Transparency of Carbon-Neutral Labels: Evidence from a Choice Experiment*

Begüm Özdemir Oluk †‡

begum.ozdemiroluk@eawag.ch

begumozdemiroluk.github.io

Please click here for the most recent version of the paper.

November 4, 2024

*I express my gratitude to Stefano Carattini, Ivana Logar, and Sabrina Eisenbarth for their valuable guidance. I thank Fabian Dvorak, the Choice Experiment Clinics participants at the IVM Institute in the VU Amsterdam, and the 9th Workshop on Experimental Economics for the Environment participants for their helpful comments and suggestions. I acknowledge the support from Eawag: Swiss Federal Institute of Aquatic Science and Technology.

†SIAW-HSG and Department of Economics, University of St. Gallen, Switzerland

‡Eawag: Swiss Federal Institute of Aquatic Science and Technology Department of Environmental Social Sciences, Dübendorf, Switzerland

Abstract

This paper examines the effect of transparency in carbon-neutral labeling on consumer willingness to pay. Carbon-neutral labels indicate that a product's CO_2 emissions have been offset (compensated) outside the company and/or directly reduced within it. Although CO_2 offsets are generally viewed as less effective than CO_2 reductions, most labels on the market lack transparency regarding the proportion of CO_2 offset versus CO_2 reduction. This study empirically investigates whether consumers are willing to pay for transparency on carbon-neutral labels and explores consumers' valuation of CO_2 reductions versus CO_2 offsets. Using a discrete choice experiment survey among UK tea consumers, I compare willingness to pay for standard versus transparent carbon-neutral labels. The control group saw a standard label without detailed breakdowns, while the treatment groups saw transparent labels. The first treatment group saw a transparent label showing 95% CO_2 offsetting and 5% CO_2 reduction, and the second treatment group saw a label with a 50%-50% split. The findings show no evidence of WTP for transparency on carbon-neutral labels regarding shares of CO_2 offsetting and reduction.

Keywords discrete choice experiment; stated preferences; environmental valuation; climate labeling; carbon neutrality; information asymmetry; environmental transparency

JEL codes C83, C90, D12, D80, D91, Q51, Q54

1 Introduction

As the negative effects of climate change become stronger, an increasing number of companies are making their products carbon-neutral. Carbon neutrality can be achieved by reducing emissions within the company itself (e.g., through energy efficiency measures) and/or by offsetting emissions externally (e.g., by investing in reforestation projects). CO_2 offsetting involves compensating for emissions by funding external projects, such as reforestation, and these offsets are traded in carbon markets, allowing companies to earn credits used to balance out their products' CO_2 emissions. In contrast, CO_2 reduction focuses on directly lowering emissions within the company through actions such as improving energy efficiency.

While carbon-neutral labels can include only CO_2 offsets, only CO_2 reductions, or a combination of both, these activities are not equivalent. How carbon neutrality is achieved–whether through CO_2 offsets, CO_2 reductions, or a combination of both–matters because there are growing debates in the literature and the media about CO_2 offsets (The Guardian, 2023b). These debates involve both ethical concerns such as moral licensing and putting a price on nature (Aldred, 2012; Hyams and Fawcett, 2013; Dorner, 2019) and practical concerns that CO_2 offsets are not as effective as direct reductions in achieving real atmospheric CO_2 decreases (Schneider et al., 2015; Becken and Mackey, 2017; Calel et al., 2021; Trencher et al., 2024).

Carbon-neutral labels generally lack transparency about the share of CO_2 offset versus CO_2 reduction the label entails. This issue is highlighted by recent developments about the concerns about quality of CO_2 offsets, including the EU's proposed ban on unverified generic environmental claims like "climate-neutral" (European Parliament, 2023), the lawsuit against Delta Air Lines over misleading carbon neutrality claims (The Guardian, 2023a), and ClimatePartner's decision to withdraw

its carbon-neutral labels (ClimatePartner, 2023). It is thus an empirical question whether consumers are aware of these concerns and value transparency on carbon-neutral labels. This paper addresses this question by examining how transparency regarding the share of CO_2 offsetting and CO_2 reduction on carbon-neutral labels affects consumers' willingness to pay (WTP).

This pre-registered study[1] leverages a discrete choice experiment (DCE), a stated preference method, for eliciting consumers' preferences and WTP for carbon-neutral labels. An online survey is conducted among 1,337 tea consumers in the UK using a split-sample approach. There is one control group and two treatment arms. The control group saw the standard carbon-neutral label without any information on CO_2 offsetting and CO_2 reduction shares on the label. The treatment groups saw a transparent version of the carbon-neutral label, which includes information on the shares of CO_2 offsetting and CO_2 reduction. Specifically, the first treatment group saw a carbon-neutral label indicating a composition of 95% CO_2 offsetting and 5% CO_2 reduction, and the second treatment group saw a carbon-neutral label displaying an equal division, with 50% CO_2 offsetting and 50% CO_2 reduction.

I chose to leverage a DCE for several reasons. Primarily, my study investigates whether transparency influences consumer WTP for carbon-neutral labels while ensuring that potential biases, such as hypothetical bias, inherent in stated preference studies do not differentially affect the subsamples. So, DCE is well-suited for this purpose as it allows for the controlled manipulation of labels across balanced subsamples. Secondly, transparent carbon-neutral labels with proportions of CO_2 offsetting and reduction are not currently available in the market. Conducting a randomized controlled trial (RCT) would necessitate complex partnerships with certifiers, companies,

[1]Ozdemir Oluk, Begum. 2024. "Consumer Preferences for Transparent Carbon-Neutral Labels: A Choice Experiment." AEA RCT Registry. August 02. https://doi.org/10.1257/rct.12520-2.0

and retailers, making the experimental design highly challenging. Finally, DCE offers distinct advantages, such as capturing consumer trade-offs among various attributes and eliciting marginal willingness to pay (MWTP) (Hanley et al., 1998), while being less susceptible to biases like "yea-saying" compared to contingent valuation methods (CVM) (Adamowicz et al., 1994).

The main research questions addressed in this study are: (i) Do consumers value transparent carbon-neutral labels more than standard ones? and (ii) Do consumers value CO_2 reductions more than CO_2 offsets? Transparency on carbon-neutral labels–specifically displaying the shares of CO_2 offsetting and CO_2 reduction–can have a twofold effect on consumer demand for carbon-neutral products. On one hand, it can positively influence demand by increasing trust. On the other hand, depending on the proportions of CO_2 offsetting and CO_2 reduction, transparency can either increase or decrease demand due to ethical and practical concerns about CO_2 offsetting (Carattini and Tavoni, 2016). Additionally, providing more information on the label could either clarify or confuse consumers, thereby affecting demand positively or negatively.

The findings, based on both control and treatment groups, suggest that consumers value carbon neutrality and are willing to pay a premium for carbon-neutral labels on tea products. However, no statistically significant differences were found between WTP for transparent compared to standard carbon-neutral labels. Similarly, no statistically significant difference was found between the transparent label indicating 95% CO_2 offsetting and 5% CO_2 reduction and the label indicating 50% CO_2 offsetting and 50% CO_2 reduction.

The findings of this study have the following policy implications. Regulating labels is important, as discussed in the literature[2], and aligning with the intentions of poli-

[2]Information asymmetry can distort market mechanisms, leading to adverse selection and the selection of lower-quality goods (Akerlof, 1970). Additionally, a body of theoretical literature suggests that, under certain assumptions and in the presence of information asymmetry, competition

cymakers such as the EU, which plans to ban unverified generic environmental claims like "climate-neutral" (European Parliament, 2023). However, based on the findings of this paper–without telling participants whether CO_2 offsets or CO_2 reductions are better–consumers do not 'naturally' pay more for transparency on carbon-neutral labels. Therefore, the policy implications of this study would imply clear communication of (i) both the share of CO_2 offsets and CO_2 reductions on the label and (ii) an explanation of why offsets are considered 'inferior' to direct reductions with consumers.

This study contributes to the following lines of the existing literature. Firstly, it adds to the literature focusing on consumers' valuation of climate labels (Drichoutis et al., 2016; Akaichi et al., 2017; Feucht and Zander, 2018; Grebitus et al., 2013; Onozaka and McFadden, 2011; Bek, 2022). While many stated preference studies and a few revealed preference studies have focused on consumers' WTP, to my knowledge, none of them focused on the effect of transparency in carbon-neutral labels. Secondly, this paper contributes to the extensive literature on the economics of CO_2 offsetting (Blasch and Farsi, 2014; Ziegler et al., 2012; Brouwer et al., 2008; MacKerron et al., 2009; Carattini and Tavoni, 2016; Chen et al., 2018). Similar to this study, a few others have examined how consumers value CO_2 offsets compared to CO_2 reductions (Bek, 2022; Roemer et al., 2023). However, these studies treated CO_2 offsets and reductions as different levels of the same attribute or as different attributes in choice experiments, meaning that each participant saw all versions of the labels. In contrast, my study uses separate subsamples (treatment and control groups) for each version of carbon-neutral labels, ensuring that each participant sees only one type of label where subsamples are balanced in terms of socioeconomic characteristics. This de-

can favor environmental labels that persist in the market (Brécard, 2014; Heyes and Martin, 2017; Brécard, 2017; Heyes and Martin, 2018; Poret, 2019)

sign allows for a clearer causal interpretation of the effects of transparency and the type of information (i.e., 95% offset, 5% reduction, and 50% offset, 50% reduction). Additionally, my study uses a significantly larger sample size, enhancing statistical power and ensuring representativeness, which makes the findings more generalizable to the broader consumer population.

Furthermore, this study connects to several other lines of literature. First, it adds to the literature on bringing non-normative pro-social behaviors to normative (Sparkman and Walton, 2017; Carattini et al., 2022; Mortensen et al., 2019; Kraft-Todd et al., 2018) by investigating whether increased transparency in environmental labeling can enhance the adoption of green products like carbon-neutral labels, which currently remain a niche market. Second, in a setting where the sellers have more information than the buyers, exploring whether consumers value transparent labels more compared to standard labels also adds to the literature on information asymmetry(Akerlof, 1970; Brounen and Kok, 2011; Brunnschweiler et al., 2021).

The remainder of the paper is outlined as follows: Section 2 provides a background and a review of the relevant literature. Section 3 details the methodology, including the survey and DCE design (3.1), data (3.2), and the empirical approach (3.3). Section 4 presents the results, which include preferences and WTP estimates (4.1), the underlying mechanisms of consumers' preferences (4.2), and the next steps (4.3), and the conclusion (5).

2 Background

This section introduces carbon-neutral labels and why they are interesting to study. Next, it examines factors that could derive the demand for carbon-neutral labels. Then, it highlights the knowledge gap regarding the demand for transparency on the label and discusses the study's implications for corporate social responsibility and policy-making.

The theory of environmental externalities underlines the importance of market-based mechanisms in addressing climate change. Without accounting for externalities like climate impacts, the social benefits of producing a product may be overestimated (Hanley et al., 1997). Integrating the cost of climate change into a product's private marginal cost enables prices to reflect the true social impact. This internalization can be achieved through mechanisms such as Pigouvian taxes, which aim to reach a socially optimal level of output by correcting for negative externalities (Pigou, 1920). By incorporating social costs, true pricing reduces deadweight loss. While carbon taxes set by the policymakers aim to reach an optimal level of output, carbon-neutral labels used by the private sector offer a complementary tool to accelerate this transition.

Many companies have announced their commitment to reach carbon neutrality by 2050 or earlier, and global voluntary carbon markets are expected to increase 15 times by 2030 (potentially reaching $50 billion by 2030) and 100 times by 2050, (McKinsey & Company, 2023). However, currently, most products on the market have not yet been labeled carbon-neutral, and opting for carbon neutrality is still a niche behavior. Therefore, understanding the consumers' demand for carbon-neutral labels can help businesses and policymakers plan effective climate change mitigation.

A carbon-neutral certified product indicates that its lifecycle GHG emissions have

been reduced, and the remaining emissions have been compensated through CO_2[3] offsetting (ClimatePartner, 2023). This means that a product can achieve carbon neutrality through 100% CO_2 offsetting, 100% CO_2 reduction, or a combination of both. CO_2 offsetting and reduction are different approaches. CO_2 offsetting projects involve initiatives outside of the entity, such as reforestation (Climate Portal, 2023). Companies can buy and sell these offsets in carbon markets, allowing them to earn credits that balance (compensate) their products' CO_2 emissions. In contrast, CO_2 reduction focuses on directly lowering a company's emissions through actions such as improving energy efficiency.

Given the link between carbon neutrality and CO_2 offsetting, the economics of CO_2 offsetting offers insights into the demand for carbon-neutral labels. The guilt associated with harming the environment (Kotchen, 2009) and the warm glow or satisfaction obtained from pro-environmental behavior (Andreoni, 1990) are factors that can drive demand. At the same time, CO_2 offsets, as opposed to actual CO_2 reductions, may be increasingly less favored due to the growing database about CO_2 offsets. Media skepticism about the reliability of CO_2 offsets and concerns about corporate greenwashing may influence how consumers perceive environmental labels, including carbon-neutral labels and CO_2 offsetting. Ethical concerns associated with CO_2 offsetting include issues such as moral licensing (Dorner, 2019) and putting a price on nature (Aldred, 2012). The practical concerns include the argument that CO_2 offsetting does not result in equivalent CO_2 reductions in the atmosphere (Becken and Mackey, 2017) due to issues of additionality (Hyams and Fawcett, 2013; Schneider and Kollmuss, 2015), credibility (Bumpus and Liverman, 2008; Hooper et al., 2008), and double-counting of emissions reductions (Schneider et al., 2015). Therefore, it

[3]In this paper, "CO_2" and "GHG" are used interchangeably. The term CO_2 refers to the carbon dioxide equivalent (CO_2e), measuring the total GHG emissions of a product, expressed as an equivalent amount of CO_2.

remains a question whether consumers are aware of these concerns and value them.

While these concerns grow, generally carbon-neutral labels in the product markets do not clearly indicate the proportion of CO_2 offsetting and CO_2 reduction on the label [4]. According to the European Commission, 53.3% of environmental claims were found to be vague, misleading, or unfounded, and based on the report by Changing Markets Foundation (2023), even the most carbon-intensive food products, such as beef, are labeled as carbon-neutral. Therefore, the empirical question of how people value these labels remains when the certifiers are more transparent regarding the environmental quality of the label.

This study examines how transparency in carbon-neutral labeling affects consumer demand. Results show that while consumers are willing to pay for carbon-neutral labels, there is no significant difference in WTP between standard and transparent labels, nor between two labels with different CO_2 offset-and-reduction shares. This has implications for corporate social responsibility and policymaking. Requiring transparency on labels is crucial, as well as aligning with goals to prevent unverified claims. However, based on this study's findings, consumers may not naturally pay more for labels indicating higher CO_2 reductions. Therefore, it would be advisable for policymakers and firms to ensure that (i) labels specify the shares of CO_2 offsets and reductions and (ii) clarify why CO_2 reductions are preferable over CO_2 offsets.

