Norms, Nudging, and Sustainable Consumption

An interactive online experiment evaluating the impact of descriptive and injunctive
norms on the consumption of sustainability labelled products

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Viktor Johansson

Supervised by Katarina Nordblom

Acknowledgements

I would like to thank my supervisor, Katarina Nordblom, for valuable feedback that undoubtedly improved the thesis. I also want to thank Blomsterlandet for their cooperation with the experiment.

Abstract

Sustainable consumption is one important factor in mitigating negative environmental and societal impacts of extensive production and consumption. Behavioural interventions, such as nudging, have been suggested as potential methods in changing customer behaviour. Nudging using social norms has been studied extensively to analyse its impact on pro-environmental behaviour and this study adds to the literature by conducting an experiment evaluating the effect of descriptive and injunctive normative messages on the consumption of sustainability labelled products in an environment where behavioural change is costly. Moreover, the study explores whether the strength of a norm impacts the effectiveness of a nudge. The main results show no significant effect of the four treatments analysed in the study. Further analysis of subgroups of individuals that, based on the theoretical framework, could be more nudgeable show some support for the hypothesis that the strength of a norm matters, at least for descriptive norms. In conclusion, the results do not provide any evidence for the descriptive and injunctive normative messages being successful at generating sustainable behaviour in the environment studied.

Keywords: nudging, sustainable consumption, injunctive norm, descriptive norm, Fairtrade, KRAV, EU Organic

Contents

1 Introduction

Agricultural practices are often related to intensive production with excessive use of pesticides potentially causing both environmental damages as well as human health problems (Alavanja et al., 2004). Although governments in many ways work towards solutions through policies and regulations, issues still remain. Certain organisations try to reduce the aforementioned issues by providing certifications to producers and their goods subject to following specific guidelines and production practices. Three such certifications are KRAV, EU Organic, and Fairtrade. The organisations behind the labels aim to mitigate environmental impacts but also to benefit other aspects such as working conditions and wages, which could be achieved by regulating how production takes place (for example regarding chemical use), ensuring price premiums on sold products, or by monitoring that companies are free from discrimination (Fairtrade International, 2017; KRAV, 2019). However, despite the potential benefits of consuming certified goods, and the knowledge about the labels in countries such as Sweden being relatively widespread (Holmberg & Robertson, 2018; KRAV, 2019), the overall market share has historically been low (KRAV, 2019; Lernoud & Willer, 2017), and the question of how to make certified goods more attractive remains.

A possible alternative when trying to increase the consumption of sustainability labelled products, when government interventions might be difficult or unreasonable, could be to explore non-price interventions, such as nudges[1], that have been shown to be effective in changing behaviour among individuals in other contexts (e.g. Allcott, 2011; Andersson & von Borgstede, 2010; Demarque et al., 2015; Kuhfuss et al., 2016). Thus, the aim of this study is to explore how nudging can affect the consumption of certified goods. More specifically, the study evaluates how descriptive and injunctive normative messages[2], which have been successful in other contexts, impact the choice between a sustainability labelled product and

---

[1]In short, a nudge is a non-price intervention that alters the choice architecture for an individual by for example changing the status quo or providing information about normative behaviour.

[2]A descriptive norm is defined as the "perception of which behaviours are typically performed" and an injunctive norm is defined as the "perception of which behaviours are typically approved or disapproved". Thus, the first is based on observed behaviour and the second is based on stated preferences. (Cialdini, 2003, p. 105)

an alternative without.

Research on nudging is extensive and of special interest is the literature examining the effectiveness of normative messages on pro-environmental behaviour. Thanks to the extent of the previous literature, normative messages have been studied in many different contexts. For example, Chabe-Ferret et al. (2019) examine water consumption among farmers, Allcott (2011) analyse household energy consumption, Andersson and von Borgstede (2010) explore low- and high-cost recycling, and Demarque et al. (2015) study consumption of eco-labelled grocery store items. In general, the empirical evidence suggests a positive effect of normative messages on pro-environmental behaviour, with only a few studies showing no or negative effects. However, the studies differ compared to this thesis in certain aspects, making the study concerning purchases of sustainability labelled products interesting in several ways. In contrast to many of the previous studies, sustainability labelled products often have a price premium compared to products without, allowing for an analysis of the effectiveness of nudges on behavioural change associated with a higher cost. Furthermore, this thesis examines four different norms, descriptive and injunctive norms, that also differ in their strength[3], allowing for an evaluation of whether the strength of a norm affects the effectiveness of said norm. Lastly, it is conducted on purchases of flowers, plants, and other garden-related items, which, to the best of my knowledge, has not been studied before. Thus, in order to contribute to the existing literature, the thesis aims at answering the two following research questions:

Does nudging with descriptive and injunctive norms increase the consumption of sustainability labelled products?

How is the effectiveness of a nudge affected by the strength of a norm?

---

[3]The strength of a norm is characterised by how common a certain behaviour or opinion is. For example, if 80% of people act in a certain way or believe a certain behaviour is appropriate, it is considered a stronger norm compared to if 30% of people act in that way or believe that behaviour to be appropriate. Furthermore, although a minority behaviour might not fulfil the definition of a norm, previous research often use the term norm to define both majority and minority behaviour.

Using an interactive online experiment with real products from the company Blomsterlandet operating in the floral industry, this study evaluates the effectiveness of four types of treatments; information of an injunctive and a descriptive norm, both divided into one strong norm and one weak norm; on the preferences for sustainability labelled products compared to a control group receiving no information. The two treatments related to the descriptive norm are based on actual sales data in Blomsterlandet's physical stores during two separate three month periods. Moreover, the two treatments related to the injunctive norm are based on answers from a pilot study conducted by graduate students.

The main results show no significant impact of normative messages on sustainable behaviour compared to the control group. The nonparametric Kruskal-Wallis H test was conducted for three different dependent variables, the count (1), the sum (in SEK) (2), and the percentage (3) of sustainability labelled products purchased. A possible factor contributing to the insignificant results is the large proportion of individuals that most likely already conform to a relatively sustainable behaviour, seeing as almost 50% of the sample chose the sustainability labelled option for 80% or more of their products.

Further tests were conducted for subsamples, motivated by what factors, according to a theoretical framework from Löfgren and Nordblom (2020), that should increase the likelihood of a nudge being successful. However, although some results are shown to be statistically significant, it is not possible to conclude that the nudges have been successful in increasing the consumption of sustainability labelled products. Results from one subsample even suggest the opposite effect, that one of the nudges caused individuals to purchase less sustainability labelled products. However, there is some support for the strength of a norm being important in behavioural change, at least for the descriptive norm. Conclusively, the nudges used in this experiment show no evidence for an increase in sustainable behaviour. It is interesting to discuss the external validity of the results, since the mechanisms that impact the likelihood of a nudge being successful are most likely the same aspects and mechanisms that change depending on the environment studied. Thus, it would still be interesting to study the effect of normative messages in similar environments, varying only certain aspects in order to learn more about what drives behavioural change caused by nudging.

The thesis is structured as follows; section 2 provides a background on the sustainabil-

ity labels used in the experiment and a review of the empirical evidence concerning nudging and pro-environmental nudging with a focus on normative messages, section 3 includes an explanation of theoretical frameworks regarding the mechanisms of nudging and regarding norm-dependent utility, which are followed by the hypotheses of this study, section 4 describes the experimental design, data and methodology, section 5 presents the results and analysis, and section 6 concludes the thesis.

2   Background

2.1   Sustainability labels

Three, to the floral industry, common labels have been chosen to represent the choice of a sustainable product. The labels; Fairtrade, KRAV, and EU Organic; were chosen due to the wide-spread knowledge among the Swedish population (Holmberg & Robertson, 2018; KRAV, 2019) and due to the labels being common among products in Blomsterlandet's product range. The underlying reasons behind the labels varies, from creating better working conditions, to regulating production methods (Fairtrade International, 2017; KRAV, 2019). Having said this, this study does not aim to evaluate the effectiveness of the organisations behind the labels. Instead, this study aims at evaluating the effects of nudging on the preferences for sustainability labelled products with a focus on descriptive and injunctive norms. Products with a sustainability label are henceforth called sustainable and the alternative without is called non-sustainable. However, this definition is only for simplicity in writing and does not mean products without a label cannot be sustainable.

2.2   Background on nudging

The basic characteristics of the non-price intervention called nudging was presented in the book *Nudge: Improving Decisions About Health, Wealth and Happiness* written by Thaler and Sunstein (2009). They define a nudge as "any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives" (Thaler & Sunstein, 2009, p. 6). Thus, the important

aspects of a nudge are that the effects should be predictable, the available alternatives and the costs of choices should remain the same as before the intervention and it should be easy to avoid the nudging.

A choice architecture could be designed in several ways depending on what behaviour the nudge aims to change and what kind of heuristics that cause the current, non-desirable choices. Heuristics that can cause harmful behaviour for the individual and/or society are, to mention a few, anchoring, availability, framing, and conformity[4] (Thaler & Sunstein, 2009). A few examples of changes in the choice architecture that have been successful in altering behaviour are changes in the size of food plates, menu designs, as well as information and messages about a normative behaviour (Allcott, 2011; Dayan & Bar-Hillel, 2011; Kallbekken & Sælen, 2013).

2.3 Effectiveness of nudging

The literature on pro-environmental nudging is large and different types of nudges have been applied and evaluated in various contexts. The next section reviews some of the most important articles on pro-environmental nudging using normative messages as well as the articles most closely related to this study. For a more extensive review of studies conducted within the area of pro-environmental nudging, please see Bergquist et al. (2019) and Farrow et al. (2017).

2.3.1 Nudging in agriculture

The lack of participation of farmers in environmentally friendly agricultural practices, despite the existence of monetary incentives in the form of subsidies, is a large concern for policymakers and a potential danger to the environment. Research has examined whether farmers' behaviour can be influenced by nudges using normative messages. For example, Kuhfuss et al. (2016) conducts a discrete choice experiment on wine farmers in France

---

[4]Anchoring is basing one's decision on some specific initial information, availability means using information that can be easily recalled, framing refers to how behaviour is affected by how information is presented, and conformity refers to that individuals tend to act according to normative behaviour. (Thaler & Sunstein, 2009)

evaluating whether a descriptive norm can affect participation in an agri-environmental scheme focused on reducing herbicide use. The authors find that the introduction of a bonus, which was conditional on a minimum participation rate, increased participation in the agri-environmental scheme, and that only part of the effect came from the increased monetary incentives, thus concluding that the descriptive norm had at least a partial effect on participation.

A similar study was performed in Italy by Defrancesco et al. (2008) evaluating factors affecting adoption of different agri-environmental measures from survey data. Results indicate that information about social norms regarding agricultural practices seem to increase adoption, at least for farmers already showing pro-environmental behaviour. Interestingly, farmers that do not normally adopt agri-environmental measures are not as sensitive to social norms, suggesting that normative messages could have varying effects depending on the subject exposed to the intervention.

