

DEMAND FOR CARBON-NEUTRAL PRODUCTS

DISSERTATION
of the University of St.Gallen,
School of Management,
Economics, Law, Social Sciences,
International Affairs and Computer Science,
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

Begum Ozdemir Oluk

from

Turkey

Approved on the application of

Prof. Sabrina Eisenbarth, Ph.D.

and

Prof. Dr. Ståle Navrud

Dissertation no. 5706

Copy Blitz AG, St.Gallen 2026

DEMAND FOR CARBON-NEUTRAL PRODUCTS

DISSERTATION
of the University of St.Gallen,
School of Management,
Economics, Law, Social Sciences,
International Affairs and Computer Science,
to obtain the title of
Doctor of Philosophy in Economics and Finance

submitted by

Begum Ozdemir Oluk

from

Turkey

Approved on the application of

Prof. Sabrina Eisenbarth, Ph.D.

and

Prof. Dr. Ståle Navrud

Prof. Beatrix Eugster, Ph.D.

Dissertation no. 5706

Copy Blitz AG, St.Gallen 2026

The University of St.Gallen, School of Management, Economics, Law, Social Sciences, International Affairs and Computer Science, hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St.Gallen, June 16, 2026

The President:

Prof. Dr. Manuel Ammann

In memory of my mother.

Acknowledgements

I am sincerely grateful to my supervisors: Ivana Logar, Stefano Carattini, and Sabrina Eisenbarth. Their steady guidance and the generous time they dedicated to me were essential to this dissertation. I appreciate how much they helped me improve and supported me at every step of my PhD.

I would also like to thank my thesis committee, Prof. Beatrix Eugster, Prof. Ståle Navrud, for their time and valuable feedback, and to Prof. Reto Föllmi for his supervision during the earlier stages of this journey.

I am very grateful to my coauthor, Fabian Dvorak. I learned a lot from him, and I thank him for his expertise and support, especially during the challenging stages of my PhD.

Moreover, this dissertation was made possible by the financial support of Eawag (Swiss Federal Institute of Aquatic Science and Technology), which I acknowledge with appreciation.

I would like to thank my colleagues at Eawag and the University of St. Gallen for creating such a supportive environment. To my friends, thank you for being there during the difficult times. I am also grateful to Mehmet Kutluay for his advice and help during my transition into doctoral life.

I am deeply grateful to my mother, Figen Özdemir, who has believed in me and my academic path from day one. Her strength, determination and trust have always been with me and are a huge part of this work. I also remember my grandparents, Niyazi Oğuz and Behiye Oğuz, with gratitude for their lifelong care and support.

I am very thankful to my father and brother for their constant support and reassurance. I am also grateful to my extended family for their care, and I remember my other grandparents with gratitude.

Finally, I am deeply grateful to my husband, Can Oluk, whose belief in me sustained me through every challenge. Thank you for walking every step of this path by my side; this journey would have been impossible without you.

Dübendorf, 29.06.2026

Begum Ozdemir Oluk

Summary

This dissertation focuses on consumer demand for carbon-neutral products and the extent to which private consumption choices can contribute to climate change mitigation. As firms increasingly commit to carbon neutrality, understanding how consumers value carbon-neutral certifications has become an important empirical question.

Chapter 1 synthesizes the literature on demand for carbon-neutral and low-carbon products through a meta-analysis. The findings indicate substantial consumer willingness to pay for emissions reductions, exceeding many estimates of the social cost of carbon. Since the existing literature mainly consists of stated-preference studies that may be prone to hypothetical bias, these findings motivate an analysis based on actual market data.

Chapter 2 infers consumers' valuation of carbon-neutral labels from revealed preferences using a hedonic approach. We use weekly product data collected from Amazon marketplaces in the United States, the United Kingdom, and Germany. Exploiting the staggered introduction of carbon-neutral certifications, the findings show no evidence that carbon-neutral labels affect product prices, suggesting that the climate-mitigation potential of these labels should be interpreted with caution, in contrast to the existing literature.

Chapter 3 assesses consumers' willingness to pay for transparency in carbon-neutral labels that disclose the shares of CO₂ offsetting and reduction. Using a discrete choice experiment with UK tea consumers, it finds no evidence that transparency increases willingness to pay for carbon-neutrality or that consumers prefer CO₂ reductions over offsets. These findings provide a rationale for additional regulation of carbon-neutral labels, as the market may not effectively penalize lower-quality mitigation claims.

Zusammenfassung

Diese Dissertation untersucht die Konsumentennachfrage nach CO₂-neutralen Produkten und den Beitrag privater Konsumententscheidungen zur Minderung des Klimawandels. Da sich Unternehmen zunehmend zur CO₂-Neutralität verpflichten, ist es eine wichtige empirische Frage, wie Konsumentinnen und Konsumenten CO₂-Neutralitätszertifizierungen bewerten.

Kapitel 1 fasst die Literatur zur Nachfrage nach CO₂-neutralen sowie Produkten mit geringem Kohlenstoffausstoß mithilfe einer Meta-Analyse zusammen. Die Ergebnisse zeigen eine substanzielle Zahlungsbereitschaft der Konsumentinnen und Konsumenten für Emissionsreduktionen, die viele Schätzungen der sozialen Kosten von CO₂ übersteigt. Da die bestehende Literatur überwiegend aus Studien mit bekundeten Präferenzen besteht, die anfällig für hypothetische Verzerrungen sein können, motivieren diese Ergebnisse eine Analyse auf Basis tatsächlicher Marktdaten.

Kapitel 2 leitet die Zahlungsbereitschaft der Verbraucher:innen für CO₂-Neutralitätslabels aus offenbarten Präferenzen ab, unter Verwendung eines hedonischen Ansatzes. Wir nutzen wöchentliche Produktdaten, die von Amazon-Marktplätzen in den Vereinigten Staaten, dem Vereinigten Königreich und Deutschland erhoben wurden. Unter Ausnutzung der gestaffelten Einführung von CO₂-neutralen Zertifizierungen zeigen die Ergebnisse keine Hinweise darauf, dass CO₂-Neutralitätslabels die Produktpreise beeinflussen. Dies legt nahe, dass das klimapolitische Minderungspotenzial dieser Labels im Gegensatz zur bestehenden Literatur mit Vorsicht zu interpretieren ist.

Kapitel 3 analysiert die Zahlungsbereitschaft für Transparenz bei CO₂-Neutralitätslabels, welche die Anteile von CO₂-Kompensation und Reduktion offenlegen. Auf Basis eines Discrete-Choice-Experiments mit britischen Teekonsumentinnen und -konsumenten zeigt sich keine Evidenz dafür, dass Transparenz die Zahlungsbereitschaft für CO₂-Neutralität erhöht oder dass CO₂-Reduktionen gegenüber CO₂-Kompensationen bevorzugt werden. Diese Ergebnisse liefern eine Begründung für zusätzliche Regulierung von CO₂-neutralen Labels, da der Markt minderwertige Minderungsansprüche möglicherweise nicht wirksam sanktioniert.

Contents

1	Demand for carbon-neutral products: evidence from a meta-analysis	1
1	Introduction	2
2	Data and empirical approach	4
2.1	Study selection	4
2.2	Data	6
2.3	Empirical approach	9
3	Results	12
3.1	Descriptive evidence	12
3.2	Meta-regression results	19
4	Conclusions	22
	Appendix	39
A.1	Data collection	39
A.2	Descriptive statistics and supplementary material	59
A.3	Robustness tests	64
2	Demand for carbon-neutral products: evidence from Amazon's online marketplace	80
1	Introduction	81
2	Data and empirical approach	85
3	Empirical Evidence	93
3.1	Results from Amazon's U.S. marketplace	93
3.2	Heterogeneity analyses	98
3.3	Results from Amazon's U.K. and German marketplaces	104
4	Conclusions	107
	Appendix	115

A.1	Carbon neutrality on Amazon’s marketplace	115
A.2	Changes in treatment status	116
A.3	Carbon footprint estimation of quasi-natural experiments	130
A.4	Difference-in-difference analysis	131
A.5	Customer reviews	133
A.6	Robustness tests	137
A.7	Additional results from Amazon’s U.K. and Germany marketplaces	145
3	Transparency of carbon-neutral labels: evidence from a choice experiment	147
1	Introduction	148
2	Methodology	155
2.1	Survey and choice experiment design	155
2.2	Data	160
2.3	Empirical approach	161
3	Results	164
3.1	Transparency and willingness to pay	164
3.2	Robustness tests	168
3.3	Mechanisms	169
3.4	Competing labels	174
4	Conclusion	176
	Appendix	188
A	Survey data	188
A.1	Descriptive statistics	188
A.2	Choice data	199
B	Robustness and heterogeneity	202
A.1	Robustness tests	202
A.2	Heterogeneity	227
C	Competing labels	229
D	Confirmatory and exploratory analyses	234
E	Survey instrument	236

F	Power analysis	259
	Generative AI usage declaration	260
	Curriculum Vitae	261

List of Tables

Chapter 1

1	Paper search strategy	5
2	Summary statistics: study means	9
3	WTP estimates and 95% confidence intervals	12
4	Publication bias test	17
5	Factors associated with WTP _R for CO ₂ reductions	21
A.1	Excluded studies and rationales	40
A.2	Literature covered in the meta-analysis	42
A.3	WTP derivation strategy	46
A.4	Summary statistics: unweighted sample	60
A.5	Means of study means: WTP estimates by product category . . .	60
A.6	Breusch-Pagan test for OLS models presented in Table 5	64
A.7	Factors associated with WTP for CO ₂ reductions: robustness tests with alternative models	67
A.8	Factors associated with WTP for CO ₂ reductions: OLS model, robustness tests with additional variables	68
A.9	Factors associated with WTP for CO ₂ reductions: mixed effects model, robustness tests with additional variables	69
A.10	Factors associated with WTP for CO ₂ reductions: OLS and WLS models, robustness tests with weights	70
A.11	Factors associated with WTP for CO ₂ reductions: mixed effects and weighted mixed effects models, robustness tests with weights	71
A.12	Factors associated with WTP for CO ₂ reductions: OLS model, robustness tests with transformations of the dependent variable	72