[4] To my knowledge, Swiss Airlines is an exception, which recently started to offer CO_2 offsetting, or a combination of CO_2 offsetting and CO_2 reduction after flight purchases (Swiss International Air Lines, 2024)

3 Methodology

3.1 Survey and Choice Experiment Design

This section explains the split-sample approach, the product, the sample, the DCE design, and the questions following the DCE.

I use a split-sample approach with three different samples. All samples receive identical survey and choice experiment designs. The only difference is in the type of carbon-neutral label, as shown in Figure 1. The control group is shown a standard carbon-neutral label, stating "CO_2 neutral" only. The information on the shares of CO_2 offsetting and CO_2 reduction are not revealed to the participants [5]. While treatment groups are shown a "transparent" carbon-neutral label with additional text indicating the shares of CO_2 offsetting and CO_2 reduction. In treatment group 1, the carbon-neutral label indicates a 95% offset and a 5% reduction, while in treatment group 2, the carbon-neutral label indicates a 50% offset and a 50% reduction. Respondents in treatment groups are informed that the specific share of the combination of CO_2 offsetting and CO_2 reduction actions ensure that the tea product's lifecycle is carbon-neutral.

This study focuses on tea consumers in the UK for several reasons. Firstly, it is common for tea products in the UK to have sustainability labels, including carbon-neutral labels, making tea a relevant product to study transparency in carbon-neutral labeling. Second, food systems significantly contribute to climate change, accounting for a third of global anthropogenic GHG emissions (Crippa et al., 2021). Although a box of tea itself may not be considered CO_2-intensive, its frequent consumption can

[5] Participants in the control group are provided the following information: "Product's greenhouse gas emissions, measured in carbon equivalent has been offset (compensated) by investing in activities outside of the company, such as tree planting projects, or reduced within the company in the last five years, such as through investments in cleaner production processes; or both offset and reduced."

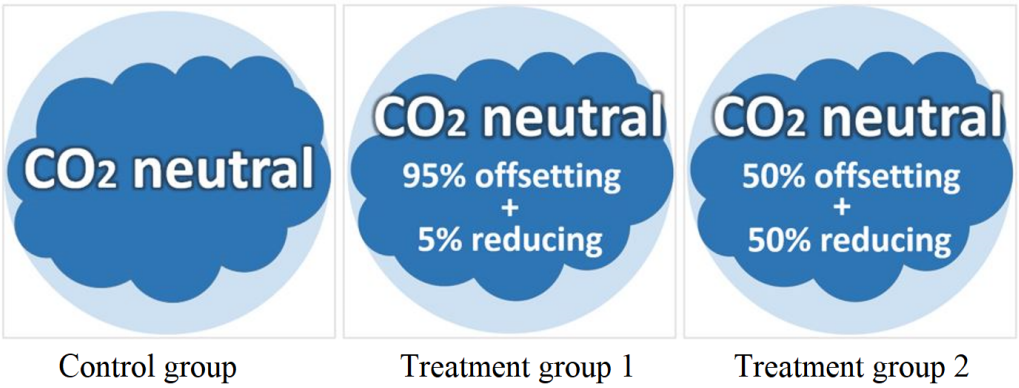


Figure 1: Carbon neutral labels shown to control and treatment groups

still result in high emissions over time. Third, the UK ranks as the country with the third-highest per capita tea consumption globally (Statista, 2023).

In the survey, which is provided in Appendix A.1, respondents are asked to consider 80 teabags in their preferred form, equivalent to approximately 200 grams of tea. The pre-registered online survey[6] is pretested on 157 respondents, and the main survey included 1,337 tea drinkers in the UK.[7]. At the beginning of the survey, screening questions are asked about age, tea consumption and tea purchasing habits. Participants under the age of 18, or those who never consume or purchase tea, are screened out.

Table 1: Choice attributes and attribute levels

Attributes	**Attribute levels**
Carbon-neutral label	Carbon-neutral label, no label
Organic label	Organic label, no label
Ethical trade label	Ethical trade label, no label
Price	£0.90, £1.90, £2.90, £3.90, £4.90, £5.90, £6.90

[6]Ozdemir Oluk, Begum. 2024. "Consumer Preferences for Transparent Carbon-Neutral Labels: A Choice Experiment." AEA RCT Registry. August 02. https://doi.org/10.1257/rct.12520-2.0

[7]Although payment was made for 1,200 responses, the survey company provided data of a total of 1,337 responses.

The DCE focuses on a 200-gram tea box with the following attributes: carbon-neutral label, organic label, ethical trade label, and price. Table 1 shows the attributes and attribute levels included in the DCE: three sustainability labels taking two levels each, the presence or absence of the label, and the price that takes seven levels ranging from £0.90 to £6.90 with £1 increments. The price levels in this survey were determined based on the average tea price in the UK, which is £2.46 for 250 grams of tea (Office for National Statistics, 2024c), as well as an online search for 100 tea products conducted on Sainsbury's, Tesco, Morrisons, Ocado, and Amazon UK during 2023.

Next, there are two choice attributes: organic and ethical trade labels, besides the carbon-neutral label. These are included because organic and fair trade labels are among the most common sustainability labels on tea products in the UK (based on an online search for 100 tea products in UK online grocery stores during 2023). Therefore, it makes the choice experiment more realistic and captures the trade-offs individuals would face in real decision-making scenarios. Organic and ethical trade do not vary across sub-samples, unlike carbon-neutral labels [8]. Participants are informed about the nature of these labels, including the carbon-neutral label, before the choice tasks. All five labels, three versions of carbon-neutral label, as well as organic, and ethical trade labels, are developed by the author and are hypothetical.

The type of tea blend (e.g., black, green, herbal) is not specified as a separate choice attribute because there is a strong preference for black tea in the UK. That is, 74% of people in the UK consume black tea, although they also drink other varieties such as green tea and herbal infusions at the same time (Tea and Infusions Association, 2022). Similarly, the packaging form (loose leaf or teabags) is not included as a

[8]The participants are informed that the organic label indicates products with only organic ingredients and no synthetic pesticides, while the ethical trade label ensures responsible labor practices and guarantees higher prices for exporters based on internationally recognized standards.

separate attribute because 97.5% of tea sold in the UK is in teabag form (Tea and Herbal Association, 2024).

I choose to leverage a DCE for several reasons. The most important reason to opt for a stated preference method rather than a revealed preference method is the fact that examples of 'transparent' carbon-neutral labels, with extra information about the shares of CO_2 offsetting and CO_2 reduction, are not available on the market. Therefore, a randomized controlled trial (RCT) would require forming partnerships with a carbon-neutral certifier willing to develop such labels, a company that meets the criteria, and also a retailer. The possibility of achieving this without having to change the main experimental design would be quite challenging. Second, there are certain advantages of using DCE compared to other valuation methods. DCE can account for consumer trade-offs among various attributes, elicit their marginal WTP (MWTP) for each attribute (Hanley et al., 2019a), and reduce susceptibility to biases such as "yea-saying", compared to contingent valuation method (CVM) (Adamowicz et al., 1994). Furthermore, I define the sustainability attributes (carbon neutral, organic, and ethical trade labels) independently in the DCE, such that there is no overlap in their meaning, while correlated attributes are a common issue in real-life scenarios (Hanley et al., 2019b), thus would be an additional challenge for a revealed preference study which aims to elicit the WTP for carbon-neutral labels in the market.

DCEs and CVMs are often criticized for being prone to hypothetical bias since participants are not required to make actual payments for their choices. To address this issue, the literature has employed various techniques such as cheap talk (Cummings and Taylor, 1999), honesty priming (Howard et al., 2017), and oath scripts (de Magistris and Pascucci, 2014). In this study, I implement cheap talk, oath scripts, and a budget reminder to mitigate the limitations of stated preference methods. In the cheap talk script, I informed the respondents that survey participants are likely to

overstate their WTP in hypothetical surveys and asked them to consider how they would feel about spending their money in a real situation. In the oath script, I asked participants to promise to provide honest responses by checking the box.

Ngene software is used to generate the DCE design, which consists of 16 different choice tasks blocked into 2^9. Figure 2 shows one of the choice cards shown to the control group. There are two tea products, and "none of the two" choice opinions [10]. The estimate parameters from the pre-test were used to create the final CE design. Please refer to Table A.16 for the respective choice design used for the survey.

The DCE is followed by questions that assess participants' choice certainty and identify protest responses. It further includes attribute non-attendance (ANA) and open-ended contingent valuation method (CVM) questions. Additionally, the survey explores participants' attitudes toward climate change and sustainability labels using Likert-scale statements, which include trust in carbon-neutral labels, confusion about carbon-neutral labels, concerns about CO_2 offsetting, climate worry, warm glow (positive emotions from climate-friendly purchases), guilt (negative feelings when not making climate-friendly choices), social approval (perceived acceptance by others), the polluter pays principle (the opinion that producers should pay for climate mitigation), time restrictions (limited time for climate-friendly choices), and financial constraints (budget limitations preventing climate-friendly purchases). Then, participants are shown all three versions of carbon-neutral labels and asked which label they trust the most and find the most confusing. It also collects information on participants' tea consumption habits and sociodemographic characteristics. The survey

[9]Therefore, each participant saw 8 choice cards.

[10]Restrictions are included to make choice cards more realistic. Alternatives with more labels had to be priced higher than alternatives with fewer labels. In addition to that, for the design of the main survey, additional restrictions are added for the lowest price level (0.9 pounds) and the highest price level (6.9 pounds) tea alternatives. If the price is 0.9 pounds, it should not have any labels; similarly, if it is 6.9 pounds, it must have all labels.

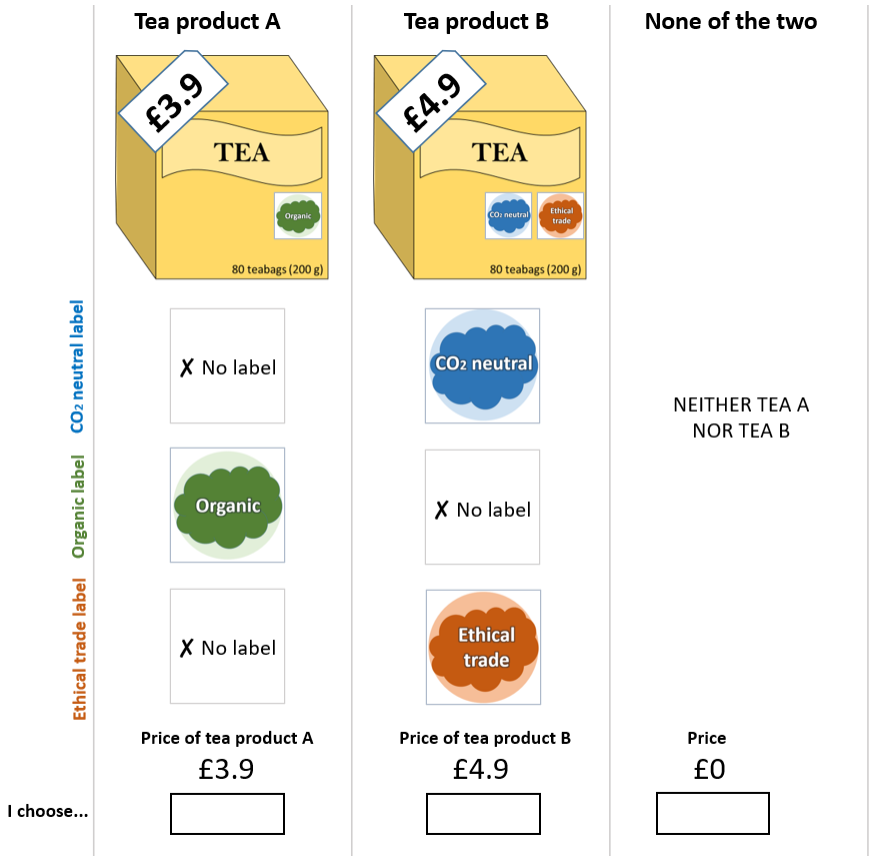


Figure 2: Illustrative choice card for sub-sample 1

concludes with feedback questions to ensure clarity and collect additional insights.

3.2 Data

The survey data were collected online in September 2024 in collaboration with a professional survey company. 54,219 respondents were initially invited to the survey. Of these, 2,366 started the survey, 580 were eliminated due to quota restrictions, and 225 were screened or dropped out[11]. Individuals who indicated that they never purchase or never drink tea, constituting 14% of the total sample, were disqualified

[11]Two participants whose treatment groups were not coded in the provided data were also excluded from the sample.

from continuing the survey[12].

Consequently, a total of 1,337 individuals successfully completed the main survey[13].

The sample mainly represents the UK population aged 18 and over, with some variations in general population statistics due to focusing on adult consumers older than 18 years old in this study. The average age of the sample is 57 years, whereas the national median age for adults is approximately 41 years (Office for National Statistics, 2022). Gender distribution is very close to the national profile, with 50% female participants in the sample compared to 51% in the population (Gov.uk, 2021b). 57% of the sample holding a post-secondary certificate (NQF Level 4) level or above, moderately higher than the national figure of 49% in the population (Gov.uk, 2021a). Additionally, 49% of the sample has an annual household income exceeding £40,000, compared to a national median of approximately £34,500 (Office for National Statistics, 2024a). Finally, the sample's employment rate of 64% is slightly below the national average of 75% (Office for National Statistics, 2024b). Therefore, these characteristics confirm that the sample largely represents the UK adult population. A summary of all detailed socioeconomic characteristics of tea consumers/drinkers is provided in Table A.1, and the balance of covariates is provided in Table A.7 in Appendix A.1.

Among the tea consumers/drinkers, almost 77% of people drink tea daily, and more than 61% purchase tea at least once a month or more frequently. The tea consumption and purchase habits of the sample are provided in more detail in Table A.3 in Appendix A.1. More than 72% of the respondents indicated that they consume

[12]Please refer to Table A.2 in Appendix A.1 for the comparison between tea drinkers/purchasers and non-tea drinkers/purchasers in terms of the three covariates (age, gender, and education) used as quota criteria to recruit a representative sample of the UK population before the screening questions.

[13]For the pre-test, 157 participants were recruited. The pre-test data is not included in the main survey data.

black tea, while the remainder consume green tea, herbal tea, and other tea blends.

3.3 Empirical Approach

To estimate consumer willingness to pay (WTP) for carbon-neutral labels, I use a mixed logit (MXL) model. The MXL model is advantageous over the multinomial logit (MNL) model because it accounts for unobserved preference heterogeneity across individuals and relaxes the independence of irrelevant alternatives (IIA) assumption of the MNL model.

According to the Random Utility Model (RUM) (Lancaster, 1966), consumers derive utility from the good's attributes. While making choices, they have to make tradeoffs between these attributes. The model assumes that consumers choose the option that provides the maximum expected utility (McFadden, 1973). Following Lancaster's random utility framework, the marginal utility of an individual derived from choosing an alternative is expressed as the summation of her marginal utility of attributes. Since a part of this utility is not observable to the researcher, the utility is modeled with both deterministic and random parts (Train, 2009).

