements for the joint degree:

Master 120 en Sciences économiques à finalité spécialisée (UCLouvain/UNamur)

and

Laurea magistrale in Economics and Political Science (Università degli Studi di Milano)

Université catholique de Louvain · Place Montesquieu 3, 1348 Louvain-La-Neuve · Belgium

Université de Namur · Rempart de la Vierge 8, 5000 Namur · Belgium

Università degli Studi di Milano · Via Festa del Perdono 7, 20122 Milano · Italy

Abstract

This thesis tests whether regional descriptive norms affect charitable giving, and whether the effect depends on whether the norm originates from the participant's own region or another. In a randomized online experiment, Belgian participants allocated a hypothetical €100 endowment after exposure to a control condition or to a descriptive norm drawn from administrative tax data and attributed to Flanders or Wallonia. In-group and out-group exposure were constructed by crossing the randomly assigned norm source with pre-treatment region of residence. The norm value was held nearly equal across regions to isolate the identity of the source from its magnitude. Exposure to a regional norm did not detectably shift donations. The in-group premium was statistically indistinguishable from zero, and regional identification did not moderate that premium. The out-group norm did not reduce giving below the control baseline; its estimated effect was positive, and the data rule out contrastive effects larger than about five euros. Participants registered the norm as differing from their prior beliefs in the expected direction, but this was not accompanied by a behavioral response. The study is an underpowered pilot, and the identity contrast is identified on the Walloon subsample, so the test concerns Walloon responses to a Flemish out-group norm; the symmetric Flemish-side test remains open. Its contribution is to show that, in this setting, an out-group norm produced no sign of active resistance.

Keywords: behavioral economics; prosocial behavior; charitable giving; descriptive norms; social identity; information provision; experimental economics; Belgium.

Data availability: The replication package, including the oTree code, Stata scripts, and anonymized analysis data, is available at https://doi.org/10.5281/zenodo.20581409.

Acknowledgments

I would first like to thank my supervisors, Giovanna d'Adda and Johannes Johnen. Their time, guidance, and support throughout this thesis were invaluable to me. From the first discussions to the final revisions, they were generous with their attention, thoughtful in their advice, and patient in helping me clarify both my ideas and the direction of this work. Their feedback shaped this thesis in many important ways, and I am deeply grateful for the trust and encouragement they offered throughout the process. I could not have asked for better support.

I owe a particular thanks to Jean Hindriks for the opportunity he gave me during my master. It was my first step into the research world, and the chance to grasp what that work really involves came at the right time. His support throughout these years has meant a great deal to me.

I am also grateful to the teachers and the people I met during my studies, who each contributed in their own way to making these years intellectually stimulating and personally meaningful. Among them, I would like to mention Yannick in particular, a friend and study partner with whom I shared much of the path toward this degree.

Finally, I thank my friends and my family for their constant support and encouragement throughout these years. And above all, I thank Margaux, my girlfriend, for her patience, her kindness, and her presence both throughout this thesis and outside of it.

Declaration on the Use of Generative AI

I used generative AI tools during the preparation of this thesis in a limited and assistive capacity. Claude was used to improve language, English phrasing, and stylistic clarity. Claude Code and Codex were used to assist with code debugging and LaTeX troubleshooting.

All research ideas, theoretical choices, experimental design decisions, empirical strategy, interpretation of results, and substantive conclusions are my own. These tools were not used to generate the research question or to determine the argument of the thesis.

All AI-assisted outputs were reviewed, edited, and validated by me. I remain fully responsible for the content of the thesis, the accuracy of the analysis, and the final wording.

Contents

List of Figures

List of Tables

Chapter 1 – Introduction

Descriptive social norms are among the most widely used tools for encouraging prosocial behavior. Telling people how much others give, recycle, or conserve can shift their own behavior toward the norm, and charities and public agencies increasingly rely on this mechanism. Yet a descriptive norm is never just a number. It is always attached to a reference group, and the behavioral weight of that group may depend on who its members are. A norm describing the giving of one's own community may carry more force than the same norm describing a group seen as different or distant. In a society organized around salient social divisions, the identity of the group behind a norm may determine whether the norm is followed, ignored, or actively resisted.

This thesis studies that question in the context of charitable giving in Belgium. It asks whether exposure to a regional descriptive norm affects how much people donate, and whether the effect depends on whether the norm originates from the participant's own region or from another. Belgium is a useful setting because its regional categories, Flemish and Walloon, are socially and politically meaningful without being artificial, so a norm attributed to one region or the other is attached to a real social identity rather than a label invented for the experiment.

The existing literature establishes that norm source can matter, but it leaves a specific question open. Studies comparing in-group and out-group norms can show that an out-group norm is followed less than an in-group norm (Hysenbelli et al., 2013; Kandul and Lanz, 2021; Reich et al., 2022), but most cannot distinguish two very different reasons for this. A weaker response to an out-group norm may mean the norm is simply less relevant, or it may mean the norm provokes active contrast, pushing giving below what it would have been with no norm at all. Separating these requires a no-norm baseline that much of the prior work does not include.

I address this with a randomized online experiment. Participants are assigned to a control condition, a Flemish norm condition, or a Walloon norm condition, with the norm drawn from administrative tax data and held nearly equal in value across regions so that any difference reflects the source rather than the magnitude. In-group and out-group exposure are constructed by crossing the randomly assigned norm source with the participant's pre-treatment region of residence. An experimental approach is necessary here because observational data cannot separate the effect of norm exposure from the pre-existing differences in generosity, beliefs, and identity that determine both who encounters a norm and how much they give. Random assignment of the norm makes treatment status independent of these characteristics, so that differences in giving across conditions can be attributed to the treatment itself. The inclusion of a no-norm control allows the design to ask not only whether an out-group norm is followed less than an in-group norm, but whether it reduces giving below the no-norm baseline.

The central result is a behavioral null. Exposure to a regional norm did not detectably shift donation amounts, the in-group premium was statistically indistinguishable from zero, and regional identification did not moderate that premium. The out-group norm did not depress giving below control; its estimated effect was positive rather than negative, and the data rule out

contrastive effects larger than approximately five euros. At the same time, the information was received and recognized as discrepant from prior beliefs in the expected direction. This belief-relevant response was not accompanied by a change in behavior. The study is an underpowered pilot, able to detect only moderately large effects, so these nulls bound the size of any effect rather than establishing its absence.

The contribution of the thesis is therefore clarifying rather than confirmatory. It does not show that regional norms change giving, and it is not powered to detect the small effects the literature reports. What it shows is that, in a setting where regional identity is meaningful, attaching a descriptive norm to a regional out-group did not generate the contrastive response that the identity-source hypothesis predicts—with the qualification that the contrast is identified among Walloon participants responding to a Flemish norm, so the test runs in one direction only. The no-norm control makes this a stronger statement than a simple comparison of in-group and out-group arms could support. The remainder of the thesis is organized as follows. Chapter 2 reviews the literature on social preferences, descriptive norms, and social identity, and develops the hypotheses. Chapter 3 describes the experimental design. Chapter 4 sets out the empirical strategy. Chapter 5 reports the results. Chapter 6 discusses their interpretation and limits, and chapter 7 concludes.

Chapter 2 – Literature Review

This chapter situates the experiment within three connected literatures. First, work on social preferences and charitable giving shows that donations are not only driven by pure altruism, but also by warm-glow motives, social image, fairness concerns, reciprocity, and conditional cooperation. This makes charitable giving potentially responsive to information about what others do. Second, the literature on descriptive social norms explains why such information may affect behavior, while also showing that norm effects depend on attention, credibility, prior beliefs, and the relevance of the reference group. Third, research on social identity and intergroup behavior suggests that the same norm may be interpreted differently when attributed to an in-group or an out-group. Together, these literatures motivate the central question of the thesis: whether regional descriptive norms affect charitable giving, and whether their effect depends on the identity source of the norm.

2.1 Social preferences and prosocial behavior

Charitable giving is a suitable setting for studying descriptive social norms because donations are not only driven by pure altruism. In the model of impure altruism, individuals may give because the act of giving itself produces private satisfaction, or "warm glow", independently of the effect of their contribution on the public good (Andreoni, 1990). Prosocial behavior may also reflect concerns for self-image, social image, and moral motivation (Bénabou and Tirole, 2006). Models of social preferences similarly show that individuals may care about fairness, reciprocity, and others' payoffs (Fehr and Schmidt, 1999; Charness and Rabin, 2002). Donation decisions are therefore socially meaningful choices. They can plausibly respond to information that changes how individuals evaluate their own behavior.

This point is especially important when prosocial behavior is conditional on beliefs about others. Evidence from public good experiments shows that many individuals are conditional cooperators. Their willingness to contribute depends on what they believe others contribute (Fischbacher et al., 2001). Social comparison can also shape prosocial behavior, because individuals may evaluate their own contribution relative to the contribution of others (Frey and Meier, 2004). This makes charitable giving a natural domain for testing descriptive social norms. If people use others' behavior as a benchmark, then information about regional donation behavior may affect their own donation choice. This raises a central issue for the present thesis. If charitable giving can depend on beliefs about others' behavior, the identity of those others may determine whether the information is behaviorally relevant. A norm attributed to an anonymous population, an in-group, or an out-group may not carry the same meaning.

2.2 Social identity and intergroup behavior

Social identity theory provides a reason why the source of social information may matter. Individuals do not only make choices as isolated decision makers. They also define themselves through social categories and group memberships (Tajfel and Turner, 1979; Tajfel and Turner, 1986). These categories can affect how people interpret the behavior of others. They can also influence which comparisons are seen as relevant. In economics, identity has been modeled as a component of preferences. Individuals may derive utility from acting in ways that are consistent with the norms and expectations attached to a social category (Akerlof and Kranton, 2000). This makes identity relevant for economic choices, including choices with social or moral content.

Experimental evidence also shows that group membership can shape prosocial behavior. Group identity affects allocation decisions, cooperation, and social preferences in economic experiments (Charness, Rigotti, et al., 2007; Y. Chen and Li, 2009). This suggests that a behavioral benchmark may not have the same effect when it is attributed to an in-group or to an out-group. Evidence on cooperation further shows that individuals often favor in-group members relative to out-group members (Balliet et al., 2014). However, in-group favoritism should not be confused with out-group hostility. Favoring one's own group does not necessarily imply a desire to harm or distance oneself from another group (Brewer, 1999). Evidence from natural groups also supports this distinction between in-group love and out-group hate (Weisel and Böhm, 2015). This distinction is central for the present thesis. A weaker response to an out-group norm may reflect lower relevance, but it may also reflect active contrast. These are different mechanisms and they imply different empirical predictions.

The contrastive prediction therefore requires a mechanism beyond in-group favoritism. Social identity theory predicts positive differentiation from out-groups (Tajfel and Turner, 1979); applied to a behavioral benchmark, differentiation implies that an individual may distance her behavior from an out-group standard in order to affirm group membership. A related channel is reactance against social information perceived as originating from a non-self-relevant or rival source. Under either mechanism, an out-group norm is not merely discounted. It becomes a benchmark to deviate from, which is the behavioral content of H3 and what distinguishes it from simple non-compliance.

Social identity therefore matters especially in the context of norms. A descriptive norm always refers to some group of others. If that group is perceived as self-relevant, the norm may provide a meaningful benchmark for behavior. If the group is perceived as distant, the norm may be ignored. If the group is perceived as an out-group, the norm may even generate resistance. The next section therefore turns to the mechanisms through which descriptive social norms can influence behavior.

2.3 Social norms: theory and mechanisms

Social norms provide one mechanism through which social information can affect behavior. A central distinction is between descriptive and injunctive norms. Descriptive norms refer to what others do. Injunctive norms refer to what others approve or disapprove of (Cialdini et al., 1990).

This thesis studies descriptive norms. The treatment does not tell participants what they should do. It gives them information about average declared charitable donations in a region. The norm is therefore presented as a behavioral benchmark, not as a moral instruction.

Descriptive norms can influence behavior because they affect empirical expectations. Individuals may use others' behavior to infer what is common, appropriate, or expected in a given situation. In Bicchieri's framework, norm compliance depends partly on beliefs about what others do and what others expect (Bicchieri and Xiao, 2009). Related work also identifies social norms as shared perceptions of the social appropriateness of actions (Krupka and Weber, 2013). The treatment used in this thesis is narrower. It does not elicit perceived appropriateness. It provides descriptive information about regional donation behavior. Its first theoretical channel is therefore informational. It may change what participants believe about how much others give.

This distinction between information and behavior is important. Information-provision experiments show that treatment effects depend on whether participants receive, understand, and believe the information they are shown (Haaland et al., 2023; Stantcheva, 2023). A descriptive norm may therefore have a first-stage effect on beliefs without producing a second-stage effect on behavior. This is especially relevant in charitable giving. A participant may learn that another group gives a certain amount on average, but still leave her own donation unchanged if the information is not salient, credible, or relevant to the decision.

Norm effects may also differ across individuals. Descriptive norms can move behavior toward a benchmark, but they can also create boomerang effects. In the energy conservation setting, households below and above the norm responded differently to descriptive feedback (Schultz et al., 2007). This possibility matters for charitable giving as well. Individuals who initially intend to give little may respond differently from individuals who already intend to give a lot. The empirical strategy therefore does not only estimate average treatment effects. It also examines whether responses differ across the donation distribution.

The theoretical prediction is therefore not that descriptive norms always increase giving. A norm message can affect behavior only if the information is processed and if the reference group is behaviorally relevant. The empirical literature confirms this mixed picture. Norm messages sometimes change prosocial behavior, but their effects depend on the domain, the outcome, and the reference group.

2.4 Social norms: empirical evidence

Empirical evidence shows that descriptive norm information can affect behavior in several domains. In a large field experiment on energy conservation, households received reports comparing their electricity use with that of similar neighbors. The treatment reduced energy consumption, which shows that social comparison can change behavior in a real-world setting (Allcott, 2011). The relevance of this evidence for the present thesis is not the domain itself, but the mechanism. Individuals may adjust their choices when they learn how their behavior compares with that of others.

Similar mechanisms have been studied in charitable giving. In a field experiment with a public

radio station, donors changed their contributions when they were informed about another donor's contribution (Shang and Croson, 2009). This suggests that giving can respond to social information. However, the effect of descriptive norms depends on the reference group used in the message. In the context of charitable donations, local descriptive norms appear more effective than more general norms (Agerström et al., 2016). This is important for the present thesis because the treatment does not only provide information about giving. It attributes this information to a regional group.

The reference-group dependence of descriptive norms is well established outside giving as well. Hotel guests reuse towels more when the norm refers to previous guests in the same room rather than to guests in general (N. J. Goldstein et al., 2008), and tax-compliance letters citing local rather than national payment norms increase payment rates at national scale (Hallsworth et al., 2017). Both results indicate that proximity and specificity of the reference group condition the behavioral force of the same descriptive information.

The broader evidence on social norms and charitable giving also suggests caution. A recent meta-analysis finds that descriptive social norms can affect charitable giving, but that the average effect is modest and heterogeneous across designs (Chapman et al., 2025). This matters for the present thesis because it sets realistic expectations. A descriptive norm may move behavior, but the literature does not imply that every credible norm message should generate a large response. The meta-analysis also suggests that effect sizes are sensitive to design features and study quality. This reinforces the need to interpret null effects carefully, rather than as simple evidence that descriptive norms never matter.

One reason for this heterogeneity is that descriptive norms differ not only in magnitude or wording. They also differ in source. The same behavioral benchmark may be interpreted differently depending on whether it is attributed to an in-group, an out-group, or a broader population. The next section therefore focuses on the literature closest to this thesis: the role of norm source and group identity in prosocial behavior.

2.5 Norm source and identity

The empirical evidence on descriptive norms suggests that the source of the norm matters. A descriptive norm is not only a number or a behavioral benchmark. It is also attached to a reference group. A norm attributed to a socially relevant group may provide a stronger behavioral signal than a norm attributed to a distant or anonymous population. This is consistent with evidence that local descriptive norms can be more effective than more general norms in charitable giving (Agerström et al., 2016). It is also consistent with the broader evidence that norm effects vary across contexts and reference groups (Chapman et al., 2025). The identity of the reference group is therefore not a secondary detail. It is part of the treatment itself.