A.13 Factors associated with WTP for CO ₂ reductions: mixed effects model, robustness tests with transformations of the dependent variable	73
A.14 Factors associated with WTP for CO ₂ reductions: OLS model, robustness tests with alternative cluster variables	74
A.15 Factors associated with WTP for CO ₂ reductions: mixed effects model, robustness tests with alternative random effects	75
A.16 Factors associated with WTP for CO ₂ reductions: OLS model, robustness tests with alternative functional form of CO ₂ reduction	76
A.17 Factors associated with WTP for CO ₂ reductions: mixed effects model, robustness tests with alternative functional form of CO ₂ reduction	77
A.18 Factors associated with WTP for CO ₂ reductions: OLS model, robustness tests with subsets of the data	78
A.19 Factors associated with WTP for CO ₂ reductions: mixed effects model, robustness tests with subsets of the data	79

Chapter 2

1 Number of treated and control products by category (United States)	90
2 Number of treated and control products by category (United Kingdom)	90
3 Number of treated and control products by category (Germany) .	91
A.1 Restrictions to define treated and control products	117
A.2 Balance	117
A.3 Timing of the changes in treatment status (United States)	118
A.3 Timing of the changes in treatment status (United States)	119
A.3 Timing of the changes in treatment status (United States)	120
A.3 Timing of the changes in treatment status (United States)	121
A.4 Timing of the changes in treatment status (United Kingdom) . . .	125
A.5 Timing of the changes in treatment status (Germany)	127
A.6 Dynamic effects (US)	131

A.7	Calendar Effects (US)	132
A.8	Keyword occurrences (US)	134
A.9	Keyword occurrences (UK)	135
A.10	Keyword occurrences (Germany)	136

Chapter 3

1	Experimental design	156
2	Choice attributes and levels	156
3	Mixed logit model in WTP space, all parameters randomized	166
4	Poe test results comparing MWTP estimates for the carbon-neutral label between experimental groups	167
5	Trust in and confusion about standard and transparent carbon-neutral labels	172
6	Mixed logit model in WTP space with interaction variables	173
A.1	Summary statistics of sociodemographics	189
A.2	Comparison of tea consumers and non-tea consumers	191
A.3	Summary statistics: tea consumption	192
A.4	Summary statistics: survey clarity, and attention	193
A.5	Summary statistics: certainty, attribute-non-attendance, consequentiality, and protest responses	194
A.6	Agreement with the list of statements (full sample)	195
A.7	Description of main interacted covariates	196
A.8	Covariate balance	197
A.10	Choice design	200
A.11	Choices: detailed information	201
A.12	Multinomial logit model in preference space estimates	204
A.13	Mixed logit model in preference space estimates, non-monetary parameters randomized	205
A.14	Mixed logit model in preference space estimates, all parameters randomized	206

A.15 Mixed logit model in preference space estimates with interaction variables	207
A.16 Poe test results comparing MWTP estimates from preference-space models	208
A.17 MWTP estimates for carbon-neutral, organic, and ethical trade labels	209
A.18 Multinomial logit model in WTP space	210
A.19 Mixed logit model in WTP space	211
A.20 Mixed logit model in WTP space with 2000 Halton draws	212
A.21 Mixed logit model in WTP space with alternative specific error components	213
A.22 Mixed logit model in WTP space with correlated error terms . .	214
A.23 Mixed logit model in WTP space, with the sample including protest responses	215
A.24 Mixed logit model in WTP space, all parameters randomized, without survey speeders	216
A.25 Mixed logit model in WTP space, excluding protest responses, survey speeders, and attention/manipulation test failers	217
A.26 Mixed logit model in WTP space, with interaction variables (unadjusted for multiple hypothesis testing)	218
A.27 Mixed logit model in WTP space (fewer interactions)	219
A.28 Mixed logit model in WTP space with interaction variables (warm glow variable added)	220
A.29 Mixed logit model in WTP space, with interaction variable for participants who misunderstand the meaning of offsetting	221
A.30 Mixed logit model in WTP space, additional interactions	222
A.31 Poe test results, estimations with various model specifications estimated in WTP space	223
A.32 Poe test results, estimations with various model specifications estimated in WTP space	224
A.33 Poe test results, estimations with different sample restrictions . .	225

A.34 Poe test results, contingent valuation method 226

A.35 Mixed logit model in WTP space 228

A.36 Mixed logit model in WTP space, with sustainability label inter-
actions, no correlated errors 230

A.37 Mixed logit model in WTP space, with sustainability label inter-
actions, correlated errors 231

A.38 Poe test results for competing label scenarios 232

A.39 Poe test results for competing label scenarios 233

List of Figures

Chapter 1

1	WTP estimates for CO ₂ reductions across studies	14
2	WTP for carbon neutrality across studies	16
3	WTP estimates and study sample size	18
A.1	WTP for carbon reductions (WTP _R) by product category	61
A.2	WTP _R and WTP _{kg} by the amount of CO ₂ emission reduction . . .	62
A.3	Residuals versus fitted values	63

Chapter 2

1	A product certified carbon neutral by Climate Partner on Amazon.com	86
2	Effect of carbon-neutral label on products' price change (United States)	94
3	Summary of results for the average effect of the carbon-neutral label on product price changes (United States)	95
4	Heterogeneous effect of the carbon-neutral label on product price changes: prices within and above the meta-analysis range (US) .	100
5	Heterogeneous effect of carbon-neutral label on product's price change: product categories (US)	101
6	Heterogeneous effect carbon-neutral label on product's price change: high and low carbon-intensive product categories (US) .	102
7	Heterogeneous effect of carbon-neutral label on product's price change: label certifiers (US)	103
8	Effect of carbon-neutral label (UK)	104

9	Effect of carbon-neutral label (Germany)	106
A.1	The effect of carbon-neutral label when removing label consistency constraint (United States)	138
A.2	Effect of carbon-neutral label using unbalanced data option for estimation (United States)	139
A.3	Effect of carbon-neutral label excluding new Climate Pledge Friendly labels and Small Business labels (United States)	140
A.4	Effect of carbon-neutral label without control variables (United States)	141
A.5	Effect of carbon-neutral label with the initial number of ratings as an additional control variable (United States)	142
A.6	Effect of carbon-neutral label using absolute price level as outcome variable (United States)	143
A.7	Effect of carbon-neutral label using the number of ratings as the outcome variable (US)	144
A.8	Effect of carbon-neutral label using the number of ratings as the outcome variable (UK)	145
A.9	Effect of carbon-neutral label using the number of ratings as the outcome variable (Germany)	146

Chapter 3

1	Example choice card presented to participants in the control group	159
2	MWTP for the carbon-neutral label in the presence or absence of other sustainability labels	175

Chapter 1: Demand for carbon-neutral products: evidence from a meta-analysis

coauthored with Stefano Carattini, Fabian Dvorak, Ivana Logar

Abstract

Corporate social responsibility and the private provision of global public goods are of key interest to economists and policymakers. As an increasing number of firms make carbon-neutral commitments, evidence on how consumers value carbon-neutral and low-carbon labels remains scattered across products, empirical approaches, and settings, limiting comparability and the ability to benchmark valuations. We address this gap with a meta-analysis of 37 studies, compiling 126 estimates of willingness to pay for CO₂ reductions from a combined sample of 29,666 participants across various geographies. We find that the implied willingness to pay for CO₂ reductions is substantial at USD 1,993 per ton of CO₂ on average and exceeds most estimates of the social cost of carbon. Meta-regressions further indicate that valuations are positively associated with the magnitude of CO₂ reductions implied by the labels and with baseline product prices, and that European studies are associated with higher valuations than studies conducted in other regions, even after controlling for GDP per capita and other study characteristics.

Keywords: carbon-neutral labels; climate change mitigation; corporate social responsibility; willingness to pay; meta-analysis; stated preferences

JEL codes: Q51; Q54; D12; C83

1 Introduction

An important question in economics relates to why people engage in pro-social behavior and to what extent society can rely on people's private motivations to ensure the provision of public goods. Climate change mitigation is such a (global) public good. While climate policy gradually expands, private behavior by individuals and firms can contribute to accelerating the transition towards a cleaner economy. Over the last few years, more and more firms decided or announced plans to make their operations, or at least part of them, carbon neutral (Rogelj et al., 2021). The main driver of these decisions was likely pressure from investors and company boards to prepare firms for a low-carbon future (Kim and Lyon, 2011).

Despite a growing literature on consumer demand for carbon-neutral and low-carbon products, the evidence remains fragmented across products, empirical methods, and geographies, and it lacks a consistent valuation metric, such as willingness to pay (WTP) per unit of carbon dioxide (CO₂) reduction. Consequently, the extent of WTP for carbon-neutral and low-carbon labels, and whether private demand is sufficient to justify their broader adoption, remains an open empirical question.