In the MXL framework, the utility U_{ij} that individual i derives from choosing alternative j:

$$U_{ij} = \boldsymbol{\beta}_i' \mathbf{X}_{ij} + \boldsymbol{\gamma}' \mathbf{Z}_{ij} + \varepsilon_{ij}$$

where β_i is a vector of individual-specific random coefficients, γ is a vector of fixed coefficients, X_{ij} is a vector of attributes associated with the random coefficients, Z_{ij} is a vector of attributes associated with the fixed coefficients, and ϵ_{ij} is an error term assumed to be independently and identically distributed (i.i.d.) type I extreme value.

The attributes included in the model are carbon-neutral, organic, ethical trade

labels, and price. I assume that the individual-specific coefficients for label attributes follow a normal distribution[14] to capture unobserved preference heterogeneity, while the price coefficient and the status quo option are fixed. The probability that individual i chooses alternative k among list of alternatives j is given by:

$$P_{ik} = \int \frac{\exp\left(\boldsymbol{\beta}_i' \mathbf{X}_{ik} + \boldsymbol{\gamma}' \mathbf{Z}_{ik}\right)}{\sum_j \exp\left(\boldsymbol{\beta}_i' \mathbf{X}_{ij} + \boldsymbol{\gamma}' \mathbf{Z}_{ij}\right)} \phi(\boldsymbol{\beta}_i \mid \boldsymbol{b}, \mathbf{W})\, d\boldsymbol{\beta}_i$$

where $\phi(\beta|b, W)$ is normal density with mean b, and covariance W.

The model is estimated in 'preference space' and the marginal willingness to pay (MWTP) for each attribute is calculated as the negative ratio of the attribute coefficient to the price coefficient(Hensher et al., 2005):

$$\text{MWTP} = -\frac{\beta_{\text{attribute}}}{\beta_{\text{price}}} \tag{1}$$

Standard errors for the MWTP estimates are derived using the Delta method.

I use (Poe et al., 2005) to test any significant differences in MWTP for the carbon-neutral label among the three samples to understand how transparency affects consumers' preferences and WTP. Please refer to Appendix C for the calculations of the minimum detectable effect sizes.

To account for multiple hypothesis testing, p-values are adjusted using the Bonferroni method based on the number of variables in each regression. Standard errors of the choice model coefficients remain unadjusted and are used to derive MWTP standard errors, while p-values for MWTP estimates are adjusted for the three tests (carbon neutral, organic, and ethical trade). Poe tests for subsample differences are

[14] I obtain the model parameters' distribution from MNL estimation. I use the coefficients as means and calculate the standard deviations as the square root of the sample size multiplied by the estimated standard errors. I use the same distribution assumption for all subsamples (control and treatment groups).

similarly adjusted for three comparisons.

For robustness purposes, I also run the MNL model, where all parameters are assumed to be fixed, and an additional MXL model, where the price is also randomized, which is included in the Appendix A.3 in Tables A.10 and A.12. When I randomize both the price and the carbon-neutral attribute, I use the Krinsky-Robb procedure to derive WTP estimates and standard errors.

To explore which factors are associated with consumers' probability of choosing carbon-neutral labeled tea, interaction terms between the carbon-neutral label and individual-specific variables are included in the model. The utility function with interactions is specified as follows:

$$U_{ij} = \boldsymbol{\beta}_i' \mathbf{X}_{ij} + \boldsymbol{\gamma}' \mathbf{Z}_{ij} + \boldsymbol{\eta}' \left(\mathrm{CN}_{ij} \times \mathbf{Y}_i \right) + \varepsilon_{ij}$$

where CN_{icj} is the carbon-neutral label attribute, Y_i is a vector of individual-specific variables as summarized in Table 2, and η is a vector of coefficients for the interaction terms.

My main variables of interest are those that could most likely drive the differences (or non-differences) between samples, namely: trust level in carbon-neutral labels, confusion level with carbon-neutral labels, concern level about CO_2 offsets, binary variable for participants who only trust in transparent carbon-neutral labels but not standard labels, binary variable for participants who are only confused with transparent carbon-neutral labels but not with standard labels. However, I also include additional variables that could drive consumer demand as interactions to check robustness.

Variable	Description	Measurement
Trust	Trust in carbon-neutral labels	7-point Likert scale (normalized)
Confusion	Confusion with carbon-neutral labels	7-point Likert scale (normalized)
Concern	Concern about CO_2 offsetting	7-point Likert scale (normalized)
Only trust in transparent labels	Trusts only transparent carbon-neutral labels but not standard labels	Binary variable (1 = Yes)
Only confused with transparent labels	Confused only by transparent carbon-neutral labels but not standard labels	Binary variable (1 = Yes)
Climate worry	Worry about climate change	7-point Likert scale (normalized)
Warm glow	Positive emotions from climate-friendly purchases	7-point Likert scale (normalized)
Guilt	Negative feelings when not making climate-friendly choices	7-point Likert scale (normalized)
Social approval	Perceived acceptance by others for climate-friendly choices	7-point Likert scale (normalized)
Polluter pays	Belief that producers should pay for climate mitigation	7-point Likert scale (normalized)
Financial constraints	Limited financial resources for climate-friendly purchases	7-point Likert scale (normalized)
Time restrictions	Limited time for climate-friendly choices	7-point Likert scale (normalized)
Age	Age of the respondent	Continuous variable
High Income	Yearly household income > £40,000 after taxes	Binary variable (1 = Yes)
Female	Gender of the respondent	Binary variable (1 = Female)

This table shows the list of covariates that interacted with the carbon neutral label to understand the factors consumers' choices for carbon neutral labels associated. For variables measured on 7-point Likert scales, higher values indicate stronger agreement, while lower values indicate stronger disagreement.

Table 2: Description of interacted covariates

4 Results

4.1 Preferences and WTP

This section presents the estimation results from the MXL model to understand the likelihood of each attribute being chosen over others, the derivation of the MWTP for each non-cost attribute, and the significant differences among samples to understand the effect of transparency on carbon-neutral labels.

Table 3 displays the estimation results from the MXL Model. The results indicate that the coefficients for the carbon-neutral label are all positive and statistically significant at the 1% level across all samples. This suggests that carbon-neutral labels increase the probability of a tea product being chosen. Furthermore, the negative coefficient for price confirms the expected decrease in the probability of choosing tea products.

The MWTP for the carbon-neutral label ranges between £0.41 and £0.67 across the three samples and is statistically significant. However, Poe test results (Poe et al., 2005) reveal no statistically significant differences between the control and treatment groups, as well as between two treatment groups, as reported under Table 4. This indicates that there is no evidence for extra WTP for transparency on carbon-neutral labels. Similarly, there is no evidence for extra WTP for CO_2 reductions as compared to CO_2 offsets. This result is robust across robustness tests as shown in Tables A.10 and A.13 in Appendix A.3.

The MWTP for the carbon-neutral label translates to an MWTP of £0.46 for reducing or offsetting 1 kg of CO_2, assuming the average CO_2 content of 200 grams of tea is 1.28 kg, based on the 'Plate up for the Planet' carbon calculator. This WTP for the carbon-neutral label in this study falls between estimates from the meta-analysis of existing literature and the hedonic difference-in-differences (DiD) approach using

Amazon UK, US, and Germany data in (Carattini et al., 2024).

Notably, the coefficient for the carbon-neutral label is positive and significant, although it is smaller in magnitude compared to the coefficients for the organic and ethical trade labels. This indicates that while consumers value the carbon-neutral label, they place a higher value on the organic and ethical trade labels. The results also show that the MWTP for the organic and ethical trade labels is approximately 1.7 to 1.9 times larger than that for the carbon-neutral label, which is consistent with existing literature. For instance, Birkenberg et al. (2021) and Bek (2022) found that fair trade labels have a higher WTP than carbon-neutral labels in the coffee market, with fair trade WTP values up to 2.4 times larger. However, both studies had small sample sizes ($N = 80$ and $N = 299$, respectively) and were conducted in Germany.

	Full Sample	**Control**	**Treatment 1**	**Treatment 2**
Status quo	-3.028***	-3.308***	-2.972***	-2.825***
	(0.104)	(0.181)	(0.175)	(0.185)
Carbon-neutral (random)	0.377***	0.440***	0.432***	0.269***
	(0.043)	(0.074)	(0.074)	(0.074)
Organic (random)	0.673***	0.634***	0.694***	0.683***
	(0.054)	(0.096)	(0.093)	(0.094)
Ethical trade (random)	0.741***	0.745***	0.715***	0.756***
	(0.054)	(0.095)	(0.094)	(0.091)
Price	-0.662***	-0.680***	-0.649***	-0.659***
	(0.026)	(0.046)	(0.044)	(0.047)
Standard deviations				
$\sigma_{CarbonNeutral}$	1.093***	1.188***	1.047***	1.061***
	(0.049)	(0.093)	(0.083)	(0.082)
$\sigma_{Organic}$	1.435***	1.421***	1.404***	1.463***
	(0.062)	(0.113)	(0.099)	(0.113)
$\sigma_{EthicalTrade}$	1.414***	1.505***	1.344***	1.412***
	(0.061)	(0.110)	(0.105)	(0.105)
Observations	10696	3568	3576	3552
Participants	1337	446	447	444
Log-likelihood	-10229.25	-3361.62	-3440.22	-3423.39

(i) This table shows the mixed logit (preference space) estimation results. All attributes, except for price and the status quo, are randomized. Normal distribution is assumed for the randomized variables.
(ii) Robust standard errors are in parentheses.
(iii) Bonferroni-corrected p-values and significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

Table 3: MXL (PS estimation)

	Full Sample	 **Control**	 **Treatment 1**	 **Treatment 2**
$\text{MWTP}_{\text{Carbon neutral}}$	0.569***	0.648***	0.666***	0.408***
	(0.065)	(0.112)	(0.117)	(0.111)
$\text{MWTP}_{\text{Organic}}$	1.017***	0.932***	1.069***	1.037***
	(0.071)	(0.123)	(0.125)	(0.126)
$\text{MWTP}_{\text{Ethical trade}}$	1.120***	1.095***	1.103***	1.148***
	(0.072)	(0.125)	(0.129)	(0.123)

(i) MWTP is estimated by dividing the negative of the attribute coefficients by the price coefficient.
(ii) Robust standard errors in brackets are obtained using the Delta method.
(iii) Bonferroni-corrected p-values and significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.
(iv) Poe test results (adjusted for three comparisons) indicate that the bilateral differences between the control and treatment groups are not statistically significant: control vs. treatment 1: Poe statistic = 0.543 (n.s.); control vs. treatment 2: Poe statistic = 0.067 (n.s.); treatment 1 vs. treatment 2: Poe statistic = 0.06 (n.s.).

Table 4: MWTP estimates (MXL, PS estimation)

4.2 Mechanisms

This section investigates the mechanisms associated with the probability of a consumer choosing a tea product with a carbon-neutral label, which is descriptive and correlational in nature.

Based on the responses of all participants after the choice experiment, the transparent label with a 50%-50% split between CO_2 reduction and offsetting is the most trusted, with 38.89% of participants indicating trust. Meanwhile, the transparent label with 95% offsetting and 5% reduction is considered the most confusing by 34.11% of participants. Table 5 summarizes participants' relative trust and confusion regarding the carbon-neutral labels based on their responses after the choice experiment. Participants were asked which label they trusted the most and which they found the

most confusing among the three label choices (standard label and two transparent labels) and a 'none' option.

The estimation results in Tables 6 and 7 provide insights into these mechanisms, among others. In the mixed logit estimations, relative confusion is significant. This indicates that participants who find at least one transparent label confusing but not the standard version of the label are significantly less likely to choose tea products with a carbon-neutral label. This effect occurs despite consumers not seeing all three labels simultaneously during the choice experiment, possibly because their confusion might have been raised by comparing the labels from the experiment with those available in the actual market while making choices. Looking at subsamples, this effect is only evident for the participants in the second treatment group, where the label indicates an equal share of CO_2 offsetting and reduction (50%-50%). In contrast, relative trust in transparent labels compared to the standard label does not show a significant association with consumers' probability of choosing a carbon-neutral label in all samples.

Having stated that 'relatively confused' participants in the second treatment group are linked with a higher probability of choosing a carbon-neutral label, the 'general confusion level' about carbon-neutral labels, general trust level for carbon-neutral labels, and concern level regarding CO_2 offsetting is not found to be linked with the probability of choosing a tea product with a carbon-neutral label.

Regarding other factors, variables such as warm glow, guilt, polluter pays principle (the belief that producers should pay for climate mitigation), and social approval do not show significant associations with the probability of choosing carbon-neutral labels. Further analysis of subsamples, based on Table A.15 in Appendix A.3, suggests that climate worry is significant for treatment group 2, who were exposed to the transparent label with 95% CO_2 offsetting and 5% reduction. This finding might

be related to how consumers interpret the high percentages (such as large number bias), possibly viewing the large shares as positive given no knowledge or pre-existing concerns about CO_2 offsetting.

Financial constraints are also significantly associated with preferences. Consumers who indicate limited financial resources for climate-friendly purchases are less likely to choose products with carbon-neutral labels. At the same time, limited time for climate-friendly choices does not appear to be linked with carbon-neutral labeled tea choices, suggesting that time constraints may not be a major barrier in this context.

Demographic factors reveal that age is positively associated with preferences for carbon-neutral labels, with older consumers showing a higher likelihood of choosing such products. Gender and high income, however, do not show significant associations. The lack of significance for high income may be partially explained by the inclusion of the financial constraints variable, which more directly captures consumers' perceived budget limitations related to purchasing climate-friendly products.

In summary, the analysis identifies relative confusion about transparent carbon-neutral labels, climate worry in specific labeling contexts, financial constraints, and age as significant factors linked to the probability of choosing a carbon-neutral label. These findings in this section are correlational in nature but support the idea that the companies should offer transparent carbon-neutral labels and make the label clear—particularly regarding why CO_2 offsets or CO_2 reductions are different and which one is preferable–to enhance consumer adoption of products with such labels.

	Full Sample		**Control**		**Treatment 1**		**Treatment 2**	
	N	Share	N	Share	N	Share	N	Share
Trust (label 1)	314	23.49%	108	24.22%	92	20.58%	114	25.68%
Trust (label 2)	330	24.68%	114	25.56%	124	27.74%	92	20.72%
Trust (label 3)	520	38.89%	178	39.91%	165	36.91%	177	39.86%
Trust (none)	274	20.49%	84	18.83%	99	22.15%	91	20.50%
Confusion (label 1)	342	25.58%	111	24.89%	130	29.08%	101	22.75%
Confusion (label 2)	456	34.11%	164	36.77%	125	27.96%	167	37.61%
Confusion (label 3)	340	25.43%	106	23.77%	117	26.17%	117	26.35%
Confusion (none)	403	30.14%	133	29.82%	147	32.89%	123	27.70%

(i) Label 1 represents the standard carbon-neutral label. Label 2 is the transparent carbon-neutral label with 95% CO_2 offsetting and 5% CO_2 reduction. Label 3 is the transparent carbon-neutral label with an equal share of CO_2 reduction and CO_2 offsetting (50%-50%).
(iI) This table is based on participants' responses to the question about which label they trust the most and which they find the most confusing among the three label choices (label 1, label 2, label 3) and a 'none' option. They are allowed for multiple selections. This question was asked later in the survey, after the choice experiment.