The closest evidence to this thesis comes from studies that connect charitable giving, group source, and social influence. Hysenbelli et al. (2013) examine whether others' opinions affect donation decisions, and whether this influence depends on group membership. Their results show that social information can affect charitable giving, but that its effect differs depending on

whether the source belongs to the in-group or the out-group. This is directly relevant for the present thesis. It suggests that people do not react only to the content of social information. They also react to the identity of the person or group from which the information comes.

Recent evidence also studies how the behavior of in-group and out-group members affects charity preferences (Reich et al., 2022). This work is useful because it focuses on charity and on the identity of the behavioral source. It shows that group membership can enter the decision process itself, rather than merely shaping attitudes measured after the decision. However, this literature does not fully settle the question addressed here. It shows that in-group and out-group sources may have different effects, but the interpretation of an out-group response depends on the benchmark used for comparison.

Evidence from public-good provision points in the same direction. Kandul and Lanz (2021) study descriptive norms from in-groups and out-groups in a public-good setting. This is relevant because public-good provision and charitable giving share a common structure. Both involve a private cost and a prosocial or collective benefit. In both settings, information about others' behavior may provide a benchmark for one's own behavior. Yet the behavioral meaning of that benchmark can change when it is attributed to an out-group rather than an in-group.

A central distinction remains. A weaker response to an out-group norm is not the same as active resistance to that norm. Social identity theory and related evidence show that in-group favoritism does not necessarily imply out-group hostility (Brewer, 1999; Weisel and Böhm, 2015). An out-group norm may therefore have several possible effects. It may be followed, if it is treated as useful information. It may be ignored, if it is seen as unrelated to the participant's own group. It may also generate contrast, if the participant distances herself from the out-group benchmark. These mechanisms imply different empirical predictions.

This distinction is central for the present thesis. Existing evidence shows that the identity source of social information can matter. Yet it remains important to separate out-group irrelevance from out-group resistance. The present design contributes a no-norm baseline that allows these two possibilities to be distinguished. The Belgian setting is useful for this purpose because regional labels are meaningful outside the experiment, while norm exposure can still be assigned experimentally.

2.6 The Belgian regional context

Belgium provides a useful setting for studying norm source and identity because regional categories are socially and politically meaningful. Belgian politics is structured around linguistic and regional divisions, and the distinction between Flanders and Wallonia is central to the country's institutional and political organisation (Deschouwer, 2012). These categories are not artificial groups created for the experiment. They are real social categories that participants can recognize and interpret.

The Belgian case is also useful because regional identity is salient without requiring a setting of open conflict. Flanders and Wallonia differ in language, political representation, economic history, and public discourse, but they remain part of the same national polity. This makes the

regional distinction strong enough to give meaning to an in-group or out-group source, while still allowing the experiment to study relatively subtle behavioral responses rather than explicit hostility.

This matters for the interpretation of descriptive norms. A norm attributed to Flanders or Wallonia may not only convey information about charitable giving. It may also signal the behavior of a group that participants perceive as more or less socially close, relevant, or comparable to themselves. In this sense, the Belgian setting allows the experiment to test whether the identity source of a norm changes its behavioral meaning.

This does not require assuming strong hostility between regions. The argument only requires that regional labels are meaningful enough for the source of the norm to be interpretable. The Belgian setting therefore allows the experiment to attach descriptive norm information to real group identities while keeping treatment assignment under experimental control. The next chapter presents the experimental design used to test this argument.

Chapter 3 – Experimental Design

3.1 Overview and hypotheses

The experiment uses a three-arm between-subjects design. Each participant is assigned to exactly one condition: a control condition, a Flemish norm condition, or a Walloon norm condition. In-group and out-group exposure are not assigned as separate treatment arms. Instead, they are constructed by crossing the assigned norm source with the participant's self-reported region of residence, which is measured before treatment exposure. The coding of in-group and out-group exposure is defined in section 3.5. The experiment is implemented in oTree (D. L. Chen et al., 2016) and deployed in French and Dutch, with an English option available.[1]

Table 3.1: Experimental arms

Arm	Information shown	Main comparison
Control	Neutral filler	Baseline
Flemish norm	SPF donation norm for Flanders	Source effect
Walloon norm	SPF donation norm for Wallonia	Source effect

Four hypotheses organize the empirical analysis. **H1** is a norm-compliance hypothesis. A descriptive norm conveys information about how much others give (Cialdini et al., 1990), so its effect on an individual's donation depends on the relationship between the norm and the individual's prior belief. A participant who learns that others give more than expected has reason to adjust upward; a participant who learns that others give less than expected has reason to adjust downward. The directional prediction of norm compliance is therefore conditional: exposure to the norm increases giving for participants whose prior belief lies below the displayed value. Because administrative giving figures are substantially higher than typical lay estimates, the norm is expected to exceed prior beliefs for most participants in this sample, so that the average effect of norm exposure is predicted to be positive. Two caveats qualify this prediction. First, the prior is elicited for the population of Belgian residents while the displayed figure is a conditional average among taxpayers claiming the deduction, so the comparison between the two is informative about perceived discrepancy rather than a strict belief-updating exercise on a single quantity. Second, the prediction presumes that the norm operates as a generic signal that others give substantially, rather than as a numerical anchor for the €100 allocation, since the annual administrative amount has no direct mapping onto the one-shot task. Both caveats are taken up in the discussion of the belief channel. This assumption is verified directly in Section 5.7. **H2** is a differential-compliance hypothesis: in-group norms increase donation amounts more than out-group norms. **H3** is an out-group-resistance hypothesis: exposure to an out-group norm reduces donation amounts below the control baseline, consistent with a contrastive identity response derived from social identity theory (Tajfel and Turner, 1979) rather than simple non-compliance. In-group favoritism does not, however, automatically imply

[1] Language availability was included to reduce participation barriers. Language choice is not used as a treatment factor in the empirical analysis.

out-group hostility (Brewer, 1999), which makes H3 a distinct and stronger claim than H2. **H4** is a moderation hypothesis: the difference between in-group and out-group norm effects increases with regional identification strength.

Table 3.2: Hypotheses and tests

Hypothesis	Claim	Main test
H1	Norm increases giving where it exceeds prior beliefs	Model 1; Section 5.7
H2	In-group norm > out-group norm	Model 3, $\beta_2 > 0$
H3	Out-group norm reduces giving below control	Model 3, $\beta_1 < 0$
H4	Identification widens in/out-group gap	Model 4, $\beta_5 > 0$

The central contribution of the design lies in H3. H2 asks whether in-group and out-group norms differ. H3 asks a stronger question: whether an out-group norm reduces giving below the no-norm baseline. This distinction is possible because the experiment includes a control condition as well as two regionally sourced norm conditions.

The rest of this chapter describes the norm stimulus and its empirical grounding (section 3.2), the participant flow (section 3.3), the sample and recruitment strategy (section 3.4), the variable construction (section 3.5), and the identification scope and design constraints (section 3.7).

3.2 Norm stimulus

The norm stimulus is drawn from administrative data. Participants in the treated conditions are shown the average annual charitable donation declared to SPF Finances, the Belgian federal tax authority, by residents of either Flanders or Wallonia. Using administrative data gives the stimulus an authoritative and verifiable source.

SPF Finances records donations that taxpayers formally declare for tax deduction purposes. In Belgium, charitable donations are deductible above a minimum threshold of €40 per year (*Dons | SPF Finances* 2023). The figures therefore measure declared giving among taxpayers who claim the deduction, not all charitable giving in the population.[2] This is a selected group, likely older, higher income, and more administratively engaged than the average resident.

The average declared donation is €309 in Flanders and €301 in Wallonia. The two euro values differ by only €8, which reduces the risk that treatment differences are driven by the absolute level of the norm. Participants are also shown the figure as a percentage of average regional income. Using Statbel income data, the donations represent approximately 1.42% of average Flemish income and 1.57% of average Walloon income (Statbel (Statistics Belgium), 2025). This contextualizes the euro amounts but does not fully hold perceived norm intensity constant.

The norm is presented as ambient social information, not as corrective feedback relative to the participant's prior belief. Prior beliefs are collected before treatment exposure and do not determine which norm is displayed. A final constraint is that the norm is an annual administrative

[2]The administrative figures exclude informal giving, donations below the deductible threshold, and eligible donations that taxpayers do not declare.

amount, while the outcome is a one-shot allocation of a hypothetical €100 endowment. The stimulus should therefore be interpreted as a signal about regional giving behavior rather than as a direct benchmark for the experimental donation amount.

3.3 Participant flow

The experiment follows a fixed sequence: consent, pre-treatment measures, treatment display, donation decision, post-decision checks, and debrief. Participants cannot navigate backwards between screens. This ordering ensures that region, identity strength, out-group closeness, prior beliefs, and demographic controls are measured before treatment exposure.

The pre-treatment block first records region of residence and province. Region is used to construct in-group and out-group exposure after random assignment. Participants then report regional identification strength and closeness to the other region on 0–10 scales. These variables are measured before treatment because exposure to a Flemish or Walloon norm could itself affect reported identification or perceived closeness.

Participants also report their prior belief about average annual charitable giving in Belgium. The belief question is framed nationally rather than regionally to avoid making the Flemish-Walloon comparison salient before treatment. The pre-treatment block also includes an instructed-response attention check (Oppenheimer et al., 2009; Berinsky et al., 2013) and three filler questions about charitable giving. The filler items create a short buffer between the prior-belief question and the treatment display. They are not used as analytical variables.

Participants are then randomly assigned to the control condition, the Flemish norm condition, or the Walloon norm condition. Treated participants see the regional donation norm and its administrative source in a highlighted information box. Control participants see neutral filler text about charitable organizations before the same donation decision. Participants then complete the main donation task, allocating a hypothetical €100 endowment between themselves and a recognized charity.

After the donation decision, participants complete manipulation and quality checks. Treated participants report the norm figure they saw and rate its credibility on a 0–10 scale. Participants also report whether they looked up external information during the experiment. The survey ends with a debrief shown after all outcome and post-decision measures have been collected.

3.4 Sample and recruitment

The target sample was determined by both statistical power considerations and recruitment constraints. There is no directly comparable experiment combining the same Belgian regional setting, online delivery, between-subjects norm-versus-control structure, and continuous donation outcome. The sample-size target was therefore calibrated against the small-to-moderate effects reported in the charitable-giving norm literature, especially the norm-versus-control estimates reported by Chapman et al. (2025). This benchmark is used to set realistic expectations rather than to assume a large treatment effect.

An a priori power calculation conducted in G*Power (Faul et al., 2007) indicates that detecting an effect of approximately 0.20 standard deviations with 80% power and $\alpha = 0.05$ would require a substantially larger sample than was feasible for this thesis; the calculation is reported in Appendix B.2.[3] The in-group versus out-group comparison and the identity-moderation specification require even larger samples because they split the treated sample and estimate interaction effects. The experiment is therefore interpreted as a pilot. Its purpose is to provide preliminary evidence, assess whether the norm stimulus is credible in a Belgian online sample, and inform the design of a larger follow-up study. Null results, especially in interaction models, should therefore be interpreted in light of limited statistical power (Gelman and Carlin, 2014).

Participants were recruited online through personal networks, university networks, and social media, in French, Dutch, and English. The target population is Belgian adults, with eligibility based on self-reported region of residence. Because recruitment relies partly on personal and professional networks, the sample is not expected to be representative of the Belgian adult population. Regional balance between Flanders and Wallonia is also unlikely to be fully achieved in practice. This does not threaten internal validity, since treatment assignment occurs within the recruited sample, but it limits precision in regional subgroup analyses and external validity.

The primary analysis sample is defined using exclusion criteria fixed before analysis. The hypotheses, exclusion rules, and primary specifications were fixed in the thesis proposal before data collection but not deposited in a public registry; "pre-specified" is used in this thesis in that sense. Incomplete responses are excluded, as are participants who fail the embedded attention check or report looking up external information during the experiment. These exclusions remove responses that weaken the interpretation of the treatment contrast. The realized sample, exclusion flow, and regional composition are reported in Chapter 5.

3.5 Variables

The experiment collects thirty variables, organized into six groups. The full variable list, including definitions, scales, and coding rules, is provided in Appendix B. This section summarizes the variables used in the main analysis.

The primary outcome is the donation amount. It is recorded as an integer between 0 and 100 and represents the number of euros the participant allocates to charity out of the hypothetical endowment.

The treatment structure is captured by a set of assignment and derived treatment variables. Models 3 and 4 use a compact identity-source parameterization based on *treat_norm* and *ingroup_norm_recipient*. The first identifies any norm exposure; the second equals one for treated participants whose norm source matches their region of residence and zero otherwise. Together, these variables identify the out-group norm effect and the in-group premium. A direct *outgroup_norm_recipient* indicator is also constructed for robustness tables that display the in-group and out-group coefficients separately. These derived variables are constructed after randomization by crossing the assigned norm source with the participant's self-reported region.

[3]The calculation uses a two-sided independent-means test with equal group sizes.

Control participants serve as the omitted reference category in the main specifications.

The pre-treatment battery includes the main moderators and belief measures. Regional identification strength and closeness to the other region are measured on 0–10 scales. Prior belief about average annual charitable giving in Belgium is recorded as a self-reported euro amount. These variables are collected before treatment exposure to avoid post-treatment contamination.

Standard demographic controls are also collected before treatment. These include region of residence, province, gender, age group, education level, and income bracket. Region of residence has a special role because it is both a descriptive covariate and the basis for the in-group and out-group classification.

Two variables define the main exclusion criteria. The first identifies participants who fail the embedded attention check. The second identifies participants who report having searched for external information during the experiment.

Post-decision variables are used for manipulation checks and descriptive mechanism analysis. Norm recall is recorded using a bracket format, from which a binary recall accuracy indicator is derived. Treated participants also rate the credibility of the norm and report the direction of their surprise relative to prior beliefs on a signed five-point scale. Page time on the treatment display screen is recorded automatically and used as a complementary attention indicator.

3.6 Identification strategy

Random assignment ensures that treatment status is independent of pre-treatment individual characteristics in expectation. The design identifies three main quantities: the effect of exposure to a regional descriptive norm, the in-group premium relative to the out-group norm response, and the heterogeneity of this premium by regional identification strength.

The first quantity is the intent-to-treat effect of norm exposure on donation amounts. It captures the effect of being randomly assigned to receive a regional norm, regardless of whether the participant fully reads, remembers, or believes the information. Some treated participants may not process the stimulus carefully, and some may find it insufficiently credible. These participants remain part of the treatment group. The intent-to-treat estimate therefore corresponds to the effect of assignment to norm information, not the effect of full compliance with the information. Low attention or low credibility would tend to attenuate the estimated effect relative to the effect among participants who actually process and believe the stimulus.

The second quantity is the in-group premium relative to the out-group norm response. In-group and out-group status is constructed by crossing the randomly assigned norm source with the participant's region of residence, which is measured before treatment. Within a given region, whether a participant receives a norm from their own region or from another region is determined by random assignment. This is the identifying variation behind the in-group premium. The interpretation is strongest in specifications that restrict the sample to the Walloon subsample, where the in-group/out-group contrast is identified by random assignment within a single region and baseline regional differences cannot confound the comparison.

This comparison is intended to capture the role of norm source identity. However, it should not be interpreted as a perfectly pure source-label effect. The Flemish and Walloon euro amounts are almost identical, which reduces the risk that responses are driven by differences in the absolute level of the norm. Yet participants also see the norm expressed as a percentage of average regional income, and the Walloon percentage is slightly higher. The in-group versus out-group estimate should therefore be interpreted as the effect of exposure to a regionally sourced descriptive norm whose euro value is nearly constant, rather than as the effect of source identity holding every aspect of the message perfectly fixed.

The third quantity is the moderation of the in-group premium by regional identification strength. Identification strength is measured before treatment exposure, so it cannot itself be affected by the norm stimulus. The moderation analysis asks whether participants with stronger pre-existing regional identification show a larger in-group premium. Because identification strength is not randomly assigned, the level of identification should not be interpreted as exogenous. However, treatment assignment is random within levels of identification, which allows the interaction term to be interpreted as a heterogeneous treatment effect by a pre-treatment characteristic. Measurement error in self-reported identification would likely attenuate the moderation term toward zero, making heterogeneous treatment effects harder to detect.