A recent study by Chabe-Ferret et al. (2019) use social comparisons to nudge farmers into reducing water consumption using "smart meters"[5]. The treatment group in the field experiment were sent information about their own as well as the average water consumption of neighbouring farms. Moreover, farmers with better than average water consumption (i.e. below average amount consumed) received a positive message. Results display no significant effects of the treatment on cumulative water consumption over time compared to the control group. However, there is some evidence that farmers that had very high relative water consumption lowered their consumption (significant at a 10% level). In contrast, evidence also suggest that farmers with no previous water consumption increased their consumption (again only significant at a 10% level). These results are very interesting, as they display potential problems with nudging and how the use of norms can have what is called a "boomerang-effect"[6] (Sherif et al., 1958).

---

[5]The smart meters collected information about water consumption in real time, allowing the researchers to distribute information about individual and average water consumption.

[6]A boomerang-effect is defined as adoption of an opposite behaviour of what was intended by an intervention.

2.3.2   Nudging energy consumption

The research on how nudging can affect energy consumption is extensive and covers for example households, workplaces and public spaces. Allcott (2011) conducts a field experiment with a sample of almost 600 000 households around the US with the help of Home Energy Reports delivered by the company OPOWER. The reports included information about a household's relative energy consumption compared to both efficient[7] neighbours and all neighbours. These descriptive norms were combined with injunctive norms in the form of comments; Great, Good, Below Average; paired with happy emoticons (for "Great" and "Good") which depended on the household's relative energy consumption. Results find significant Average Treatment Effects supporting the use of norms to increase energy conservation.

Further research on energy consumption was conducted by Carrico and Riemer (2011) where they evaluate the effect of two interventions; Peer education and Feedback; on energy use in the workplace. Results indicate that provision of information about energy use and conservation can lead to reductions in energy consumption, but in contrast to much previous research on social norms, they find no significant impact of perceived descriptive or injunctive norms elicited through surveys to the employees.

Oceja and Berenguer (2009) conducted two field experiments evaluating the effect of descriptive and injunctive norms on energy conservation behaviour in public spaces. In practice, they exposed subjects using the public restroom to one of two descriptive norms; entering the restroom with the lights on or off; and measured to what extent an individual leaves the lights on given the initial state. Results of the first study suggest that people are affected by the descriptive norm, since individuals who entered with the lights on (off) also more often left the room with the lights on (off). In their second experiment, the researchers also included injunctive norms in the form of four different messages aimed at inducing energy conservation behaviour. Interestingly, the descriptive norm once again had a significant effect on behaviour, and furthermore, only one of the four injunctive norms had a significant effect on behaviour compared to the control group with no message.

---

[7]Defined as the 20th percentile of the comparison group.

2.3.3   Nudging recycling

The majority of research conducted on recycling has been done through surveys eliciting self-reported recycling behaviour or intentions to recycle. For example, Andersson and von Borgstede (2010) evaluate the importance of knowledge and norms on recycling that incur both low and high costs to individuals. The authors elicited knowledge about environmental consequences, knowledge of disposal, personal and social norms as well as current recycling behaviour. Results suggest that knowledge of disposal was the strongest predictor for both low-cost and high-cost recycling followed by personal norms and social norms, respectively. Moreover, comparing low-cost and high-cost recycling, the same variables had a stronger impact on high-cost recycling.

One field experiment on recycling was conducted by Schultz (1999) where he analysed four treatments on kerbside recycling, of which two were individual feedback and group feedback. The data collected during the experiment enabled three different dependent variables; participation frequency, amount recycled and contamination in the garbage bins. Results indicate that the two treatments with feedback information had significant effects in both increasing participation and the recycling amount.

2.3.4   Nudging green consumption

Research on nudging has also been conducted on consumption decisions that we might face on a daily basis, such as purchases of convenience or shopping goods. Chen and Deng (2016) did, with the help of surveys, try to analyse green consumption and what effect for example attitudes and norms toward such behaviour had on intended purchasing decisions. Results indicate that norms had a significant positive relationship with intention of green consumption both before and after controlling for knowledge of green products. Moreover, when analysing convenience and shopping goods separately, the effect of norms was similar for both groups.

Further research related to consumption behaviour was run by Groot et al. (2013) where they, through a field experiment, measured the effect of personal and injunctive normative messages on plastic bag usage in a supermarket. Results suggest that both personal and

injunctive norms significantly decreased the plastic bag usage compared to the initial state. Moreover, combining both types of messages did not increase the effectiveness of the nudge, suggesting that simply providing more information about normative behaviour does not necessarily increase the effectiveness at the same rate.

Perhaps the most closely related research to this study is an article from Demarque et al. (2015) examining the effect of descriptive norms on the consumption of eco-labelled goods. The authors ran two lab experiments using an artificial online shopping software consisting of both standard and ecological products. To nudge pro-environmental behaviour the experiment used three types of descriptive norms, two "strong" norms and one "weak" norm, based on results from a pilot study. The difference between a strong and a weak norm was the framing, where the strong norm described the share of people who had purchased "at least one ecological product" (70%) and the weak norm described the share of people who had purchased "one ecological product" (9%). Results from both experiments show a significant increase in purchases of ecological products for both strong and weak norms. However, results are ambiguous between the two experiments regarding whether the strong norm performed significantly better than the weak norm.

Previous research on pro-environmental nudging has, as presented above, in general generated positive effects, implying that providing information about norms normally affect the behaviour of individuals in a beneficial way to the environment. Although there are many similarities of this study compared to the studies conducted before, it also differs in a few important ways. Firstly, compared to agriculture, farmers have to consider factors such as uncertainty regarding future income, whereas for sustainability labelled product purchases there are no such uncertainties. Secondly, comparing to especially energy conservation, individuals or households have clear incentives to change behaviour due to possibilities of reduced costs which is in contrast to the purchase of a sustainable product, which often instead has a higher price. Thirdly, compared to recycling behaviour, the costs of recycling are often relatively low, and most likely lower relative to the price premium for sustainable products, which could make behavioural change easier. Finally, regarding the most closely related study from Demarque et al. (2015), the current experiment is conducted on different types of products and it also evaluates injunctive norms. Moreover, this study

ensures that behavioural change is costly by only studying sustainable products with a price premium, a feature that is unclear in the article by Demarque et al. (2015). Thus, the study of descriptive and injunctive norms and their effects on the demand for sustainable products can contribute in a good way to the existing research.

3   Theoretical Framework

3.1   Effectiveness of a nudge

Löfgren and Nordblom (2020) propose a theoretical framework regarding the factors and variables that will affect the effectiveness of nudging in a decision environment. The authors develop a model divided into two parts, which are important in deciding both when a nudge will be effective, but also which type of nudge that is most likely to succeed.

3.1.1   Inattentiveness

The first step regards whether decisions are made inattentively or attentively, i.e. how much attention an individual pays to the decision that he/she is making. Löfgren and Nordblom (2020) suggest that this could be thought of as whether an individual uses System 1 or System 2[8], a concept made famous by Kahneman (2003). An attentive choice will activate System 2, which requires cognitive effort and thus is more costly. However, the benefit of making an attentive choice is that the choice made is the correct one (given the information available). In contrast, making an inattentive choice will be easier, less costly, as it relies on System 1, with the downside being that the choice made might not give the favoured outcome. Following the derivations from Löfgren and Nordblom (2020), but slightly modifying the choices to fit the current context of sustainable and non-sustainable products, the choice of whether to make an attentive or an inattentive choice can be modelled as follows:

---

[8]System 1 can be thought of as our intuitions and System 2 as our reasoning.

Firstly, the expected[9] utility of making an attentive choice can be described as:

$$E[U_{att}] = E[max\{U(a_{ST}); U(a_{NS})\}] - \varsigma \tag{1}$$

where

$E[U(a_{ST})]$ is the expected utility of making a sustainable choice,

$E[U(a_{NS})]$ is the expected utility of making a non-sustainable choice, and

ς is the cost of making an attentive choice.

In contrast to the attentive choice, an inattentive choice does not cause the individual to exert any cognitive effort. However, the expected utility of making an inattentive choice is proposed to be affected by an individual's confidence in making the optimal choice inattentively. This confidence factor can be affected both by actual knowledge, but also by over-confidence in one's ability to make the correct choice. The expected utility of an inattentive choice is thus modelled by Löfgren and Nordblom (2020) as follows:

$$E[U_{in}] = \theta E[max\{U(a_{ST}); U(a_{NS})\}] + (1 - \theta)E[min\{U(a_{ST}); U(a_{NS})\}] \tag{2}$$

where $\theta \in [0, 1]$ is the confidence parameter, i.e. a subjective probability of making the most optimal choice in the inattentive choice.

Whether an individual makes the choice inattentive will thus be given by $E[U_{in}] > E[U_{att}]$ which can be derived more precisely as:

$$\varsigma > (1 - \theta)E[|U(a_{ST}) - U(a_{NS})|] \tag{3}$$

Clearly, the cost of making an attentive choice must be larger than the expected benefit in utility given by the attentive choice for an individual to make the choice inattentively[10]. Löfgren and Nordblom (2020) also suggest that the probability of making an inattentive choice will decrease with the importance of the choice (i.e. less important choices are more likely to be made inattentively), which is quite intuitive.

---

[9]$E[U_{att}]$ is the expectation since one cannot *ex ante* know the actual utility experienced.

[10]For a more in depth explanation, see Löfgren and Nordblom (2020).

3.1.2   Preference nudging

Löfgren and Nordblom (2020) argue that only an inattentive choice can be nudged. An inattentive choice can further be explained by the idea that there are two different types of utilities for each individual. The first utility occurs when a choice is to be made, and thus an individual tries to maximize his/her choice utility. However, when the individual actually consumes the good or the behaviour is carried out, the individual gains instead an experience utility which might differ from the choice utility. Using this idea, Löfgren and Nordblom (2020) tries to explain what factors will influence the choice and experience utility, as well as what factors will impact the effectiveness of a nudge.

Consider again an individual which is given the choice between a sustainable (a_{ST}) and a non-sustainable (a_{NS}) item. An individual maximises the choice utility (V) but gains an experience utility (U), Löfgren and Nordblom (2020) model the relationship as:

$$V(a_k) = \theta E[U(a_k)] + (1-\theta)\mu_{a_k} \tag{4}$$

where

$\theta \in [0,1]$ is the confidence parameter (as in equation 2),

$E[U(a_k)]$ is the expected experience utility for product k, and

μ_{a_k} denotes preference-irrelevant attributes[11] for choice a_k.

The choice will then be determined by evaluating the difference between the choice utilities:

$$V(a_{ST}) - V(a_{NS}) = \theta E[U(a_{ST}) - U(a_{NS})] + (1-\theta)(\mu_{a_{ST}} - \mu_{a_{NS}}) \tag{5}$$

Thus, if the expression is positive, an individual will purchase the sustainable product, and vice versa.

Moreover, Löfgren and Nordblom (2020) divide nudges into two different types, pure nudges and preference nudges[12]. Pure nudges are thought to be most effective in an environment of low importance and low confidence, while preference nudges are thought to be

---

[11]A preference-irrelevant attribute could for example be where in a shop a product is located.