Table 5: Confusion with and trust for each carbon-neutral label

	MNL I	MXL I	MXL II	MXL III	MXL IV	MXL V
Status quo	-2.005*** (0.082)	-3.028*** (0.104)	-3.028*** (0.104)	-3.030*** (0.104)	-3.030*** (0.104)	-3.036*** (0.104)
Carbon-neutral	0.187*** (0.035)	0.377*** (0.043)	0.377*** (0.043)	0.514** (0.168)	-0.253 (0.232)	0.094 (0.265)
Organic	0.616*** (0.034)	0.673*** (0.054)	0.673*** (0.054)	0.673*** (0.054)	0.673*** (0.054)	0.670*** (0.054)
Ethical trade	0.620*** (0.034)	0.741*** (0.054)	0.741*** (0.054)	0.740*** (0.054)	0.742*** (0.054)	0.739*** (0.054)
Price	-0.463*** (0.016)	-0.662*** (0.026)	-0.662*** (0.026)	-0.662*** (0.026)	-0.663*** (0.026)	-0.663*** (0.026)
Interactions						
Carbon-neutral x Trust	–	–	–	-0.131 (0.192)	0.104 (0.197)	-0.049 (0.214)
Carbon-neutral x Confusion	–	–	–	-0.314 (0.185)	-0.259 (0.182)	-0.034 (0.179)
Carbon-neutral x Concern	–	–	–	0.364 (0.201)	0.441 (0.198)	0.145 (0.223)
Carbon-neutral x Only confused with transparent labels	–	–	–	–	-0.381*** (0.077)	-0.361*** (0.078)
Carbon-neutral x Only trust in transparent labels	–	–	–	–	-0.055 (0.082)	-0.081 (0.083)
Carbon-neutral x Climate worry	–	–	–	–	–	0.465 (0.217)
Carbon-neutral x Warm glow	–	–	–	–	–	0.422 (0.252)
Carbon-neutral x Guilt	–	–	–	–	–	0.249 (0.211)
Carbon-neutral x Social approval	–	–	–	–	–	-0.030 (0.243)
Carbon-neutral x Producer pays	–	–	–	–	–	-0.236 (0.195)
Carbon-neutral x Limited resources	–	–	–	–	–	-0.682*** (0.184)
Carbon-neutral x Lack of time	–	–	–	–	–	0.012 (0.078)
Carbon-neutral x Female	–	–	–	–	0.031 (0.076)	0.012 (0.078)
Carbon-neutral x Age	–	–	–	–	0.012*** (0.003)	0.011*** (0.003)
Carbon-neutral x High income	–	–	–	–	0.147 (0.079)	0.133 (0.078)
SD						
$\sigma_{\text{Carbon Neutral}}$	–	1.093*** (0.049)	1.091*** (0.049)	1.091*** (0.049)	1.054*** (0.048)	1.034*** (0.048)
σ_{Organic}	–	1.435*** (0.062)	1.437*** (0.062)	1.437*** (0.062)	1.432*** (0.062)	1.435*** (0.062)
$\sigma_{\text{Ethical Trade}}$	–	1.414*** (0.061)	1.415*** (0.061)	1.415*** (0.061)	1.417*** (0.061)	1.422*** (0.061)
σ_{Price}	–	–	0.734*** (0.022)	–	–	–
Observations	10,696	10,696	10,696	10,696	10,696	10,696
Participants	1,337	1,337	1,337	1,337	1,337	1,337
Log-likelihood	-10,229.25	-10,229.25	-10,229.25	-9,163.05	-9,151.42	-9,132.86

Robust standard errors are in parentheses.
Bonferroni-corrected p-values and significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.
Model Descriptions: MNL I is a Multinomial Logit (MNL) model.
MXL I is a Mixed Logit (MXL) model with randomized label attributes, fixed price.
MXL II is a MXL model with all attributes randomized except for the status quo.
MXL III, IV, and V are MXL models with randomized label attributes, fixed price, and including interaction terms.

Table 6: Comparison of results across six models for the full sample

	Full Sample	**Control**	**Treatment 1**	**Treatment 2**
Status quo	-3.036*** (0.104)	-3.306*** (0.180)	-2.978*** (0.173)	-2.825*** (0.184)
Carbon-neutral (random)	-0.253 (0.232)	-0.512 (0.427)	-0.220 (0.404)	-0.115 (0.367)
Organic (random)	0.673*** (0.054)	0.634*** (0.096)	0.693*** (0.092)	0.687*** (0.094)
Ethical Trade (random)	0.742*** (0.054)	0.749*** (0.095)	0.716*** (0.094)	0.759*** (0.092)
Price	-0.663*** (0.026)	-0.680*** (0.046)	-0.649*** (0.044)	-0.658*** (0.047)
Interactions				
Carbon-neutral x Trust	0.104 (0.197)	0.466 (0.366)	0.204 (0.324)	-0.269 (0.336)
Carbon-neutral x Confusion	-0.259 (0.182)	0.208 (0.330)	-0.575 (0.303)	-0.320 (0.302)
Carbon-neutral x Concern for carbon offsets	0.441 (0.198)	0.094 (0.390)	0.472 (0.334)	0.744 (0.313)
Carbon-neutral x Only confused with transparent labels	-0.381*** (0.077)	-0.403* (0.143)	-0.219 (0.134)	-0.480*** (0.127)
Carbon-neutral x Only trust in transparent labels	-0.055 (0.082)	0.092 (0.146)	-0.090 (0.144)	-0.134 (0.141)
Carbon-neutral x Female	0.031 (0.076)	0.047 (0.144)	0.066 (0.126)	-0.014 (0.127)
Carbon-neutral x Age	0.012*** (0.003)	0.010 (0.005)	0.013* (0.004)	0.011 (0.004)
Carbon-neutral x High income	0.147 (0.079)	0.188 (0.146)	0.166 (0.134)	0.123 (0.139)
Standard deviations				
$\sigma_{CarbonNeutral}$	1.054*** (0.048)	1.147*** (0.091)	1.001*** (0.084)	0.995*** (0.078)
$\sigma_{Organic}$	1.432*** (0.062)	1.425*** (0.114)	1.398*** (0.099)	1.465*** (0.114)
$\sigma_{EthicalTrade}$	1.417*** (0.061)	1.510*** (0.110)	1.351*** (0.105)	1.419*** (0.105)
Number of Observations	10696.00	3568.00	3576.00	3552.00
Number of Participants	1337.00	446.00	447.00	444.00
Log Likelihood	-9151.415	-3012.032	-3037.105	-3094.164

Table shows the Mixed Logit (MXL) model estimations with randomized label attributes, fixed price, and status quo.
Robust standard errors are reported in parentheses to the right of the estimates.
Bonferroni-corrected p-values and significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

Table 7: MXL (PS estimation)

4.3 Next Steps

This section provides further information about survey data and presents the planned future work, which includes robustness tests to ensure the reliability of the findings.

The first robustness check relates to the certainty levels. The data indicates that participants were generally confident in their choices, with over 83% selecting a certainty level of 6 or higher on a scale from 0 to 10. The certainty level will interact with the label attributes to check if the results change. Examining this interaction is important because participants' confidence in their choices may affect their WTP.

The second robustness check relates to the perceived consequentiality of the sur-

vey. More than 54% of respondents believed their answers could potentially influence tea prices or labeling policies (see Table A.4 in Appendix A.1). This perception of consequentiality suggests that participants took the survey seriously, although it raises concerns about potential biases due to attempts to influence outcomes. To address these concerns, I will compare the WTP estimates between those who perceived the survey as consequential and those who did not, also examining differences across subsamples. Comparing these groups is crucial to determine if perceived policy impact influences stated preferences, potentially introducing bias.

The third robustness check relates to attribute non-attendance. As shown in Table A.5 in Appendix A.1 , attribute non-attendance was notable: 33% did not consider the carbon-neutral label, 34% organic label, 24% ethical trade, and 20% price attribute, whereas 20% considered all attributes. These findings highlight that participants assigned different levels of importance to attributes, emphasizing the need for robustness checks regarding attribute non-attendance. I will interact binary indicators for attribute non-attendance with label attributes to evaluate whether non-attendance to attributes influences my results.

Fourth, I will run robustness checks using sub-samples that exclude observations from participants who did not pass the attention and manipulation checks. The attention and manipulation checks yielded mixed results. While 62% of participants correctly identified the number of labels on the choice cards, 21% of the control group misinterpreted the carbon-neutral label, compared to 11% and 13% in the treatment groups. Additionally, 49% accurately defined CO_2 offsetting, while others either chose incorrect definitions or did not remember (see Appendix A.1, Table A.4). Excluding these participants may be important to ensure data quality by removing responses that may not reflect true preferences.

Fifth, I will conduct a robustness check related to survey completion time. The

average survey duration was slightly longer than anticipated, totaling approximately 14 minutes and 5 seconds, with the choice experiment taking about 3 minutes and 5 seconds. I will exclude speeders as part of these robustness checks.

Sixth, I will validate the findings using open-ended contingent valuation (CVM) questions to compare these distributions with WTP estimates from the choice experiment. Validating the findings with open-ended CVM questions helps to cross-verify the differences between samples and assess the consistency between different valuation methods.

Seventh, I will run mixed logit models in 'WTP space' rather than ' preference space', as well as mixed latent class models. Running these alternative models allows for different specifications that may capture preference heterogeneity differently, providing robustness to the results.

Finally, I plan to estimate heterogeneous WTP across categories such as income levels, employment status, gender, and age to determine whether significant differences exist among these groups.

5 Conclusion

In the growing criticisms for carbon-neutral labels and CO_2 offsets, this study examined whether transparency in carbon-neutral labeling affects consumer WTP. Using a discrete choice experiment with a control group and two treatment groups, the findings reveal that while UK tea consumers are willing to pay a premium for carbon-neutral labels, there is no statistically significant difference in WTP between standard labels and those that transparently display the shares of CO_2 offsetting and CO_2 reduction. Additionally, no significant difference was observed between labels indicating different shares of CO_2 offsetting and reduction (95% offsetting + 5% reduction vs.

50% offsetting + 50% reduction). These results suggest that there is no evidence that CO_2 offsetting and reduction composition of carbon-neutral labels influences consumers' WTP.

These findings have important implications for both policymakers and companies. Label regulation and transparency are important due to recent developments and concerns about CO_2 offsetting. However, based on this study's findings, consumers might not naturally value transparency on carbon-neutral labels. Policymakers can advocate for transparency by disclosing CO_2 offsetting and reduction shares on labels and also consider additional measures, such as educational campaigns, to inform consumers clearly about the differences between CO_2 offsetting and CO_2 reduction and their environmental impacts.

References

Adamowicz, W., J. Louviere, and M. Williams (1994). Combining revealed and stated preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management 26*(3), 271–292.

Akaichi, F., R. M. Nayga Jr, and L. L. Nalley (2017). Are there trade-offs in valuation with respect to greenhouse gas emissions, origin and food miles attributes? *European Review of Agricultural Economics 44*(1), 3–31.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics 84*(3), 488–500.

Aldred, J. (2012). The ethics of emissions trading. *New Political Economy 17*(3), 339–360.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal 100*(401), 464–477.

Becken, S. and B. Mackey (2017). What role for offsetting aviation greenhouse gas emissions in a deep-cut carbon world? *Journal of Air Transport Management 63*, 71–83.

Bek, D. (2022). Pricing sustainable shipping of coffee: Consumersâ preferences and willingness to pay for emission reductions and offsets. *Junior Management Science*.

Birkenberg, A., M. E. Narjes, B. Weinmann, and R. Birner (2021). The potential of carbon neutral labeling to engage coffee consumers in climate change mitigation. *Journal of Cleaner Production 278*, 123621.

Blasch, J. and M. Farsi (2014). Context effects and heterogeneity in voluntary carbon offsetting–a choice experiment in Switzerland. *Journal of Environmental Economics and Policy 3*(1), 1–24.

Brécard, D. (2014). Consumer confusion over the profusion of eco-labels: Lessons from a double differentiation model. *Resource and Energy Economics 37*, 64–84.

Brécard, D. (2017). Consumer misperception of eco-labels, green market structure and welfare. *Journal of Regulatory Economics 51*, 340–364.

Brounen, D. and N. Kok (2011). On the economics of energy labels in the housing market. *Journal of Environmental Economics and Management 62*(2), 166–179.

Brouwer, R., L. Brander, and P. Van Beukering (2008). "a convenient truth": air travel passengersâ willingness to pay to offset their co_2 emissions. *Climatic Change 90*, 299–313.

Brunnschweiler, C., I. Edjekumhene, and P. Lujala (2021). Does information matter? transparency and demand for accountability in ghana's natural resource revenue management. *Ecological Economics 181*, 106903.

Bumpus, A. G. and D. M. Liverman (2008). Accumulation by decarbonization and the governance of carbon offsets. *Economic Geography 84*(2), 127–155.

Calel, R., J. Colmer, A. Dechezleprêtre, and M. Glachant (2021). Do carbon offsets offset carbon?

Carattini, S., F. Dvorak, I. Logar, and B. Özdemir Oluk (2024). Demand for carbon-neutral products.

Carattini, S., K. Gillingham, X. Meng, and E. Yoeli (2022). Peer-to-peer solar and social rewards: Evidence from a field experiment. CESifo Working Paper 10173, Center for Economic Studies and ifo Institute (CESifo), Munich.

Carattini, S. and A. Tavoni (2016). How green are green economists? *CSLF Articles* (9).

Changing Markets Foundation (2023). Feeding us greenwash. `http://changingmarkets.org/wp-content/uploads/2023/03/Feeding-Us-Greenwash-web.pdf`. Accessed: 2023-06-10.

Chen, N., Z.-H. Zhang, S. Huang, and L. Zheng (2018). Chinese consumer responses to carbon labeling: Evidence from experimental auctions. *Journal of Environmental Planning and Management 61*(13), 2319–2337.

Climate Portal (2023). Carbon offsets. `https://climate.mit.edu/explainers/carbon-offsets`. Accessed: 2024-02-19.

ClimatePartner (2023). Carbon neutral. what does that actually mean? `https://www.climatepartner.com/en/knowledge/glossary/carbon-neutral`. Accessed: 2024-02-19.

Crippa, M., E. Solazzo, D. Guizzardi, F. Monforti-Ferrario, F. N. Tubiello, and A. Leip (2021). Food systems are responsible for a third of global anthropogenic ghg emissions. *Nature Food 2*(3), 198–209.

Cummings, R. G. and L. O. Taylor (1999). Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. *American Economic Review 89*(3), 649–665.

de Magistris, T. and S. Pascucci (2014). The effect of the solemn oath script in hypothetical choice experiment survey: A pilot study. *Economics Letters 123*(2), 252–255.

Djimeu, E. W. and D.-G. Houndolo (2016). Power calculation for causal inference in social science: sample size and minimum detectable effect determination. *Journal of Development Effectiveness 8*(4), 508–527.

Dorner, Z. (2019). A behavioral rebound effect. *Journal of Environmental Economics and Management 98*, 102257.