The design does not identify all possible mechanisms behind treatment effects. A significant effect of norm exposure would be consistent with several channels, including information receipt, identity-based conformity, social desirability, or contrastive responses to an out-group source. Prior beliefs, reported surprise, recall, and credibility measures help interpret which channels are most plausible, but they are not independently randomized. They are therefore used as descriptive mechanism evidence rather than as causal mediation estimates.

The design also does not identify the counterfactual effect of the same numerical norm attributed to a neutral, non-regional source. Since the treatment arms vary both the regional source and the associated administrative figure, isolating a pure source effect relies on the near equality of the Flemish and Walloon euro values as an approximation. This approximation is useful but imperfect. The empirical strategy therefore reports specifications that separate Flemish and Walloon norm arms, in-group and out-group exposure, and sensitivity checks related to norm magnitude and regional composition.

3.7 Constraints in identification strategy

Several design constraints qualify the interpretation of the estimates. They do not invalidate the randomized comparison, but they define what the experiment can and cannot show.

First, the donation decision is hypothetical. This limits the interpretation of absolute donation levels and external validity, even though all randomized groups face the same task.

Second, the norm stimulus and outcome use different units. The stimulus is an annual administrative donation amount, while the outcome is a one-shot allocation from a hypothetical €100 endowment. The treatment is therefore better understood as a signal about regional giving behavior than as a direct numerical benchmark for the donation task.

Third, the SPF Finances data capture declared tax-deductible donations, not all charitable giving. The displayed figures reflect a selected population of taxpayers who claim the deduction and may overstate typical giving among Belgian residents. This matters for the interpretation of prior beliefs, credibility, and surprise-direction analyses, but the main intent-to-treat comparison still rests on randomized exposure to the same type of administrative stimulus.

Fourth, the in-group and out-group analysis is constrained by regional composition. Brussels participants are excluded from the primary identity specifications because no Brussels-sourced norm is presented, making their out-group status a design artifact rather than a clean identity contrast. If the Flemish subsample is too small for precise within-region inference, the primary in-group and out-group specifications are estimated on the Walloon subsample, with pooled and Brussels-inclusion specifications reported as robustness checks.

Fifth, the donation slider initializes at the midpoint of the scale, at 50 out of 100. Salient defaults are known to anchor allocation decisions (Johnson and D. Goldstein, 2003), including charitable donations specifically (Altmann et al., 2019; Goswami and Urminsky, 2016), and a midpoint default may compress responses toward 50 in every arm. Because the default is identical across conditions it does not threaten the internal validity of the treatment contrasts, but it may attenuate measured treatment effects. Section 5.10 reports the share of responses left exactly at the default and a robustness check excluding them, and section 6.4 discusses the implications for interpreting the null.

Sixth, the control and treatment screens do not differ only in normative content. Control participants see neutral filler text, while treated participants see the regional norm in a highlighted information box. The design therefore cannot fully separate the effect of receiving highlighted additional information from the effect of the regional norm itself. The in-group and out-group comparison remains cleaner on this dimension, since both treated arms share the same box format and differ primarily in the regional source.

A further limitation is that experimenter demand effects cannot be fully ruled out. The national-level belief question, the filler items, and the non-evaluative presentation of the norm reduce this concern, but participants may still infer that the study concerns charitable giving, regional identity, or social information (Zizzo, 2010; Quidt et al., 2018).

The online convenience sample also limits external validity. The results should therefore be interpreted as evidence from the recruited sample rather than as population-representative estimates for Belgian adults (Henrich et al., 2010).

Chapter 4 – Empirical Strategy

4.1 Identification logic

The empirical strategy follows directly from the experimental design described in chapter 3. Random assignment identifies the causal effect of exposure to a regional descriptive norm. In-group and out-group exposure are constructed by interacting assigned norm source with pre-treatment region of residence. The main estimands are therefore the pooled norm effect, the separate Flemish and Walloon norm effects, the out-group effect, the in-group premium, and moderation by regional identification strength.

Region of residence is a key pre-treatment covariate and enters the pooled specifications (Models 1 and 2) as a control, absorbing baseline differences between Flemish and Walloon participants. Models 3 and 4 are estimated on the Walloon subsample, where region is constant and therefore drops out. The treatment-effect estimates do not rely on these controls for consistency, since assignment is random; the controls serve to absorb residual variation and improve precision.

Table 4.1: Confirmatory vs exploratory analyses

Status	Analyses
Primary	Model 1, Model 3
Secondary	Model 4
Exploratory	Belief gap/underestimation, quantile regressions, credibility heterogeneity
Robustness	Attrition, Brussels inclusion, exclusion rules

4.2 Primary specifications

4.2.1 Model 1: Baseline pooled norm OLS

$$D_i = \alpha + \beta_1 \text{Norm}_i + \beta_2 X_i + \varepsilon_i \tag{4.1}$$

The first specification tests whether exposure to a regional descriptive norm affects donation amounts on average. D_i is the donation amount of participant i, measured from 0 to 100. Norm_i is a binary indicator equal to one if participant i receives either the Flemish or Walloon norm, and zero if the participant is assigned to the control condition. X_i is a vector of pre-treatment demographic controls, including age group, gender, education, and income bracket.

The coefficient β_1 identifies the pooled intent-to-treat effect of norm exposure relative to the control condition. A positive estimate would be consistent with norm compliance, given that the displayed norm exceeds prior beliefs for the large majority of participants (Section 5.7), while a null estimate would indicate that exposure to regional norm information does not shift donations on average in this sample.

4.2.2 Model 2: Disaggregated treatment arms

$$D_i = \alpha + \beta_1 \text{NormFL}_i + \beta_2 \text{NormWA}_i + \beta_3 X_i + \varepsilon_i \tag{4.2}$$

The second specification separates the two treatment arms. NormFL_i is equal to one for participants assigned to the Flemish norm, and NormWA_i is equal to one for participants assigned to the Walloon norm. The control group is the omitted category.

The coefficients β_1 and β_2 identify the average effects of the Flemish and Walloon norms relative to the control condition. This specification tests whether the two norm sources generate different average responses before introducing the participant's own regional identity. Since the two euro values are nearly identical, large differences between β_1 and β_2 would suggest that the regional source or its perceived meaning matters. A formal test of $\beta_1 = \beta_2$ is reported alongside the estimates.

4.2.3 Model 3: In-group and out-group norm effects

The third specification estimates the out-group norm effect and the in-group premium, with the control group as the omitted category:

$$D_i = \alpha + \beta_1 \text{Norm}_i + \beta_2 (\text{Norm}_i \times \text{Ingroup}_i) + \beta_3 X_i + \varepsilon_i \tag{4.3}$$

Norm_i equals one for treated participants and zero for control participants. Ingroup_i equals one when the assigned norm source matches the participant's region of residence. The interaction $\text{Norm}_i \times \text{Ingroup}_i$ therefore equals one only for in-group norm recipients. The model is estimated on the Walloon subsample, where random assignment to the Walloon (in-group) or Flemish (out-group) norm identifies the contrast with adequate cell sizes. Brussels participants and the small Flemish subsample are excluded from the primary specification for the reasons set out in section 3.7. The pooled Flanders + Wallonia specification and the Brussels-inclusion specification are reported as sensitivity checks in section 4.5.

Each coefficient maps directly onto a substantive quantity. β_1 is the out-group norm effect relative to control. β_2 is the in-group premium relative to the out-group norm effect. The in-group norm effect relative to control is $\beta_1 + \beta_2$. The four hypotheses are tested as follows. H1 predicts a positive average norm effect because the displayed norm exceeds prior beliefs for almost all participants (section 5.7); it is assessed by the pooled coefficient in Model 1 alongside the belief-channel evidence. H2 (differential compliance) is tested by $\beta_2 > 0$. H3 (out-group resistance) is tested by $\beta_1 < 0$ as a one-sided test on a single coefficient. A negative β_1 indicates that out-group norm exposure reduces donations below the no-norm baseline, which is the contrastive identity response posited by H3.

This parameterization is useful because differential compliance is read directly from the interaction term, while out-group resistance is read directly from the out-group norm coefficient β_1.

4.2.4 Model 4: Moderation by identification strength

The fourth specification introduces regional identification strength as a continuous moderator, centered at the sample mean. To ensure that the moderation coefficient isolates identity-based heterogeneity in the in-group premium, the specification includes all lower-order terms of the triple interaction (Brambor et al., 2006): the main effect of identification, the interaction of identification with norm exposure, and the in-group interaction itself. Because in-group status is defined only for treated participants, the in-group main effect coincides with the norm-by-in-group interaction, and the identification-by-in-group term coincides with the triple interaction; the specification below is therefore the full hierarchy that the design admits.

$$\begin{aligned} D_i = \alpha &+ \beta_1 \text{Norm}_i + \beta_2(\text{Norm}_i \times \text{Ingroup}_i) \\ &+ \beta_3 \text{ID}_i + \beta_4(\text{Norm}_i \times \text{ID}_i) \\ &+ \beta_5(\text{Norm}_i \times \text{Ingroup}_i \times \text{ID}_i) + \beta_6 X_i + \varepsilon_i \end{aligned} \tag{4.4}$$

ID_i is the regional identification strength score (0–10), centered at the sample mean. With this centering, β_1 is the out-group norm effect for a participant with average identification, and β_2 is the in-group premium at average identification. β_3 captures the association between identification and giving among control participants; it is descriptive rather than causal, since identification is not randomly assigned. β_4 captures whether identification moderates the response to the out-group norm. The coefficient of interest is β_5, which captures whether regional identification widens the in-group premium relative to the out-group norm response.

H4 (identity moderation) is tested by $\beta_5 > 0$. Omitting the lower-order terms would force the triple interaction to absorb any moderation of the general norm response by identification and any direct association between identification and giving, conflating identity-source moderation with these channels.

Measurement error in self-reported identification would likely attenuate β_5 toward zero, making heterogeneous treatment effects harder to detect. The interaction coefficient is therefore interpreted cautiously.

4.2.5 Note on controls

In all specifications, X_i includes gender, age group, education level, and income bracket, encoded as indicator variables. These pre-treatment controls are included to improve precision. Because treatment is randomly assigned, the treatment effect estimates do not rely on these controls for consistency.

Models 1 and 2 include region as a control because they are estimated on the full analysis sample. Models 3 and 4 are estimated on the Walloon subsample and therefore do not include region as a control, since region is constant within the estimation sample. In the pooled Flanders + Wallonia sensitivity specification reported in section 4.5, region is included as a control to absorb baseline differences between Flemish and Walloon participants.

Prior beliefs are included in secondary specifications that examine informational surprise. They

are not treated as a complete control for the informational channel. Instead, they help assess whether donation responses vary with the signed gap between the displayed norm and the participant's baseline expectation.

4.3 Secondary specifications: informational surprise

To examine the role of informational surprise, I construct two belief-channel measures. The signed gap is the displayed norm minus the participant's prior belief about average charitable giving in Belgium, so positive values indicate that the norm exceeded the prior. The underestimation indicator equals one when the participant's prior belief lies below the displayed norm.

The secondary specifications regress donation on these belief-channel measures:

$$D_i = \alpha + \beta_1 \text{Gap}_i + \beta_2 X_i + \varepsilon_i, \quad \text{treated sample}, \tag{4.5}$$

$$D_i = \alpha + \beta_1 \text{Underestimate}_i + \beta_2 X_i + \varepsilon_i, \quad \text{treated sample}, \tag{4.6}$$

$$D_i = \alpha + \beta_1 \text{Norm}_i + \beta_2 (\text{Norm}_i \times \text{Underestimate}_i) + \beta_3 X_i + \varepsilon_i. \tag{4.7}$$

The first two specifications ask whether donation amounts among treated participants vary with the size or sign of the informational surprise. The third specification uses the full analysis sample and tests whether norm exposure has a different association with donation for participants whose prior belief lay below the displayed norm. These analyses are interpreted as descriptive evidence on the informational channel. They do not fully isolate the informational mechanism, because prior beliefs are not experimentally manipulated and may be correlated with other participant characteristics.

4.4 Quantile regression

OLS estimates average treatment effects, but norm effects may not be uniform across the donation distribution. Participants who would otherwise donate little, moderate amounts, or almost the full endowment may respond differently to the same normative information. I therefore estimate quantile regressions (Koenker and Bassett, 1978) to recover treatment effects at different points of the conditional donation distribution. This allows me to detect heterogeneous responses that may be hidden by average effects.

The theoretical motivation follows directly from the structure of the experiment. Low donors may be nudged upward if the norm provides a salient benchmark for higher giving. Medium donors may be less responsive if their preferred donation is already close to what the norm suggests. High donors may respond differently again, including possible downward adjustment if the norm makes lower giving appear socially acceptable. This pattern is related to the boomerang effect documented by Schultz et al. (2007), where descriptive norms can move behavior in different

directions depending on the individual's baseline position.

The out-group resistance mechanism may operate differently. If the response is driven by source identity rather than numerical anchoring, there is no clear theoretical reason to expect it to be concentrated only among low, medium, or high donors. A low donor may reduce giving further after exposure to an out-group norm, but a high donor may also react negatively if the norm source is perceived as identity-relevant. Quantile regression is therefore useful because it lets me examine where in the donation distribution out-group responses are strongest without imposing that structure in advance.

I estimate quantile regression equivalents of the in-group-premium specification and the moderation specification at the 25th, 50th, and 75th percentiles, using bootstrap standard errors. The specifications follow the OLS models presented in section 4.2. The difference is that each coefficient is estimated separately at quantile τ, rather than as an average effect over the full conditional distribution.

Model 3:

$$Q_\tau(D_i \mid \mathcal{Z}_i) = \alpha_\tau + \beta_1(\tau)\text{Norm}_i + \beta_2(\tau)(\text{Norm}_i \times \text{Ingroup}_i) + \beta_3(\tau)X_i + \varepsilon_i \tag{4.8}$$

Model 4:

$$\begin{aligned} Q_\tau(D_i \mid \mathcal{Z}_i) = \alpha_\tau &+ \beta_1(\tau)\text{Norm}_i + \beta_2(\tau)(\text{Norm}_i \times \text{Ingroup}_i) \\ &+ \beta_3(\tau)\text{ID}_i + \beta_4(\tau)(\text{Norm}_i \times \text{ID}_i) \\ &+ \beta_5(\tau)(\text{Norm}_i \times \text{Ingroup}_i \times \text{ID}_i) + \beta_6(\tau)X_i + \varepsilon_i \end{aligned} \tag{4.9}$$

where $Q_\tau(D_i \mid \mathcal{Z}_i)$ denotes the τ-th conditional quantile of the donation distribution, with $\mathcal{Z}_i$ collecting the right-hand-side variables in the corresponding specification and $\tau \in \{0.25, 0.50, 0.75\}$. All variables are defined as in section 4.2. The quantile estimates are not interpreted as substitutes for the OLS estimates, but as complementary evidence on distributional heterogeneity.

For each quantile, I focus on the out-group coefficient and the in-group premium. $\beta_1(\tau)$ captures the out-group norm effect at quantile τ. A negative estimate is consistent with out-group resistance at that point of the conditional donation distribution. $\beta_2(\tau)$ captures the in-group premium relative to the out-group norm at that quantile. In Model 4, $\beta_5(\tau)$ captures whether regional identification strength amplifies the in-group premium at that point of the distribution.

The quantile analysis is exploratory. The main theoretical prediction is that the in-group premium is positive, and that this premium increases with regional identification strength. The theory does not precisely predict whether these effects should be strongest among low, median, or high donors. Comparing coefficients across the 25th, 50th, and 75th percentiles therefore helps identify whether the average treatment effects mask stronger responses in particular parts of the donation distribution.

4.5 Falsification and robustness

The following diagnostic specifications use direct in-group and out-group treatment indicators. This notation is used for robustness reporting because these checks display the two treatment arms separately, while the primary Models 3 and 4 report the equivalent out-group coefficient and in-group premium.

4.5.1 Attrition analysis

Differential dropout across treatment arms could introduce selection into the analysis sample. I therefore test whether completion rates differ systematically by treatment condition. The dependent variable is a binary completion indicator equal to one if the participant completed the full experiment and zero otherwise:

$$C_i = \alpha + \beta_1 \text{IngroupNorm}_i + \beta_2 \text{OutgroupNorm}_i + \beta_3 X_i + \varepsilon_i \tag{4.10}$$

Under non-differential attrition, β_1 and β_2 should be close to zero. If completion differs by treatment status, the results section reports the extent of the imbalance and discusses the implications for the analysis sample.