This paper addresses this question. It collects evidence from the literature on consumers' WTP for products that are presented as either carbon neutral or low carbon, as certified by labels. The literature that we cover includes 37 studies, providing 126 WTP observations, and an overall sample of 29,666 participants. From this body of evidence, it is possible to estimate average WTP for CO₂ reductions, including carbon neutrality, across studies, geographies, and samples, and compare it with the range of estimates that economists have provided for the social cost of carbon (SCC), with the goal of understanding whether and to what extent consumers privately internalize the climate externality. Further, with the tools of meta-analysis, it is possible to determine the product and study features that are associated with higher or lower WTP for CO₂ reductions.

The meta-analysis indicates a large, positive WTP for carbon-neutral and

low-carbon products: at USD 1,993 per ton of CO₂, exceeding the distribution of current carbon prices and many estimates of the SCC. We also find a positive and significant association between the magnitude of CO₂ reductions and WTP, suggesting that study participants are responsive to the amount of CO₂ reduction indicated on the labels. Higher product prices are associated with higher WTP, implying that the relative cost of CO₂ reductions may matter. Finally, studies conducted in Europe are associated with higher WTP than studies in other regions (Asia, America, and Africa), even after controlling for GDP per capita.

This paper advances the existing literature through three main contributions. First, we consolidate the empirical evidence on consumer demand for carbon-neutral and low-carbon products, which is dispersed across products, preference elicitation methods, geographies, and valuation metrics (e.g., price premia, product-level WTP, or WTP per unit of emissions). We compile WTP estimates from the existing literature and convert them to a common unit, WTP per kilogram of CO₂ reduced. This synthesis yields a comparable set of valuation measures that allows for direct benchmarking against the SCC and enables meta-regression analyses of study and contextual characteristics associated with WTP.

Second, we provide the first meta-analysis focusing exclusively on carbon-neutral and low-carbon products and, in doing so, isolate the implicit value of CO₂ reductions. Unlike prior reviews or meta-analyses that pool a broad set of sustainability labels (see Tully and Winer, 2014; Li and Kallas, 2021; Andreani and Sogari, 2023; Piracci et al., 2024), our focus on climate labels enables consistent synthesis across different carbon claims.

Finally, we relate to the literature on meta-analysis methods (Nelson and Kennedy, 2009a) and stated preference valuation (Johnston et al., 2017; List and Gallet, 2001) by examining the factors that are associated with WTP. Using meta-regression models to identify key contextual and design features can provide practical guidance for designing future research and effective carbon labels. For example, the positive association between the amount of CO₂ reductions and WTP suggests that labels that communicate the magnitude of emissions

reductions may be more effective.

The remainder of the paper is organized as follows. Section 2 describes the meta-analysis database and empirical approach. Section 3 reports the main empirical results and robustness checks. Section 4 concludes with a discussion of implications and avenues for future research.

2 Data and empirical approach

This section outlines the data and empirical approach used in the meta-analysis. We begin by describing the study selection process and inclusion criteria. We then present the construction of the meta-analysis dataset and the derivation of WTP measures for CO₂ reductions. Finally, we describe the empirical strategy used to summarize the literature and analyze the factors that correlates with WTP.

2.1 Study selection

This section describes the scope and inclusion criteria of the meta-analysis and then explains how we identified the studies that meet these criteria.

Our scope includes studies that elicit consumers' valuations of emissions reductions through real or hypothetical product purchases and focus on climate labels indicating carbon footprint information, carbon reduction, or carbon neutrality.¹ Following standard practice in meta-analysis to ensure comparability (Smith and Pattanayak, 2002; Nelson and Kennedy, 2009b), we restrict our attention to studies and observations from which a WTP estimate for a quantifiable CO₂ reduction can be derived. In defining the scope of emissions, we treat reductions reported as CO₂ and greenhouse gases expressed in CO₂ equivalents as comparable (IPCC, 2018; Greenhouse Gas Protocol, 2004). We also use the term “CO₂ reduction” broadly to include abatement, offsets, and CO₂ capture, consistent with evidence that consumers do not differentiate between direct reductions and offsets in valuation (Ozdemir-Oluk, 2025).

¹We do not focus on studies that value environmental or social responsibility attributes, such as energy efficiency, fair trade, organic, and reduced water footprint, unless they also value climate labels.

To identify studies that meet these criteria, we proceeded in two ways. First, we conducted keyword searches on Google Scholar, EconPapers (RePEc), EconLit, and ProQuest to capture both published studies and working papers. Second, we used backward and forward citation searches based on the studies identified in the first step. Table 1 outlines the search strategy, including the databases visited and the keywords used.

Period	Databases & Search engines	Search terms
Jan 2021 - Jun 2021	Google Scholar Scopus EconPapers ProQuest	Combination of words such as “carbon footprint,” “carbon neutral,” “climate-friendly,” “low carbon,” “label,” “valuation,” “experiment,” “survey,” “stated preference,” “revealed preference”
Sep 2022 - Oct 2022	Google Scholar Scopus EconPapers ProQuest	(Carbon footprint label OR carbon label OR carbon neutral label OR climate-friendly OR carbon reduction OR low carbon OR carbon trust label) AND (stated preferences OR revealed preferences OR choice experiment OR contingent valuation OR field experiment OR lab experiment OR auction experiment) AND (environmental valuation OR Willingness to Pay)
July 2023	Google Scholar (2,780) Scopus (32) EconPapers (81) EconLit (6) ProQuest (549)	(“carbon footprint label” OR “carbon neutral label” OR “carbon-neutral label” OR “low carbon label” OR “food miles” OR “product miles” OR “transportation distance”) AND (“Willingness to Pay” OR “Willingness to Accept” OR “stated preferences” OR “revealed preferences” OR “choice experiment” OR “contingent valuation” OR “field experiment” OR “lab experiment” OR “auction experiment” OR “hedonic” OR “environmental valuation” OR “non-market valuation”)

Table 1: Paper search strategy

Multiple searches were conducted during the years 2021 and 2022. During the final search (July 2023), all of the search outputs, for which we specify the number of results in parentheses, were reviewed. In addition to the searches, we also checked papers cited in a review article by Rondoni and Grasso (2021). Backward citations of relevant papers’ titles were checked by searching for the word “carbon,” while forward citations were checked using combinations of the following words: “carbon,” “label,” “willingness,” “kilometers,” and “miles.”

Our initial sample includes 83 studies. Among the selected studies, we narrow the scope to those that enable us to derive or make assumptions regarding the amount of CO₂ reduction. Furthermore, we exclude studies that value carbon

transparency reporting constant baseline emissions, as these do not allow us to infer WTP for emissions reductions. Instead, we only rely on study designs that vary emissions levels and derive WTP from the incremental differences between reported levels, such as the difference in valuation between low and high CO₂ footprints.

Applying these exclusion criteria, which are described in more detail in Appendix A.1.1, we focus only on studies that allow for the quantification of CO₂ reductions and associated WTP measures. The excluded studies and the reasons for their exclusion are summarized in Table A.1. As a result, our final dataset consists of 37 studies.

2.2 Data

This section describes the meta-analysis database and outlines the procedures used to derive WTP measures for CO₂ reductions from the selected studies.

The literature that meta-analysis covers a broad set of products, which we classify into dairy and eggs, fruits and vegetables, meat, non-food items, oil and grain, snacks, and water and drinks. Accordingly, the outcome of interest captures the ‘marginal value of CO₂ reductions’ communicated through emissions-related information (such as carbon-neutral or low-carbon labels and carbon footprint information), rather than the value of the low-carbon product.

In the main analyses, the dataset consists of 126 WTP observations drawn from 37 studies²; these 126 observations are obtained by aggregating multiple WTP estimates within each study whenever they share the same product category and covariate values (see Appendix Section A.1.3 for details).

Studies leverage a range of valuation methods. These include four contingent valuation (CV) surveys and 29 discrete choice experiments (DCEs) based on stated preferences, two laboratory experiments and one field experiment based on revealed preferences, as well as one study that combines a DCE with a field experiment. The underlying sample comprises 29,666 participants. Appendix

²Cerroni et al. (2019a) and Cerroni et al. (2019b) are counted as a single study because they analyze the same underlying sample.

Table A.2 summarizes the studies included in the meta-analysis, reporting the valued products, study countries, valuation methods, number of WTP observations, and the type of emissions information valued.

To facilitate comparisons across studies, we construct four WTP measures that capture consumers' valuations of CO₂ reductions at four different levels of standardization. First, the measure WTP_R refers to the WTP for CO₂ reductions, which may vary both between and within studies. This value is either directly obtained or derived from studies and is normalized to 2020 USD. Second, WTP_{kg} represents the WTP per 1 kg of CO₂ reduction, calculated by dividing WTP_R by the amount of CO₂ reduction in kilograms, and is expressed in USD. This WTP measure is useful for comparing with estimates of the SCC. Third, WTP_{CN} denotes the WTP for achieving carbon neutrality of a product. It is calculated by multiplying WTP_{kg} by the baseline CO₂ emissions of the product, and is also expressed in USD. Finally, WTP_{CN%} represents the markup for carbon neutrality. Specifically, it is WTP_{CN} normalized by the product price and thus captures the proportion of a product's price that consumers would be willing to pay as an additional premium for carbon neutrality. From each study, we extract one of the four WTP measures and, for comparability, derive additional measures by converting this estimate into alternative metrics, depending on the information reported in each study.

To derive consistent estimates of WTP_R, which form the basis for all standardized measures, we apply a common set of rules. First, when studies report WTP for a specific CO₂ reduction amount, we use the reported estimate directly. When the reduction magnitude is not explicitly stated, we infer emissions levels using third-party product life-cycle carbon calculators, such as "MyEmissions" and "Plate up for the Planet", and retain only observations for which implied reductions can be quantified.