Drichoutis, A. C., J. L. Lusk, and V. Pappa (2016). Elicitation formats and the WTA/WTP gap: A study of climate neutral foods. *Food Policy 61*, 141–155.

European Parliament (2023). EU to ban greenwashing and improve consumer information on product durability. `https://www.europarl.europa.eu/news/en/press-room/20230918IPR05412/eu-to-ban-greenwashing-and-improve-consumer-information-on-product-durability`. Accessed: 2023-11-06.

Feucht, Y. and K. Zander (2018). Consumers' preferences for carbon labels and the underlying reasoning. a mixed methods approach in 6 european countries. *Journal of Cleaner Production 178*, 740–748.

Gov.uk (2021a). Education and training statistics for the UK: 2021. `https://explore-education-statistics.service.gov.uk/find-statistics/education-and-training-statistics-for-the-uk`. Accessed on 3 November 2024.

Gov.uk (2021b). Male and female populations. `https://www.`

ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/male-and-female-populations/latest/. Accessed on 3 November 2024.

Grebitus, C., B. Steiner, and M. Veeman (2013). Personal values and decision making: evidence from environmental footprint labeling in Canada. *American Journal of Agricultural Economics 95*(2), 397–403.

Hanley, N., D. MacMillan, R. E. Wright, C. Bullock, I. Simpson, D. Parsisson, and B. Crabtree (1998). Contingent valuation versus choice experiments: estimating the benefits of environmentally sensitive areas in scotland. *Journal of agricultural economics 49*(1), 1–15.

Hanley, N., J. F. Shogren, and B. White (1997). *Environmental Economics: In Theory and Practice*. London: Macmillan.

Hanley, N., R. E. Wright, and V. Adamowicz (1998). Using choice experiments to value the environment. *Environmental and resource economics 11*, 413–428.

Hensher, D. A., J. M. Rose, and W. H. Greene (2005). *Applied Choice Analysis (2nd ed.)*. Cambridge University Press.

Heyes, A. and S. Martin (2017). Social labeling by competing NGOs: A model with multiple issues and entry. *Management Science 63*(6), 1800–1813.

Heyes, A. and S. Martin (2018). Inefficient NGO labels: Strategic proliferation and fragmentation in the market for certification. *Journal of Economics & Management Strategy 27*(2), 206–220.

Hooper, P., B. Daley, H. Preston, and C. Thomas (2008). An assessment of the potential of carbon offset schemes to mitigate the climate change implications of

future growth of UK aviation. Final OMEGA Project Report. Centre for Air Transport and the Environment, Manchester Metropolitan University.

Howard, G., B. E. Roe, E. C. Nisbet, and J. F. Martin (2017). Hypothetical bias mitigation techniques in choice experiments: Do cheap talk and honesty priming effects fade with repeated choices? *Journal of the Association of Environmental and Resource Economists 4*(2), 543–573.

Hyams, K. and T. Fawcett (2013). The ethics of carbon offsetting. *Wiley Interdisciplinary Reviews: Climate Change 4*(2), 91–98.

International Initiative for Impact Evaluation (3ie) (2016). Power calculation for causal inference in social science: Sample size and minimum detectable effect determination [excel spreadsheet]. `https://www.3ieimpact.org/evidence-hub/publications/working-papers/power-calculation-causal-inference-social-science-sample`.

Kotchen, M. J. (2009). Voluntary provision of public goods for bads: A theory of environmental offsets. *The Economic Journal 119*(537), 883–899.

Kraft-Todd, G. T., B. Bollinger, K. Gillingham, S. Lamp, and D. G. Rand (2018). Credibility-enhancing displays promote the provision of non-normative public goods. *Nature 563*, 245–248.

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of political economy 74*(2), 132–157.

MacKerron, G. J., C. Egerton, C. Gaskell, A. Parpia, and S. Mourato (2009). Willingness to pay for carbon offset certification and co-benefits among (high-) flying young adults in the uk. *Energy policy 37*(4), 1372–1381.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (Ed.), *Frontiers in Econometrics*, pp. 105–142. Academic Press New York, New York, NY, USA.

McKinsey & Company (2023). A blueprint for scaling voluntary carbon markets to meet the climate challenge. https://www.mckinsey.com/capabilities/sustainability/our-insights/a-blueprint-for-scaling-voluntary-carbon-markets-to-meet-the-climate-challenge. Accessed: March 2023.

Mortensen, C. R., R. Neel, R. B. Cialdini, C. M. Jaeger, R. P. Jacobson, and M. M. Ringel (2019). Trending norms: A lever for encouraging behaviors performed by the minority. *Social Psychological and Personality Science 10*(2), 201–210.

Office for National Statistics (2022). UK population estimates, mid-2021. https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2021. Accessed on 3 November 2024.

Office for National Statistics (2024a). Average household income, UK: financial year ending 2023. https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/householddisposableincomeandinequality/financialyearending2023. Accessed on 3 November 2024.

Office for National Statistics (2024b). Employment in the UK: October 2024. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/employmentintheuk/october2024. Accessed on 3 November 2024.

Office for National Statistics (2024c, February). RPI: Ave price - Tea bags, per 250g. https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/cznq/mm23. Release date: 14 February 2024.

Onozaka, Y. and D. T. McFadden (2011). Does local labeling complement or compete with other sustainable labels? A conjoint analysis of direct and joint values for fresh produce claim. *American Journal of Agricultural Economics 93*(3), 693–706.

Pigou, A. C. (1920). *The Economics of Welfare*. London: Macmillan.

Poe, G. L., K. L. Giraud, and J. B. Loomis (2005). Computational methods for measuring the difference of empirical distributions. *American Journal of Agricultural Economics 87*(2), 353–365.

Poret, S. (2019). Label wars: Competition among NGOs as sustainability standard setters. *Journal of Economic Behavior & Organization 160*, 1–18.

Roemer, N., G. C. Souza, C. Tröster, and G. Voigt (2023). Offset or reduce: How should firms implement carbon footprint reduction initiatives? *Production and Operations Management 32*(9), 2940–2955.

Schneider, L. and A. Kollmuss (2015). Perverse effects of carbon markets on hfc-23 and sf6 abatement projects in russia. *Nature Climate Change 5*(12), 1061–1063.

Schneider, L., A. Kollmuss, and M. Lazarus (2015). Addressing the risk of double counting emission reductions under the UNFCCC. *Climatic Change 131*(4), 473–486.

Sparkman, G. and G. M. Walton (2017). Dynamic norms promote sustainable behavior, even if it is counternormative. *Psychological Science 28*(11), 1663–1674.

Statista (2023). Global per capita tea consumption by country. Accessed: 2024-02-19.

Swiss International Air Lines (2024). Sustainable choices. Accessed: 2024-10-09.

Tea and Herbal Association (2024). Tea facts. https://www.tea.co.uk/tea-facts. Accessed: 2024-02-19.

Tea and Infusions Association (2022). UK Tea Census Report 2022. Accessed: 2024-02-19.

The Guardian (2023a). Delta air lines lawsuit carbon neutrality. https://www.theguardian.com/environment/2023/may/30/delta-air-lines-lawsuit-carbon-neutrality-aoe. Accessed: 2023-11-06.

The Guardian (2023b, January). Revealed: more than 90% of rainforest carbon offsets by biggest provider are worthless, analysis shows. Accessed: 2024-11-01.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

Trencher, G., S. Nick, J. Carlson, and M. Johnson (2024). Demand for low-quality offsets by major companies undermines climate integrity of the voluntary carbon market. *Nature Communications 15*(1), 6863.

Ziegler, A., J. Schwarzkopf, and V. H. Hoffmann (2012). Stated versus revealed knowledge: Determinants of offsetting CO_2 emissions from fuel consumption in vehicle use. *Energy Policy 40*, 422–431.

Appendix

A Survey Data

A.1 Descriptive Statistics

This section presents the descriptive statistics of the main survey data. Table A.1 includes the age, gender, education, employment, and income of the participants, while Table A.2 shows the socio-demographic comparison between tea drinkers/purchasers and never tea drinkers/purchasers. Table A.3 presents participants' tea consumption habits, including their frequency of tea drinking, purchasing habits, and preferences for tea blends. Table A.4 shows the survey clarity, attention checks, and completion time, while Table A.5 presents the levels of certainty, attribute non-attendance, consequentiality, and protest responses of participants. Table A.6 presents the level of agreement with various statements using a Likert scale. Finally, Table A.7 shows covariate balance.

	Full		Control		Treatment 1		Treatment 2	
	N	Share	N	Share	N	Share	N	Share
	1337	100.00%	446	16.68%	447	16.72%	444	16.60%
Age								
18 - 34 years	411	30.74%	136	30.49%	134	29.98%	141	31.76%
35 - 54 years	477	35.68%	157	35.20%	161	36.02%	159	35.81%
55+ years	449	33.58%	153	34.30%	152	34.00%	144	32.43%
Gender								
Male	652	48.77%	226	50.67%	210	46.98%	216	48.65%
Female	675	50.49%	217	48.65%	232	51.90%	226	50.90%
Non-binary	9	0.67%	3	0.67%	5	1.12%	1	0.23%
I prefer not to say	1	0.07%	-	0.00%	-	0.00%	1	0.23%
Education								
Primary school	3	0.22%	2	0.45%	1	0.22%	-	0.00%
Secondary school: High school or equivalent	364	27.23%	132	29.60%	125	27.96%	107	24.10%
Post-secondary vocational training (up to 1 year)	55	4.11%	16	3.59%	27	6.04%	12	2.70%
Post-secondary vocational training (2 and more years)	159	11.89%	44	9.87%	58	12.98%	57	12.84%
Post-secondary academic below-degree level qualification (up to 1 year)	82	6.13%	32	7.17%	23	5.15%	27	6.08%
Post-secondary academic below-degree level qualification (2 and more years)	173	12.94%	49	10.99%	50	11.19%	74	16.67%
Bachelors or equivalent first degree qualification (e.g., BA, BSc, BEng)	311	23.26%	111	24.89%	95	21.25%	105	23.65%
Masters or equivalent higher degree level qualification (e.g., MA, MSc, MBA)	140	10.47%	48	10.76%	48	10.74%	44	9.91%
PhD or equivalent doctoral level qualification (e.g., PhD)	39	2.92%	8	1.79%	19	4.25%	12	2.70%
None of above	11	0.82%	4	0.90%	1	0.22%	6	1.35%

This figure displays the number of participants (N) and their share of the total sample.

Table A.1: Summary statistics of sociodemographics

	Full		Control		Treatment 1		Treatment 2	
	N	Share	N	Share	N	Share	N	Share
	1337	100.00%	446	16.68%	447	16.72%	444	16.60%
Employment								
Full or part time employment	768	57.44%	248	55.61%	260	58.17%	260	58.56%
Self-employed	95	7.11%	32	7.17%	29	6.49%	34	7.66%
Unemployed	83	6.21%	25	5.61%	28	6.26%	30	6.76%
Retired	267	19.97%	90	20.18%	88	19.69%	89	20.05%
Looking after family or home	79	5.91%	30	6.73%	27	6.04%	22	4.95%
Full-time student	32	2.39%	14	3.14%	11	2.46%	7	1.58%
None of above	13	0.97%	7	1.57%	4	0.89%	2	0.45%
Annual household income after taxes								
Under £10,000	76	5.68%	23	5.16%	22	4.92%	31	6.98%
£10,000 - £19,999	180	13.46%	51	11.43%	64	14.32%	65	14.64%
£20,000 - £29,999	252	18.85%	85	19.06%	88	19.69%	79	17.79%
£30,000 - £39,999	175	13.09%	70	15.70%	48	10.74%	57	12.84%
£40,000 - £49,999	140	10.47%	44	9.87%	55	12.30%	41	9.23%
£50,000 - £59,999	103	7.70%	32	7.17%	34	7.61%	37	8.33%
£60,000 - £69,999	69	5.16%	25	5.61%	16	3.58%	28	6.31%
£70,000 - £79,999	50	3.74%	15	3.36%	18	4.03%	17	3.83%
£80,000 - £89,999	43	3.22%	10	2.24%	21	4.70%	12	2.70%
£90,000 - £99,999	31	2.32%	12	2.69%	9	2.01%	10	2.25%
£100,000 - £129,999	69	5.16%	21	4.71%	27	6.04%	21	4.73%
£130,000 or more	49	3.66%	16	3.59%	15	3.36%	18	4.05%
No answer	100	7.48%	42	9.42%	30	6.71%	28	6.31%
Environmental organization membership								
Member	149	11.14%	46	10.31%	55	12.30%	48	10.81%

This figure displays the number of participants (N) and their share.

Table A.1: Summary statistics of sociodemographics

Covariate	**Tea drinkers/purchasers** (N = 1,339)	**Non-tea drinkers** (N = 80)	**Non-tea purchasers** (N = 142)
Age	Mean: 47.54 (SD: 16.52)	Mean: 53.65 (SD: 14.71)	Mean: 48.79 (SD: 17.64)
Female, n (%)	677 (50.49%)	50 (62.5%)	58 (40.85%)
High education, n (%)	665 (49.59%)	25 (31.25%)	57 (40.14%)

This table displays the summary statistics (mean and standard deviation or proportion) for tea drinkers/purchasers and never-tea drinkers, including age (a continuous variable), female (a binary variable coded as 1 for females and 0 for males, non-binary, or 'prefer not to say' categories), and high education (a binary variable coded as 1 for those with a post-secondary academic qualification below a degree level which is 2 years or more).

Table A.2: Comparison between tea drinkers/purchasers and never tea drinkers/purchasers

	Full		Control		Treatment 1		Treatment 2	
	N	Mean	N	Mean	N	Mean	N	Mean
	1337	100.00%	446	16.68%	447	16.72%	444	16.60%
Tea drinking frequency								
Daily	1033	77.26%	337	75.56%	345	77.18%	351	79.05%
Once a week	140	10.47%	53	11.88%	50	11.19%	37	8.33%
Once every two weeks	62	4.64%	19	4.26%	22	4.92%	21	4.73%
Once a month	37	2.77%	11	2.47%	10	2.24%	16	3.60%
Several times a year	65	4.86%	26	5.83%	20	4.47%	19	4.28%
Tea purchase frequency								
Once a week	253	18.92%	95	21.30%	83	18.57%	75	16.89%
Once every two weeks	273	20.42%	79	17.71%	99	22.15%	95	21.40%
Once a month	507	37.92%	174	39.01%	161	36.02%	172	38.74%
Several times a year	262	19.60%	84	18.83%	87	19.46%	91	20.50%
Once a year	42	3.14%	14	3.14%	17	3.80%	11	2.48%
Preferred tea blend								
Black tea	963	72.03%	318	71.30%	327	73.15%	318	71.62%
Green tea	213	15.93%	76	17.04%	62	13.87%	75	16.89%
Herbal tea	105	7.85%	36	8.07%	39	8.72%	30	6.76%
Other	56	4.19%	16	3.59%	19	4.25%	21	4.73%
Regular grocery shopping								
Carbon-neutral label	168	12.57%	54	12.11%	61	13.65%	53	11.94%
Organic label	321	24.01%	109	24.44%	111	24.83%	101	22.75%
Fair trade label	486	36.35%	168	37.67%	161	36.02%	157	35.36%
No sustainability label	252	18.85%	80	17.94%	82	18.34%	90	20.27%
Other label	403	30.14%	130	29.15%	135	30.20%	138	31.08%
No knowledge about label	26	1.94%	15	3.36%	4	0.89%	7	1.58%
Price & amount	N	Mean	N	Mean	N	Mean	N	Mean
Price paid for tea (£)	1060	7.79	348	10.21	349	5.72	363	7.46
Tea amount (grams)	290	158.18	91	182.65	98	155.33	101	138.90
Tea amount (teabags)	1045	123.47	349	119.62	348	129.46	348	121.34

This figure displays the number of participants (N), along with either their share of the total sample for binary variables or the mean for continuous variables.