4.5.2 Placebo norm source

I use a placebo classification test to assess whether the in-group and out-group coding mechanically captures pre-existing differences in generosity. The test is conducted among control participants only, who receive no regional norm information. A fictitious norm source is randomly assigned to each control participant, and a pseudo in-group indicator is constructed by crossing this placebo source with the participant's region of residence:

$$D_i = \alpha + \beta_1 \text{PseudoIngroup}_i + \beta_2 X_i + \varepsilon_i \tag{4.11}$$

Since control participants receive no actual norm information, PseudoIngroup_i should not predict donation amounts. A statistically significant β_1 would suggest that the in-group classification is correlated with pre-existing differences in generosity, which would weaken the interpretation of Model 3.

4.5.3 Credibility robustness check

A possible concern with the in-group and out-group comparison is that participants may perceive one regional norm as more credible than the other. The credibility check augments the in-group and out-group specification with the post-decision credibility rating and its interactions with both treatment indicators:

$$
\begin{aligned}
D_i = \alpha &+ \beta_1 \text{IngroupNorm}_i + \beta_2 \text{OutgroupNorm}_i \\
&+ \beta_3 \text{Credibility}_i + \beta_4 (\text{IngroupNorm}_i \times \text{Credibility}_i) \\
&+ \beta_5 (\text{OutgroupNorm}_i \times \text{Credibility}_i) + \beta_6 X_i + \varepsilon_i
\end{aligned}
\tag{4.12}
$$

The diagnostic of interest is the linear restriction $\beta_4 = \beta_5$, which tests whether credibility moderates the in-group versus out-group differential. Because credibility is measured after treatment, this check is interpreted as supplementary evidence rather than as a causal moderation estimate (Montgomery et al., 2018).

4.5.4 Brussels inclusion robustness check

The primary in-group and out-group specifications exclude Brussels participants for the reasons set out in section 3.7. As a robustness check, I re-estimate Models 3 and 4 including Brussels participants coded as out-group recipients, since no Brussels-sourced norm is presented. Stability across this specification indicates that the exclusion choice does not materially affect the conclusions.

Chapter 5 – Results

This chapter presents the empirical results of the experiment. It documents the realized sample, randomization balance, attrition, manipulation checks, the four main models, and the robustness specifications that anchor the interpretation. The central finding is that exposure to a regional descriptive norm does not detectably shift donation amounts in this sample, at the precision afforded by the design. This holds regardless of whether the norm originates from the participant's own region or another, and regardless of the strength of regional identification. The treatment was received and recognized as discrepant from priors in the expected direction, but this informational response was not accompanied by detectable changes in donation behavior.

5.1 Sample

The raw export contained 501 rows.[4] Before defining the response-flow sample, 76 unused oTree/session placeholder observations created after participant ID 425 were removed. These rows were not treated as launched survey responses and are therefore excluded from the attrition and exclusion counts.

The launched-response sample therefore consists of 425 observations. Of these, 103 were excluded according to the pre-specified data-quality rules: 86 because the survey was incomplete, 9 because the lookup-information item was missing, 5 because the attention check was failed, and 3 because the participant reported looking up external information during the survey. The final analysis sample contains 322 participants.

Table 5.1: Analysis sample by region of residence and treatment arm

Region	Control	Flemish norm	Walloon norm	Total
Flanders	9	9	4	22
Wallonia	93	80	93	266
Brussels	13	10	11	34
Total	115	99	108	322

The sample is dominated by Walloon residents (266 of 322). The Flemish cells are too small for within-region inference, which is why the in-group and out-group contrast is identified on the Walloon subsample, where random assignment to the Walloon (in-group) or Flemish (out-group) norm produces adequate cell sizes.

5.2 Descriptive statistics

The donation outcome has mean 42.5 euros and standard deviation 29.2 across the analysis sample, with median 41.5. 4.7% of participants donate zero and 9.0% donate the full hundred-euro

[4] oTree can create session rows that are not completed survey launches, for example when participant links are generated but not used. These rows are removed before computing attrition and exclusion counts.

endowment. The distribution is approximately symmetric rather than the strongly right-skewed pattern typical of dictator-game data.

One candidate explanation for this shape is the response interface: the donation slider initializes at 50, and 21.1% of analysis-sample participants left it exactly at the default. The implications of default anchoring are taken up in sections 5.10 and 6.4.

A descriptive norm could compress the distribution of giving even without shifting its mean, by pulling donations toward the displayed value. The data show no such compression. The standard deviation of donations is 28.8 in the control group and 29.5 in the norm group. A Brown–Forsythe test does not reject equality of variances across the two groups ($W = 0.17$, $p = 0.68$), and a two-sample Kolmogorov–Smirnov test does not reject equality of the full distributions ($D = 0.08$, $p = 0.77$). The norm did not move the location or the spread of giving. One caveat applies to the variance comparison: if the midpoint default compressed responses toward 50 in every arm, a treatment-induced compression could be masked, so the absence of a variance difference is most informative under the assumption that default anchoring operated similarly across conditions.

Differences in raw means across treatment arms are small in the Walloon identity-comparison sample. Control participants donate 41.7 euros on average ($N = 93$), in-group recipients donate 43.1 euros ($N = 93$), and out-group recipients donate 44.2 euros ($N = 80$). The out-group mean is the largest of the three, the opposite direction from the contrastive prediction of H3. Figure 5.1 shows the corresponding empirical distributions by arm.

Regional identification has mean 6.6 out of 10 (SD 2.3) in the analysis sample, indicating that participants are meaningfully attached to their region. Closeness to the other Belgian region averages 3.7. Prior beliefs about average annual charitable giving in Belgium are recorded with a mean of 110 euros across the analysis sample after outlier cleaning, far below the administrative figure of around 300 euros displayed in the norm conditions. The norm is therefore informationally meaningful for most participants: they do not already know the value being displayed.

Prior beliefs outside the 0–1000 euro range were treated as out-of-range and excluded from belief-channel analyses. This rule affects six analysis-sample observations, including four treated observations.

Whether the small raw differences across arms survive controlled estimation is examined in the remainder of the chapter.

5.3 Randomization and attrition

Randomization assigned participants to the control, Flemish norm, or Walloon norm condition with equal probability; balance is assessed with region included as a control given its central role in the design. A balance regression of the treatment indicator on pre-treatment covariates, with region fixed effects, fails to reject joint independence ($F(15, 298) = 0.96$, $p = 0.50$). Within the Walloon sample, where the in-group and out-group contrast is identified, a balance regression of the in-group indicator on the same covariates also fails to reject ($F(15, 154) = 0.59$, $p = 0.88$). Sparse demographic categories with fewer than five observations are collapsed for these joint

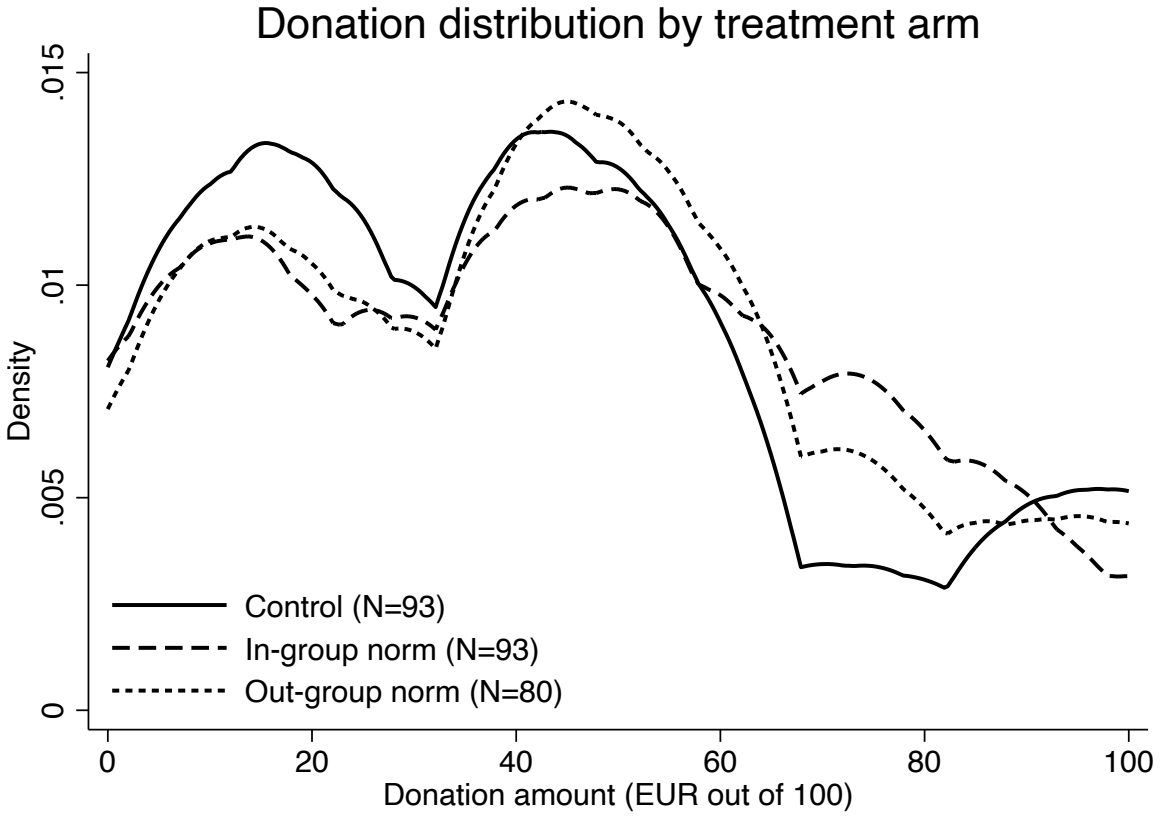

Figure 5.1: Distribution of donation amounts by treatment arm. Kernel-density estimates on the Walloon identity-comparison sample. The three distributions overlap substantially across the full support, consistent with the absence of a treatment shift documented in the regression analysis.

tests, since singleton cells can mechanically inflate the statistic (Bruhn and McKenzie, 2009).[5] The continuous pre-treatment covariates are similarly balanced across arms. Random assignment therefore produced balanced cells on observed characteristics.

Differential attrition is not detected. A regression of a binary completion indicator on the pooled treatment indicator returns $p = 0.19$, and the corresponding regression on the disaggregated Flemish and Walloon treatment indicators returns a joint $p = 0.42$. Participants exposed to the norm conditions did not drop out at statistically distinguishable rates from control participants. The behavioral null is therefore unlikely to be explained by differential selection into the analysis sample.

All 86 incomplete responses have a recorded donation decision in the raw export, so the available screen-level proxy places them after treatment display rather than before treatment exposure. Among these incomplete responses, 24 were assigned to control and 62 to a norm condition. This pattern is consistent with the attrition regressions above: dropout is not cleanly concentrated before treatment, but it is not statistically differential across arms. Because donations are recorded for the incomplete responses, a completion-free robustness check is feasible in principle;

[5]The diagnostic balance tests collapse the sparse "other" categories for gender and education into the reference categories. These collapsed variables are used only for balance diagnostics, not for outcome regressions.

it is left to the replication package.

5.4 Manipulation checks

The manipulation checks suggest that the behavioral null is unlikely to be driven by complete failure of exposure to the treatment. Among treated participants, 63% correctly recall the bracket containing the displayed norm value. This is below the 70% pre-specified benchmark but substantially above chance, and recall accuracy does not differ between Flemish and Walloon norm recipients. The median participant spent 36 seconds on the norm display page (mean 43 seconds), indicating substantial exposure to the stimulus.

Perceived credibility of the displayed norm has mean 5.12 on the 0–10 scale with standard deviation 2.19. A regression of credibility on the in-group indicator within the Walloon treated subsample returns $\beta = -0.14$ (SE 0.34, $p = 0.69$): participants do not find the Walloon norm more or less credible than the Flemish norm. Credibility is therefore neither concentrated nor differentially distributed across norm sources.

The surprise check provides the clearest evidence that the manipulation was received and recognized as discrepant from priors. A regression of the post-treatment surprise-direction measure on the signed prior-to-norm gap, on the treated sample, returns a coefficient of -0.0026 with $p < 0.001$ and $R^2 = 0.12$. Participants whose prior beliefs were below the displayed norm value reported the norm as larger than expected; those whose priors were above the displayed value reported it as smaller than expected. This measure is a self-reported direction of surprise rather than an elicited posterior belief; it establishes that the information was received and recognized as discrepant from priors, not that participants' quantitative beliefs about giving were durably revised. The behavioral null in Models 1 through 4 is therefore unlikely to reflect a wholesale failure of treatment processing in the analysis sample.

5.5 Models 1 and 2: pooled and disaggregated norm effects

Model 1 regresses donation on a pooled treatment indicator with demographic controls (region, gender, age group, education, income bracket) on the full analysis sample. The coefficient is +3.95 euros (SE 3.27, $p = 0.23$). The pooled norm effect is positive but statistically indistinguishable from zero. H1 is not supported. As noted in Section 3.7, the pooled coefficient bundles the effect of receiving highlighted additional information with the effect of the regional norm content, since the control screen shows unhighlighted filler text; the in-group versus out-group contrast in Models 3 and 4 is clean on this dimension.

Model 2 separates the Flemish and Walloon treatment arms with the same controls. The Flemish norm coefficient is +4.21 (SE 3.86, $p = 0.28$) and the Walloon norm coefficient is +3.72 (SE 3.85, $p = 0.34$). A Wald test of equality between the two coefficients yields $F(1, 302) = 0.01$ with $p = 0.90$. The two regional norm sources produce statistically indistinguishable average effects on donations, providing no evidence of a magnitude-driven difference between the Flemish and Walloon norms before the identity dimension is introduced. The identity dimension of the contrast is examined in the next specifications.

5.6 Models 3 and 4: identity contrasts and moderation

Models 3 and 4 are estimated on the Walloon analysis sample ($N = 266$), where random assignment to either the Walloon (in-group) or Flemish (out-group) norm identifies the contrast within a single region. Controls are gender, age group, education, and income bracket.

Model 3 returns an out-group norm coefficient of +3.85 euros (SE 4.42, $p = 0.39$) and an in-group premium of −0.87 euros (SE 4.44, $p = 0.84$). Three hypotheses can be assessed from this specification. The data therefore do not support **H2**: the in-group premium is not statistically distinguishable from zero. The out-group coefficient is positive, not negative.[6] The data therefore provide no evidence supporting **H3**. The 95% confidence interval on the out-group coefficient runs from approximately -€5 to +€13, which rules out large contrastive effects but is consistent with smaller effects of either sign.

Model 4 adds identification strength and its interactions following the full hierarchical specification of section 4.2.4. Identification strength is mean-centered, so that the in-group premium is interpreted at average identification rather than at a score of zero, which no participant approaches. The main effect of identification is +0.17 (SE 1.12, $p = 0.88$), the interaction of identification with norm exposure is −0.07 (SE 1.73, $p = 0.97$), and the triple interaction—the H4 test—returns −0.19 (SE 1.79, $p = 0.92$). The data therefore do not support **H4**: regional identification strength does not detectably widen the gap between in-group and out-group norm responses, and this conclusion holds in a specification in which the triple interaction is not contaminated by lower-order channels.

Table 5.2 reports the estimates from Models 1 through 4.

5.6.1 Minimum detectable effects

The null results in Models 3 and 4 should be interpreted in light of the precision of the realized design. At $\alpha = 0.05$ and 80% power, the minimum detectable effect is approximately 12.42 euros for the in-group versus out-group differential, 12.37 euros for the out-group versus control effect, and 11.92 euros for the in-group versus control effect.

For the pooled norm effect in Model 1, the corresponding minimum detectable effect is approximately 9.2 euros.

The design can therefore rule out moderately large contrastive effects, but not smaller forms of out-group resistance. The 95% confidence interval on the out-group coefficient rules out contrastive effects more negative than approximately −4.85 euros. Since H3 is a stronger claim than differential compliance, the result should be interpreted as evidence against large behavioral backlash in this sample rather than definitive evidence that out-group resistance is absent at all possible magnitudes.