[12]Löfgren and Nordblom (2020, p. 6) define a pure nudge as "an alteration of a preference-irrelevant attribute in an inattentive choice situation" and a preference nudge as "an alteration of expected utility in an inattentive choice situation, without altering the attentive choice". For example, changing the order of items

most effective in an environment of low importance but high confidence. This can be seen clearly from equation 5 by realising that the first expression, $\theta E[U(a_{ST}) - U(a_{NS})]$, is affected by a preference nudge and the second expression, $(1-\theta)(\mu_{a_{ST}} - \mu_{a_{NS}})$, is affected by a pure nudge. Thus, as this study uses preferences nudges, an increase in the confidence level will increase the nudgeability of an individual in two ways. Firstly, remembering equation 3, a higher confidence increases the chance of making a choice inattentively. Secondly, a higher confidence puts more weight on the first expression in equation 5 which is affected by preferences nudges.

3.2   Norm-dependent utility

The theoretical framework above attempts to model the factors and variables that affect the successfulness of nudges in general. This section instead tries to explain how specifically normative messages could impact the behaviour and decisions of individuals. Recent developments and discoveries in behavioural economics on what drives individuals' decisions in different settings have sparked researchers to try to model this behaviour. The idea that people adhere to norms and care about what other people do has been discussed for a long time. For example, Smith (1759/2002) wrote in his book The Theory of Moral Sentiments that behaviour can be altered when evaluated from the perspective of an "impartial spectator", which could be seen as the collective norm.

Furthermore, a famous experiment by Asch (1955) suggests that individuals are affected by what others do and often behave in the same way, something called conformity. In his simple experiment subjects were placed into a group where individuals, one by one, were supposed to compare and match the length of a line with one of three lines. All subjects but one were instructed before the experiment to guess the wrong answer. Results showed that as the group size increased (up to a certain size), subjects were more likely to guess the wrong answer, implying that individuals wanted to adhere to the norm. The idea of conformity can be seen as a factor explaining why information about a norm can cause change in behaviour.

in a menu could represent a pure nudge, and providing individuals with information about a normative behaviour (as in this study) could represent a preference nudge.

Krupka and Weber (2013) present a simple model describing the possible effect of adhering to normative behaviour. Although the article by Krupka and Weber (2013) focuses on injunctive norms, the framework presented can reasonably also be used to explain behaviour involving descriptive norms. Furthermore, it has been slightly modified to the context of flower, plant, and other garden-related purchases. Consider the possible choices of products $A = \{a_1, \ldots, a_k\}$ where k is the total number of products in for example a shop. Krupka and Weber (2013) suggests that the utility of a choice can be modelled as follows:

$$U(a_k) = P(a_k) + \gamma N(a_k) \tag{6}$$

where

$P(a_k)$ denotes the utility gained from the purchase of product a_k,

$N(a_k)$ is the utility gained from conforming to a norm of purchasing product a_k, and

γ is a parameter that measures how inclined an individual is to adhere to a norm.

In contrast to Krupka and Weber (2013), where they assume $N(a_k) \in [-1, 1]$, this experiment only studies the case where $N(a_k) \in [0, 1]$. The difference is that if $N(a_k) < 0$ it means that the normative behaviour is undesirable, and the normative messages included in this study involve behaviour that is seemed to be appropriate. Krupka and Weber (2013) define $\gamma \geq 0$, however it is also mentioned that there are situations where $\gamma < 0$[13]. Thus, as γ increases, the utility of acting according to the normative behaviour increases, and vice versa.

Whether γ will behave in the same way for both descriptive and injunctive norms is difficult to say. Cialdini and Trost (1998) suggest that descriptive norms are mostly used in an intrapersonal perspective as a guide to make the most effective or correct decision in an environment. In contrast, injunctive norms are thought to matter more in an interpersonal perspective, relating to social aspects. Moreover, the different motivating factors between descriptive and injunctive norms were studied by Jacobson et al. (2011). The results of the study suggest that descriptive norms mostly relate to the intrapersonal perspective of choosing the correct decision, while injunctive norms can have conflicting beliefs of, on one hand,

---

[13]For example, if an individual displays a contrarian behaviour, i.e. enjoys acting contrary to a normative behaviour, this person could gain utility from not adhering to a normative behaviour.

choosing the effective decision and, on the other hand, caring about interpersonal aspects such as social responsibility. Thus, according to previous research, it might be reasonable to consider that γ does not act in the same way for the two types of norms.

Assuming, or normalizing such that purchasing a non-sustainable product does not yield normative utility, i.e. $N(a_{NS}) = 0$, the model can be simplified and utility of the two possible choices, a_{ST} and a_{NS}, can be described as follows:

$$U(a_{ST}) = P(a_{ST}) + \gamma N(a_{ST})$$
$$U(a_{NS}) = P(a_{NS})$$

(7)

The assumption of no normative utility for the non-sustainable choice can be thought of as the purchase being the status-quo choice, which can be deviated from either by valuing the features of the sustainable product higher (affecting $P(a_{ST})$) or by valuing conforming to a normative behaviour (affecting $\gamma N(a_{ST})$). Thus, an individual will purchase the sustainable alternative if

$$U(a_{ST}) = P(a_{ST}) + \gamma N(a_{ST}) > P(a_{NS}) = U(a_{NS})$$

(8)

which will depend on utility gained from purchasing a specific product (affected by for example price, quality differences and intrinsic attitude towards sustainable products and/or labels) and the utility gained from conforming to normative behaviour (affected by individual preferences to conform and potentially the strength of the norm).

3.3   Hypotheses

Based on the theoretical framework and the empirical evidence on nudging using normative messages I below state three hypotheses. In section 4, the experiment and methodology used to test the hypotheses will be discussed.

Hypothesis 1

Information about a descriptive norm, related to current sustainable consumption, increases the consumption of sustainable products compared to the control group.

Hypothesis 2

Information about an injunctive norm, related to opinions about sustainable consumption,

increases the consumption of sustainable products compared to the control group.

Hypothesis 3

Strong norms, characterised by a more common behaviour or opinion, have a larger effect on consumption of sustainable products compared to weak norms.

4   Data and Methodology

4.1   Experimental design

Similar to a previous study by Demarque et al. (2015), this study develops an online shopping experience giving subjects a hypothetical sum to spend on a variety of products, both with and without sustainability labels. Each subject was placed with the task of collecting products for a value of up to 350 SEK[14] from a range of 42 products, 21 products with a sustainability label and 21 substitute products without a label. The experiment was conducted in collaboration with the company Blomsterlandet, a company selling flowers, plants, and other garden-related items, and thus the products used in the experiment were products from their product range.

The included products were chosen following three criteria; (1) the product was to be in stock during the months of April and/or May, (2) each type of product had to have alternatives with and without a sustainability label, and (3) the sustainable product had to be more expensive than its counterpart. These three criteria were added in order to make the online shop relevant for the customers, to make sure customers could choose between a sustainable and a non-sustainable product, and to be able to analyse nudging in an environment where behavioural change was costly. The products with a sustainability label had an average price premium of ≈ 12 SEK (price premiums ranged from 5 SEK to 25 SEK). The average price of a non-sustainable product was ≈ 37 SEK, meaning that the price premium was $\approx 32\%$ of the average price of a non-sustainable product. To incentivise behaviour as if it was a normal situation, Blomsterlandet agreed to a lottery in which three winners

---

[14]The choice of giving subjects 350 SEK to purchase products for was made with the help from representatives of the company. 10 people from the company each picked a shopping cart from the products included in the experiment. Their purchases averaged to an amount of 313 SEK which was rounded up to 350 SEK.

would receive the products that they had chosen. By including a pay-off that is based on the choices of the individual, chances of truthful and realistic choices increase.

Each subject was allocated a random number between 1 and 5 upon beginning the experiment representing the condition this individual was exposed to (see Table 1). An advantage of the experimental design was the possibility of controlling the contents of the experiment, keeping everything constant across subjects (such as product range, description, instructions, and funds for purchasing), except for the normative messages, making it possible to measure only the effect of the normative message on consumption behaviour.

Group	Condition
1	Control
2	Weak descriptive norm
3	Strong descriptive norm
4	Weak injunctive norm
5	Strong injunctive norm

Table 1: Experiment conditions

4.1.1 Normative messages

The purpose of the thesis is to see whether or not nudging using normative messages can affect customer decisions even when the cost of changing behaviour is high, or at least evident. The experiment included four treatment groups with four different normative messages; two descriptive and two injunctive, each divided into a strong and a weak norm. The two descriptive norms were based on actual sales figures of the products used in the experiment, and to be able to create a strong and a weak norm, the norms were based on the sales in two different three month periods[15]. The two descriptive normative messages were[16]:

- Of the products found in this shop, barely 25% were sold with a sustainability label during spring* 2019.

---

[15]March - May 2019 for the weak norm and November - January 2019/2020 for the strong norm.

[16]The normative messages are here translated from Swedish to English by the author of this thesis since the experiment was conducted in Swedish.

> * Based on the sales in all of Blomsterlandet's shops during March, April, and May 2019.

- Of the products found in this shop, almost 70% were sold with a sustainability label during winter* 2019/2020.

  > * Based on the sales in all of Blomsterlandet's shops during November, December, and January 2019/2020.

In contrast to the descriptive norm, the injunctive norm could not be based on information available at the time and thus had to be created. This was achieved through a pilot study in which 42 graduate students conducted the experiment. In addition to the control questions asking for gender, age, income, et cetera, a question was asked to incur an injunctive normative behaviour[17]. The following two messages were then created to represent the two injunctive norms:

- Barely 24% of respondents* thought that you should always buy products with a sustainability label despite often being more expensive than the alternative without a label.

- Almost 93% of respondents* thought that you should buy products with a sustainability label despite often being more expensive than the alternative without a label.

  > * Based on 42 answers in a pilot study conducted by graduate students at the School of Business, Economics and Law at the University of Gothenburg.

The normative messages include the wording "barely" and "almost" to amplify the difference between the norms and highlight the strength of the norm, weak and strong, increasing the chance that the norm is identified as weak and strong, respectively. Another option would have been to use a more neutral wording, and there is merit to both approaches. In Figure 1

---

[17]The question asked was: *"Do you think people should buy products with sustainability labels despite often being more expensive than the alternative without a label?"* and had three possible answers: "Yes, always", "Yes, sometimes", and "No". 39 students (92.8%) answered either "Yes, always" or "Yes, sometimes", thus making up the strong injunctive norm. 10 students (23.8%) answered "Yes, always", thus making up the weak injunctive norm. Only 3 students answered "No".

you can see a screenshot of the experiment, which an individual in treatment group 4 (weak injunctive norm) would have observed. The only factor that changed between groups was the message given at the top of the page.

Figure 1: Screenshot of the experiment (with the weak injunctive normative message)

4.1.2 Control questions

An advantage with the experimental design used in this study is the flexibility in collecting additional information about subjects. Thus, it is possible to analyse if certain characteristics could affect the effectiveness of the treatment. For example, there might be attitudinal differences, previous behavioural differences, et cetera, that could make it more or less likely that an individual changes behaviour due to the normative message. Moreover, the control questions allow, at least to some extent, an evaluation of whether the randomisation was successful, i.e. if groups consist of similar types of people. Subjects were asked to answer 13 questions related to both socio-demographic factors and attitudes towards sustainability and sustainable products. All questions can be found in Appendix B together with the response alternatives given to the subjects.