Table A.3: Summary statistics: Tea consumption

	Full		Control		Treatment 1		Treatment 2	
	N	Mean	N	Mean	N	Mean	N	Mean
Survey completion time								
Choice experiment time (min.)	1337	3.05	446	4.63	447	1.85	444	2.67
Full survey time (min.)	1337	14.05	446	14.57	447	11.42	444	16.17
Device								
Smartphone	754	56.39%	242	54.26%	262	58.61%	250	56.31%
Tablet	42	3.14%	12	2.69%	13	2.91%	17	3.83%
Desktop	541	40.46%	192	43.05%	172	38.48%	177	39.86%
Attention and manipulation								
The number of labels								
One label	65	4.86%	28	6.28%	20	4.47%	17	3.83%
Two labels	337	25.21%	104	23.32%	119	26.62%	114	25.68%
Three labels (correct resp.)	827	61.85%	281	63.00%	267	59.73%	279	62.84%
Not remember.	108	8.08%	33	7.40%	41	9.17%	34	7.66%
Percentage info. on label								
Yes (correct for samples 2 & 3)	567	42.41%	95	21.30%	242	54.14%	230	51.80%
No (correct for sample 1)	273	20.42%	163	36.55%	51	11.41%	59	13.29%
I do not remember.	497	37.17%	188	42.15%	154	34.45%	155	34.91%
Definition of CO_2 offsetting								
Correct resp.	653	48.84%	222	49.78%	227	50.78%	204	45.95%
False resp. (CO_2 reduction def.)	478	35.75%	156	34.98%	150	33.56%	172	38.74%
Not remember.	206	15.41%	68	15.25%	70	15.66%	68	15.32%
Survey clarity								
Clear instructions								
No.	10	0.75%	-	0.00%	5	1.12%	5	1.13%
Yes.	1327	99.25%	446	100.00%	442	98.88%	439	98.87%
Confusion with survey								
No.	1246	93.19%	424	95.07%	419	93.74%	403	90.77%
Yes.	91	6.81%	22	4.93%	28	6.26%	41	9.23%

This figure displays the number of participants (N), along with either their share of the total sample for binary variables or the mean for continuous variables.

Table A.4: Summary statistics: survey clarity, and attention

	Full		Control		Treatment 1		Treatment 2	
	N	Share	N	Share	N	Share	N	Share (%)
Certainty level (0-10)								
0 (very uncertain) -2	11	0.82%	3	0.67%	5	1.12%	3	0.68%
2 - 4	38	2.84%	13	2.91%	15	3.36%	10	2.25%
4 - 6	177	13.24%	59	13.23%	62	13.87%	56	12.61%
6 - 8	533	39.87%	180	40.36%	164	36.69%	189	42.57%
8 - 10 (very certain)	578	43.23%	191	42.83%	201	44.97%	186	41.89%
Attribute non-attendance								
Not considered: CN label	443	33.13%	135	30.27%	154	34.45%	154	34.68%
Not considered: Organic label	453	33.88%	129	28.92%	154	34.45%	170	38.29%
Not considered: Ethical trade l.	318	23.78%	98	21.97%	108	24.16%	112	25.23%
Not considered: Price	267	19.97%	104	23.32%	73	16.33%	90	20.27%
Considered all attributes	413	30.89%	145	32.51%	143	31.99%	125	28.15%
Consequentiality								
Yes (policy and price impact).	720	53.85%	231	51.79%	257	57.49%	232	52.25%
No (no impact).	617	46.15%	215	48.21%	190	42.51%	212	47.75%
Consistent SQ responses								
Total responses	31	2.32%	9	2.02%	10	2.24%	12	2.70%
The products were too expensive.	9	0.67%	4	0.90%	2	0.45%	3	0.68%
I oppose one or more of the labels.[P]	3	0.22%	1	0.22%	-	0.00%	2	0.45%
Insufficient information was provided about the labels or the products.[P]	7	0.52%	2	0.45%	2	0.45%	3	0.68%
I prefer to spend money on other social and environmental responsibility projects.	3	0.22%	1	0.22%	2	0.45%	-	0.00%
I disagree with the way the choice question was asked. [P]	2	0.15%	-	0.00%	-	0.00%	2	0.45%
Other reason	7	0.52%	1	0.22%	4	0.89%	2	0.45%

This table displays the number of participants (N) and their share. Protest responses indicated by [P].

Table A.5: Summary statistics: certainty, attribute non-attendance, consequentiality, and protest responses

Table A.6: Agreement to the list of statements - full sample

Statement	**SD**	**MD**	**SltD**	**N**	**SltA**	**MA**	**SA**
I worry about climate change	6.13	5.46	7.70	13.16	27.97	22.51	17.05
Limited financial resources prevent me from buying climate-friendly products	6.06	4.56	8.53	16.98	23.41	21.84	18.62
Lack of time prevents me from buying climate-friendly products	16.53	12.79	19.45	24.46	12.94	8.68	5.16
My positive emotions increase when I choose climate-friendly products	8.38	6.36	9.20	32.16	22.44	13.31	8.15
I feel guilty when I buy conventional products	15.11	11.14	18.77	22.14	16.90	8.60	7.33
Most people approve of my choice of climate-friendly products	5.83	3.66	8.23	47.20	15.78	12.27	7.03
Producers are responsible for climate change mitigation costs	2.39	3.59	8.45	24.01	23.64	21.77	16.16
I trust carbon-neutral labels	5.98	5.98	12.72	31.64	25.28	13.31	5.09
I am confused about carbon-neutral labels	5.09	8.23	14.81	26.85	29.32	10.17	5.53
I am concerned about carbon offsets	5.39	4.64	13.46	30.89	25.36	11.67	8.60
Carbon offsetting reduces carbon emissions	5.76	4.79	12.49	31.34	27.75	12.42	5.46
Carbon offsetting allows producers to continue polluting	1.94	2.92	8.15	30.44	30.67	15.26	10.62
Carbon offsetting is a misleading sense of relief	2.62	1.65	6.28	27.52	31.79	19.00	11.14
Carbon offsetting is a form of greenwashing	1.87	1.94	6.51	33.81	27.90	15.78	12.19

Note: SD - Strongly Disagree, MD - Mostly Disagree, SltD - Slightly Disagree, N - Neutral, SltA - Slightly Agree, MA - Mostly Agree, SA - Strongly Agree.

Covariate	C	T 1	T 2	C vs. T1	C vs. T2	T1 vs. T2
Trust	0.61	0.60	0.61	0.77	0.11	-0.65
	(0.20)	(0.22)	(0.21)	p=0.44	p=0.92	p=0.52
Confusion	0.61	0.59	0.60	1.02	0.93	-0.10
	(0.20)	(0.21)	(0.21)	p=0.31	p=0.35	p=0.92
Concern for carbon offsets	0.63	0.62	0.62	1.30	1.13	-0.20
	(0.20)	(0.22)	(0.21)	p=0.20	p=0.26	p=0.84
Only trust in carbon-neutral labels	0.57	0.57	0.54	-0.10	0.94	1.03
	(0.50)	(0.50)	(0.50)	p=0.92	p=0.35	p=0.30
Only confused with transparent labels	0.45	0.38	0.50	2.20	-1.27	-3.48
	(0.50)	(0.49)	(0.50)	p=0.03	p=0.20	p=0.00
Climate worry	0.71	0.68	0.70	1.47	0.41	-1.05
	(0.23)	(0.25)	(0.24)	p=0.14	p=0.68	p=0.29
Warm glow	0.63	0.60	0.60	2.19	2.25	0.05
	(0.21)	(0.24)	(0.23)	p=0.03	p=0.03	p=0.96
Guilt	0.54	0.52	0.52	1.22	1.48	0.27
	(0.25)	(0.25)	(0.25)	p=0.22	p=0.14	p=0.79

This table continues on the next page, where table notes are provided.

Table A.7: Covariate balance

Covariate	C	T 1	T 2	C vs. T1	C vs. T2	T1 vs. T2
Social approval	0.63	0.61	0.59	1.53	3.06	1.47
	(0.19)	(0.21)	(0.21)	p=0.13	p=0.00	p=0.14
Producer pays principle	0.71	0.69	0.71	1.16	0.21	-0.94
	(0.20)	(0.22)	(0.21)	p=0.25	p=0.83	p=0.35
Limited resources	0.71	0.68	0.70	2.09	0.63	-1.43
	(0.23)	(0.25)	(0.24)	p=0.04	p=0.15	p=0.15
Lack of time	0.52	0.48	0.51	2.14	0.75	-1.38
	(0.24)	(0.24)	(0.25)	p=0.03	p=0.46	p=0.17
Female	0.49	0.52	0.51	-0.97	-0.67	0.30
	(0.50)	(0.50)	(0.50)	p=0.33	p=0.50	p=0.77
High income	0.49	0.50	0.48	-0.50	0.27	0.77
	(0.50)	(0.50)	(0.50)	p=0.62	p=0.79	p=0.44
Age	47.74	47.94	46.92	-0.19	0.74	0.93
	(16.46)	(16.46)	(16.69)	p=0.85	p=0.46	p=0.35

This table shows the covariate balance. C represents the control group, T1 represents treatment group 1, and T2 represents treatment group 2. The first three columns show the means and standard deviations for each group, while the last three columns present the results of t-tests, including t-values and associated p-values. The table includes the following variables: trust (trust in carbon-neutral labels), confusion (confusion regarding carbon-neutral labels), concern (concerns about CO_2 offsetting), warm glow (positive emotions from purchasing climate-friendly products), guilt (guilt associated with choosing conventional products), social image (perceived social approval of buying climate-friendly products), polluter pays (belief that producers are responsible for covering the cost of climate change), time restriction (limited time preventing the choice of climate-friendly products), and financial constraint (financial limitations preventing the purchase of climate-friendly products). Trust, confusion, concern, warm glow, guilt, social approval, producer pays, time restriction, and financial constraint are measured on a 7-point Likert scale and normalized by dividing each score by seven. Additionally, there are binary variables: only trust in transparent labels (trusts at least one transparent label but not the standard label) and only confused with transparent labels (confused by at least one transparent label but not the standard label). Demographic variables include age (age of participants), high income (annual household income exceeding £40,000 after taxes), and female participants.

Table A.7: Covariate balance

A.2 Choice data

This section presents the choice experiment data. Table A.8 summarizes the overall choice design, including the different blocks, choice scenarios (cards), and respective attribute levels. Table A.9 details the specific choices available to participants within each block, presenting the number of alternatives (Alt A, Alt B, Alt SQ) and their respective distribution percentages for each choice card.

Block	Choice card	Price AltA	Price AltB	CN AltA	CN AltB	Organic AltA	Organic AltB	ET AltA	ET AltB
1	1	3.9	4.9	No	Yes	Yes	No	No	Yes
1	2	2.9	6.9	No	Yes	No	Yes	No	Yes
1	3	2.9	3.9	No	No	Yes	No	No	Yes
1	4	4.9	4.9	No	No	No	Yes	Yes	No
1	5	0.9	4.9	No	Yes	No	Yes	No	Yes
1	6	6.9	0.9	Yes	No	Yes	No	Yes	No
1	7	2.9	0.9	Yes	No	Yes	No	No	No
1	8	4.9	2.9	Yes	Yes	Yes	No	No	Yes
2	1	1.9	5.9	Yes	No	No	Yes	No	Yes
2	2	1.9	4.9	No	Yes	Yes	No	Yes	Yes
2	3	0.9	1.9	No	Yes	No	Yes	No	Yes
2	4	1.9	1.9	No	No	Yes	No	No	Yes
2	5	6.9	3.9	Yes	No	Yes	No	Yes	Yes
2	6	5.9	1.9	Yes	No	Yes	Yes	No	Yes
2	7	3.9	2.9	Yes	No	No	No	No	No
2	8	5.9	5.9	No	Yes	Yes	No	No	No

Table A.8: Choice design

Block	Subsample	N Alt A	N Alt B	N Alt SQ	Total	AltA	AltB	AltSQ
1	1	318	188	168	674	47.18%	27.89%	24.93%
1	2	362	165	147	674	53.71%	24.48%	21.81%
1	3	402	171	101	674	59.64%	25.37%	14.99%
1	4	226	203	245	674	33.53%	30.12%	36.35%
1	5	308	251	115	674	45.70%	37.24%	17.06%
1	6	173	332	169	674	25.67%	49.26%	25.07%
1	7	337	282	55	674	50.00%	41.84%	8.16%
1	8	110	447	117	674	16.32%	66.32%	17.36%
2	1	433	155	75	663	65.31%	23.38%	11.31%
2	2	506	103	54	663	76.32%	15.54%	8.14%
2	3	204	420	39	663	30.77%	63.35%	5.88%
2	4	258	342	63	663	38.91%	51.58%	9.50%
2	5	123	313	227	663	18.55%	47.21%	34.24%
2	6	77	537	49	663	11.61%	81.00%	7.39%
2	7	149	296	218	663	22.47%	44.65%	32.88%
2	8	186	147	330	663	28.05%	22.17%	49.77%

Table A.9: Choices: detailed information

A.3 Robustness Tests

This section shows the robustness test results for the estimations in Section 4 and Section 4.2. Tables A.10 and A.11 report the estimation results of the only-attribute MNL model and the corresponding WTP estimates, respectively. Similarly, Tables A.12 and A.13 present the only-attribute MXL model results and the estimated WTPs, where both the attributes and the price are randomized. These results are shown for the full sample as well as for all subsamples.

To explore the underlying mechanisms that are associated with consumer's probability of choosing carbon-neutral labeled tea, additional estimations are conducted. Tables A.14 and A.15 display the estimation outputs for all three subsamples and the full sample, incorporating variables related to trust, confusion, and concern, along with additional interaction terms listed in the tables.