Second, to isolate the marginal WTP for the carbon-related information, we net out the value of the unlabeled baseline when required. For example, when WTP is reported for a carbon-neutral product, we subtract the WTP or price of the corresponding unlabeled alternative; if an unlabeled WTP estimate is

unavailable, we use the conventional product price as a proxy.

Third, following Haab and McConnell (2002), we treat mean and median WTP estimates as equivalent under the assumption of linear utility and symmetric error terms.

Fourth, when marginal WTP is not reported but discrete choice model coefficients are available, we compute marginal WTP as the ratio of the climate-label coefficient to the price coefficient:

$$\text{Marginal WTP} = -\frac{\beta_{\text{climate label}}}{\beta_{\text{cost}}}.$$

We do not compute WTP in cases where both cost and climate parameters are random, as the implied WTP cannot be summarized by a single ratio without additional distributional assumptions, or when climate-label is interacted with multiple categorical or ordinal variables, as such specifications substantially reduce the effective sample size and yield highly imprecise WTP estimates. A very detailed, study-by-study description of the WTP derivations is provided in Table A.3 in Appendix A.1.3.

Table 2 reports summary statistics of the study-level means for the WTP measures and covariates used in the analyses. It reports key study characteristics and contextual variables, such as product prices, valuation methods, sample sizes, publication status, and the shares of observations that rely on CO₂ emission assumptions or WTP derivations conducted by us rather than directly taken from the studies. Table A.4 in Appendix A.2 reports summary statistics for the full sample of WTP observations at the estimate level (i.e., 126 WTP estimates). Detailed variable definitions are provided in Appendix Section A.1.3.

	N	Mean	Std. Dev.	Min	Max
WTP _R (USD)	37	1.30	1.43	0.00	6.85
WTP _{kg} (USD)	37	1.99	3.94	0.02	23.73
WTP _{CN} (USD)	37	15.01	49.18	0.00	296.92
WTP _{CN%} (%)	37	251.71	464.05	0.06	1786.64
CO ₂ reduction (kg)	37	3.66	7.65	0.00	39.43
Product CO ₂ emissions (kg)	37	7.41	11.80	0.04	43.33
Carbon neutral label	37	0.16	0.37	0.00	1.00
Colored label	37	0.12	0.29	0.00	1.00
Price (USD)	37	4.82	4.80	0.09	22.15
Stated pref. method	37	0.90	0.28	0.00	1.00
In-person	37	0.33	0.47	0.00	1.00
Sample size	37	574	652	19	3085
Publication	37	0.81	0.40	0.00	1.00
Study year	37	2015	3.77	2008	2021
GDP per capita (100 USD)	37	407.84	182.90	5.84	935.47
Europe	37	0.64	0.48	0.00	1.00
CO ₂ reduction assump.	37	0.62	0.49	0.00	1.00
WTP derivation	37	0.38	0.48	0.00	1.00

Table 2: Summary statistics: study means

This table displays the number of observations (N), and the summary statistics of the study means. Standard deviations are provided in parentheses. The WTP_R denotes (un-standardized) WTP for CO₂ reductions. WTP_{kg} is the WTP for 1 kg CO₂ reduction, WTP_{CN} is the WTP for carbon-neutrality, and WTP_{CN%} is the markup for carbon neutrality (proportion of a product's price that consumers would be willing to pay extra for carbon neutrality). For detailed variable descriptions, see Section A.1.3.

2.3 Empirical approach

This section outlines our empirical approach for the meta-analysis. We organize the analysis in three steps. First, we report the mean of study-level mean WTP_R, WTP_{kg}, WTP_{CN}, and WTP_{CN%}. This approach assigns equal weight to each study and summarizes the central tendency in the literature without being driven by differences in the number of WTP observations across studies. Second, we describe the distribution of WTP measures across studies and compare WTP_{kg} with the SCC. Third, we assess publication bias in the WTP measures reported in the literature, which could distort the findings if more precise estimates are more likely to be reported or published (Nelson and Kennedy, 2009b; Stanley

and Doucouliagos, 2014).

To evaluate publication bias, we use funnel asymmetry–precision-effect test with weighted least squares (FAT-PET-WLS) method (Stanley and Doucouliagos, 2014). The FAT-PET-WLS procedure estimates a weighted least squares (WLS) meta-regression to account for heteroskedasticity in the error term, where weights are typically based on the inverse of the squared standard error. However, because fewer than half of the studies report sufficient information to recover standard errors, we follow a common alternative approach in the meta-analysis literature and use the square root of the study sample size as a proxy for precision (Florax et al., 2005; Tuncel and Hammitt, 2014; Mattmann et al., 2016; Penn and Hu, 2019).

For the publication bias analysis, we aggregate observations at the sample level within each study, yielding 52 observations from 37 studies. This aggregation accounts for the fact that multiple estimates may originate from the same underlying sample. The estimated model is:

$$Y_{ij}\sqrt{n_{ij}} = \beta_0\sqrt{n_{ij}} + \beta_1 + v_{ij}.$$

Here, Y_{ij} denotes the WTP estimate for sample i from study j , and $\sqrt{n_{ij}}$ is the precision proxy (square root of sample size). In this specification, β_0 captures the underlying economic effect (PET), while β_1 provides the test for publication bias (FAT). The core test is the null hypothesis $H_0 : \beta_1 = 0$; a statistically significant non-zero intercept indicates funnel-plot asymmetry consistent with publication selection. Standard errors are clustered at the study level.

We then turn to a regression analysis to understand which factors are associated with WTP_R , including the amount of CO_2 reduction, product price, stated versus revealed preference methods, in-person studies, study sample size, published versus unpublished studies, study year, GDP per capita, and studies conducted in Europe versus elsewhere. We control for observations that required assumptions regarding the amount of CO_2 reduction or where we made additional calculations based on the information provided in the studies to derive

WTP estimates.

We estimate ordinary least squares (OLS) models with standard errors clustered at the study level. We focus on WTP_R because it is the non-standardized measure of marginal WTP for CO₂ reductions (in 2020 USD) and is typically directly available, whereas the remaining WTP measures are mainly derived from it. The dependent variable WTP_R is transformed using the inverse hyperbolic sine function, which allows for zero or negative WTP values while reducing the influence of extreme observations. Because the inverse hyperbolic sine transformation implies that marginal effects vary with the level of WTP_R and coefficients do not map directly into dollar changes, we focus on the sign and statistical significance of the estimated associations. We use OLS with standard errors clustered by study.

To assess the robustness of our results on factors associated with WTP estimates, we perform a series of robustness checks that address potential sources of bias. First, we examine robustness to alternative specifications, including models that control for additional label characteristics such as carbon-neutral and colored labels. Second, we examine sensitivity to alternative weighting schemes. Because standard meta-analyses may be driven by studies contributing many estimates, we estimate a weighted mixed-effects model assigning equal weight to each study. Conversely, to account for differences in statistical precision, we also estimate specifications using sample size as a weighting factor. Third, we account for unobserved heterogeneity and within-cluster correlation. Given the meta-analytical nature of the data, errors may be correlated within studies and product categories. We therefore estimate mixed-effects models with random effects at both the study and product-category levels. In addition, we apply two-way clustering of standard errors by product category and country, along with country-specific random effects, to account for broader spatial and contextual dependencies. Fourth, we address potential model misspecification and omitted variable bias. We test the robustness of our results to alternative functional form assumptions for both the dependent variable WTP_R and the CO₂ reduction variable. Finally, to mitigate concerns about measurement error, we re-estimate

the baseline model on a restricted sample that excludes observations requiring assumptions about CO₂ reductions or derived WTP values. This restriction reduces potential bias arising from constructed data.

3 Results

In this section, we first present the average WTP measures, their distribution across studies, and our assessment of publication bias. We then turn to the meta-regressions to explore the factors associated with WTP for CO₂ reductions.

3.1 Descriptive evidence

We begin by summarizing the WTP_R and WTP_{kg} estimates. Figure 1 shows the distribution of WTP_R and WTP_{kg} across studies, where a logarithmic x-axis is used for a better representation of the distribution of observations. Table 3 provides mean WTP estimates and their corresponding bootstrap 95% confidence intervals. The average WTP_R of the study averages is USD 1.30 [0.89, 1.79]. The average WTP_{kg} is USD 1.99 [1.06, 3.45] per kg of CO₂ reduction, or USD 1,993 per ton of CO₂.

Measure	Mean	95% CI
WTP _R (USD)	1.30	[0.89, 1.79]
WTP _{kg} (USD)	1.99	[1.06, 3.45]
WTP _{CN} (USD)	15.01	[4.35, 32.91]
WTP _{CN%} (%)	251.71	[119.02, 412.26]

Table 3: WTP estimates and 95% confidence intervals

95% confidence intervals are computed using the percentile bootstrap with 10,000,000 resamples. The mean of WTP estimates for the full sample of 37 studies is obtained by averaging the study-specific WTP estimates and bootstrapping. The non-standardized measure WTP_R denotes WTP for carbon reductions. WTP_{kg} is WTP per 1 kg carbon reduction, WTP_{CN} is WTP for carbon-neutrality, and WTP_{CN%} is the markup (proportion of a product's price that consumers would be willing to pay extra) for carbon neutrality.