4.1.3 Limitations

Despite the experimental design having many advantages as described above, it does not come without issues of its own. One common limitation is self-selection bias, which is unavoidable when distributing the experiment through E-mail. In contrast to many field experiments, where it is difficult for subjects to opt-out or avoid to answer, only individuals who want to conduct the experiment chooses to do so. In this case, they had to self-select into the experiment by clicking a link distributed through E-mail. If the individuals have certain characteristics that affect the dependent variable and have a different sensitivity towards the treatments, then the results will not represent the population.

Other limitations relate to money and incentives to make truthful choices. The experiment allocated a fictitious sum of 350 SEK to all individuals up to which they could collect products from the shop. Thus, they use no out-of-pocket money which could affect the choices made. To try to counteract this, as mentioned above, a lottery was added to ensure that potential pay-off was based on a subject's choices. However, hypothetical bias cannot be ruled out and the extent of it is difficult to estimate.

4.2 Econometric strategy

4.2.1 Data collection and sample

Data was collected from a sample consisting of members of Blomsterlandet's customer club since the experiment was distributed through their weekly newsletter[18]. The newsletter, including a link to the experiment, was sent to approximately 250 000 individuals, and a total of 1019 answers (0.4%) were collected during a period of 13 days (14 April - 26 April). Although the response rate is low, it should be noted that the E-mails were not sent out with the main purpose of informing about the experiment, and moreover, the link was located at the end of the newsletter. The low response rate is not optimal, but it is difficult to say whether the individuals who conducted the experiment have certain characteristics that could affect the results. The experiment collected information in two parts; the first part collected information regarding the subject's preferred purchases and the second part collected information regarding socio-demographic factors and attitudes towards sustainability and sustainable consumption.

4.2.2 Dependent variables

It is clear that sustainable consumption is the area of interest. However, it is not entirely clear how to measure sustainable consumption as it can be done in several ways. This study evaluates three different dependent variables, which could add to the robustness of the results. The first one is simply the count of products purchased with a sustainability label for each individual. However, this variable could leave out relevant information, since there are different price premiums for different types of products. Thus, this could mean that even though a certain group purchases a fewer amount of sustainable products, their monetary value could still be equal or greater depending on what type of products are purchased. To control for this type of error, a second dependent variable is created, which sums the sustainable products purchased in monetary terms (SEK). The second dependent variable could come with its own issues however, as it is possible that individuals utilise different

---

[18]The experiment was first sent out in the newsletter for week 16 and a reminder was given in the newsletter for week 17.

amounts of the full 350 SEK which is at their disposal. The third and final dependent variable created, in an attempt to minimise both problems, is a variable of the percentage of funds spent on sustainable products in relation to total funds spent. The third variable should be most representative of how individuals spend their funds and of their choices to purchase either sustainable or non-sustainable products.

To summarise, the three dependent variables are:

(1) The count of products purchased with a sustainability label.

(2) The sum (in SEK) of products purchased with a sustainability label.

(3) The percentage of funds spent on products with a sustainability label in relation to total funds spent.

4.2.3   Descriptive statistics

A total of 1019 subjects conducted the online experiment and were randomly allocated a treatment (or no treatment). Due to the randomisation not forcing even groups, the five groups have a differing amount of subjects. However, each group holds approximately 20% of the total sample. From a graphical analysis of histograms of different variables collected through the control questions, the treatment groups appear to have similar distributions concerning these variables. For example, groups display similar income levels and attitudes towards sustainable consumption, two variables that likely could be positively correlated with sustainable purchases (see Figure 3 and 4 in Appendix A). Preventing any group from being over-represented concerning these variables is key to being able to measure only the effect of the treatment[19].

Moreover, Table 2 shows that the distribution of gender, mean household size, and percentage of garden and balcony/patio availability are comparable across groups. The similarities between the groups further support an appropriate comparison and evaluation of the treatments. An interesting feature of the data, which can be seen from Table 2, is the fact

---

[19]Analysing the correlation between these variables and sustainable purchases in the experiment show a positive correlation for agreement regarding customer responsibility (around 0.42) and close to no correlation for income level (around 0.05).

	Variable	Total	Control	Descriptive		Injunctive	
				Weak	Strong	Weak	Strong
	Observations	1019	218	199	225	198	179
	Perc. female subj.	93.3%	91.7%	92.0%	95.6%	93.9%	93.3%
	Mean househ. size (max. size 5)	2.40	2.29	2.36	2.44	2.54	2.39
	Perc. garden and balcony/patio access	70.8%	69.3%	69.3%	68.4%	74.2%	73.2%
(1)	Mean	4.9	4.8	4.9	5.0	5.2	5.0
	SD	2.5	2.5	2.4	2.5	2.7	2.5
	Median	6	5	5	6	6	6
(2)	Mean	238.6	235.7	235.9	237.7	242.0	242.7
	SD	112.7	112.3	106.9	115.0	117.9	111.6
	Median	279.4	272.0	269.4	288.5	301.1	288.6
(3)	Mean	70.5	69.3	69.9	70.2	71.4	72.0
	SD	33.0	32.7	31.5	33.5	34.5	32.7
	Median	84.1	80.6	82.1	84.3	88.2	85.3

(1): Count of sustainable products, (2): Sum (in SEK) of sustainable products,

(3): Percentage sustainable products

Table 2: Overall summary statistics

that 93% of all subjects in the experiment were female. How this could affect the results and the effectiveness of the different treatments can only be speculated around. However, the large proportion of women conducting the study is in line with the distribution among Blomsterlandet's customer club, as 87% consists of women. Moreover, the subjects of interest should arguably be people that commonly act within the context that the study is conducted in, meaning that the large proportion of women only represents the actual environment to a higher degree. Overall, the sample collected arguably consists of the individuals of interest, as these are the individuals that would normally act in the environment and thus would be

the recipients of potential interventions.

Age is fairly evenly distributed between the groups (see Figure 5 in Appendix A) with the most common age group being people aged 61-70. However, the size of the age groups are fairly even except for individuals aged 30 or below. Table 2 also presents some basic information about the three dependent variables used in the analysis. Although the mean value and the standard deviation are not the best measures considering the distribution (see for example Figure 2 in the next section), it could still be interesting to receive information about these statistics. A better measure is the median value, which is also presented. Overall, there seems to be some difference between the treatment groups, but the differences are in most cases not substantial, and whether there are statistically significant differences remains to be seen. A better representation of the overall distribution per group can be seen later in Table 3.

4.2.4   Non-normality

Next, in order to choose the appropriate econometric strategy, the distribution of the sample has to be discussed. A graphical visualisation of the data and the Shapiro-Wilk test will support the argumentation below regarding the question of normality.

The data concerns customer purchases of flowers, plants, and other garden-related items and the variables of interest are related to the amount of sustainable purchases a customer makes. Thus, in order for the normality assumption to be a reasonable assumption, it should not be common among customers to purchase very few or very many products with a sustainability label. However, this seems an unreasonable assumption, since purchases will depend much on an individual's attitude towards sustainability and sustainable products. Therefore, given that a subject does not value sustainability labelled items, it is not unlikely that this subject purchases very few of those products, and vice versa. This behaviour is exactly what can be observed in Figure 2, displaying that almost 50% of subjects in the study had a shopping cart consisting of no less than 80% sustainability labelled products. In contrast, 15% of the sample had no more than 20% of their shopping cart being sustainable products.

Shapiro-Wilk tests for the three dependent variables (see section 4.2.2) reported p-values

Distribution of the proportion of sustainable purchases in a subject's cart
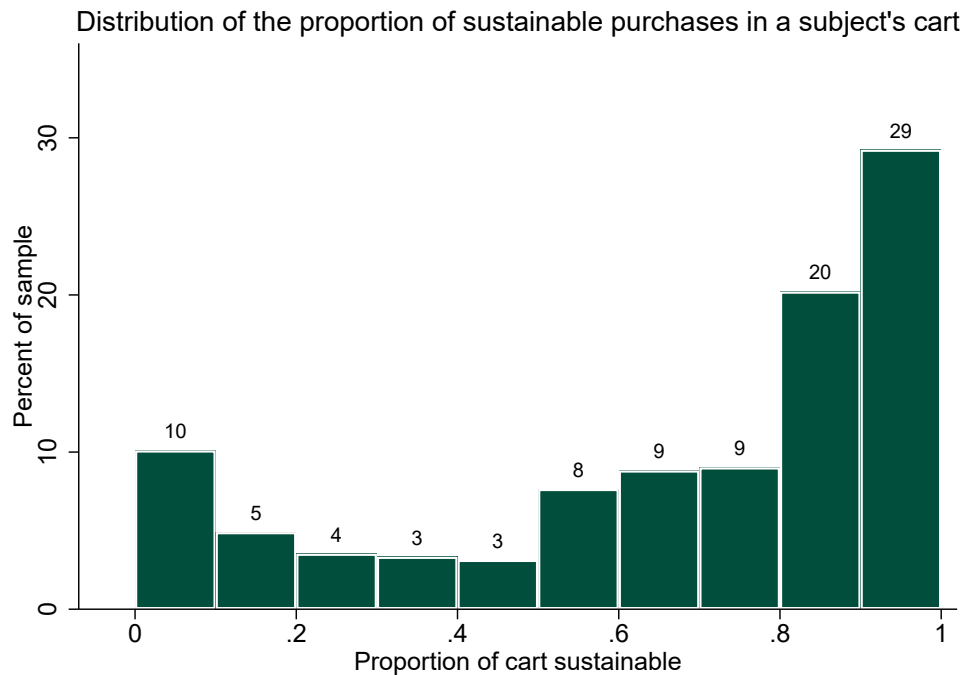
Figure 2: Overall customer behaviour regarding sustainable purchases

of close to zero (0.00000), with W = 0.942 for (1), W = 0.839 for (2), and W = 0.911 for (3). Thus, the tests for normality support the motivation given before about the data unlikely being normally distributed. In the following section, the appropriate hypothesis testing, given the current data set, will be explained.

4.2.5   Hypothesis testing

Since the normality assumption cannot reasonably be assumed to be fulfilled, a test has to be chosen which accounts for the non-normal distribution. In line with previous research dealing with similar data (e.g. Demarque et al., 2015), the analysis is conducted using a non-parametric test called the Kruskal-Wallis H test which is a rank test developed by Kruskal and Wallis (1952). Rank tests assign a value of 1 to N[20] (N being the total number of observations in the sample) based on the magnitude of the variable evaluated.

---

[20]If two or more observations have the same value, i.e. a tie in rank, it is solved by allocating an average rank to the observations involved.

More technically, calculate a value H according to the formula (Conover, 1999):

$$H = \frac{1}{S^2} \left\{ \sum_{i=1}^{k} \frac{R_i{}^2}{n_i} - \frac{N(N+1)^2}{4} \right\} \tag{9}$$

where

$$S^2 = \frac{1}{N-1} \left(\sum_{all\ ranks} R(X_{ij})^2 - N\frac{(N+1)^2}{4} \right) \tag{10}$$

and

k = the number of treatment groups, including control,

n_i = the number of observations in group i,

$N = \sum n_i$ is the total number of observations,

$R(X_{ij})$ = the rank for observation j in group i, and

$R_i = \sum R(X_{ij})$ = the sum of ranks in group i.

Furthermore, Conover (1999) lists four assumptions to be fulfilled:

1. All samples are random samples from their respective populations.

2. In addition to independence within each sample, there is mutual independence among the various samples.

3. The measurement scale is at least ordinal.

4. Either the k population distribution functions are identical, or else some of the populations tend to yield larger values than other populations do.