	Full Sample	**Control**	**Treatment 1**	**Treatment 2**
Status quo	-2.005***	-2.105***	-1.989***	-1.926***
	(0.082)	(0.145)	(0.140)	(0.139)
Carbon-neutral	0.187***	0.171***	0.217***	0.175***
	(0.035)	(0.061)	(0.061)	(0.062)
Organic	0.616***	0.606***	0.637***	0.605***
	(0.034)	(0.059)	(0.059)	(0.059)
Ethical trade	0.620***	0.646***	0.614***	0.602***
	(0.034)	(0.058)	(0.059)	(0.057)
Price	-0.463***	-0.477***	-0.456***	-0.456***
	(0.016)	(0.028)	(0.027)	(0.028)
Observations	10696	3568	3576	3552
Participants	1337	446	447	444
Log-likelihood	-10229.25	-3361.62	-3440.22	-3423.39

Robust standard errors are in parentheses.
Bonferroni-corrected p-values and significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

Table A.10: MNL (PS estimation)

	Full Sample	 **Control**	 **Treatment 1**	 **Treatment 2**
$\text{MWTP}_{\text{Carbon neutral}}$	0.405***	0.359**	0.476***	0.383**
	(0.074)	(0.125)	(0.130)	(0.132)
$\text{MWTP}_{\text{Organic}}$	1.331***	1.272***	1.397***	1.327***
	(0.067)	(0.111)	(0.116)	(0.120)
$\text{MWTP}_{\text{Ethical trade}}$	1.341***	1.356***	1.346***	1.320***
	(0.071)	(0.119)	(0.126)	(0.123)

(i) MWTP is estimated by dividing the negative of the attribute coefficients by the price coefficient.
(ii) Robust standard errors in brackets are obtained using the Delta method.
(iii) Bonferroni-corrected p-values and significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$
(iv) Poe test results (adjusted for three comparisons) indicate that the bilateral differences between the control and treatment groups are not statistically significant: control vs. treatment 1: Poe statistic = 0.758 (n.s.); control vs. treatment 2: Poe statistic = 0.527 (n.s.); treatment 1 vs. treatment 2: Poe statistic = 0.337 (n.s.).

Table A.11: MWTP estimates (MNL, PS estimation)

	Full Sample	**Control**	**Treatment 1**	**Treatment 2**
Status quo	-3.028***	-3.308***	-2.972***	-2.825***
	(0.104)	(0.181)	(0.175)	(0.185)
Carbon-neutral (random)	0.377***	0.440***	0.432***	0.269***
	(0.043)	(0.074)	(0.074)	(0.074)
Organic (random)	0.673***	0.634***	0.694***	0.683***
	(0.054)	(0.096)	(0.093)	(0.094)
Ethical trade (random)	0.741***	0.745***	0.715***	0.756***
	(0.054)	(0.095)	(0.094)	(0.091)
Price (random)	-0.662***	-0.680***	-0.649***	-0.659***
	(0.026)	(0.046)	(0.044)	(0.047)
Standard deviations				
$\sigma_{CarbonNeutral}$	1.093***	1.188***	1.047***	1.061***
	(0.049)	(0.093)	(0.083)	(0.082)
$\sigma_{Organic}$	1.435***	1.421***	1.404***	1.463***
	(0.062)	(0.113)	(0.099)	(0.113)
$\sigma_{EthicalTrade}$	1.414***	1.505***	1.344***	1.412***
	(0.061)	(0.110)	(0.105)	(0.105)
σ_{Price}	0.734***	0.773***	0.680***	0.735***
	(0.022)	(0.039)	(0.050)	(0.038)
Observations	10696	3568	3576	3552
Participants	1337	446	447	444
Log-likelihood	-10229.25	-3361.62	-3440.22	-3423.39

Robust standard errors are in parentheses.
Bonferroni-corrected p-values and significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

Table A.12: MXL (PS estimation)

	Full Sample	 **Control**	 **Treatment 1**	 **Treatment 2**
$\text{MWTP}_{\text{Carbon neutral}}$	0.253***	0.334***	0.332***	0.091
	(0.066)	(0.109)	(0.109)	(0.114)
$\text{MWTP}_{\text{Organic}}$	0.966***	0.967***	1.020***	0.965***
	(0.056)	(0.090)	(0.103)	(0.101)
$\text{MWTP}_{\text{Ethical trade}}$	1.100***	1.129***	1.104***	1.103***
	(0.058)	(0.096)	(0.098)	(0.100)

(i) MWTP estimates and the standard errors are estimated by the Krinsky-Robb procedure.
(ii) Bonferroni-corrected p-values and significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.
(iii) Poe test results (adjusted for three comparisons(indicate that the bilateral differences between the control and treatment groups are not statistically significant: control vs. treatment 1: Poe statistic = 0.494 (n.s.); control vs. treatment 2: Poe statistic = 0.056 (n.s.); treatment 1 vs. treatment 2: Poe statistic = 0.068 (n.s.).

Table A.13: WTP estimates (MXL, PS estimation)

	Full Sample	**Control**	**Treatment 1**	**Treatment 2**
Status quo	-3.028***	-3.304***	-2.973***	-2.828***
	(0.104)	(0.180)	(0.175)	(0.184)
Carbon-neutral (random)	0.419**	0.165	0.591**	0.375
	(0.167)	(0.300)	(0.265)	(0.283)
Organic (random)	0.673***	0.634***	0.693***	0.683***
	(0.054)	(0.097)	(0.093)	(0.094)
Ethical trade (random)	0.740***	0.745***	0.714***	0.755***
	(0.054)	(0.095)	(0.094)	(0.091)
Price	-0.662***	-0.680***	-0.649***	-0.659***
	(0.026)	(0.046)	(0.044)	(0.047)
Interactions				
Carbon-neutral x Trust	-0.131	0.328	-0.026	-0.629
	(0.192)	(0.343)	(0.324)	(0.319)
Carbon-neutral x Confusion	-0.314	0.064	-0.572	-0.368
	(0.185)	(0.340)	(0.300)	(0.312)
Carbon-neutral x Concern	0.364	0.055	0.324	0.805
	(0.201)	(0.394)	(0.330)	(0.321)
Standard deviations				
$\sigma_{CarbonNeutral}$	1.091***	1.188***	1.042***	1.050***
	(0.049)	(0.092)	(0.083)	(0.080)
$\sigma_{Organic}$	1.437***	1.422***	1.408***	1.469***
	(0.062)	(0.113)	(0.099)	(0.114)
$\sigma_{EthicalTrade}$	1.415**	1.506***	1.347***	1.416***
	(0.061)	(0.110)	(0.105)	(0.105)
Observations	10696	3568	3576	3552
Participants	1337	446	447	444
Log-likelihood	-10229.25	-3361.62	-3440.22	-3423.39

Robust standard errors are in parentheses.
Bonferroni-corrected p-values and significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

Table A.14: MXL (PS estimation)

	Full Sample	**Control**	**Treatment 1**	**Treatment 2**
Status quo	-3.036*** (0.104)	-3.312*** (0.181)	-2.980*** (0.173)	-2.831*** (0.184)
Carbon-neutral (random)	0.094 (0.265)	-0.071 (0.506)	-0.102 (0.501)	0.395 (0.376)
Organic (random)	0.670*** (0.054)	0.633*** (0.097)	0.690*** (0.093)	0.679*** (0.094)
Ethical Trade (random)	0.739*** (0.054)	0.747*** (0.095)	0.715*** (0.094)	0.751*** (0.092)
Price	-0.663*** (0.026)	-0.681*** (0.046)	-0.650*** (0.044)	-0.660*** (0.047)
Interactions				
Carbon-neutral x Trust	-0.049 (0.214)	0.264 (0.403)	0.031 (0.358)	-0.424 (0.363)
Carbon-neutral x Confusion	-0.034 (0.179)	0.610 (0.329)	-0.428 (0.300)	-0.221 (0.304)
Carbon-neutral x Concern for carbon offsets	0.145 (0.223)	-0.152 (0.425)	0.244 (0.383)	0.408 (0.368)
Carbon-neutral x Only confused with transparent labels	-0.361*** (0.078)	-0.371 (0.146)	-0.220 (0.134)	-0.508*** (0.126)
Carbon-neutral x Only trust in transparent labels	-0.081 (0.083)	0.054 (0.146)	-0.153 (0.149)	-0.118 (0.138)
Carbon-neutral x Climate worry	0.465 (0.217)	0.017 (0.369)	0.169 (0.377)	1.400*** (0.358)
Carbon-neutral x Warm glow	0.422 (0.252)	0.936 (0.470)	0.165 (0.461)	-0.116 (0.415)
Carbon-neutral x Guilt	0.249 (0.211)	0.331 (0.410)	0.367 (0.385)	0.356 (0.349)
Carbon-neutral x Social approval	-0.030 (0.243)	-0.142 (0.470)	0.446 (0.406)	-0.607 (0.420)
Carbon-neutral x Producer pays	-0.236 (0.195)	-0.367 (0.406)	-0.232 (0.310)	-0.103 (0.318)
Carbon-neutral x Limited resources	-0.682*** (0.184)	-0.337 (0.354)	-0.476 (0.295)	-1.179*** (0.302)
Carbon-neutral x Lack of time	0.012 (0.078)	-0.020 (0.153)	0.069 (0.129)	-0.030 (0.129)
Carbon-neutral x Female	0.012 (0.078)	-0.020 (0.153)	0.069 (0.129)	-0.030 (0.129)
Carbon-neutral x Age	0.011*** (0.003)	0.009 (0.005)	0.013** (0.004)	0.010 (0.004)
Carbon-neutral x High income	0.133 (0.078)	0.188 (0.146)	0.166 (0.134)	0.123 (0.139)
Standard deviations				
$\sigma_{CarbonNeutral}$	1.034*** (0.048)	1.110*** (0.087)	0.984*** (0.085)	0.950*** (0.076)
$\sigma_{Organic}$	1.435*** (0.062)	1.430*** (0.114)	1.411*** (0.100)	1.466*** (0.113)
$\sigma_{EthicalTrade}$	1.422*** (0.061)	1.518*** (0.109)	1.358*** (0.105)	1.423*** (0.105)
Number of Observations	10696.00	3568.00	3576.00	3552.00
Number of Participants	1337.00	446.00	447.00	444.00
Log Likelihood	-9132.8575	-3003.0381	-3089.5178	-3021.8149

Robust standard errors are reported in parentheses.
Bonferroni-corrected p-values and significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

Table A.15: MXL (PS estimation)

B Survey

A.1 Survey instrument

This survey is part of a research project conducted by Eawag, the Swiss Federal Institute of Aquatic Science and Technology, and the University of St. Gallen in Switzerland. The project studies the role of sustainability labels on people's preferences for tea products. By participating in this survey, you are contributing to improving our understanding of this topic. Your participation is voluntary, your responses are anonymous and will be kept confidential. You can opt out anytime without providing a reason, in which case your responses will not be stored. You will receive a payment from the survey company after submitting a completed survey. The survey is expected to take 10 minutes.

While you can return to previous pages in the survey, you will not be able to change your answers. We would like to ask you to read each question carefully. Please refrain from seeking information online or from any other source while answering the survey. There is no right or wrong answer, we are simply interested in knowing your opinion.

Please indicate your consent to participate in this study by selecting the appropriate option.

☐ I consent to participate in this study.

☐ I do not consent to participate in this study.

1 What is your year of birth?

[4 digit number]

2 What is your gender?

- Male.
- Female.
- Non-binary.
- I prefer not to say.

3. What is the highest level of education you have completed?

- Primary school
- Secondary school: High school or equivalent
- Post-secondary vocational training (2 and more years)
- Post-secondary vocational training (up to 1 year)
- Post-secondary academic below-degree level qualification (2 and more years)
- Post-secondary academic below-degree level qualification (up to 1 year)
- Bachelors or equivalent first degree qualification (e.g., BA, BSc, BEng)
- Masters or equivalent higher degree level qualification (e.g., MA, MSc, MBA)
- PhD or equivalent doctoral level qualification (e.g., PhD)
- None of above

4 How frequently do you purchase tea?

- Once a week
- Once every two weeks
- Once a month
- Several times a year
- Once a year
- Never

5 How frequently do you drink tea?

- Daily
- Once a week
- Once every two weeks
- Once a month
- Several times a year
- Never

If you are participating in the survey on a smartphone, please turn it to a **horizontal** orientation for the best display of the upcoming survey section.

Imagine that a box of tea containing **80 teabags** (roughly 200 grams) as shown in this figure is your preferred tea type-be it black, green, or herbal tea-and in your favored packaging type, either teabags or loose leaf. In the following questions, you will be asked to choose between various versions of this tea, differentiated by the presence or absence of carbon-neutral, organic, and ethical trade labels, and price. Note that the tea product and the labels are hypothetical and designed for the purpose of this survey.

This specific **carbon-neutral (CO_2 neutral) label** indicates that the product's greenhouse gas emissions, measured in carbon equivalents, have been:

- **offset** (compensated) by investing in activities **outside of the company** such as tree planting projects; **or**
- **reduced within the company** in the last five years, such as through investments in cleaner production processes; **or**
- both **offset** and **reduced**.

This specific **organic label** indicates that the product contains **only organic ingredients** and no synthetic pesticides.

This specific **ethical trade label** indicates that the product is produced following **responsible labor practices**, which guarantee higher prices for exporters based on internationally recognized standards.

If you are participating in the survey on a smartphone, please turn it to a **vertical orientation** for the best display of the remainder part of survey.

In the next questions, you will be asked to choose between two tea products with certain labels and a given price and 'none of the two' option. Note that the questions are hypothetical, i.e. you are not required to pay for your choices at any point in the survey. All you have to do is to indicate your most preferred option as if you were choosing between such products in a supermarket.

We would like to inform you that people are likely to overstate their willingness to pay for a product in a survey and would not pay the stated amount in real life. Please consider how you would feel spending your money on such products in a real-life situation, and answer accordingly.

Please also consider your budget limitations. Depending on the amount you choose to spend on tea, you will have less money available for other products.

Please check the box if you agree with the statement below.

I **understand** the importance of providing truthful answers as if I was making a decision in a real-life setting and promise to provide honest and accurate responses to the questions that follow.

☐

6 Imagine a box containing 80 teabags or 200 grams of loose tea. In the following questions, please indicate which option you prefer the most.

If you are participating in the survey on a smartphone, please keep it in a vertical orientation for the best display.