The bound on contrastive effects can also be stated as a formal equivalence-style test (Lakens,

[6] H3 is directional, predicting a negative out-group norm effect. The point estimate is positive (+3.85), so it runs opposite to the predicted direction. A one-sided test in the hypothesized direction therefore cannot approach significance, and the relevant question is not whether the coefficient is significantly negative but whether the confidence interval admits a contrastive effect of meaningful size. The lower bound of approximately -€5 indicates that it does not.

Table 5.2: OLS regression results

	Model 1	Model 2	Model 3	Model 4
Norm exposure	+3.95		+3.85	+3.85
	(3.27)		(4.42)	(4.46)
Flemish norm		+4.21		
		(3.86)		
Walloon norm		+3.72		
		(3.85)		
Norm × in-group			−0.87	−0.87
			(4.44)	(4.46)
Identification				+0.17
				(1.12)
Norm × identification				−0.07
				(1.73)
Norm × in-group × identif.				−0.19
				(1.79)
Tests (p-value)				
$\beta_{\text{FL}} = \beta_{\text{WA}}$	—	0.90	—	—
In-group premium = 0 (H2)	—	—	0.84	—
Out-group coefficient = 0	—	—	0.39	—
Triple interaction = 0 (H4)	—	—	—	0.92
N	322	322	266	266

Notes: Robust standard errors in parentheses. Models 1 and 2 estimated on the full analysis sample with controls for region, gender, age group, education, and income bracket. Models 3 and 4 estimated on the Walloon analysis sample with controls for gender, age group, education, and income bracket. Identification strength is centered at the sample mean. Model 4 includes the full lower-order hierarchy for the triple interaction. In Models 3 and 4, Norm exposure is the out-group effect in the Walloon sample, and Norm × in-group is the in-group premium; the in-group effect relative to control is their sum.

2017). Because H3 is directional, the relevant question is whether out-group effects more negative than a smallest effect size of interest can be rejected. I set the smallest effect size of interest at 0.20 standard deviations of the donation outcome, approximately 5.8 euros, the small-effect benchmark also used in the a priori power calculation. A one-sided test of the null hypothesis that the out-group coefficient is less than or equal to -5.8 euros yields $t = 2.19$, $p = 0.015$. Contrastive out-group effects of 0.20 standard deviations or larger can therefore be rejected at the five percent level. This formalizes the reading of the confidence interval: the data do not merely fail to support H3; they actively reject backlash at or above the small-effect benchmark. The statement is one of directional non-inferiority rather than full equivalence. A further caution is that the rejection is partly a consequence of the positive point estimate: with the same standard error and a point estimate of zero, backlash at the smallest effect size of interest could not be rejected ($t \approx 1.31$). The non-inferiority result should therefore be read jointly with the confidence interval, as a bound that the realized data happen to support, rather than as a standalone finding robust to sampling variation in the sign of a noisy estimate. Because the point estimate is positive, the data remain consistent with positive effects of moderate size, and full two-sided equivalence to zero is neither established nor the relevant claim for H3.

5.7 The belief channel

Sections 5.5 and 5.6 document that the norm did not detectably move donations. This section asks whether the norm carried information, and whether the behavioral null survives once the informational content of the treatment is taken into account.

The treatment delivered the informational discrepancy that a norm-compliance hypothesis requires. The displayed norm exceeded the participant's prior belief for 190 of 203 treated participants, or 93.6%.[7] Prior beliefs averaged about €113 against a displayed norm of about €305 (€301 in Wallonia, €309 in Flanders). For almost the entire treated sample, the norm therefore conveyed that others give more than the participant expected. This matters for interpreting the null. If the sample contained a large share of participants who overestimated giving, a flat average effect could mask offsetting responses, with underestimators adjusting upward and overestimators downward. This is not the case here. With only thirteen participants overestimating, the behavioral null cannot be explained by such cancellation. The norm pointed in the same direction for nearly everyone, and donations still did not move. One caveat applies throughout this section: the prior refers to the population of residents while the displayed norm refers to declarers, so "underestimation" is measured across two different statistical objects. The implications of this for interpreting the surprise evidence are discussed in Section 6.3.

The belief-channel regressions confirm this directly. Table 5.3 reports three specifications. In the first, donation is regressed on the signed gap between the displayed norm and the prior belief, among treated participants. The coefficient is -0.001 ($p = 0.96$): the size of the informational surprise does not predict donation. In the second, donation is regressed on a dummy for having underestimated giving. The coefficient is $+1.88$ ($p = 0.73$): participants who learned that others give more than they thought did not give detectably more than the few who already overestimated. In the third, estimated on the full sample, norm exposure is interacted with the underestimation dummy. Neither the norm effect for the small overestimating group ($+0.36$, $p = 0.94$) nor the additional effect among underestimators ($+3.91$, $p = 0.42$) is distinguishable from zero.

The informational channel of the norm therefore operated while no behavioral response was detected. Participants registered the norm as differing from their prior beliefs in the expected direction, but neither the magnitude of the belief gap nor the fact of having underestimated giving translated into changed donations. These belief-conditioned tests are underpowered, and the small number of overestimators means the symmetric prediction that overestimators reduce giving cannot be tested. Read together with the average nulls of Models 1 through 4, however, the pattern is consistent across every margin examined: the norm was informative but did not shift behavior in this sample.

[7] The denominator is the 203 treated participants with a non-missing prior belief that survived outlier cleaning; four treated participants are dropped from the belief-channel analysis for a missing or out-of-range prior. The cleaning rule excludes prior beliefs outside the 0–1000 euro range. Underestimation is defined as prior_belief_clean $<$ norm_score.

Table 5.3: Belief channel: donation, belief gap, and underestimation

	(1) Treated	(2) Treated	(3) Full sample
Signed norm–prior gap	−0.001		
	(0.009)		
Underestimated (prior < norm)		1.88	
		(5.42)	
Norm exposure			0.36
			(4.94)
Norm × underestimated			3.91
			(4.82)
Controls	Yes	Yes	Yes
N	203	203	322

Robust standard errors in parentheses. Columns 1 and 2 are estimated on treated participants and include region, gender, age group, education, and income controls. Column 3 is estimated on the full analysis sample; norm exposure is interacted with an indicator for having underestimated giving, coded zero for control participants. No coefficient is statistically distinguishable from zero.

5.8 Quantile regression

Quantile regressions at the 25th, 50th, and 75th percentiles of the conditional donation distribution confirm that the OLS null is not concentrated at a particular point of the distribution. Bootstrap point estimates for the out-group coefficient and in-group premium are small and statistically null at every quantile, and tests of the in-group premium fail to reject (all $p > 0.50$). There is no evidence of a boomerang effect concentrated among high donors, in the sense of Schultz et al. (2007). The full quantile panel is reported in Appendix C.

5.9 Falsification

Two falsification tests provide no evidence that the behavioral null is mechanically generated by the coding strategy or by imbalance in pre-treatment beliefs. The placebo specification randomly assigns a fictitious norm source to Walloon control participants and constructs a placebo in-group indicator by crossing this pseudo-source with the participant's region of residence. A regression of donation on the placebo indicator on the Walloon control subsample returns a coefficient of −2.94 (SE 6.31, $p = 0.64$). There is therefore no evidence that the in-group/out-group coding is capturing pre-existing generosity differences among control participants.

The second test uses prior beliefs about average Belgian giving as a placebo outcome. Since prior beliefs are measured before treatment exposure, they cannot be affected by the norm. A regression of prior belief on the disaggregated treatment indicators returns $F(2, 296) = 0.93$ with $p = 0.39$. Treatment does not predict pre-measured beliefs, supporting the integrity of the prior elicitation.

5.10 Robustness

The behavioral null is stable across alternative specifications. Four standard robustness specifications of Model 3 are reported in Table 5.4. Removing demographic controls leaves the in-group

and out-group coefficients essentially unchanged. Pooling Flanders with Wallonia produces nearly identical point estimates with marginally larger standard errors. Coding Brussels participants as out-group recipients does not move the conclusion. Including Walloon participants who were originally excluded under the pre-specified rules produces the same null. Across these specifications, neither coefficient is statistically distinguishable from zero, and equality tests never reject.

A separate credibility robustness check examines whether perceived credibility moderates the in-group versus out-group differential among treated Walloon participants. Interacting the in-group indicator with credibility returns a Wald test of $F(1, 154) = 0.04$ ($p = 0.84$) on the interaction differential. This provides no evidence that the absence of an identity-source effect is explained by differential trust in the norm source.

A further robustness check addresses the midpoint default of the donation slider. 21.1% of analysis-sample participants recorded a donation of exactly 50, the value at which the slider initializes. Re-estimating Model 3 excluding these observations returns an out-group coefficient of +5.23 (SE 5.86) and an in-group premium of −1.57 (SE 5.83), substantively unchanged from the primary estimates. An extensive-margin regression of an indicator for moving the slider away from the default on the Model 3 treatment terms returns −0.03 for the out-group norm coefficient ($p = 0.70$) and +0.06 for the in-group premium ($p = 0.37$), providing no evidence that norm exposure affected whether participants engaged with the decision at all. The behavioral null is therefore not an artifact of default-stickiness, although the interface plausibly reduced the sensitivity of the outcome measure.

Accounting for the bounded outcome with a two-limit Tobit leaves the conclusions unchanged. The pooled norm coefficient is +3.86 (SE 3.66) in Model 1. In the Walloon Model 3 specification, the Tobit out-group coefficient is +3.76 (SE 4.97) and the in-group premium is −1.66 (SE 4.97).

Table 5.4: Direct-indicator robustness checks for Model 3

	(1) No controls	(2) Pooled F+W	(3) + Brussels	(4) Incl. excluded
In-group norm	+1.43	+1.82	+2.60	+2.30
	(4.34)	(4.03)	(3.92)	(3.76)
Out-group norm	+2.54	+3.54	+5.27	+2.81
	(4.47)	(4.27)	(3.88)	(4.01)
Equality test (p)	0.85	0.69	0.45	0.91
N	266	288	322	313

Notes: Robust standard errors in parentheses. This table reports the direct in-group and out-group coefficients for readability across sensitivity specifications. Column (1) estimates the Wallonia-primary specification without demographic controls. Column (2) pools Flanders and Wallonia with full controls. Column (3) uses the full analysis sample and codes Brussels participants as out-group recipients. Column (4) includes Walloon residents with non-missing outcome and controls regardless of exclusion status, reinstating observations excluded for incomplete survey status, missing or failed attention-check status, missing lookup status, or self-reported information lookup.

5.11 Heterogeneity

Two heterogeneity analyses complete the chapter. Income bracket midpoint is positively associated with donation amounts in the Walloon subsample at the 10% level ($p \approx 0.06$), but its interactions with the in-group and out-group indicators are statistically null. Richer participants donate marginally more in absolute terms but do not respond differently to norm source. A region-specific regression on the 22 Flemish participants is reported descriptively in Appendix D and is not used for inference; the small sample size precludes valid standard errors.[8]

5.12 Synthesis

This chapter reports a behavioral null in the present sample. Exposure to a regional descriptive norm does not detectably shift donation amounts. The pooled norm effect is positive but statistically indistinguishable from zero. The Flemish and Walloon norms produce indistinguishable effects. Within the Walloon subsample, the in-group premium is statistically null and the out-group coefficient is positive rather than negative. Identification strength does not moderate the in-group premium. The behavioral null is stable across robustness checks, falsification tests, quantiles of the donation distribution, and heterogeneity analyses.

The behavioral null is unlikely to reflect a complete failure of exposure, although recall and credibility indicate the first stage operated at less than full strength (Section 5.4). Section 5.4 documents that participants registered the norm as differing from their prior beliefs in the expected direction with high statistical significance: participants whose priors were below the displayed norm reported the norm as larger than expected. Section 5.7 documents that this registered surprise was not accompanied by detectable changes in donation. In this sample, the treatment had a detectable effect on stated surprise but not on donation behavior. This pattern, with detectable information receipt alongside a null on donation behavior, frames the discussion that follows.

The precision limits set out in Section 5.6.1 apply to all of these conclusions: the design provides evidence against moderately large contrastive effects while leaving smaller effects open. Their implications are discussed in Chapter 6.

[8] With only 22 Flemish participants, the region-specific regression is retained as a transparency check rather than as a basis for statistical inference.

Chapter 6 – Discussion

6.1 Interpreting the behavioral null

This experiment produced a behavioral null. Exposure to a regional descriptive norm did not detectably shift donation amounts, the in-group premium was statistically indistinguishable from zero, and identification strength did not moderate the premium. This null should be read against the precision of the design. The minimum detectable effects at conventional power are approximately €12 across the three contrasts of interest (section 5.6.1). The 95% confidence interval on the out-group coefficient runs from approximately -€5 to +€13. The design therefore rules out moderately large effects but is silent on smaller ones.

Two features of the results make the null more informative than a single average effect would be. The quantile estimates show that the null holds at the 25th, 50th, and 75th percentiles of the donation distribution, with no significant out-group coefficient or in-group premium at any quantile and no evidence of a boomerang pattern among high donors (section 5.8). The null is therefore not an averaging artifact concealing offsetting responses at the tails. The robustness specifications show that the null is also not specification-dependent. It holds without demographic controls, in the pooled Flanders and Wallonia sample, with Brussels coded as out-group, and when previously excluded participants are reinstated (section 5.10). Within the range of effects that the design can detect, the evidence is consistently null across the distribution and across reasonable specification choices.

The null is consistent with several explanations that the present data cannot distinguish. It is consistent with the meta-analytic finding that descriptive norm effects in charitable giving are modest and heterogeneous across designs (Chapman et al., 2025). It is consistent with the possibility that an institutional aggregate norm is not behaviorally relevant for a one-shot anonymous allocation, since the reference group of the norm is distant from the decision context. It is also consistent with the possibility that a true effect of plausible magnitude exists but falls below the precision of this study. These readings are not mutually exclusive, and the data do not separate them.

Experimenter demand deserves one further remark, concerning its likely direction. The norm box appears directly above the donation slider, and a participant inferring the study's purpose would most plausibly infer that generosity is expected after seeing an average of roughly 300 euros. Any demand pressure in this design therefore points upward. This has two implications. The null on average norm effects is, if anything, strengthened: donations failed to move despite a plausible upward demand component. Conversely, the small positive out-group point estimate could reflect demand rather than information, which is a further reason not to interpret its sign substantively (Zizzo, 2010; Quidt et al., 2018).

6.2 Norm source, identity, and out-group resistance

The design was constructed specifically to test for contrastive out-group effects, and this is where it makes its clearest contribution. Existing work compares compliance with in-group and out-group norms, but rarely includes a no-norm control that can separate a weak out-group response from an actively contrastive one. This experiment includes that control. It can therefore ask not only whether the out-group norm is followed less than the in-group norm, but whether it pushes donations below the no-norm baseline.

It does not. Within the Walloon sample, the out-group coefficient is positive and small (+3.85), and the in-group premium is statistically indistinguishable from zero (-0.87, $p = 0.84$). The disaggregated arms reduce the concern that this pattern is driven by displayed euro magnitude: the Flemish and Walloon norms produce statistically indistinguishable average effects ($p = 0.90$). The H3 prediction of contrastive resistance is not supported, and the point estimate runs in the opposite direction. The out-group coefficient is not merely insignificant. It is positive. The 95% confidence interval rules out contrastive effects larger than approximately €5 in absolute magnitude.

A formal one-sided equivalence test rejects contrastive effects at or beyond 0.20 standard deviations at the five percent level (section 5.6.1). As noted there, this rejection leans on the positive point estimate and should be read as a data-supported bound rather than a sampling-robust result.

This is more than a simple failure to reject. The design was built to test for out-group backlash in a setting where regional identity is meaningful. Identification in this sample is relatively high, with a mean of 6.6 on a 0–10 scale. The absence of contrast is therefore unlikely to be explained by a generally regionally indifferent sample. Even so, an out-group norm produced no hint of reduced giving relative to no norm. This is consistent with the established distinction between in-group favoritism and out-group hostility (Brewer, 1999; Weisel and Böhm, 2015). The present data provide no indication that an out-group norm depresses giving, even in a setting where regional identity is meaningful. The result places a bound on contrastive resistance, and the bound does not include the large backlash that H3 anticipated.