The first two assumptions should be fulfilled given the randomisation of treatments conducted in the experiment and the data collection. The third assumption is fulfilled given the dependent variables, which are cardinal and continuous. The last assumption is the most difficult to prove, as it is difficult to test whether it is fulfilled or not. However, it seems reasonable that the different groups, given the randomisation, should have had a similar distribution had there not been any treatments since all subjects come from the same source. This should be further supported by the large sample size in this experiment, reducing the chance of certain groups exhibiting different behaviour regardless of a treatment. Thus, it

seems reasonable that any difference in distribution between the groups should be caused by a treatment.

The sampling distribution of H is approximately chi-squared with degrees of freedom $k-1$ and thus a large value of H would suggest a difference between the groups. The null hypothesis and alternative hypothesis of the test are as follows:

H_0 : All of the k sample distribution functions are identical.

H_1 : At least one of the groups tends to yield larger observations than at least one of the other groups.

Thus, if we can reject the null hypothesis, it would suggest that at least one treatment causes individuals to change behaviour. However, the Kruskal-Wallis H test cannot determine which treatment(s) is causing the change in behaviour. Therefore, in the event that the null hypothesis is rejected, further tests have to be conducted to gain insight into which treatment(s) caused the effect. If the null hypothesis cannot be rejected, no further tests are conducted.

Dunn's test, developed by Dunn (1964), builds on the results collected from the Kruskal-Wallis H test by using the mean values of the ranks given in the first test. Specifically, a pairwise comparison between group i and k will yield a value z_{ik} as follows (Dinno, 2015):

$$z_{ik} = \frac{y_{ik}}{\sigma_{ik}} \tag{11}$$

where

$y_{ik} = \overline{R_i} - \overline{R_k}$ with $\overline{R_i} = \frac{R_i}{n_i}$, and

σ_{ik} is the standard deviation of y_{ik}, defined as:

$$\sigma_{ik} = \sqrt{\left\{ \frac{N(N+1)}{12} - \frac{\sum_{s=1}^{r} \tau_s{}^3 - \tau_s}{12(N-1)} \right\} \left(\frac{1}{n_i} + \frac{1}{n_k} \right)} \tag{12}$$

where the second expression within the curly braces corrects for ties,

r is the number of tied ranks, and

τ_s is the number of observations tied at the sth specific tied value.

In order to correctly adjust for the multiple comparisons made, which increase the probability of finding a significant result, statistical software allows for a range of options, including Bonferroni adjustment, Šidák (1967) adjustment, and Holm (1979) adjustment. The adjustments through the statistical software allow for a direct interpretation of the p-values, having been corrected for the increase in probability of making a Type I error (i.e. incorrectly rejecting the null hypothesis). A Šidák adjustment is used here. Finally, the null hypothesis for each pairwise comparison is "that the probability of observing a random value in the first group that is larger than a random value in the second group equals one half" (Dinno, 2015, p. 294). Thus, if the null hypothesis is rejected, it suggests that there is a difference in behaviour between the two groups being compared. However, and importantly, the test does not tell the direction of the behavioural change. This has to be analysed through a comparison of the distributions.

5   Results and Analysis

5.1   Full sample

To get an idea of how the treatments have affected the behaviour of individuals concerning sustainable and non-sustainable purchases, Table 3 provides some descriptive statistics of the three dependent variables. As can be seen, the distribution between groups does not differ much, perhaps suggesting unsuccessful nudges. It is also clear that many individuals are conforming to the sustainable behaviour, and from the distribution of the control group, it seems as though a large proportion of subjects act sustainable even without a treatment. This could clearly have an effect on the successfulness of the treatments since the amount of subjects that can change behaviour is reduced.

The main part of the results regard the nonparametric tests testing the hypothesis of a difference between the groups. The first test is conducted using dependent variable (1), i.e. the count of sustainable products purchased. As we can see from Table 4, the null hypothesis of no difference between the groups cannot be rejected (p-value = 0.3226). However, as argued before, a better measure of sustainable purchases is the sum, in monetary terms,

N = 1019				Percentile		
(1)	Mean	SD	10th	25th	Median	75th	90th
Control	4.8	2.5	1	3	5	7	8
Weak des.	4.9	2.4	1	4	5	7	8
Strong des.	5.0	2.5	1	3	6	7	8
Weak inj.	5.2	2.7	0	4	6	7	8
Strong inj.	5.0	2.5	0	4	6	7	8
(2)							
Control	235.7	112.3	34.9	169.6	272.0	333.4	343.6
Weak des.	235.9	106.9	49.9	189.6	269.4	323.5	344.2
Strong des.	237.7	115.0	34.9	169.6	288.5	338.4	344.2
Weak inj.	242.0	117.9	0	179.6	301.1	334.4	347.7
Strong inj.	242.7	111.6	0	189.5	288.6	329.2	344.2
(3)							
Control	69.3	32.7	10.4	49.2	80.6	100	100
Weak des.	69.9	31.5	17.7	55.6	82.1	100	100
Strong des.	70.2	33.5	10.0	53.9	84.3	100	100
Weak inj.	71.4	34.5	0	54.3	88.2	100	100
Strong inj.	72.0	32.7	0	57.2	85.3	100	100

(1): Count of sustainable products as dependent variable.

(2): Sum (in SEK) of sustainable products as dependent variable.

(3): Percentage of sustainable products as dependent variable.

Table 3: Distribution of dependent variables for the full sample.

of the items purchased with a sustainability label. However, from Table 4 it is clear that the null hypothesis cannot be rejected in this case either (p-value = 0.6161). Finally, testing with perhaps the most appropriate measure of sustainable purchases, the percentage spent on sustainable products in relation to total funds spent, the null hypothesis still cannot be rejected (p-value = 0.6231). Thus, no evidence supports the hypotheses of the normative

messages being effective ways to change sustainable behaviour, or that a stronger norma-
tive behaviour is more effective.

N = 1019		Rank Sum		
Treatment Group	Obs.	(1)	(2)	(3)
Control	218	107070.00	108750.50	108075.00
Weak descriptive	199	97627.50	97085.50	97840.50
Strong descriptive	225	115664.50	115783.50	115329.50
Weak injunctive	198	107870.50	105072.50	105054.00
Strong injunctive	179	91457.50	92998.00	93391.00
Chi-squared with ties		4.673	2.660	2.621
P > chi-squared		0.3226	0.6161	0.6231

(1): Count sustainable products as dependent variable

(2): Sum (in SEK) sustainable products as dependent variable

(3): Percentage sustainable products as dependent variable

Table 4: Kruskal-Wallis equality-of-populations rank test among the full sample.

One can only theorise why the results are insignificant, but one obvious possible reason is
the price premium that subjects had to pay when choosing the sustainable option. However,
this explanation seems unreasonable considering the large proportion of people that, judg-
ing by the control group, already act sustainable regardless of any treatment. Perhaps the
more probable explanation is that the existence of a large amount of people that already act
sustainable reduced the possibility of measuring an effect of the treatments.

5.2   Importance

Although not the main purpose of the thesis, the insignificant results from the previous
section leaves the question of whether there are features that could increase the likelihood
of a nudge being successful in altering behaviour. Thus, the subsequent analyses could still
be interesting to conduct, and could potentially contribute to the empirical evidence on what
environments a nudge will likely be most successful in. Based on the theory proposed by

Löfgren and Nordblom (2020), the likelihood of a preference nudge (the type of nudge used here) being successful will be affected by the importance of a choice and an individual's confidence that the choice they make inattentively is the preferred one.

The first parameter, importance, is according to the theory from Löfgren and Nordblom (2020) thought to have a negative relationship with the likelihood of a nudge being successful, i.e. a less important choice is thought to increase the likelihood of a nudge being successful. One group that might view the choice between a sustainable product and a non-sustainable product as unimportant are individuals who do not believe that customers have a responsibility to make environmentally friendly choices or choices that impact workers' working conditions. However, from the control questions asking these two questions, only 21 respondents answered that they did not believe customers had a responsibility to at least one of the questions. Thus, the sample size is not large enough to make an appropriate analysis.

Another group that could see the choice between a sustainable and a non-sustainable product as less important is a group consisting of individuals with a higher than average income. The reason for this group potentially considering the choice less important is because the cost of changing behaviour is lower in relative terms compared to lower income individuals. The group was based on higher than average household income[21] within the sample, and was adjusted for the number of household members as well as the number of children below the age of 18[22]. Thus, a Kruskal-Wallis H test was conducted for this group, and from Table 5 it can be seen that the results are not quite statistically significant (p-value = 0.1696, 0.1877, and 0.1583, respectively for each dependent variable). Further pairwise comparisons should, as explained before, only be made when the results are statistically significant.

---

[21]Household income is arguably the most appropriate measure for purchases regarding flowers, plants, and other garden-related items as they are most often purchased for a household and not an individual.

[22]Information on disposable income was collected in brackets (see appendix B), and thus the measure was based on the mean income for these brackets, and adjusted based on an equivalence scale from Statistics Sweden (2019). Although full information on income levels are not observed, the measure should still be appropriate to divide the full sample into high and low income individuals.

N = 431		Rank Sum		
Treatment Group	Obs.	(1)	(2)	(3)
Control	100	19483.50	19412.50	19585.00
Weak descriptive	95	20482.50	20096.00	19905.50
Strong descriptive	83	18423.50	19445.50	19283.00
Weak injunctive	78	18795.00	18029.00	18412.50
Strong injunctive	75	15911.50	16113.00	15910.00
Chi-squared with ties		6.425	6.157	6.604
P > chi-squared		0.1696	0.1877	0.1583

(1): Count sustainable products as dependent variable

(2): Sum (in SEK) sustainable products as dependent variable

(3): Percentage sustainable products as dependent variable

Table 5: Kruskal-Wallis equality-of-populations rank test among above average income customers.

5.3   Confidence

The second parameter, confidence, is suggested to positively correlate with the successfulness of a preference nudge. A group that reasonably could have a higher confidence in making the correct choice between a sustainable and a non-sustainable product are individuals who often purchase sustainable goods, or at least pay attention to sustainability labels when shopping. Thus, from the control questions asking how often the respondent thinks about sustainability labels when shopping, either garden-related products or other types of products, about 49% of the sample answered "Always" or "Most often" to both questions[23]. This group can be analysed separately, with answers "Always" or "Most often" being proxies for a higher confidence regarding a choice between sustainable and non-sustainable products. Whether this is a good proxy for confidence is debatable, but it could reasonably be thought that individuals who often think about sustainability labels also posses a greater knowledge (or at least to the individual subjective knowledge) about those sort of products. Another interpretation could be that the subsample consists of individuals with pro-environmental

---

[23]One was aimed at garden-related items and one was aimed at other types of items, see Appendix B.