To benchmark these estimates, we compare the average WTP_{kg} with the SCC. The SCC is commonly used to define the appropriate level at which

carbon should be priced (Aldy et al., 2021), and cost-effectiveness estimates are generally in a similar range (e.g., Stiglitz et al., 2017; IMF, 2019). This comparison enables us to assess the extent to which consumers voluntarily internalize the climate externalities. To put our estimate of USD 1,993 per ton of CO₂ into perspective, the SCC during the Obama administration was around USD 40 per ton of CO₂ (IWG on Social Cost of Carbon, 2010, 2013), while under the Biden administration it was at USD 51 per ton of CO₂ (IWG on Social Cost of Carbon, 2016; IWG on Social Cost of Greenhouse Gases, 2021) for several years before being raised to USD 190 per ton of CO₂ (Environmental Protection Agency, 2023) towards the end of that administration and before the start of a new administration. The economic literature points, however, to potentially higher SCC values than these policy benchmarks, with considerable dispersion in estimates (see, e.g., Tol, 2011; Pindyck, 2013; Howarth et al., 2014; Pezzey, 2019; Aldy et al., 2021; National Academies of Sciences, Engineering, and Medicine, 2017; Rennert et al., 2022, for reviews and discussions). While some of these figures are in the thousands, most often they are in the low hundreds, hence much lower than the average WTP_{kg} that the meta-analysis provides.

Carbon prices around the world also vary widely. They have generally kept increasing over the last few years, but only a few countries, such as a few Nordic countries, Switzerland, and Uruguay, have carbon prices above USD 130 per ton of CO₂ (World Bank, 2025), about a fifteenth of the average WTP_{kg} derived from the literature that the meta-analysis covers.

A few interesting observations emerge from Figure 1. First, there is substantial variation in WTP_R and WTP_{kg} estimates, both between and within studies. Our regressions in Section 3.2 explore potential sources of this variation from a correlational perspective. Second, there appears to be a positive relationship between WTP_R and the amount of CO₂ reductions, as indicated by the more numerous larger circles to the right of the mean (vertical line).³

³In contrast, the relationship is negative for WTP_{kg} by construction because this measure is defined as WTP_R divided by the amount of CO₂ reduction.

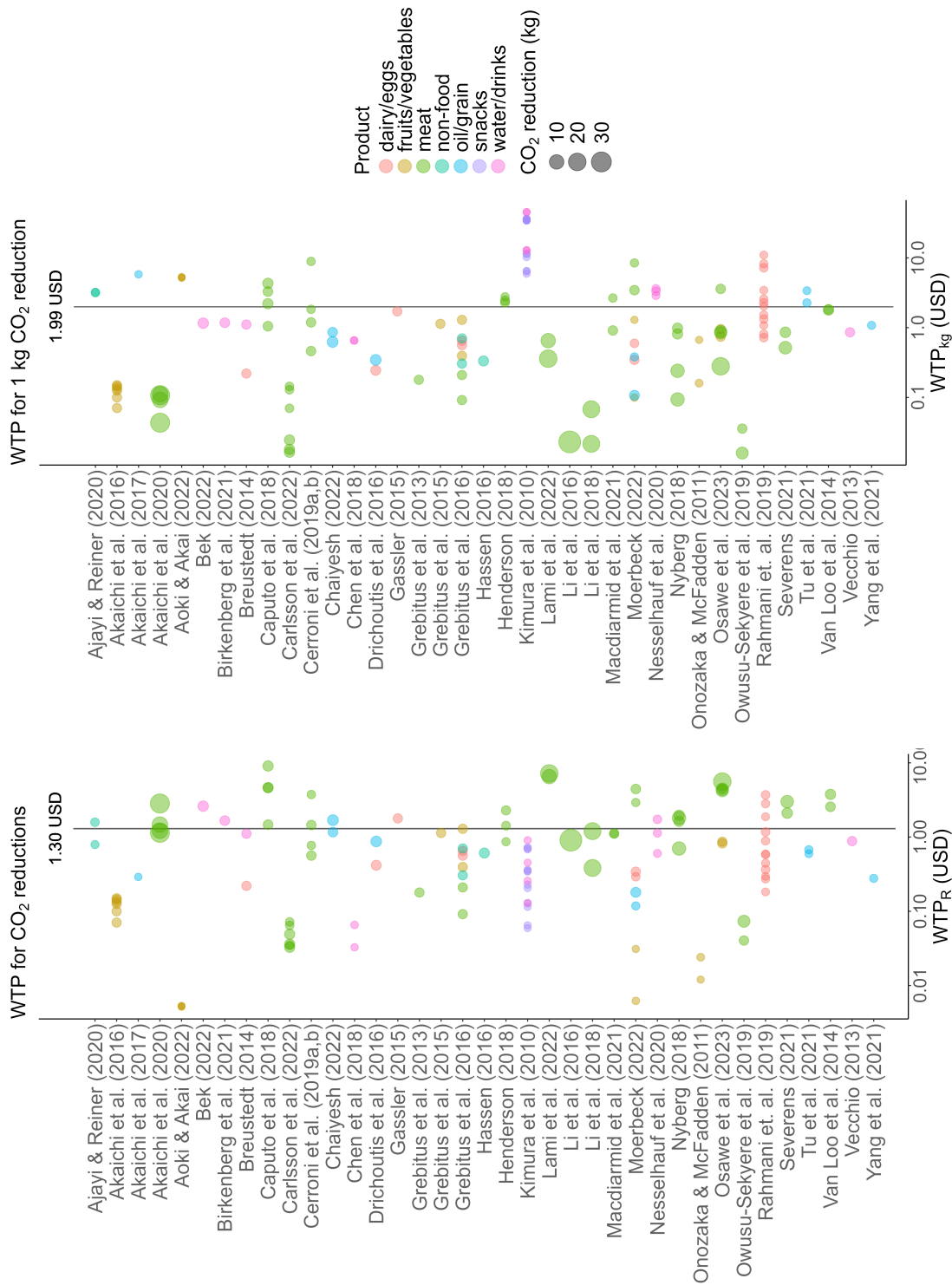


Figure 1: WTP estimates for CO₂ reductions across studies

A logarithmic axis (base 10) is used to create this figure. The vertical lines represent the mean of study means. The left graph displays WTP_R (non-standardized WTP for CO₂ reductions) across studies, where the size of each circle represents the amount of CO₂ reduction in kilograms. The right-hand graph shows WTP_{kg} (WTP for 1 kg CO₂ reduction), which is calculated by dividing WTP_R by the amount of CO₂ reduction. Both WTP_R and WTP_{kg} are measured in 2020 USD.

Figure A.2 in Appendix A.2, which plots WTP measures against CO₂ reductions, confirms this association. Third, WTP_R and WTP_{kg} estimates vary across products. Table A.5 and Figure A.1 in Appendix A.2 further confirm heterogeneity in WTP_R across product categories.

Next, we discuss two alternative WTP measures: WTP_{CN} and WTP_{CN%}, based on the mean of study means. WTP_{CN} of USD 14.99 [4.35, 32.86] represents the WTP to make a product carbon neutral. Because we aggregate a variety of products with different prices, we also report WTP_{CN%}, the markup (percentage premium over product price that consumers are willing to pay) for carbon neutrality, which equals 252% [119, 412].

Figure 2 shows the distribution of WTP_{CN} and WTP_{CN%} measures across studies and provides several insights. First, substantial variation across studies in WTP estimates remains. Second, there appears to be a positive relationship between WTP for carbon neutrality and product CO₂ footprint, which is the relevant measure in this context because carbon neutrality corresponds to fully reducing (or offsetting) a product’s baseline emissions rather than achieving a specific amount of emissions reduction. Third, as in the case of WTP_R and WTP_{kg} we also observe heterogeneity in WTP_{CN} and WTP_{CN%} across product categories.

As the next step, we assess publication bias using the FAT-PET-WLS specification. Table 4 summarizes the results. The intercept β_1 for WTP_{kg} is 7.81 [−20.82, 36.44], which is statistically insignificant, indicating that there is no evidence of publication bias. Figure 3, which plots WTP measures against the square root of the study sample size, does not suggest a relationship between estimated WTP magnitudes and sample size, consistent with the regression-based findings.

We also test for publication bias for the remaining three WTP measures, WTP_R, WTP_{CN}, and WTP_{CN%} as shown in Table 4, and find statistically insignificant intercepts, indicating no evidence of publication bias for these outcomes either. Overall, the FAT diagnostics do not indicate publication bias and do not change our conclusions.