An important qualification is that the test is one-directional. The contrast is identified on the Walloon subsample, so the evidence concerns Walloon participants exposed to a Flemish norm. Regional identity in Belgium is asymmetric: regionalist and autonomist sentiment is politically far more prominent in Flanders than in Wallonia (Deschouwer, 2012), so the direction in which contrastive resistance is theoretically most plausible is precisely the direction this study could not test, given the small Flemish subsample. The absence of backlash among Walloon participants, whose mean regional identification of 6.6 indicates that the sample is not regionally indifferent, is informative, but it bounds out-group resistance in one direction only. A Flemish-side replication is therefore not merely a robustness extension but the natural next test of the identity-source hypothesis, and it is listed accordingly in Section 6.5.

The contribution, as set out in Chapter 1, is to clarify rather than to confirm: it shows that, in this setting, attaching a norm to a real regional out-group does not generate the contrastive response the identity-source hypothesis would predict. This distinction matters because a simple

comparison between in-group and out-group treatment arms, without a no-norm control, could not separate weak compliance from active resistance.

6.3 Information receipt without behavioral compliance

The behavioral result is one half of the experiment's evidence. The other half concerns whether the treatment carried information at all. The surprise-direction regression shows that it did. Among treated participants, the signed distance between the prior belief and the displayed norm predicts the reported direction of surprise ($\beta = -0.0026$, $p < 0.001$, $R^2 = 0.12$, section 5.4). Participants whose priors were below the displayed norm reported the norm as larger than expected, and those whose priors were above it reported the norm as smaller. This measure is a self-reported direction of surprise rather than an elicited posterior belief; it establishes that the information was received and recognized as discrepant from priors, not that participants' quantitative beliefs about giving were durably revised.

A further qualification concerns the objects being compared. The pre-treatment belief question asks about the average annual donation of a person living in Belgium, a population mean that includes non-donors. The displayed norm is the average declared donation among taxpayers who claim the tax deduction, a conditional mean within a selected group that Section 3.2 already characterizes as older, higher-income, and more administratively engaged than the average resident. A participant whose prior of roughly €110 is accurate for the population could therefore report, correctly and without revising any belief about typical giving, that declarers give "more than expected." Under this reading, the registered surprise reflects recognition that the statistic refers to a different object rather than an informational shock about the behavior of typical others, and the absence of a behavioral response is consistent with rational updating rather than a failure of compliance. The surprise measure cannot distinguish these interpretations. This reinforces the case, developed in Section 6.5, for eliciting incentivized posterior beliefs about the same quantity as the prior in future work. Evidence that correcting misperceived norms can shift behavior when the corrected belief concerns the decision-relevant reference group (Bursztyn et al., 2020) underlines why the object of the elicited belief matters, and work distinguishing norm perception from norm internalization (Tankard and Paluck, 2016) provides a framework for the gap between registered discrepancy and behavioral response observed here.

The texture of the manipulation checks qualifies this. Recall of the displayed bracket was 63%, below the 70% benchmark set in advance, so a non-trivial share of treated participants did not retain the exact figure. Perceived credibility averaged 5.12 on a 0–10 scale, near the midpoint rather than high, and did not differ across norm sources. The information was therefore received, but it was not retained perfectly and was found only moderately credible. This matters for interpretation. The first stage operated, but not at full strength, which is one reason a second-stage behavioral effect may be harder to produce.

The donation-side evidence is weaker still and should be read with care. Among treated participants, the signed norm-prior gap does not predict donation amounts ($\beta = -0.001$, $p = 0.96$, $N = 203$, section 5.7). Three caveats apply. This is descriptive evidence rather than a causal mediation estimate, because the belief gap is not experimentally manipulated.

The regression is restricted to treated participants and has limited precision. And the pattern of detectable information receipt alongside an undetectable donation effect is consistent with information-provision experiments that document first-stage belief movement without detectable second-stage behavioral change (Haaland et al., 2023). Whether this reflects a genuine decoupling between information and behavior, moderate credibility dampening the behavioral response, or simply the limited precision of the study, cannot be settled with the present data. It is best treated as a candidate mechanism for future work rather than as a finding of this study.

6.4 Limitations and external validity

The internal validity of the treatment contrast rests on random assignment, supported by the balance tests, the placebo test on control participants, the absence of differential attrition, and the within-Walloon identification of the identity contrast (sections 5.3 and 5.9). The main remaining constraints concern precision, the decision setting, and external validity, and several are set out in section 3.7.

First, the study is underpowered for small effects. The minimum detectable effects are around €12 across the main contrasts (Section 5.6.1), with the interpretive consequences already drawn there and in Section 6.1. Second, the sample is regionally unbalanced and recruited online through personal and university networks. The identity-source specifications are estimated primarily within the Walloon sample, while the Flemish subsample remains too small for within-region inference, and the estimates describe the recruited participants rather than the Belgian adult population (Henrich et al., 2010). Third, the decision is a hypothetical online allocation. This makes the experiment feasible at thesis scale, but it removes the reputational, observability, and repeated-interaction channels through which identity-based giving may operate. Participants made a private one-shot decision, without an audience whose judgement could make conformity to, or contrast with, the norm socially consequential. Fourth, the norm is an annual administrative amount, while the outcome is a one-shot allocation from a hypothetical €100 endowment, so the stimulus functions as a broad signal about regional giving rather than as a direct numerical benchmark for the task.

A fifth consideration concerns the response interface. The donation slider initialized at the midpoint of 50 euros in all conditions, and the realized donation distribution is approximately symmetric around a mean of 42.5, an atypical shape for dictator-style allocations, which are usually right-skewed. A salient midpoint default can anchor choices (Johnson and D. Goldstein, 2003) and has been shown to affect charitable allocations specifically (Altmann et al., 2019; Goswami and Urminsky, 2016). Two implications follow. First, default anchoring may compress donations toward the midpoint in every arm, attenuating any treatment effect; the range of true effects consistent with the observed null, including the bound on contrastive out-group effects, should be read with this attenuation in mind. Stated more bluntly: the minimum detectable effects of Section 5.6.1 are computed on an outcome that may itself be compressed by the default, so the effective detection threshold in terms of latent generosity is plausibly larger than the nominal €12, and the €5 bound on contrastive effects is correspondingly less tight in latent terms. Second, because the anchor is constant across arms, it cannot generate spurious treatment

differences and does not threaten the internal comparison. The robustness analysis in section 5.10 shows that excluding exact-50 responses leaves the out-group coefficient and in-group premium substantively unchanged, so the null is not driven solely by responses left at the default. Future implementations should randomize or remove the slider's starting position.

A further limitation is that the design does not fully observe participants' prior relationship to charitable giving. The survey included charitable-giving filler items, but it did not collect a validated measure of past giving, volunteering, association membership, or attachment to the selected charity. These factors may matter for treatment response. Habitual donors or participants already involved in associative life may interpret a regional donation norm differently from participants for whom charitable giving is less central. Similarly, regional identification does not necessarily imply attachment to the charity itself. If the charity is perceived as politically or socially salient, neutral, or distant, this may interact with the regional source of the norm in ways the present design cannot isolate.

These constraints do not undermine the internal comparison, since all treatment arms share the same online setting, donation task, and administrative norm source. They do limit generalization. The result is best read as evidence that, in an anonymous and hypothetical online setting, a distant institutional norm attributed to a regional in-group or out-group did not shift giving among the recruited participants. It does not establish that regional norms are behaviorally inert in settings with real stakes, social observability, or more proximate reference groups.

6.5 Directions for future research

These results and limitations point to several changes for future work, of which only the first concerns sample size.

The first is power. A larger, publicly pre-registered study should be sized to detect smaller effects than this thesis could identify. In particular, it should have enough power to distinguish very small negative out-group effects from zero, which the present design cannot. Public pre-registration of the hypothesis tests would also protect the analysis against post-hoc reframing.

Related to power but conceptually distinct, a follow-up study should secure a Flemish subsample large enough to identify the contrast in the opposite direction. Because regionalist identity is politically stronger in Flanders, the Flemish-side test is the more demanding test of out-group resistance, and the present design leaves it open.

The second is the norm stimulus, and it is a design change rather than a matter of scale. The norm here is an annual administrative amount of around €300, while the decision is a one-shot allocation from a hypothetical €100 endowment. Expressing the norm in the same unit as the decision would give it a direct numerical interpretation for the donation task and would test whether the absence of a behavioral response reflects the distance between the stimulus and the outcome.

The third is the reference group. The meta-analytic evidence suggests that proximate referents such as friends and family produce stronger norm effects than distant ones (Chapman et al., 2025). Pairing a regional norm with a more proximate reference group would test whether

reference-group proximity is what the present design was missing.

The fourth concerns the donation environment. A future study could use an incentivised donation decision or vary whether the donation is private or socially observable. This would test whether regional norm effects require the reputational or self-image channels that are weak in an anonymous hypothetical allocation.

A further extension would measure donor profile and charity attachment more directly. Past giving, volunteering, association membership, and baseline attachment to the charity could identify subgroups for whom the norm is more or less behaviorally relevant. Varying the charity, or using charities that differ in political or regional salience, would also help separate the effect of the regional norm source from the fit between that source and the object of giving.

A future design should also elicit posterior beliefs about regional giving after treatment, ideally with incentives for accuracy (Haaland et al., 2023), so that belief revision can be measured directly rather than inferred from self-reported surprise, in the spirit of norm-correction designs that elicit and update beliefs about a single well-defined quantity (Bursztyn et al., 2020).

The present design produced a behavioral null in the precision range it could test. The directions above describe what a future study would need to change to test the smaller effects that remain open.

Chapter 7 – Conclusion

This thesis asked whether regional descriptive norms affect charitable giving, and whether the effect depends on the identity source of the norm. Using a randomized online experiment with a no-norm control and two regionally sourced norm conditions, it finds no detectable effect of norm exposure on donation amounts. The out-group coefficient is not negative, the in-group premium is statistically indistinguishable from zero, and regional identification does not moderate that premium. The information was received and recognized as discrepant from prior beliefs in the expected direction, though recall was imperfect and credibility moderate, and this informational response was not accompanied by a detectable change in behavior.

The contribution is clarifying rather than confirmatory. The experiment does not establish that regional descriptive norms change giving, and it is not powered to detect the small effects that the literature reports. What it does is separate two possibilities that a simple comparison of in-group and out-group arms cannot distinguish. By including a no-norm control, the design can ask not only whether an out-group norm is followed less than an in-group norm, but whether it actively reduces giving relative to no norm. It does not. The point estimate runs in the opposite direction, and the confidence interval rules out contrastive effects larger than approximately €5. In a setting where regional identity is meaningful, an out-group norm produced no sign of the backlash that the identity-source hypothesis would predict. This is consistent with the established distinction between in-group favoritism and out-group hostility.

These results carry a modest practical caution rather than a policy recommendation. The experiment shows that participants noticed the informational content of the norm: reported surprise moved clearly with the prior-to-norm gap. It also shows that this information receipt was not accompanied by a behavioral response in the present setting. Information campaigns built on descriptive norms should not be assumed to convert belief-relevant information into changed giving as a matter of course. The study is an underpowered pilot in a hypothetical online setting, however, and it does not support a stronger claim than this. It cannot establish that regional norms are ineffective in settings with real stakes, social observability, or more proximate reference groups.

A larger, publicly pre-registered study, sized for smaller effects, could test what this one could not. The design used here, with its no-norm control and pre-treatment identity measures, provides a structure for that work. The behavioral null reported in this thesis is best read not as evidence that norm source does not matter, but as evidence that, in this setting and at this precision, large identity-driven effects on giving did not appear.

Appendix A – Survey Screens

A.1 oTree experiment screens

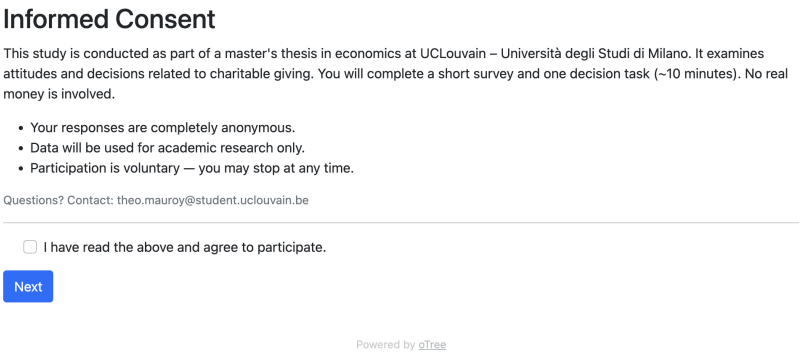

Figure A.1: Informed consent screen.

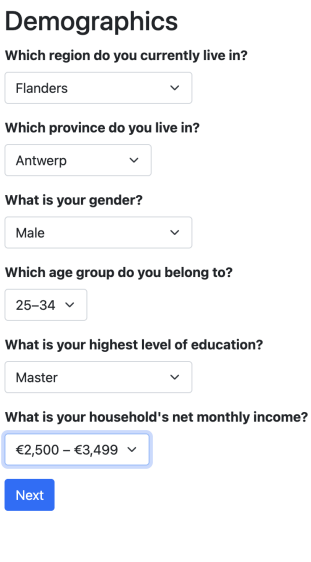

Figure A.2: Pre-treatment demographic questions.

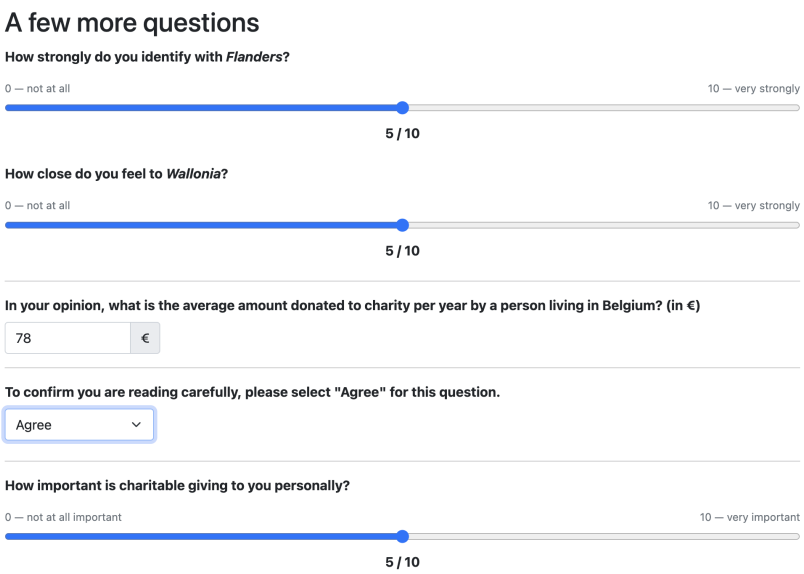

Figure A.3: Pre-treatment regional identification, closeness, prior belief, and attention-check items.

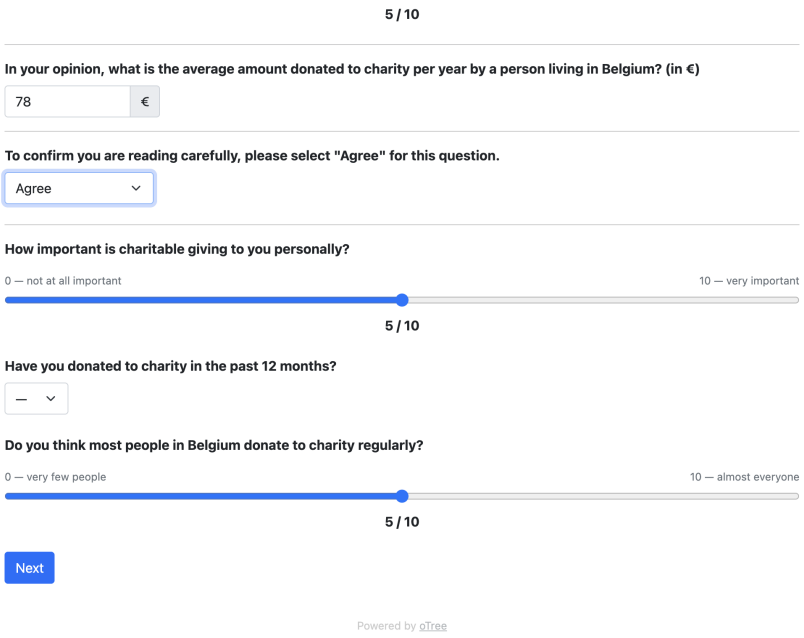

Figure A.4: Pre-treatment belief, attention-check, and charitable-giving filler items.