N = 431			Percentile				
(2)	Mean	SD	10th	25th	Median	75th	90th
Control	229.6	111.4	59.9	139.8	259.4	326.8	343.8
Weak des.	248.6	100.2	69.9	204.5	274.4	329.2	344.4
Strong des.	258.9	105.0	99.8	204.6	309.4	339.3	344.3
Weak inj.	259.9	104.4	69.8	234.5	303.9	338.4	349.2
Strong inj.	241.9	115.9	0	184.6	274.4	334.4	344.2
(3)							
Control	68.2	32.3	16.9	47.6	76.1	100	100
Weak des.	73.6	29.2	20.6	60.0	84.4	100	100
Strong des.	76.5	30.3	29.4	62.9	91.4	100	100
Weak inj.	77.5	30.4	19.4	73.1	88.6	100	100
Strong inj.	71.7	33.4	0	58.3	82.4	100	100

(2): Sum (in SEK) of sustainable products as dependent variable.

(3): Percentage of sustainable products as dependent variable.

Table 6: Distribution of dependent variables among above average income participants by group.

preferences which could affect behaviour, perhaps due to individuals wanting to uphold an image of being sustainable compared to the rest of the sample, which could affect the effectiveness of the treatments. It is also possible that these individuals have a higher willingness to pay for behavioural change because they either value the sustainable aspect of the product to a higher degree or draw more utility from conforming to a sustainable behaviour (see the discussion on norm-dependent utility in section 3.2).

In Table 7 we see the results from the Kruskal-Wallis H test for the subgroup of individuals defined as "confident". Interestingly, there are signs of statistical significance (p-value = 0.0816, 0.0242, and 0.0779, respectively), which suggests that we can reject the null hypothesis of all groups having the same behaviour, i.e. one or several groups seem to purchase a different amount of sustainable products compared to one or several other groups. Multiple pairwise comparisons, using Dunn's test, have to be made in order to gain insight into which

group(s) are showing a different behaviour. The test only provides information about which groups are different from each other, but not how the two groups are behaving differently. However, it is possible to compare the distribution of values if there are any groups that behave differently to get an idea of how the treatments have affected behaviour.

N = 497		Rank Sum		
Treatment Group	Obs.	(1)	(2)	(3)
Control	106	25770.50	26453.50	26225.50
Weak descriptive	93	19978.00	19164.00	19939.50
Strong descriptive	113	29459.50	30341.00	30200.50
Weak injunctive	96	25579.50	24984.00	24380.00
Strong injunctive	89	22965.50	22810.50	23007.50
Chi-squared with ties		8.288	11.222	8.403
P > chi-squared		0.0816	0.0242	0.0779

(1): Count sustainable products as dependent variable

(2): Sum (in SEK) sustainable products as dependent variable

(3): Percentage sustainable products as dependent variable

Table 7: Kruskal-Wallis equality-of-populations rank test among "confident" customers.

Interestingly, the difference in behaviour does not seem to be between any treatment group and the control group, but instead between the three treatments; strong descriptive, weak injunctive, and strong injunctive; individually against the weak descriptive norm. In Table 8 we can see the p-values from the comparisons between the weak descriptive norm and the four remaining conditions (see the full output in Table 12, 13, and 14 in Appendix C). The results, suggesting a weak statistically significant difference for the injunctive norms and a quite strong statistically significant difference for the strong descriptive norm towards the weak descriptive norm, has to be further analysed to determine in what way they differ. Since there was no significant difference compared to the control condition, the results should reasonably be a combination of an increase and a decrease in the groups compared.

Thus, further investigation is needed[24].

N = 497	Weak descriptive norm		
Compared with	(1)	(2)	(3)
Control	0.5589	0.1533	0.3753
Strong descriptive	0.0953	0.0094	0.0304
Weak injunctive	0.0571	0.0465	0.2180
Strong injunctive	0.1764	0.0879	0.1444

(1): Count sustainable products as dep. var.

(2): Sum (in SEK) sustainable products as dep. var.

(3): Percentage sustainable products as dep. var.

Table 8: P-values for Dunn's test comparing the weak descriptive norm with the remaining four conditions.

From Table 9 it can be seen that quite a large proportion of people spent all funds on sustainable products. Furthermore, it does appear that the significant results are, as expected, a combination of a decrease at one end and an increase at the other. Judging by the mean, median, and the 10th and 25th percentile, the weak descriptive norm has caused a decrease in sustainable behaviour compared to the control group, which in combination with an increase in the behaviour from the other treatment groups has generated the statistically significant results. Therefore, it is not possible to say that any treatment was successful, seeing as the results are mainly driven by an unsuccessful weak descriptive norm, causing the opposite effect of what was intended.

The results suggesting that the weak descriptive norm caused an opposite behaviour than what was intended by the nudge is something that has been observed before. It was observed in the article by Chabe-Ferret et al. (2019) studying the effect of nudging on farmers' water consumption. Sherif et al. (1958) observed the behaviour already in 1958 and coined it the "boomerang-effect". Interestingly, based on the results from this study, the

---

[24]Notice that the investigation below is only conducted for dependent variable (2) and (3) because they are arguably the most relevant and will allow for most variation.

N = 497			Percentile				
(2)	Mean	SD	10th	25th	Median	75th	90th
Control	285.1	81.2	189.5	254.4	321.4	339.2	344.3
Weak des.	266.2	83.3	154.7	233.7	289.5	333.5	344.3
Strong des.	295.8	66.1	219.5	269.4	323.6	339.3	344.3
Weak inj.	288.5	77.2	189.5	246.9	319.4	339.3	349.1
Strong inj.	292.3	68.8	209.7	269.4	318.5	339.3	344.3
(3)							
Control	83.5	30.0	54.3	73.9	93.5	100	100
Weak des.	79.5	24.4	50.0	68.6	86.1	100	100
Strong des.	87.3	18.7	65.2	79.7	97.2	100	100
Weak inj.	84.3	22.8	54.3	71.9	95.7	100	100
Strong inj.	87.0	19.3	68.6	82.4	94.1	100	100

(2): Sum (in SEK) of sustainable products as dependent variable.

(3): Percentage of sustainable products as dependent variable.

Table 9: Distribution of dependent variables among "confident" participants by group.

same type of behaviour did not occur for the weak injunctive norm, despite the percentages in the normative messages (i.e. weakness of the norm) being almost identical. Whether this means that injunctive norms are better at minimising an unintended behaviour from nudges is difficult to say. However, the question is interesting and should be evaluated further in the future.

It might not actually be surprising to see a negative effect of the weak descriptive norm as the actual majority behaviour is the opposite of what the message is informing about. Thus, the result could be driven by an idea that people want to act according to the majority behaviour, which could be connected to the theoretical framework on norm-dependent utility explained in section 3.2. Perhaps by realising that few individuals act in a sustainable way, it causes the individual to feel better about not actually displaying a sustainable behaviour. This could also be an explanation for why the weak injunctive norm did not have

the same effect. The weak injunctive norm, informing about 24% of people believing the sustainable behaviour is appropriate, does not explicitly translate to the rest, i.e. 76% of people, believing that the non-sustainable behaviour is appropriate.

Another explanation is related to the discussion in section 3.2 about intra- and inter-personal perspectives. As suggested by Cialdini and Trost (1998), descriptive norms are used as information about making the most effective decision while an injunctive norm also relates to a certain social responsibility. Thus, since the weak descriptive norm shows a majority behaviour in the opposite direction of what was intended, individuals might use this information to argue for the unintended behaviour being the most effective. In contrast, although the weak injunctive norm might show that the intended behaviour is not the majority behaviour, there is also a social responsibility regarding sustainable consumption that individuals care about. Nonetheless, these are only theories about the mechanisms underlying the results, and must be studied to a greater extent in different contexts in order to draw any clear conclusions.

5.4   Non-sustainable customers

As stated before, a very large proportion of the sample acts in a sustainable way, and by observing the control group it seems as though a large proportion of the sample would have acted sustainable regardless of any treatment. Thus, it would be interesting to analyse a group consisting instead of individuals that would normally not act in a sustainable way, which is also arguably the group of interest for policymakers and companies, since these individuals have the possibility of changing their behaviour.

However, the issue is with identifying these individuals since one cannot *ex post* identify what behaviour is caused by the treatments and what behaviour is not. One potential identification method would be to use the control questions asked at the end of the experiment. More specifically, perhaps it would be reasonable to consider the opposite group of the group analysed in section 5.3, i.e. the 51% of the sample that answered either "Sometimes" or "Never" to the question of whether one thinks about sustainability labels when shopping. By observing the distribution among these individuals, it seems to be more focused around a less sustainable behaviour, which would support the motivation that this group has pos-

sibilities for behavioural change. For example, comparing the median values in Table 10 to those in Table 9 and Table 3, we see smaller values for all treatment groups.

N = 522			Percentile				
(2)	Mean	SD	10th	25th	Median	75th	90th
Control	188.9	117.8	0	91.9	207.0	301.6	338.5
Weak des.	209.3	118.0	0	84.8	236.7	319.3	339.4
Strong des.	179.0	124.0	0	69.8	198.8	294.0	338.4
Weak inj.	198.2	132.4	0	49.9	251.6	314.1	339.2
Strong inj.	193.6	123.9	0	99.0	226.6	304.4	336.8
(3)							
Control	55.9	34.3	0	28.4	61.1	85.9	100
Weak des.	61.4	34.6	0	25.5	71.0	92.8	100
Strong des.	53.0	36.3	0	19.7	61.1	85.7	100
Weak inj.	59.2	39.0	0	16.4	75.9	94.3	100
Strong inj.	57.1	36.3	0	27.6	66.6	91.4	100

(2): Sum (in SEK) of sustainable products as dependent variable.

(3): Percentage of sustainable products as dependent variable.

Table 10: Distribution of dependent variables among "non-sustainable" participants by group.

However, as can be seen from Table 11, no statistically significant results can be found from the Kruskal-Wallis H test, and thus the null hypothesis cannot be rejected. Again, it is only possible to theorise on causes of the results. In the analysis of the results containing the full sample, it was argued that it was quite a low possibility of the price premiums being the main cause of the insignificant results. However, for this sample it might seem more reasonable, i.e. the cost of changing behaviour might have been too high for an individual to choose the sustainable option instead of the non-sustainable. This could also be affected by the preferences for sustainable products among the subsample, since a lower pro-sustainable attitude might mean a lower willingness to pay for behavioural change. This could be explained by the theoretical framework on norm-dependent utility proposed by Krupka and Weber

(2013) in section 3.2 where a parameter γ identified personal preferences to conforming to a normative behaviour. This parameter could reasonably be affected by the type of normative behaviour that is to be conformed to, and an individuals preferences to that context. In other words, individuals with low initial preferences towards sustainable consumption could, according to the theoretical framework, gain less utility from acting in accordance with the normative behaviour, thus decreasing the chance of a nudge being successful.

Moreover, the insignificant results could possibly also be explained by the confidence parameter in the theoretical framework proposed by Löfgren and Nordblom (2020) explained in section 3.1. The confidence among these individuals, in contrast to the subgroup in the previous section, might be low, suggesting that a pure nudge would have been more successful in changing behaviour.

N = 522		Rank Sum		
Treatment Group	Obs.	(1)	(2)	(3)
Control	112	28176.50	28350.00	28290.00
Weak descriptive	106	29390.00	29730.00	29507.00
Strong descriptive	112	27952.00	27308.00	27219.00
Weak injunctive	102	28291.50	27611.00	28025.00
Strong injunctive	90	22693.00	23504.00	23462.00
Chi-squared with ties		3.884	3.959	4.225
P > chi-squared		0.4219	0.4115	0.3764

(1): Count sustainable products as dependent variable

(2): Sum (in SEK) sustainable products as dependent variable

(3): Percentage sustainable products as dependent variable

Table 11: Kruskal-Wallis equality-of-populations rank test among "non-sustainable" customers.