CHOICE CARD 1/8

	Tea product A	Tea product B	None of the two
			
CO2 neutral label	✗ No label	CO2 neutral	NEITHER TEA A NOR TEA B
Organic label	Organic	✗ No label	
Ethical trade label	✗ No label	Ethical trade	
	Price of tea product A £3.9	Price of tea product B £4.9	Price £0
I choose...	☐	☐	☐

CHOICE CARD 2/8

	Tea product A	Tea product B	None of the two
			
CO_2 neutral label	✗ No label	CO_2 neutral	NEITHER TEA A NOR TEA B
Organic label	✗ No label	Organic	
Ethical trade label	✗ No label	Ethical trade	
	Price of tea product A £2.9	Price of tea product B £6.9	Price £0
I choose...	☐	☐	☐

CHOICE CARD 3/8

	Tea product A	Tea product B	None of the two
			
CO_2 neutral label	✗ No label	✗ No label	NEITHER TEA A NOR TEA B
Organic label	Organic	✗ No label	
Ethical trade label	✗ No label	Ethical trade	
	Price of tea product A £2.9	Price of tea product B £3.9	Price £0
I choose...	☐	☐	☐

CHOICE CARD 4/8

	Tea product A	Tea product B	None of the two
			
CO_2 neutral label	✗ No label	✗ No label	NEITHER TEA A NOR TEA B
Organic label	✗ No label	Organic	
Ethical trade label	Ethical trade	✗ No label	
	Price of tea product A £4.9	Price of tea product B £4.9	Price £0
I choose...	☐	☐	☐

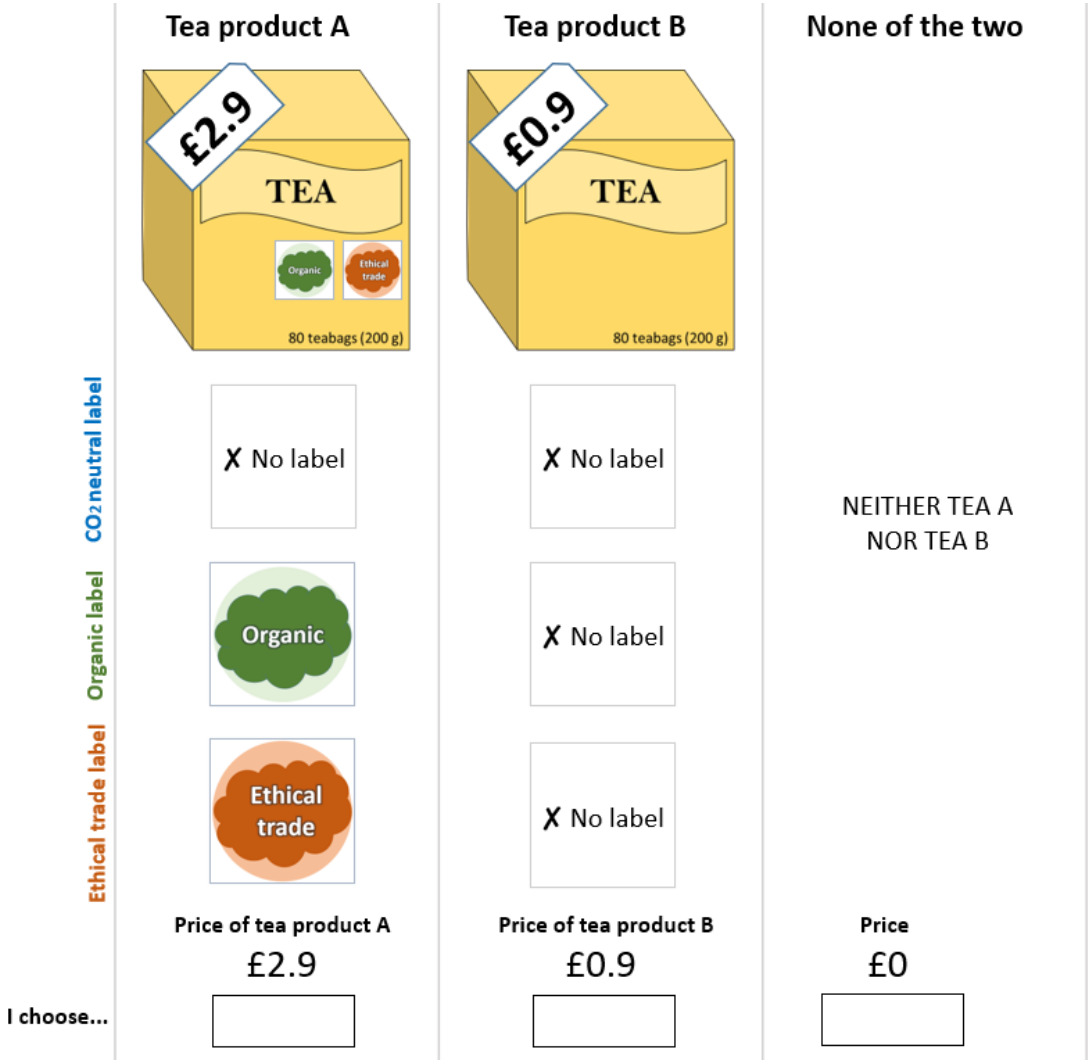
CHOICE CARD 5/8

	Tea product A	Tea product B	None of the two
			
CO_2 neutral label	✗ No label	CO_2 neutral	NEITHER TEA A NOR TEA B
Organic label	✗ No label	Organic	
Ethical trade label	✗ No label	Ethical trade	
	Price of tea product A	Price of tea product B	Price
	£0.9	£4.9	£0
I choose...	☐	☐	☐

CHOICE CARD 6/8

	Tea product A	Tea product B	None of the two
			
CO2 neutral label	CO2 neutral	✗ No label	NEITHER TEA A NOR TEA B
Organic label	Organic	✗ No label	
Ethical trade label	Ethical trade	✗ No label	
	Price of tea product A	Price of tea product B	Price
	£6.9	£0.9	£0
I choose...	☐	☐	☐

CHOICE CARD 7/8



CHOICE CARD 8/8

	Tea product A	Tea product B	None of the two
			
CO₂ neutral label	CO₂ neutral	CO₂ neutral	NEITHER TEA A NOR TEA B
Organic label	Organic	✗ No label	
Ethical trade label	✗ No label	Ethical trade	
	Price of tea product A £4.9	Price of tea product B £2.9	Price £0
I choose...	☐	☐	☐

7 How certain are you about your choices? Please use the slider below to indicate your level of certainty.

Very uncertain (0) ——slider——(10) Very certain

8. Could you please indicate the main reason why you always chose the "neither tea A nor tea B" option? (This question will only be shown to the participants who always chose the "none of the two" option.)

- The products were too expensive.
- I oppose one or more of the labels.
- Insufficient information was provided about the labels or the products.
- I prefer to spend money on other social and environmental responsibility projects.
- I disagree with the way the choice question was asked.
- Other reason, please specify: ...

9. Can you shortly describe how you made your choices?

... (Open-ended)

10. Which tea characteristics did you **not consider** when making your choices? You can choose one, more than one, or none of the characteristics.

- Carbon neutral label
- Organic label
- Ethical trade label
- Price

- None of above (I considered all tea characteristics).

11. What was your main reason for not considering this tea characteristic/these tea characteristics? (This question will only be shown to those who have not chosen the "None of above" option in the previous question .)

...

12. What is the maximum amount you would be willing to pay for **the carbon-neutral label** (shown on the previous choice cards) **in addition to the cost of the tea product?**

£...

13. Which of the following can be considered **carbon offsetting**?

- **Compensating emissions** by investing in activities outside of the company, such as tree planting projects.
- **Reducing emissions** within the company, such as through investments in cleaner production processes.
- I do not remember.

14. How many different **tea product labels** appear on the choice cards?

- 1 label
- 2 label
- 3 labels
- I do not remember.

15. In this survey, did carbon-neutral label include any percentages (%) of carbon offsetting and carbon reduction?

- Yes.
- No.
- I do not remember.

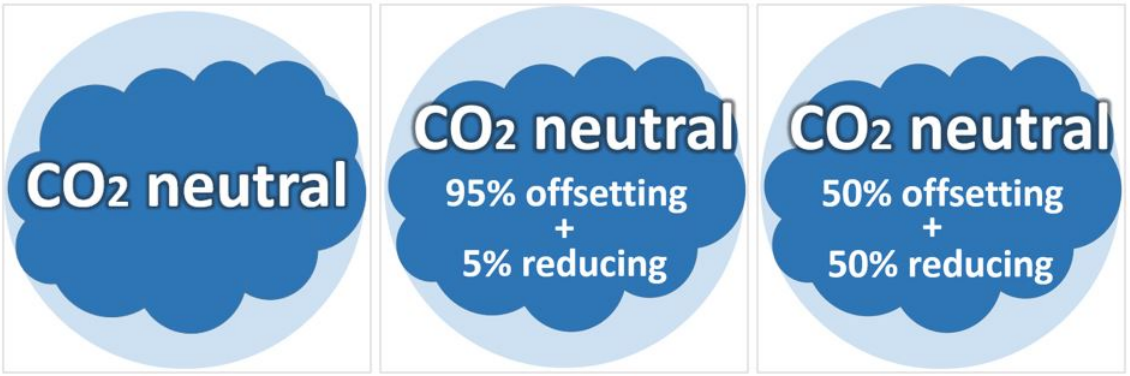
16. Please indicate to which degree you agree or disagree with the following statements.

Statement	Strongly Disagree	Mostly Disagree	Slightly Disagree	Neutral	Slightly Agree	Mostly Agree	Strongly Agree
I worry about climate change.							
Limited financial resources prevent me from buying climate-friendly products instead of the conventional ones.							
Lack of time prevents me from buying climate-friendly products instead of conventional ones.							
My positive emotions increase when I choose climate-friendly products over conventional ones.							
I feel guilty when I buy conventional products instead of climate-friendly ones.							
Most people who are important to me approve of my choice of climate-friendly products over the conventional ones.							
Producers, not consumers, are responsible for covering climate change mitigation costs.							

17. Remember that carbon offsetting involves compensating emissions by investing in projects outside of the company, such as tree planting projects, while carbon reductions occur within the company, such as investments in cleaner production processes. Emission offsetting, reduction, or a combination of both ensures that the entire life-cycle of a product is carbon-neutral.

Please indicate to which degree you agree or disagree with the following statements.

Statement	Strongly Disagree	Mostly Disagree	Slightly Disagree	Neutral	Slightly Agree	Mostly Agree	Strongly Agree
I trust carbon neutral labels.							
I am confused about carbon-neutral labels.							
I am concerned about carbon offsetting.							



Label 1 Label 2 Label 3

Please look at the different types of hypothetical labels shown above carefully, and answer the questions below. You may choose one or more label options, or none.

18. Which label(s) do you trust the most?

- Label 1
- Label 2
- Label 3
- None

19. Which label(s) do you find the most confusing?

- Label 1
- Label 2
- Label 3
- None

20. Please indicate to which degree you agree or disagree with the following statements.

Statement	Strongly Disagree	Mostly Disagree	Slightly Disagree	Neutral	Slightly Agree	Mostly Agree	Strongly Agree
Carbon offsetting effectively reduces carbon emissions.							
Carbon offsetting puts a price tag on emissions, thereby allowing producers to continue polluting.							
Carbon offsetting generates a misleading sense of relief, without encouraging further efforts to reduce emissions.							
Carbon offsetting is a form of greenwashing.							

Please now consider your usual **real-life** grocery shopping.

21. Which type of tea do you consume the most? Please select only one.

- Black tea (e.g., English Breakfast, Early Grey tea)
- Green tea
- Herbal tea
- Other, please specify: ...

22. How much do you typically pay for tea?

- £...
- I do not know.

23. Which tea quantity do you usually buy?

- I usually buy ... **grams** of tea.
- I usually buy ... **teabags**.
- I do not know.

24. Please indicate which sustainability labels the grocery products you buy have.

- Carbon neutral label
- Organic label
- Fair trade label
- No sustainability label
- I do not know
- Other, please specify: ...

Finally, a few questions about yourself.

25. What is your approximate annual household income after taxes?

- Under £10,000
- £10,000 - £19,999
- £20,000 - £29,999
- £30,000 - £39,999
- £40,000 - £49,999
- £50,000 - £59,999

- £60,000 - £69,999
- £70,000 - £79,999
- £80,000 - £89,999
- £90,000 - £99,999
- £100,000 - £129,999
- £130,000 or more

26. What is your current employment status?

- Full or part time employment
- Self-employed
- Unemployed
- Retired
- Looking after family or home
- Full-time student
- None of above

27. Are you a member of any environmental organization?

- Yes
- No

Now, we want to learn about your opinions and experience with this survey.

28. Do you think that your responses in this survey will influence tea product labeling or pricing policies?

- Yes
- No

29. Is there anything about the tea choices that you find confusing or unclear?

- Yes. Please specify what was confusing or unclear: ...
- No.

30.Were the instructions clear to you?

- Yes
- No. Please specify what was unclear: ...

31. Do you have any further comments or feedback about the survey?

... [open-ended]

Thank you for your participation in this survey. Your responses are very valuable to us and contribute to the research project by Eawag, the Swiss Federal Institute of Aquatic Science and Technology, and the University of St.Gallen, Switzerland.

If you have any questions or concerns about the survey please contact Begüm Özdemir Oluk (begum.ozdemiroluk@eawag.ch).

Please click "continue" to submit the survey and receive your payment.

A.2 Choice design

This section includes detailed information on the choice design created for the main survey on Ngene using the MNL model. Table A.16 shows the details of the attributes and combinations for each choice situation, respectively.

Block	Choice Situation	Price AltA (£)	CN AltA	Org AltA	ET AltA	Price AltB (£)	CN AltB	Org AltB	ET AltB
1	1	3.9	No	Yes	No	4.9	Yes	No	Yes
1	2	2.9	No	No	No	6.9	Yes	Yes	Yes
1	3	2.9	No	Yes	No	3.9	No	No	Yes
1	4	4.9	No	No	Yes	4.9	No	Yes	No
1	5	0.9	No	No	No	4.9	Yes	Yes	Yes
1	6	6.9	Yes	Yes	Yes	0.9	No	No	No
1	7	2.9	Yes	Yes	No	0.9	No	No	No
1	8	4.9	Yes	Yes	No	2.9	Yes	No	Yes
2	1	1.9	Yes	No	No	5.9	No	Yes	Yes
2	2	1.9	No	Yes	Yes	4.9	Yes	No	Yes
2	3	0.9	No	No	No	1.9	Yes	Yes	Yes
2	4	1.9	No	Yes	No	1.9	No	No	Yes
2	5	6.9	Yes	Yes	Yes	3.9	No	No	Yes
2	6	5.9	Yes	Yes	No	1.9	No	Yes	Yes
2	7	3.9	Yes	No	No	2.9	No	No	No
2	8	5.9	No	Yes	No	5.9	Yes	No	No

AltA refers to tea alternative A (on the left of the choice card), and AltB refers to tea alternative B (on the right of the choice card).
D error = 0.287, A error = 0.447, B estimate = 73.188, S estimate = 29.530.

Table A.16: Main survey choice design

C Power Analysis

In this section, I discuss the ex-ante power calculations for my main analysis to understand the consumers' WTP differences between samples. This analysis involves three bilateral comparisons, each with subsamples of 400 participants. The objective of the power calculation is to determine the minimum detectable effect size (MDEs) for WTP differences between subsamples, using the formula from Djimeu and Houndolo (2016) and the power calculation sheet from International Initiative for Impact Evaluation (3ie) (2016):

$$\text{MDEs} = \frac{(t_1 + t_2) \times \text{sd}(y)}{\sqrt{p \times (1-p) \times n}}$$

where sd(y) is the pooled total standard deviation of the estimated effect on the outcome variable, p is the proportion of the study that is randomly assigned to the treatment group, n is the sample size, t1 is the t-value corresponding to the significance level (0.05) of the test, t2 is the t-value corresponding to the power of the design (0.80), and MDEs is the minimum detectable effect size. I assume sd(y) follows a truncated normal distribution, and the mean and standard deviation of the distribution are based on both the literature (Bek, 2022; Carattini et al., 2024) and the my expectations regarding the differences between the subsamples.

I find that the MDEs is 0.18 pounds for the difference between sample 3 and samples 1 and 2, which corresponds to 18% of the standard deviation. Furthermore, the MDEs for the difference between samples 1 and 2 is 0.36 pounds, which also corresponds to 18% of the standard deviation.