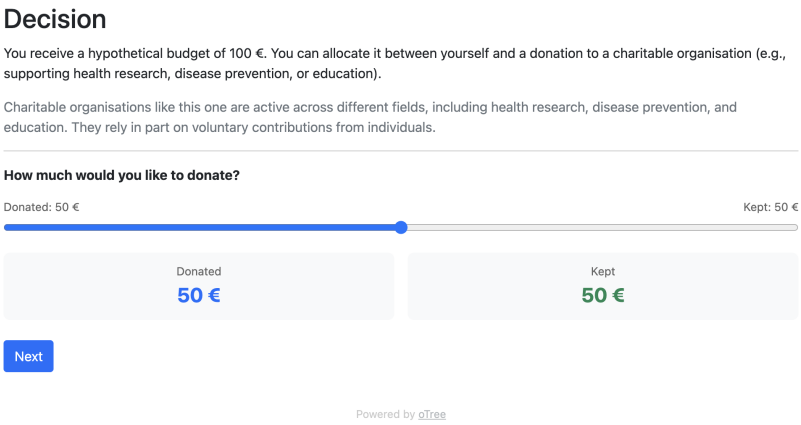

Figure A.5: Donation decision screen in the control condition.

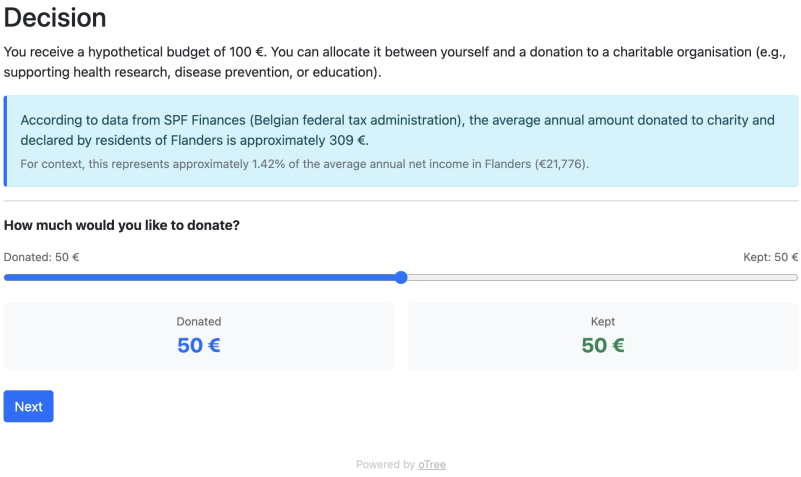

Figure A.6: Donation decision screen in the Flemish norm condition.

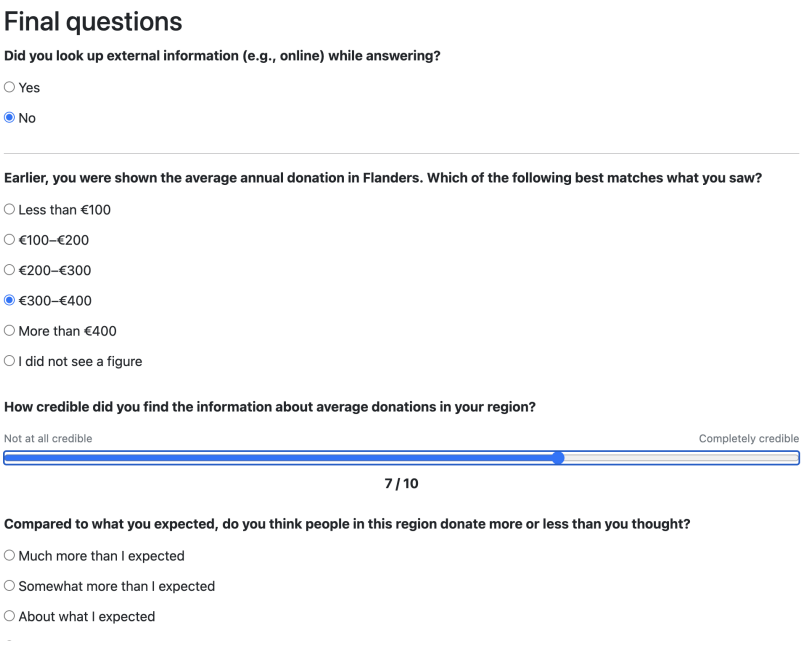

Figure A.7: Post-decision external-information and norm-recall questions for treated participants.

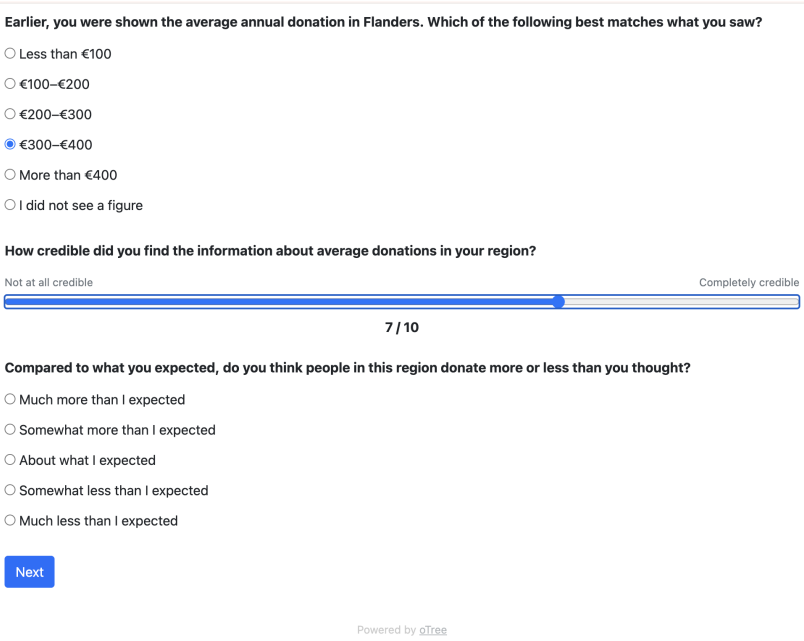

Figure A.8: Post-decision norm recall, credibility, and belief-update questions.

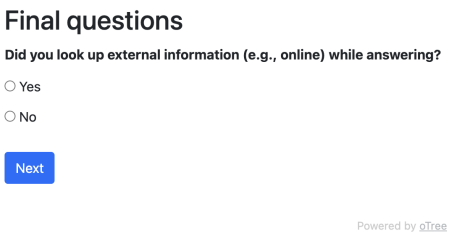

Figure A.9: Post-decision external-information question.

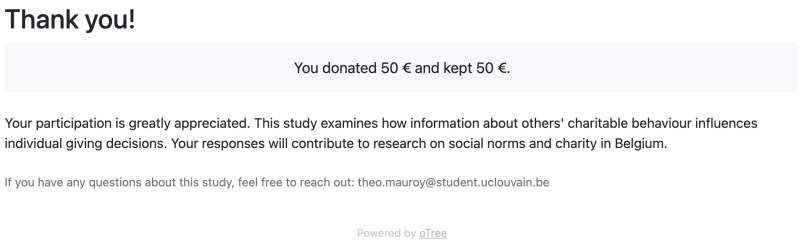

Figure A.10: Debrief and thank-you screen.

Appendix B – Variable Documentation

Table B.1: Variable Summary

Variable	Type	Scale	Timing	Role
Outcome				
donation	Continuous	0–100	Post-treatment	Primary outcome
Treatment				
norm_condition	Categorical	3 levels	Assignment	Treatment arm
norm_region	Categorical	FL/WA/null	Assignment	Derived
norm_score	Integer	301/309/null	Assignment	Robustness
treat_norm	Binary	0/1	Assignment	Model 1
treat_norm_flanders	Binary	0/1	Assignment	Model 2
treat_norm_wallonia	Binary	0/1	Assignment	Model 2
ingroup_norm	Binary	0/1/null	Derived	Models 3–4
Pre-treatment moderators and beliefs				
identity_strength	Continuous	0–10	Pre-treatment	Model 4 moderator
outgroup_closeness	Continuous	0–10	Pre-treatment	Secondary
prior_belief	Integer	Euros	Pre-treatment	Belief decomposition
Demographic controls				
region	Categorical	FL/WA/BXL	Pre-treatment	Stratification
province	Categorical	—	Pre-treatment	Descriptive
gender	Categorical	—	Pre-treatment	Control
age_group	Categorical	Brackets	Pre-treatment	Control
education_level	Categorical	—	Pre-treatment	Control
income_bracket	Categorical	Brackets	Pre-treatment	Control
Data quality				
attention_check_passed	Binary	0/1	Pre-treatment	Exclusion
looked_up_info	Binary	0/1	Post-treatment	Exclusion
Post-decision measures				
norm_page_time	Continuous	Seconds	Post-treatment	Descriptive
norm_recall_bracket	Categorical	6 levels	Post-treatment	Manipulation check
recall_correct	Binary	0/1/null	Post-treatment	Manipulation check
norm_credibility	Continuous	0–10	Post-treatment	Heterogeneity
belief_update_direction	Ordinal	−2 to +2	Post-treatment	Belief decomposition

Note: FL = Flanders, WA = Wallonia, BXL = Brussels. ingroup_norm is null for control participants. norm_score, norm_credibility, recall_correct, and belief_update_direction are null for control participants where indicated. All 0–10 scales follow Stantcheva (2023) and Haaland et al. (2023).

B.1 Variable descriptions

The experiment exports thirty variables. These are organized below by function: outcome, treatment, moderators and beliefs, demographic controls, data quality checks, and post-decision measures.

B.1.1 Outcome variable

D_i is the donation amount entered by participant i. It is a self-reported integer between 0 and 100, representing euros allocated to charity out of a hypothetical €100 endowment. It is the primary dependent variable in all regression specifications. It is continuous by construction and bounded at both ends, which motivates the quantile regression and robustness checks discussed in chapter 4.

B.1.2 Treatment variables

Six variables capture the treatment structure. Together they allow the full disaggregation of treatment arms and the derivation of in-group and out-group status.

norm_condition is a categorical variable with three levels: control, norm_FL, and norm_WA. It records the treatment arm to which participant i was randomly assigned. It is the primary treatment indicator.

norm_region records the region whose norm was displayed to the participant. It takes the values Flanders, Wallonia, or null for control participants. It is derived from norm_condition.

norm_score records the actual norm value shown. It takes the value 309 for Flemish norm recipients, 301 for Walloon norm recipients, and null for control participants. It is used in the norm magnitude robustness check in section 4.5.

treat_norm is a binary indicator equal to 1 if the participant received any normative information and 0 for control participants. It is the treatment indicator in the baseline specification.

treat_norm_flanders is a binary indicator equal to 1 if the participant was assigned to the Flemish norm condition. It is the treatment indicator for the Flemish arm in the disaggregated specification.

treat_norm_wallonia is a binary indicator equal to 1 if the participant was assigned to the Walloon norm condition. It is the treatment indicator for the Walloon arm in the disaggregated specification.

ingroup_norm is the underlying coding variable that records whether a treated participant's norm source matches their region of residence. It equals 1 if the participant is from Flanders or Wallonia and the norm received originates from their region of residence, 0 if the norm originates from the other region, and is null for control participants and for Brussels participants. It is derived post-randomization by crossing norm_region with region. Brussels participants are coded as null because no Brussels norm is presented and their out-group status is therefore by design rather than by identity contrast. The two regression-ready indicators derived from this variable, ingroup_norm_recipient and outgroup_norm_recipient, are defined below.

ingroup_norm_recipient is a binary indicator equal to 1 if the participant is treated and ingroup_norm = 1, and 0 otherwise (control participants and out-group recipients). It is null for Brussels participants, who are excluded from the primary identity-source specifications. Together with treat_norm, it forms the $\text{Norm}_i \times \text{Ingroup}_i$ interaction used to estimate the in-group premium in Models 3 and 4. It is also reported directly as IngroupNorm_i in some robustness checks for readability.

outgroup_norm_recipient is a binary indicator equal to 1 if the participant is treated and ingroup_norm = 0, and 0 otherwise (control participants and in-group recipients). It is null for Brussels participants in the primary specifications and equal to 1 in the Brussels-inclusion robustness check (section 4.5). In the compact Model 3 parameterization, the out-group effect is read from Norm_i, with $\text{Ingroup}_i = 0$. The direct OutgroupNorm_i indicator is retained for robustness tables that display in-group and out-group coefficients separately.

B.1.3 Pre-treatment moderators and beliefs

identity_strength measures the participant's strength of identification with their region of residence. It is recorded on a 0–10 scale, where 0 indicates no identification and 10 indicates very strong identification. It is collected before treatment exposure. It enters the moderation specification as the primary continuous moderator of the norm effect. In the moderation specifications it enters mean-centered: the sample mean is subtracted, so the variable equals zero at average identification. This does not change the scale or the data. It sets the point at which the out-group coefficient and in-group premium are interpreted, so that those coefficients describe the norm effect and premium for a participant with average regional identification.

outgroup_closeness measures the participant's perceived closeness to the other main Belgian region. It is recorded on a 0–10 scale. It is collected before treatment exposure. It is used in secondary heterogeneity analyses. It is not included in the primary specifications.

prior_belief records the participant's stated belief about average annual charitable giving in Belgium before treatment exposure. It is entered as an integer in euros. Prior beliefs outside the 0–1000 euro range are treated as out-of-range and excluded from belief-channel analyses; this rule affects six observations in the analysis sample. The cleaned variable is used to compute the signed gap $\text{norm\_score}_i - \text{prior\_belief}_i$ for treated participants and an underestimation indicator equal to one when the prior belief lies below the displayed norm. Positive signed gaps indicate that the displayed norm exceeded the participant's prior belief. These measures are used in supplementary specifications to describe the informational channel of the treatment.

B.1.4 Demographic controls

region records the participant's self-reported region of residence. It takes the values Flanders, Wallonia, or Brussels. It is collected on the first substantive screen, before any identity priming. It is used to derive ingroup_norm and as a stratification variable in robustness checks.

province records the participant's self-reported province within their region. It provides geographic granularity for descriptive analyses.

gender records the participant's self-reported gender. It is included as a demographic control in all primary specifications.

age_group records the participant's age bracket. Age is collected in brackets rather than as a continuous variable to reduce respondent burden. It is included as a demographic control in all primary specifications.

education_level records the participant's highest completed level of education. It is a categorical variable. It is included as a demographic control in all primary specifications.

income_bracket records the participant's self-reported monthly household income bracket. It is included as a demographic control in all primary specifications. It also enters secondary heterogeneity analyses, since higher-income participants may respond differently to a norm expressed as an absolute annual euro figure.

B.1.5 Data quality variables

attention_check_passed is a binary indicator equal to 1 if the participant selected the pre-specified response on the embedded attention check item and 0 otherwise. Participants with attention_check_passed = 0 are excluded from the primary analysis sample.

looked_up_info is a binary indicator equal to 1 if the participant self-reports having searched for external information during the experiment. Participants with looked_up_info = 1 are excluded from the primary analysis sample. Together with attention_check_passed, it defines the two pre-specified exclusion criteria.

B.1.6 Post-decision measures

norm_page_time records the time in seconds spent on the norm display page, captured automatically by oTree. It is used descriptively to assess whether participants were exposed to the stimulus for a meaningful duration. It is not used as an exclusion criterion. Its distribution is reported alongside the manipulation check results.

norm_recall_bracket records the participant's selection from a set of brackets describing the norm value they were shown: less than €100, €100–€200, €200–€300, €300–€400, more than €400, or I did not see a figure. It is collected from all participants. Control participants are expected to select the last option.

recall_correct is a binary indicator derived from norm_recall_bracket. It equals 1 if the participant selected the €300–€400 bracket, which contains both €301 and €309, and 0 otherwise. It is null for control participants. It serves as the primary manipulation check indicator.

norm_credibility records the treated participant's rating of the credibility of the norm information they received. It is measured on a 0–10 scale and is null for control participants. It is used in exploratory heterogeneity analyses to assess whether responses differ by perceived credibility.

belief_update_direction records the participant's reported direction of surprise in response to the norm. It is measured on a signed five-point scale: –2 (much less than expected), –1

(somewhat less than expected), 0 (about as expected), 1 (somewhat more than expected), 2 (much more than expected). It is collected from treated participants only. It is null for control participants who received no norm. A positive value indicates that the norm was higher than the participant's prior belief. It is used in the belief channel decomposition alongside prior_belief and norm_score, but it is not an elicited posterior belief.