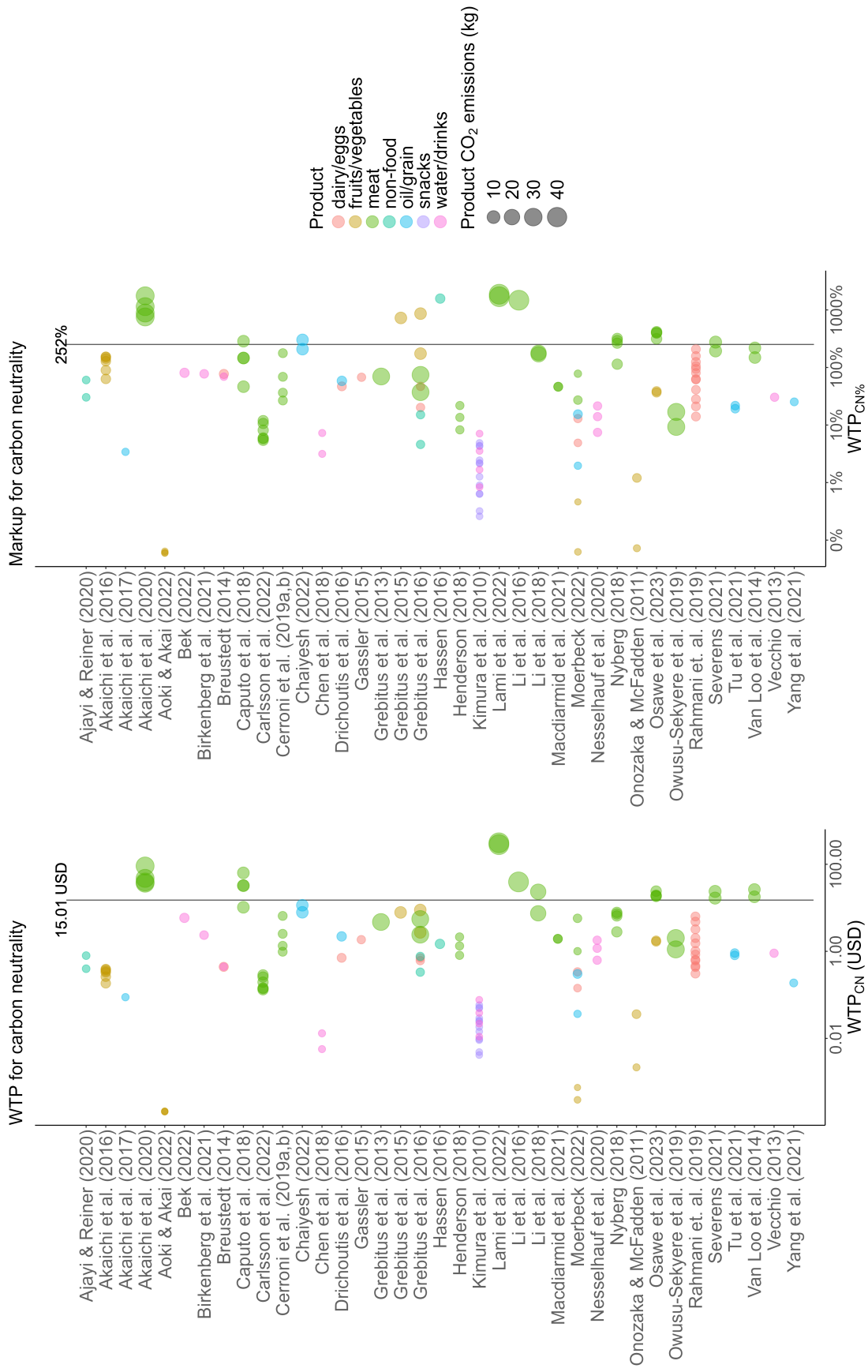


Figure 2: WTP for carbon neutrality across studies

A logarithmic x-axis (base 10) is used to create this figure. The left graph displays WTP_{CN} (WTP for carbon neutrality). The right graph displays WTP_{CN}%, which is the proportion of the product's price that consumers would be willing to pay extra for carbon neutrality.

Measure	N	β_1 [95% CI]
WTP _{kg} (USD)	52	7.81 [-20.82, 36.44]
WTP _R (USD)	52	11.74 [-3.29, 26.77]
WTP _{CN} (USD)	52	116.35 [-231.81, 464.52]
WTP _{CN%} (%)	52	-640.71[-5392.20, 4110.78]

Table 4: Publication bias test

β_1 is the intercept in the FAT-PET-WLS regression $Y_i\sqrt{n_i} = \beta_0\sqrt{n_i} + \beta_1 + \epsilon_i$; where $H_0 : \beta_1 = 0$ is used as a test of whether or not there is publication selection. N is the number of study-sample effect sizes. Confidence intervals use study-clustered robust standard errors.

Overall, if taking the estimates from the literature at face value, the analysis provided in this section points to a very strong WTP for reductions in CO₂ emissions that are an order of magnitude larger than most estimates of the SCC and current levels of policy stringency.

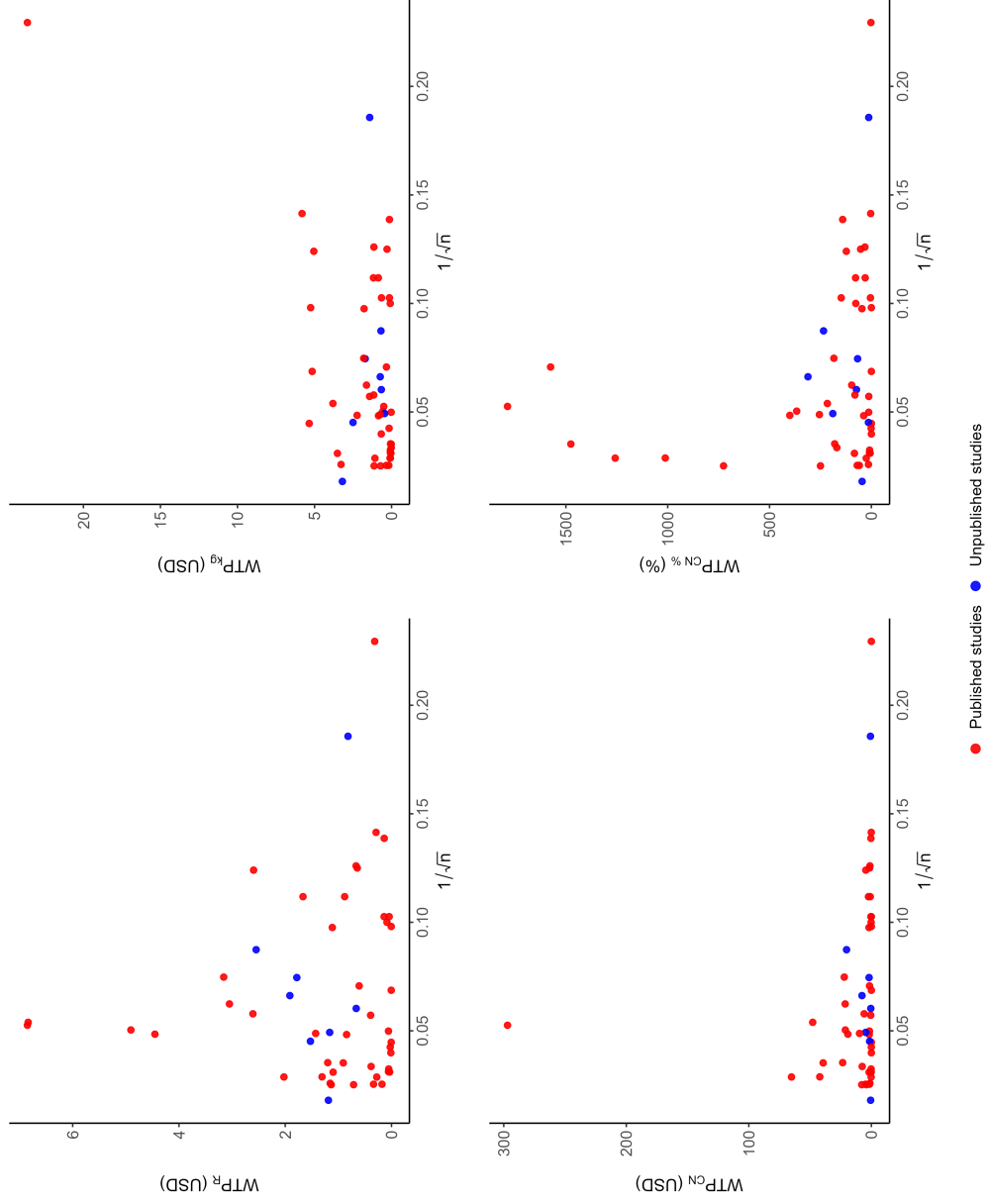


Figure 3: WTP estimates and study sample size

This plot displays WTP estimates against $1/\sqrt{n}$ (inverse square root of the sample size). To maintain readability, extreme observations are excluded using 95th-percentile cutoffs on both axes. subsample is no longer analyzed. The non-standardized measure WTP_R denotes WTP for carbon reductions; WTP_{kg} is WTP per 1 kg carbon reduction; WTP_{CN} is WTP for carbon neutrality; and WTP_{CN}% is the share of the product price consumers would pay extra for carbon neutrality.

3.2 Meta-regression results

In this section, we turn to meta-regressions to try to determine the main factors, including methodological, that are associated with WTP for CO₂ reductions. This analysis is correlational in spirit, yet informative in contributing to addressing our overarching question about the real-world demand for climate certifications, including carbon neutrality.

Our independent variables in the first model (OLS I) include key product characteristics that may explain variation in WTP_R, namely the amount of CO₂ reduction, which is z-scored, to understand whether study participants value greater contributions to climate mitigation; and product price, which is also z-scored, to assess the proportionality of WTP_R to product price. We z-score continuous covariates to improve numerical stability and ensure that variables measured in different units enter the regression on comparable scales.

In the second model⁴ (OLS II) we add study characteristics, including a dummy variable for stated preference methods, to account for potential biases such as hypothetical bias that may arise in hypothetical survey settings, a dummy variable for in-person studies to account for potential differences relative to computer-based or online studies, a variable indicating the sample size of the study, which is z-scored, a dummy variable for published studies, to account for potential biases in comparison to unpublished studies, and study year, which is z-scored, to account for potential secular trends in public awareness regarding climate change. Contextual controls include gross domestic product (GDP) per capita, which is z-scored, to account for the economic situation of consumers in the study country, and a dummy variable for studies conducted in Europe to control for geography-specific effects, including potential variation in environmental friendliness.