6 Conclusions

This study aimed to answer the question of whether descriptive and injunctive normative messages could increase the consumption of sustainability labelled products, despite a change in behaviour being costly. Through an interactive online experiment involving real products, four treatments consisting of different normative messages were compared to a control group. The main results showed no statistical significance regarding the effect of the treatments compared to the control group.

The insignificant results for the full sample asked questions of whether there could be subgroups, supported by the theoretical framework, that would be more susceptible to the normative messages. Firstly, a group of above average income individuals were analysed to see whether lower importance regarding the choice between a sustainable and a non-sustainable product could increase the likelihood of the nudges being successful. However, results were not quite statistically significant and any conlusions regarding the impact of importance are difficult to draw.

Secondly, using a proxy for confidence and analysing a subgroup of "confident" customers, results gave some support for a difference between treatment groups. The observed difference was between the weak descriptive norm and the three remaining treatments. Further investigation suggested a combination of a "boomerang-effect" (i.e. a negative effect on sustainable behaviour) for the weak descriptive norm and an increase in sustainable behaviour for the other treatments. However, it is wise to be cautious when interpreting the results since there was no statistically significant difference between any of the treatments and the control group, thus not supporting the two first hypotheses of the thesis. The difference between the weak and the strong descriptive norms show some support for the third hypothesis of the strength of a normative behaviour affecting the successfulness of the nudge. However, and interestingly, the same relationship could not be observed for the weak and strong injunctive norms.

The change in results when examining the subgroup of "confident" participants is an interesting feature of its own. From the theoretical framework developed by Löfgren and Nordblom (2020), confidence in one's ability to make the most optimal choice is proposed to

increase the likelihood that a preference nudge is successful. Thus, the change in results between the full sample and the sample of "confident" customers could support this idea. However, it is impossible to rule out pro-environmental preferences as a driving force of the significant results, which could be similar to the results from Defrancesco et al. (2008) showing normative information being effective only on farmers already showing pro-environmental behaviour. Disentangling the confidence parameter with aspects such as pro-environmental behaviour is reasonably a difficult task, and could be an avenue for future research to explore.

Finally, an analysis consisting of "non-sustainable" customers was conducted in an attempt to analyse individuals that could change behaviour, seeing as this group is arguably the group of interest. However, results were insignificant and could possibly be explained by several factors, and perhaps a relationship between the factors. For example, the high cost of changing behaviour cannot be ruled out as a contributing factor. Additionally, the group's overall preferences towards sustainable consumption could also impact the successfulness of the nudges.

It is interesting to discuss the external validity of the results, and also compare and connect to similar studies that find different results. Overall, external validity might be difficult to achieve in studies where mechanisms affecting the treatments are the same mechanisms that will change depending on what kind of environment that is studied. Thus, parameters such as importance and confidence could change for each decision environment that is studied, and most likely also for each individual within a given sample.

For example, comparing to the study by Demarque et al. (2015), it is reasonable to assume that individuals make choices regarding grocery store items more frequently, and often also choices regarding the same type of item (e.g. bananas, coffee, soap), compared to an environment consisting of flowers and plants. Thus, it is possible that an increase in the frequency of a decision can be related to the confidence an individual has of one's ability to make the optimal choice inattentively. This idea could for example be driven by how an individual can learn from experience, as well as a decreased chance of the individual forgetting about his/her experience when faced with the same choice again.

Another aspect that could explain the difference between this study and the study by

Demarque et al. (2015) are the prices, and perhaps mainly the price premiums explored in this study. It is not entirely clear whether the eco-labelled products used in Demarque et al. (2015) had price premiums compared to non-labelled products, and it is reasonable to think that price premiums could impact the importance of a decision, which was the other factor suggested by Löfgren and Nordblom (2020) to affect the successfulness of a nudge.

Much of the discussion consists mainly of theories about potential mechanisms related to either normative messages or nudging in general and should be interpreted with caution. However, theories allow for new avenues to be explored in order to test whether or not there is any support in practice. Thus, there are many directions in which future research can go, and many questions still remain unanswered about the effectiveness of nudging. For example, researchers can explore the confidence parameter and the impact of importance for the successful of nudging. Furthermore, more research is needed to analyse nudging in environments where behavioural change is costly, which was one purpose of this study. Many pieces of important research are still missing in order to complete the puzzle regarding nudging interventions.

References

Alavanja, M., Hoppin, J., & Kamel, F. (2004). Health effects of chronic pesticide exposure: Cancer and neurotoxicity* 3. *Annual Review of Public Health*, *25*, 155–197 (cited on page 1).

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, *95*(9-10), 1082–1095 (cited on pages 1, 2, 5, 7).

Andersson, M., & von Borgstede, C. (2010). Differentiation of determinants of low-cost and high-cost recycling. *Journal of Environmental Psychology*, *30*(4), 402–408 (cited on pages 1, 2, 8).

Asch, S. E. (1955). Opinions and social pressure. *Scientific American*, *193*(5), 31–35 (cited on page 13).

Bergquist, M., Nilsson, A., & Schultz, W. P. (2019). A meta-analysis of field-experiments using social norms to promote pro-environmental behaviors. *Global Environmental Change, 2019, Vol. 59*, *59* (cited on page 5).

Carrico, A. R., & Riemer, M. (2011). Motivating energy conservation in the workplace: An evaluation of the use of group-level feedback and peer education. *Journal of Environmental Psychology*, *31*(1), 1–13 (cited on page 7).

Chabe-Ferret, S., Le Coent, P., Reynaud, A., Subervie, J., & Lepercq, D. (2019). Can we nudge farmers into saving water? evidence from a randomised experiment. *European Review Of Agricultural Economics*, *46*(3), 393–416 (cited on pages 2, 6, 35).

Chen, K., & Deng, T. (2016). Research on the green purchase intentions from the perspective of product knowledge. *Sustainability*, *8*(9), 1–16 (cited on page 8).

Cialdini, R. B. (2003). Crafting normative messages to protect the environment. *Current Directions in Psychological Science*, *12*(4), 105–109 (cited on page 1).

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. In D. T. Gilbert & S. T. Fiske (Eds.), *The handbook of social psychology: Vol. 2* (pp. 151–192). Boston, McGraw-Hill. (cited on pages 14, 37).

Conover, W. J. (1999). *Practical nonparametric statistics*. Wiley. (cited on page 26).

Dayan, E., & Bar-Hillel, M. (2011). Nudge to nobesity ii: Menu positions influence food orders. *Judgment and Decision Making*, *6*(4), 333–342 (cited on page 5).

Defrancesco, E., Gatto, P., Runge, F., & Trestini, S. (2008). Factors affecting farmers' participation in agri-environmental measures: A northern italian perspective. *Journal of Agricultural Economics*, *59*(1), 114–131 (cited on pages 6, 41).

Demarque, C., Charalambides, L., Hilton, D. J., & Waroquier, L. (2015). Nudging sustainable consumption: The use of descriptive norms to promote a minority behavior in a realistic online shopping environment. *Journal of Environmental Psychology*, *43*(100), 166–174 (cited on pages 1, 2, 9, 10, 16, 25, 41, 42).

Dinno, A. (2015). Nonparametric pairwise multiple comparisons in independent groups using dunn's test. *The Stata Journal*, *15*(1), 292–300 (cited on pages 27, 28).

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics*, *6*(3), 241–252 (cited on page 27).

Fairtrade International. (2017). *Fairtrade standard for flowers and plants*. https://files.fairtrade.net/standards/Flowers-and-Plants-Standard-HL-EN.pdf. (cited on pages 1, 4)

Farrow, K., Grolleau, G., & Ibanez, L. (2017). Social norms and pro-environmental behavior: A review of the evidence. *Ecological Economics*, *140*, 1–13 (cited on page 5).

Groot, J., Abrahamse, W., & Jones, K. (2013). Persuasive normative messages: The influence of injunctive and personal norms on using free plastic bags. *Sustainability*, *5*(5), 1829–1844 (cited on page 8).

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70 (cited on page 28).

Holmberg, U., & Robertson, F. (2018). *Kundkännedomsundersökning 2018* [Kantar SIFO]. Kantar SIFO. https://fairtrade.se/wp-content/uploads/2020/04/Fairtrade-K%C3%A4nnedomsunders%C3%B6kning-rapport-2017.pdf. (cited on pages 1, 4)

Jacobson, R., Mortensen, C., & Cialdini, R. (2011). Bodies obliged and unbound: Differentiated response tendencies for injunctive and descriptive social norms. *Journal of personality and social psychology*, *100*, 433–48 (cited on page 14).

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics †. *American Economic Review*, *93*(5), 1449–1475 (cited on page 10).

Kallbekken, S., & Sælen, H. (2013). 'nudging' hotel guests to reduce food waste as a win–win environmental measure. *Economics Letters*, *119*(3), 325–327 (cited on page 5).

KRAV. (2019). *Hållbarhetsrapport 2019*. https://www.krav.se/wp-content/uploads/2020/04/h-rapport_2019_webb.pdf. (cited on pages 1, 4)

Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, *11*(3), 495–524 (cited on pages 14, 38).

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*(260), 583–621 (cited on page 25).

Kuhfuss, L., Préget, R., Thoyer, S., & Hanley, N. (2016). Nudging farmers to enrol land into agri-environmental schemes: The role of a collective bonus. *European Review of Agricultural Economics*, *43*(4), 609–636 (cited on pages 1, 5).

Lernoud, J., & Willer, H. (2017). The organic and fairtrade market 2015 (J. Lernoud & H. Willer, Eds.). In J. Lernoud & H. Willer (Eds.), *The world of organic agriculture 2017*. Frick and Bonn, FiBL; IFOAM - Organics International. (cited on page 1).

Löfgren, Å., & Nordblom, K. (2020). A theoretical framework of decision making explaining the mechanisms of nudging. *Journal of Economic Behavior & Organization*, *174*, 1–12 (cited on pages 3, 10–12, 31, 39, 40, 42).

Oceja, L., & Berenguer, J. (2009). Putting text in context: The conflict between pro-ecological messages and anti-ecological descriptive norms. *Spanish Journal of Psychology*, *12*(2), 657–666 (cited on page 7).

Schultz, P. W. (1999). Changing behavior with normative feedback interventions: A field experiment on curbside recycling. *Basic and Applied Social Psychology*, *21*(1), 25–36 (cited on page 8).

Sherif, M., Taub, D., & Hovland, C. I. (1958). Assimilation and contrast effects of anchoring stimuli on judgments. *Journal of Experimental Psychology*, *55*(2), 150–155 (cited on pages 6, 35).

Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal dis-
      tributions. *Journal of the American Statistical Association*, *62*(318), 626–633 (cited
      on page 28).

Smith, A. (1759/2002). *Adam Smith: The theory of moral sentiments* (K. Haakonssen, Ed.).
      Cambridge University Press. (cited on page 13).

Statistics Sweden. (2019). *Statistikens framställning: Inkomster och skatter*. https://www.
      scb.se/contentassets/894d4afa900f4123aece341cd10e59e2/he0110_staf_2018_jl190918.
      pdf. (cited on page 31)

Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth,
      and happiness* (Rev. and expanded ed.). New York: Penguin Books. (cited on pages 4,
      5).