B.2 A priori power calculation

Table B.2: A priori G*Power calculation for a two-sample mean comparison

Parameter	Value
Test family	t tests
Statistical test	Means: difference between two independent means (two groups)
Type of power analysis	A priori: compute required sample size given α, power, and effect size
Tail(s)	Two
Effect size d	0.20
α error probability	0.05
Power $(1 - \beta)$	0.80
Allocation ratio N_2/N_1	1
Noncentrality parameter δ	2.8071338
Critical t	1.9629867
Degrees of freedom	786
Required sample size, group 1	394
Required sample size, group 2	394
Required total sample size	788
Actual power	0.8005931

Appendix C – Quantile Regression Estimates

Section 5.8 reports that the behavioral null holds across the donation distribution. This appendix gives the full quantile panel behind that statement. Both models are estimated on the Walloon analysis sample ($N = 266$) using bootstrapped quantile regression with 200 replications. Controls are gender, age group, education, and income bracket, matching the OLS specifications in section 4.2. Coefficients are in euros. The estimates confirm that neither the out-group coefficient nor the in-group premium is concentrated at a particular point of the distribution, and that no boomerang pattern appears among high donors.

Table C.1: Model 3, out-group effect and in-group premium by quantile

	Q25	Q50	Q75
Norm exposure	+1.00	0.00	+3.17
	(6.01)	(5.03)	(9.03)
Norm × in-group	+3.00	+1.50	+5.50
	(6.68)	(5.31)	(8.18)
In-group premium test (p)	0.67	0.78	0.57
Controls	Yes	Yes	Yes
N	266	266	266
Pseudo R^2	0.091	0.118	0.100

Bootstrap standard errors in parentheses (200 replications). No coefficient is significant at any quantile. The in-group premium is statistically null at all three points of the distribution.

Table C.2: Model 4, identity moderation of the in-group premium by quantile

	Q25	Q50	Q75
Norm exposure	+3.01	0.00	+5.08
	(5.92)	(5.16)	(9.21)
Norm × in-group	+2.98	+2.46	+5.30
	(6.56)	(5.95)	(8.93)
Identification	-1.33	0.00	+1.05
	(1.45)	(1.07)	(2.08)
Norm × identification	-0.27	0.00	-0.32
	(2.23)	(2.19)	(3.33)
Norm × in-group × identification	-1.22	+0.75	-0.90
	(2.80)	(2.40)	(3.64)
H4 test: triple interaction (p)	0.66	0.76	0.80
Controls	Yes	Yes	Yes
N	266	266	266
Pseudo R^2	0.100	0.120	0.101

Bootstrap standard errors in parentheses (200 replications). Model 4 includes the full lower-order hierarchy for the triple interaction. The identity moderation of the in-group premium is statistically null at every quantile, consistent with the OLS result in section 5.6. Median coefficients equal to zero are printed as 0.00 to avoid giving numerical minus signs substantive meaning.

Appendix D – Region-Specific Estimates for Flanders

Section 5.11 notes that the Flemish subsample is too small for inference and is reported descriptively only. This appendix gives that regression. It repeats the Model 3 specification on the 22 Flemish analysis-sample participants, with the same controls as the Walloon primary specification. It is included for transparency, so a reader can see the Flemish data were examined. It is not a basis for any claim in the thesis, and the in-group versus out-group contrast is identified on the Walloon sample throughout.

Table D.1: Model 3 on the Flemish subsample (descriptive only)

	Donation
In-group norm	-25.53
	(9.20)
Out-group norm	-17.29
	(19.51)
Constant	+100.15
	(28.18)
Controls	Yes
N	22
R^2	0.78

Robust standard errors in parentheses. Controls: gender, age group, education, income bracket. No inference should be drawn from this table.

These estimates carry no inferential weight. With 22 observations and 13 covariates, the model has 8 residual degrees of freedom, which is what drives the high R^2 and the apparent significance of the in-group coefficient. The point estimates are also implausible in magnitude. A -25.53 in-group effect is several times larger than anything in the Walloon sample, and the constant of 100.15 sits at the top of the donation scale. The table is reported for completeness, not because the Flemish contrast is identified.

Bibliography

Agerström, Jens, Rickard Carlsson, Linda Nicklasson, and Linda Guntell (Feb. 1, 2016). "Using Descriptive Social Norms to Increase Charitable Giving: The Power of Local Norms". In: *Journal of Economic Psychology* 52, pp. 147–153. ISSN: 0167-4870. DOI: 10.1016/j.joep.2015.12.007.

Akerlof, George A. and Rachel E. Kranton (Aug. 1, 2000). "Economics and Identity". In: *The Quarterly Journal of Economics* 115.3, pp. 715–753. ISSN: 0033-5533. DOI: 10.1162/003355300554881.

Allcott, Hunt (2011). "Social Norms and Energy Conservation". In: *Journal of Public Economics* 95.9–10, pp. 1082–1095. DOI: 10.1016/j.jpubeco.2011.03.003.

Altmann, Steffen, Armin Falk, Paul Heidhues, Rajshri Jayaraman, and Marrit Teirlinck (2019). "Defaults and Donations: Evidence from a Field Experiment". In: *The Review of Economics and Statistics* 101.5, pp. 808–826. DOI: 10.1162/rest_a_00774.

Andreoni, James (1990). "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving". In: *The Economic Journal* 100.401, pp. 464–477. DOI: 10.2307/2234133.

Balliet, Daniel, Junhui Wu, and Carsten K. W. De Dreu (2014). "Ingroup Favoritism in Cooperation: A Meta-Analysis". In: *Psychological Bulletin* 140.6, pp. 1556–1581. DOI: 10.1037/a0037737.

Bénabou, Roland and Jean Tirole (2006). "Incentives and Prosocial Behavior". In: *The American Economic Review*. DOI: 10.1257/AER.96.5.1652.

Berinsky, Adam J., Michele F. Margolis, and Michael W. Sances (2013). "Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys". In: *American Journal of Political Science*. DOI: 10.1111/ajps.12081.

Bicchieri, Cristina and Erte Xiao (2009). "Do the Right Thing: But Only If Others Do So". In: *Journal of Behavioral Decision Making* 22.2, pp. 191–208. DOI: 10.1002/bdm.621.

Brambor, Thomas, William Roberts Clark, and Matt Golder (2006). "Understanding Interaction Models: Improving Empirical Analyses". In: *Political Analysis* 14.1, pp. 63–82. DOI: 10.1093/pan/mpi014.

Brewer, Marilynn B. (1999). "The Psychology of Prejudice: Ingroup Love and Outgroup Hate?" In: *Journal of Social Issues* 55.3, pp. 429–444. ISSN: 1540-4560. DOI: 10.1111/0022-4537.00126.

Bruhn, Miriam and David McKenzie (2009). "In Pursuit of Balance: Randomization in Practice in Development Field Experiments". In: *American Economic Journal: Applied Economics* 1.4, pp. 200–232. DOI: 10.1257/app.1.4.200.

Bursztyn, Leonardo, Alessandra L. González, and David Yanagizawa-Drott (2020). "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia". In: *American Economic Review* 110.10, pp. 2997–3029. DOI: 10.1257/aer.20180975.

Chapman, Cassandra M, Jessica Spence, Lucas Dixon, Aimee E Smith, and Matthew J Hornsey (2025). "Social Norms and Charitable Giving: A Meta-Analytic Review". In: *Personality and Social Psychology Bulletin*, pp. 1–21. DOI: 10.1177/01461672251383996.

Charness, Gary and Matthew Rabin (2002). "Understanding Social Preferences with Simple Tests". In: *Quarterly Journal of Economics*. DOI: 10.1162/003355302760193904.

Charness, Gary, Luca Rigotti, and Aldo Rustichini (2007). "Individual Behavior and Group Membership". In: *American Economic Review* 97.4, pp. 1340–1352. DOI: 10.1257/aer.97.4.1340.

Chen, Daniel L., Martin Schonger, and Chris Wickens (2016). "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments". In: *Journal of Behavioral and Experimental Finance* 9.C, pp. 88–97.

Chen, Yan and Sherry Xin Li (2009). "Group Identity and Social Preferences". In: *American Economic Review* 99.1, pp. 431–457. DOI: 10.1257/AER.99.1.431.

Cialdini, Robert B., Raymond R. Reno, and Carl A. Kallgren (1990). "A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places". In: *Journal of Personality and Social Psychology* 58.6, pp. 1015–1026. DOI: 10.1037/0022-3514.58.6.1015.

Deschouwer, Kris (2012). *The Politics of Belgium: Governing a Divided Society*. 2nd ed. London: Palgrave Macmillan. DOI: 10.1007/978-1-137-27338-6.

Dons | SPF Finances (2023).

Faul, Franz, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner (2007). "G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences". In: *Behavior Research Methods* 39.2, pp. 175–191. DOI: 10.3758/BF03193146.

Fehr, Ernst and Klaus M. Schmidt (1999). "A Theory of Fairness, Competition, and Cooperation". In: *The Quarterly Journal of Economics* 114.3, pp. 817–868. DOI: 10.1162/003355399556151.

Fischbacher, Urs, Simon Gächter, and Ernst Fehr (2001). "Are People Conditionally Cooperative? Evidence from a Public Goods Experiment". In: *Economics Letters* 71.3, pp. 397–404. DOI: 10.1016/S0165-1765(01)00394-9.

Frey, Bruno S and Stephan Meier (Dec. 1, 2004). "Social Comparisons and Pro-social Behavior: Testing "Conditional Cooperation" in a Field Experiment". In: *American Economic Review* 94.5, pp. 1717–1722. ISSN: 0002-8282. DOI: 10.1257/0002828043052187.

Gelman, Andrew and John Carlin (Nov. 2014). "Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors". In: *Perspectives on Psychological Science* 9.6, pp. 641–651. ISSN: 1745-6916. DOI: 10.1177/1745691614551642.

Goldstein, Noah J., Robert B. Cialdini, and Vladas Griskevicius (2008). "A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels". In: *Journal of Consumer Research* 35.3, pp. 472–482. DOI: 10.1086/586910.

Goswami, Indranil and Oleg Urminsky (2016). "When Should the Ask Be a Nudge? The Effect of Default Amounts on Charitable Donations". In: *Journal of Marketing Research* 53.5, pp. 829–846. DOI: 10.1509/jmr.15.0001.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart (Mar. 1, 2023). "Designing Information Provision Experiments". In: *Journal of Economic Literature* 61.1, pp. 3–40. ISSN: 0022-0515. DOI: 10.1257/jel.20211658.

Hallsworth, Michael, John A. List, Robert D. Metcalfe, and Ivo Vlaev (2017). "The Behavioralist as Tax Collector: Using Natural Field Experiments to Improve Tax Compliance". In: *Journal of Public Economics* 148, pp. 14–31. DOI: 10.1016/j.jpubeco.2017.02.003.

Henrich, Joseph, Steven J. Heine, and Ara Norenzayan (2010). "Most People Are Not WEIRD". In: *Nature* 466.7302, p. 29. DOI: 10.1038/466029a.

Hysenbelli, Dorina, Enrico Rubaltelli, and Rino Rumiati (Nov. 2013). "Others' Opinions Count, but Not All of Them: Anchoring to Ingroup versus Outgroup Members' Behavior in Charitable Giving". In: *Judgment and Decision Making* 8.6, pp. 678–690. ISSN: 1930-2975. DOI: 10.1017/S193029750000471X.

Johnson, Eric J. and Daniel Goldstein (2003). "Do Defaults Save Lives?" In: *Science* 302.5649, pp. 1338–1339. DOI: 10.1126/science.1091721.

Kandul, Serhiy and Bruno Lanz (Aug. 2021). "Public Good Provision, in-Group Cooperation and out-Group Descriptive Norms: A Lab Experiment". In: *Journal of Economic Psychology* 85, p. 102382. ISSN: 01674870. DOI: 10.1016/j.joep.2021.102382.

Koenker, Roger and Gilbert Bassett (1978). "Regression Quantiles". In: *Econometrica* 46.1, pp. 33–50. DOI: 10.2307/1913643.

Krupka, Erin L. and Roberto A. Weber (2013). "Identifying Social Norms Using Coordination Games: Why Does Dictator Game Sharing Vary?" In: *Journal of the European Economic Association* 11.3, pp. 495–524. DOI: 10.1111/jeea.12006.

Lakens, Daniël (2017). "Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses". In: *Social Psychological and Personality Science* 8.4, pp. 355–362. DOI: 10.1177/1948550617697177.

Montgomery, Jacob M., Brendan Nyhan, and Michelle Torres (2018). "How Conditioning on Posttreatment Variables Can Ruin Your Experiment and What to Do about It". In: *American Journal of Political Science* 62.3, pp. 760–775. DOI: 10.1111/ajps.12357.

Oppenheimer, Daniel M., Tom Meyvis, and Nicolas Davidenko (2009). "Instructional Manipulation Checks: Detecting Satisficing to Increase Statistical Power". In: *Journal of Experimental Social Psychology* 45.4, pp. 867–872. DOI: 10.1016/j.jesp.2009.03.009.

Quidt, Jonathan de, Johannes Haushofer, and Christopher Roth (2018). "Measuring and Bounding Experimenter Demand". In: *American Economic Review* 108.11, pp. 3266–3302. DOI: 10.1257/aer.20171330.

Reich, Lars M., Luisa A. M. Mahr, Martina Vacondio, and Afreen S. Khalid (May 27, 2022). "The Effect of Ingroup vs. Outgroup Members' Behavior on Charity Preference: A Drift-Diffusion Model Approach". In: *Frontiers in Psychology* 13, p. 854747. ISSN: 1664-1078. DOI: 10.3389/fpsyg.2022.854747. PMID: 35712167.

Schultz, P. Wesley, Jessica M. Nolan, Robert B. Cialdini, Noah J. Goldstein, and Vladas Griskevicius (2007). "The Constructive, Destructive, and Reconstructive Power of Social Norms". In: *Psychological Science* 18.5, pp. 429–434. DOI: 10.1111/j.1467-9280.2007.01917.x.

Shang, Jen and Rachel Croson (Oct. 1, 2009). "A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods". In: *The*

Economic Journal 119.540, pp. 1422–1439. ISSN: 0013-0133. DOI: 10.1111/j.1468-0297.2009.02267.x.

Stantcheva, Stefanie (2023). "How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible". In: *Annual Review of Economics* 15, pp. 205–234.

Statbel (Statistics Belgium) (2025). *Revenus Fiscaux — Revenus Moyens Par Habitant Par Région, 2023*.

Tajfel, Henri and John C. Turner (1979). "An Integrative Theory of Intergroup Conflict". In: *The Social Psychology of Intergroup Relations*. Ed. by William G. Austin and Stephen Worchel. Monterey, CA: Brooks/Cole, pp. 33–47.

– (1986). "The Social Identity Theory of Intergroup Behavior". In: *Psychology of Intergroup Relations*. Ed. by Stephen Worchel and William G. Austin. Chicago: Nelson-Hall, pp. 7–24.

Tankard, Margaret E. and Elizabeth Levy Paluck (2016). "Norm Perception as a Vehicle for Social Change". In: *Social Issues and Policy Review* 10.1, pp. 181–211. DOI: 10.1111/sipr.12022.

Weisel, Ori and Robert Böhm (2015). "Ingroup Love and Outgroup Hate in Intergroup Conflict between Natural Groups". In: *Journal of Experimental Social Psychology* 60, pp. 110–120. DOI: 10.1016/j.jesp.2015.04.008.

Zizzo, Daniel John (2010). "Experimenter Demand Effects in Economic Experiments". In: *Experimental Economics* 13.1, pp. 75–98. DOI: 10.1007/s10683-009-9230-z.