In the third model (OLS III) we add control variables for observations requiring assumptions about the amount of CO₂ reductions, and for observations we derived from studies ourselves (as opposed to those directly reported in the

⁴Breusch–Pagan tests reported in Table A.6 indicate heteroskedasticity in the OLS specifications, while Figure A.3 provides complementary visual diagnostics. Accordingly, we report standard errors clustered at the study level.

original study), to account for the need to make additional calculations and interpret the results from the original findings.⁵

Table 5 shows that coefficients remain largely unchanged across different model specifications, in particular for estimates that are statistically significant, which we discuss first. The estimates confirm a positive and significant association between the magnitude of CO₂ reductions and WTP_R.⁶ Figure A.2 in Appendix A.2 provides a visual illustration confirming this relationship. This finding is consistent with the expectation that consumers value larger emissions reductions more and suggests that study participants are responsive to the magnitude of CO₂ reductions rather than treating carbon labels as purely symbolic. Other interesting associations also emerge. For instance, a higher product price is associated with a higher WTP. That is, a price increase for achieving comparable reductions in CO₂ emissions may be more easily accepted by study participants when it represents a smaller share of the overall product price, so that the relative price increase is more muted. Additionally, studies conducted in Europe (compared to Africa, America, and Asia) also seem to be associated with a higher WTP, even when controlling for the GDP per capita.

We do not find a robust significant association of WTP_R with the following variables, neither in OLS II nor in OLS III models: stated preference studies; in-person studies; sample size of the study; published studies; study year; and GDP per capita. Similarly, in OLS III we do not find a robust significant association of WTP_R with the control variables for CO₂ assumptions and WTP derivations.

Appendix A.3 includes our battery of robustness tests and a detailed discussion thereof. The main takeaway is that also in robustness tests the significance of the abovementioned coefficients shows little sensitivity to specification changes, which we summarize shortly here.

First, we present the estimates of alternative models, including the mixed-effects and weighted mixed-effects models, in Table A.7. Second, we incorporate

⁵Appendix A.1.3 provides more details on carbon reduction assumptions and WTP derivations.

⁶We use WTP_R in the regression analysis, as it is the unstandardized measure of marginal WTP and provides a more appropriate basis for examining correlates of valuation across studies. Using WTP_{kg} instead of WTP_R as the outcome variable would mechanically generate a negative association with the amount of CO₂ reduction because WTP_{kg} is defined as WTP_R divided by the amount of CO₂ reduction (in kilograms).

	OLS I	OLS II	OLS III
Intercept	0.75*** (0.08)	0.51 (0.33)	0.58* (0.32)
CO ₂ reduction	0.12** (0.05)	0.11*** (0.04)	0.11*** (0.04)
Price	0.38*** (0.05)	0.35*** (0.05)	0.35*** (0.05)
Stated pref. method		0.00 (0.24)	-0.02 (0.25)
In-person		-0.08 (0.24)	-0.10 (0.23)
Sample size		-0.05 (0.08)	-0.06 (0.08)
Publication		0.01 (0.23)	0.00 (0.22)
Study year		0.04 (0.09)	0.02 (0.10)
GDP per capita		0.06 (0.08)	0.07 (0.08)
Europe		0.37** (0.17)	0.33* (0.18)
CO ₂ reduction assumpt.			-0.09 (0.12)
WTP derivation			0.07 (0.15)
R ²	0.40	0.50	0.50
Adj. R ²	0.39	0.46	0.45
Num. obs.	126	126	126
Number of obs.	126	126	126
Adjusted-R ²	0.39	0.46	0.45
AIC	217.05	209.92	212.91
BIC	228.40	241.12	249.78
Log Likelihood	-104.53	-93.96	-93.45

***p<0.01; **p<0.05; *p<0.1.

Table 5: Factors associated with WTP_R for CO₂ reductions

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use an OLS model with clustered standard errors by study. The dependent variable is the (unstandardized) WTP for CO₂ emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

two additional variables: a dummy variable for colored labels, which are used in the literature to distinguish between higher and lower carbon footprint options, and a dummy variable for carbon-neutral certifications, as opposed to carbon reductions and carbon footprint certifications in Tables A.8 and A.9. Third, we show WLS (weighted least squares) and weighted mixed-effects models in Tables A.10 and A.11. Fourth, we compare alternative transformations of the dependent variable in Tables A.12 and A.13. Fifth, we run the OLS model with two-way clustered errors: studies and product categories or countries in Table A.14. Sixth, we conduct mixed effects regressions with random effects

for studies and either product categories or countries in Table A.15. Seventh, we include the square of the z-scored CO₂ emission reduction variable in Tables A.16 and A.17. Finally, we show regressions with different subsets of the sample that exclude observations requiring CO₂ reduction assumptions, WTP derivations, or both, in Tables A.18 and A.19.

In sum, the meta-regressions highlight that emissions reductions, product prices, and studies conducted in Europe are positively associated with variation in WTP for CO₂ reductions, and the results remain stable across a wide range of robustness checks.

4 Conclusions

Assessing the demand for carbon-neutral products is crucial to determining the potential for voluntary CO₂ reductions in the private sector. While carbon-neutral products have increasingly become available, they often remain a niche market segment, driven largely by investors' pressure and corporate preparations for future policy tightening. This paper synthesizes the existing empirical evidence on WTP for carbon-neutral and low-carbon labels through a comprehensive meta-analysis of 37 studies, 126 WTP observations from total of 29,666 participants.

Our findings indicate a substantial WTP for carbon reductions. The average implied valuation of USD 1.99 per kilogram of CO₂ reduction (USD 1,993 per ton) is an order of magnitude larger than current carbon prices and exceeds most estimates of the SCC. Taking these estimates at face value suggests that consumers are willing to privately internalize climate externalities to a significant degree when presented with labeled products. Furthermore, we find that WTP is sensitive to the magnitude of CO₂ reductions, suggesting that labels which communicate the size of CO₂ emissions reductions may be more effective. We also find a positive association between WTP and product price. Finally, valuations are significantly higher in European samples than in other regions.

These results have important policy and research implications. From a

policy perspective, the strong positive WTP suggests that carbon-neutral and low-carbon labels can unlock significant private funding for climate mitigation. If consumers are indeed willing to pay premiums well above the SCC, voluntary labeling could act as a powerful complement to mandatory carbon pricing. However, the sheer magnitude of these estimates, far exceeding the costs of abatement in many sectors, may point to the likely presence of hypothetical bias inherent in the stated-preference methods that dominate the literature.

Therefore, while this study establishes that the perceived value of carbon neutrality is high in survey and experimental settings, a gap may remain between these stated preferences and real-world market behavior. We contribute to closing this gap by providing a standardized metric (WTP for reducing CO₂ emissions by 1 kg) that future researchers can use to benchmark revealed-preference findings against the baseline established here. Future research could focus on validating these high valuations in uncontrolled market environments to determine how much of this stated demand is reflected in an actual market, a task we undertake in a companion study using marketplace data (Carattini et al., 2025).

References

- Adalja, A., J. Hanson, C. Towe, and E. Tselepidakis (2015). An examination of consumer willingness to pay for local products. *Agricultural and Resource Economics Review* 44(3), 253–274.
- Ajayi, V. and D. Reiner (2020). Are consumers willing to pay for industrial decarbonisation? Evidence from a discrete choice experiment on green plastics. Cambridge Working Paper in Economics 20110, No EPRG 2033.
- Akaichi, F., S. de Grauw, P. Darmon, and C. Revoredo-Giha (2016). Does fair trade compete with carbon footprint and organic attributes in the eyes of consumers? Results from a pilot study in Scotland, the Netherlands and France. *Journal of Agricultural and Environmental Ethics* 29(6), 969–984.
- 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.
- Akaichi, F., C. Revoredo Giha, K. Glenk, and J. M. Gil (2020). How consumers in the UK and Spain value the coexistence of the claims low fat, local, organic and low greenhouse gas emissions. *Nutrients* 12(1), 120.
- Aldy, J. E., M. J. Kotchen, R. N. Stavins, and J. H. Stock (2021). Keep climate policy focused on the social cost of carbon. *Science* 373(6557), 850–852.
- Andreani, G. and G. Sogari (2023). The effect of sustainability labels on consumers' willingness-to-pay: Systematic review and meta-analysis. <https://osf.io/e9mkr/>. preprint.
- Aoki, K. and K. Akai (2022). Testing hypothetical bias in a choice experiment: An application to the value of the carbon footprint of mandarin oranges. *PLOS One* 17(1), e0261369.

- Asioli, D., C. Bazzani, and R. M. Nayga (2018). Consumers' valuation for lab produced meat: An investigation of naming effects. 2018 Agricultural & Applied Economics Association Annual Meeting, Paper no. 274066.
- Asioli, D., J. Fuentes-Pila, S. Alarcón, J. Han, J. Liu, J.-F. Hocquette, and R. M. Nayga (2022). Consumers' valuation of cultured beef burger: A multi-country investigation using choice experiments. *Food Policy* 112, 102376.
- Asioli, D., X. Zhou, A. Halmemies-Beauchet-Filleau, A. Vanhatalo, D. I. Givens, A. Rondoni, and A. Turpeinen (2023). Consumers' valuation for low-carbon emission and low-saturated fat butter. *Food Quality and Preference* 108, 104859.
- Bek, U. (2022). Pricing sustainable shipping of coffee: Consumers' preferences and willingness to pay for emission reductions and offsets. *Junior Management Science* 7(3), 543–568.
- 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.
- Boehm, R., H. Kitchel, S. Ahmed, A. Hall, C. M. Orians, J. R. Stepp, A. Robbat, Jr, T. S. Griffin, and S. B. Cash (2019). Is agricultural emissions mitigation on the menu for tea drinkers? *Sustainability* 11(18), 4883.
- Boesch, I. and M. Weber (2012). Processor's preferences and basic differentiation strategies for potatoes, milk, and wheat in Switzerland. *Journal of Agricultural & Food Industrial Organization* 10(1), 1–24.
- Breustedt, G. (2014). Demand for carbon-neutral food – evidence from a Discrete Choice Experiment for milk and apple juice. In *Proceedings of the 88th Annual Conference of the Agricultural Economics Society*. Paper no. 169748.
- Broeckhoven, I., W. Verbeke, J. Tur-Cardona, S. Speelman, and Y. Hung (2021). Consumer valuation of carbon labeled protein-enriched burgers in European older adults. *Food Quality and Preference* 89, 104114.