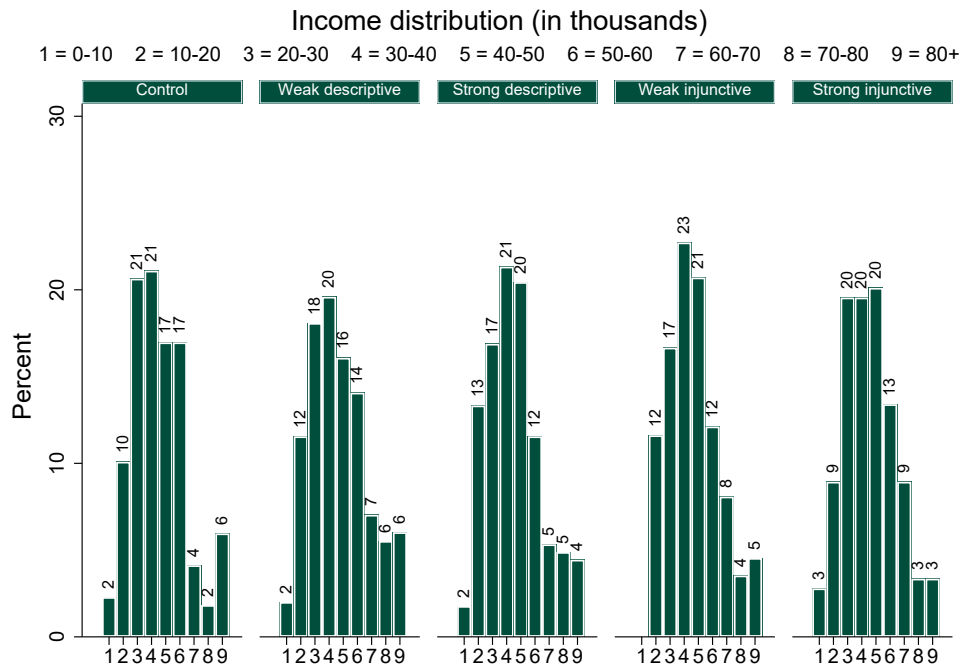# Appendix A   Descriptive statistics

Figure 3: Household income levels among participants (in thousands) per treatment group

Responsibility among customers to make environmentally friendly choices
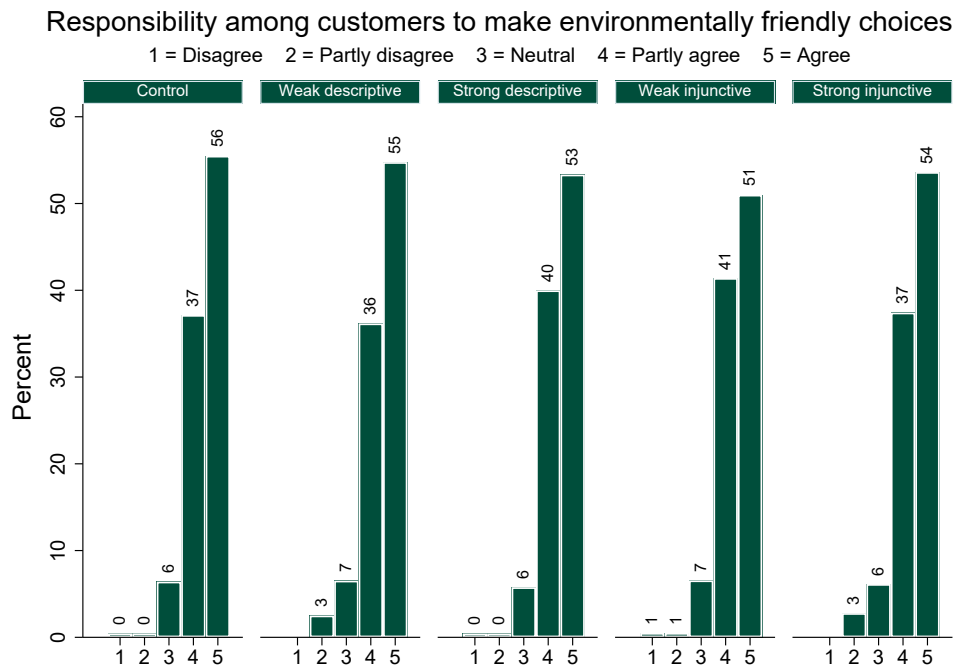1 = Disagree   2 = Partly disagree   3 = Neutral   4 = Partly agree   5 = Agree

Figure 4: How participants view customer responsibility related to environmentally friendly consumption, for each treatment group

Age distribution
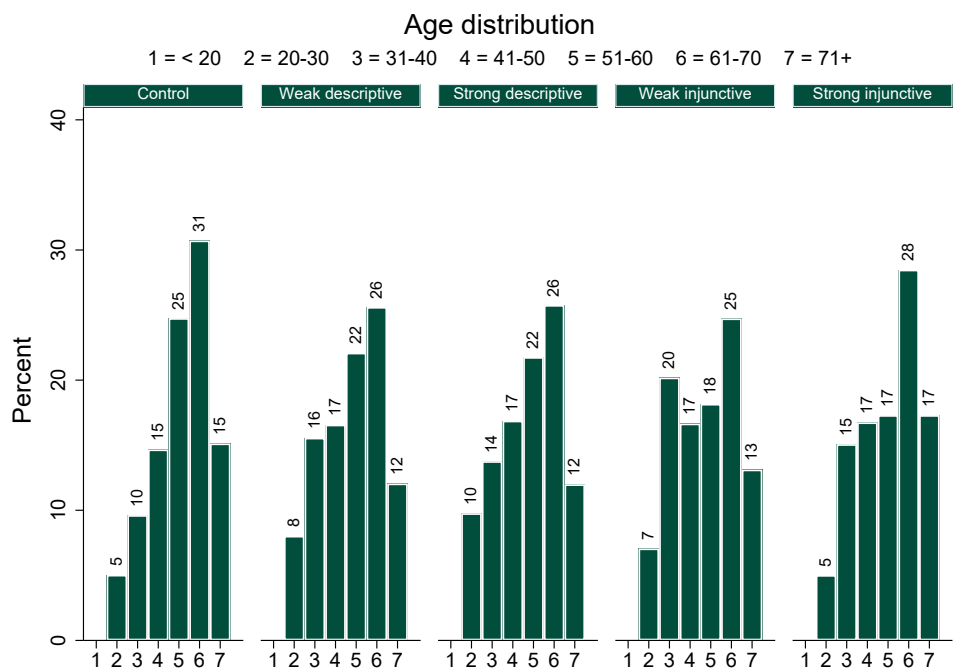1 = < 20   2 = 20-30   3 = 31-40   4 = 41-50   5 = 51-60   6 = 61-70   7 = 71+

Figure 5: Age distribution per treatment group

Appendix B   Control questions

All questions have been translated from Swedish to English by the author.

Gender

☐ Male   ☐ Female   ☐ Other

Age

☐ Below 20   ☐ 20-30   ☐ 31-40   ☐ 41-50   ☐ 51-60   ☐ 61-70   ☐ 71+

How many people live in your household? (yourself included)

☐ 1 person   ☐ 2 people   ☐ 3 people   ☐ 4 people   ☐ 5 or more people

---

What is the household's total monthly income? (in thousands of SEK, after taxes and including any subsidies)

☐ 0-10   ☐ 10-20   ☐ 20-30   ☐ 30-40   ☐ 40-50   ☐ 50-60   ☐ 60-70   ☐ 70-80   ☐ 80+

How many children below the age of 18 live in your household?

☐ 0 children   ☐ 1 child   ☐ 2 children   ☐ 3 children   ☐ 4 or more children

---

What of the items listed below do you have available?

☐ Garden and balcony/patio   ☐ Only garden   ☐ Only balcony/patio   ☐ Neither

---

What is your main occupation?

☐ Student   ☐ Working (full-time)   ☐ Working (part-time)   ☐ Unemployed

☐ Retired   ☐ Other

What is your highest achieved education level?

☐ Secondary school or lower   ☐ Upper secondary school   ☐ Bachelor/3 years at university

☐ Master/5 years at university   ☐ More than 5 years at university

I believe us customers have a responsibility of making environmentally friendly choices.

☐ Disagree   ☐ Partly disagree   ☐ Neutral   ☐ Partly agree   ☐ Agree

I believe us customers have a responsibility of making choices that positively affect workers' working conditions.

☐ Disagree   ☐ Partly disagree   ☐ Neutral   ☐ Partly agree   ☐ Agree

How often do you buy flowers, plants, or other garden-related items?

☐ Never/About once per year   ☐ Several times per year   ☐ About once per month
☐ Several times per month   ☐ Once per week or more often

When you are buying flowers, plants, or other garden-related items, how often to you think about purchasing products with sustainability labels?

☐ Always   ☐ Most often   ☐ Sometimes   ☐ Never

When you buy other products (such as groceries), how often do you think about purchasing products with sustainability labels?

☐ Always   ☐ Most often   ☐ Sometimes   ☐ Never

Appendix C   Dunn's tests

Below is the complete output from the three tests of multiple comparisons conducted using Dunn's test.

(1)	Control	Weak des.	Strong des.	Weak inj.
Weak descriptive	1.415			
	(0.5589)			
Strong descriptive	-0.924	-2.328		
	(0.8589)	(0.0953)		
Weak injunctive	-1.176	-2.520	-0.294	
	(0.7207)	(0.0571)	(0.9922)	
Strong injunctive	-0.737	-2.070	0.134	0.406
	(0.9272)	(0.1764)	(0.9973)	(0.9849)

Z-value presented with p-value in parentheses, Šidák adjusted. N = 497.

(1): Count of sustainable products as dependent variable

Table 12: Dunn's test of multiple comparisons among "confident" customers using dependent variable (1).

(2)	Control	Weak des.	Strong des.	Weak inj.
Weak descriptive	2.132			
	(0.1533)			
Strong descriptive	-0.976	-3.106		
	(0.8345)	(0.0094)		
Weak injunctive	-0.528	-2.594	0.414	
	(0.9712)	(0.0465)	(0.9842)	
Strong injunctive	-0.326	-2.359	0.600	0.187
	(0.9905)	(0.0879)	(0.9595)	(0.9961)

Z-value presented with p-value in parentheses, Šidák adjusted. N = 497.

(2): Sum (in SEK) of sustainable products as dependent variable

Table 13: Dunn's test of multiple comparisons among "confident" customers using dependent variable (2).

(3)	Control	Weak des.	Strong des.	Weak inj.
Weak descriptive	1.685			
	(0.3753)			
Strong descriptive	-1.065	-2.739		
	(0.7874)	(0.0304)		
Weak injunctive	-0.337	-1.972	0.695	
	(0.9898)	(0.2180)	(0.9386)	
Strong injunctive	-0.560	-2.158	0.448	-0.224
	(0.9664)	(0.1444)	(0.9810)	(0.9950)

Z-value presented with p-value in parentheses, Šidák adjusted. N = 497.

(3): Percentage sustainable products as dependent variable

Table 14: Dunn's test of multiple comparisons among "confident" customers using dependent variable (3).