- Caputo, V., R. M. Nayga Jr, and R. Scarpa (2013). Food miles or carbon emissions? Exploring labelling preference for food transport footprint with a stated choice study. *Australian Journal of Agricultural and Resource Economics* 57(4), 465–482.
- Caputo, V., E. J. Van Loo, R. Scarpa, R. M. Nayga Jr, and W. Verbeke (2018). Comparing serial, and choice task stated and inferred attribute non-attendance methods in food choice experiments. *Journal of Agricultural Economics* 69(1), 35–57.
- Carattini, S., F. Dvorak, I. Logar, and B. Ozdemir-Oluk (2025). Demand for carbon-neutral products. CESifo Working Paper 12232.
- Carlsson, F., M. Kataria, E. Lampi, E. Nyberg, and T. Sterner (2022). Red, yellow, or green? Do consumers' choices of food products depend on the label design? *European Review of Agricultural Economics* 49(5), 1005–1026.
- Carroll, K. A. (2018). *Three Papers on Consumer Food Choices: An Experimental Approach*. Ph.D. dissertation. The University of Wisconsin-Madison.
- Cerroni, S., V. Watson, D. Kalentakis, and J. I. Macdiarmid (2019a). Value-elicitation and value-formation properties of discrete choice experiment and experimental auctions. *European Review of Agricultural Economics*, Vol. 46, No. 1, pp. 3–27.
- Cerroni, S., V. Watson, and J. I. Macdiarmid (2019b). Consumers' rationality and home-grown values for healthy and environmentally sustainable food. *Biobased and Applied Economics*, Vol. 8, No. 2, pp. 101–132.
- Chaiyesh, T. (2022). Consumers' willingness to pay for healthiness, quality and carbon footprint reduction of bagged rice. *Global Business Review*, 1–13.
- Chang, M.-Y., J.-C. Lin, and H.-S. Chen (2023). Consumer attitudes and preferences for healthy boxed meal attributes in Taiwan: Evidence from a choice experiment. *Nutrients* 15(4), 1032.

- 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.
- Colantuoni, F., G. Cicia, T. Del Giudice, D. Lass, F. Caracciolo, and P. Lombardi (2016). Heterogeneous preferences for domestic fresh produce: Evidence from German and Italian early potato markets. *Agribusiness* 32(4), 512–530.
- Cubero Dudinskaya, E., S. Naspetti, G. Arsenos, E. Caramelle-Holtz, T. Latvala, D. Martin-Collado, S. Orsini, E. Ozturk, and R. Zanolli (2021). European consumers' willingness to pay for red meat labelling attributes. *Animals* 11(2), 556.
- Cuong, O. Q., M. Connor, M. Demont, B. O. Sander, and K. Nelson (2022). How do rice consumers trade off sustainability and health labels? Evidence from Vietnam. *Frontiers in Sustainable Food Systems* 6, 1010161.
- De-Magistris, T., A. Gracia, and R. M. Nayga Jr (2013). On the use of honesty priming tasks to mitigate hypothetical bias in choice experiments. *American Journal of Agricultural Economics* 95(5), 1136–1154.
- De Marchi, E., V. Caputo, R. M. Nayga Jr, and A. Banterle (2016). Time preferences and food choices: Evidence from a choice experiment. *Food Policy* 62, 99–109.
- 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.
- Dudinskaya, E. C., S. Naspetti, and R. Zanolli (2020). Using eye-tracking as an aid to design on-screen choice experiments. *Journal of Choice Modelling* 36, 100232.
- Echeverría, R., V. H. Moreira, C. Sepúlveda, and C. Wittwer (2014). Willingness to pay for carbon footprint on foods. *British Food Journal* 116(2), 186–196.

- Environmental Protection Agency (2023). Report on the social cost of greenhouse gases: Estimates incorporating recent scientific advances. Technical report. U.S. Environmental Protection Agency.
- Erraach, Y., S. Sayadi, and C. Parra-López (2017). Measuring preferences and willingness to pay for sustainability labels in olive oil: Evidence from Spanish consumers. 15th Congress of the European Association of Agricultural Economists.
- Feucht, Y. and K. Zander (2017). Consumers' attitudes on carbon footprint labelling: Results of the SUSDIET project. Thünen Working Paper No. 78.
- 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.
- Florax, R. J., C. M. Travisi, and P. Nijkamp (2005). A meta-analysis of the willingness to pay for reductions in pesticide risk exposure. *European Review of Agricultural Economics* 32(4), 441–467.
- Flowers from the Farm. <https://www.flowersfromthefarm.co.uk/>. Accessed: 01 March 2023.
- Gassler, B. (2015). How green is your “Grüner”? Millennial wine consumers' preferences and willingness-to-pay for eco-labeled wine. *Jahrbuch der Österreichischen Gesellschaft für Agrarökonomie* 24, 131–140.
- GlobalProductPrices.com. Prices of goods and services. <https://www.globalproductprices.com/>. Accessed: 01 March 2023.
- 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.

- Grebitus, C., B. Steiner, and M. Veeman (2015). The roles of human values and generalized trust on stated preferences when food is labeled with environmental footprints: Insights from Germany. *Food Policy* 52, 84–91.
- Grebitus, C., B. Steiner, and M. M. Veeman (2016). Paying for sustainability: A cross-cultural analysis of consumers' valuations of food and non-food products labeled for carbon and water footprints. *Journal of Behavioral and Experimental Economics* 63, 50–58.
- Greenhouse Gas Protocol (2004). The Greenhouse Gas Protocol: A corporate accounting and reporting standard. Technical report, World Resources Institute and World Business Council for Sustainable Development.
- Haab, T. C. and K. E. McConnell (2002). *Valuing Environmental and Natural Resources: The Econometrics of Non-market Valuation*. Edward Elgar Publishing.
- Harrison, G. W. and J. A. List (2004). Field experiments. *Journal of Economic Literature* 42(4), 1009–1055.
- Hassen, A. S. (2016). Consumers' willingness to pay for environmental attributes of a cut flower in Ethiopia: A choice experiment approach. *Yildiz Social Science Review* 2(1), 31–46.
- Henderson, C. (2018). Consumer preferences for GM food labeling: A market segments analysis. Agricultural Economics and Agribusiness Undergraduate Honors Theses, University of Arkansas.
- Howarth, R. B., M. D. Gerst, and M. E. Borsuk (2014). Risk mitigation and the social cost of carbon. *Global Environmental Change* 24, 123–131.
- IMF (2019). *Fiscal Monitor: How to Mitigate Climate Change*. Washington, DC.
- IPCC (2018). Special report: Global warming of 1.5°C. Online glossary. <https://www.ipcc.ch/sr15/chapter/glossary/>, Accessed 2025-12-13.

- IWG on Social Cost of Carbon (2010). Technical support document: Social cost of carbon for regulatory impact analysis under Executive Order 12866. Interagency Working Group on Social Cost of Carbon, Washington, DC.
- IWG on Social Cost of Carbon (2013). Technical support document: Social cost of carbon for regulatory impact analysis under Executive Order 12866. Interagency Working Group on Social Cost of Carbon, Washington, DC.
- IWG on Social Cost of Carbon (2016). Technical support document: Technical update of the social cost of carbon for regulatory impact analysis under Executive Order 12866. Interagency Working Group on Social Cost of Carbon, Washington, DC.
- IWG on Social Cost of Greenhouse Gases (2021). Technical support document: Social cost of carbon, methane, and nitrous oxide interim estimates under Executive Order 13990. Interagency Working Group on Social Cost of Greenhouse Gases, Washington, DC.
- Janßen, D. and N. Langen (2017). The bunch of sustainability labels - Do consumers differentiate? *Journal of Cleaner Production* 143, 1233–1245.
- Johnston, R. J., K. J. Boyle, W. Adamowicz, J. Bennett, R. Brouwer, T. A. Cameron, W. M. Hanemann, N. Hanley, M. Ryan, R. Scarpa, et al. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists* 4(2), 319–405.
- Kim, E.-H. and T. Lyon (2011). When does institutional investor activism increase shareholder value?: The Carbon Disclosure Project. *The B.E. Journal of Economic Analysis & Policy* 11(1).
- Kim, H., L. A. House, and T.-K. Kim (2016). Consumer perceptions of climate change and Willingness to Pay for mandatory implementation of low carbon labels: The case of South Korea. *International Food and Agribusiness Management Review* 19(4), 129–144.

- Kimura, A., Y. Wada, A. Kamada, T. Masuda, M. Okamoto, S.-i. Goto, D. Tsuzuki, D. Cai, T. Oka, and I. Dan (2010). Interactive effects of carbon footprint information and its accessibility on value and subjective qualities of food products. *Appetite* 55(2), 271–278.
- Kovalsky, K. L. and J. L. Lusk (2013). Do consumers really know how much they are willing to pay? *Journal of Consumer Affairs* 47(1), 98–127.
- Lami, O., F. J. Mesías, C. Balas, C. Díaz-Caro, M. Escribano, and A. Horrillo (2022). Does carbon footprint play a relevant role in food consumer behaviour? A focus on Spanish beef. *Foods* 11(23), 3899.
- Li, S. and Z. Kallas (2021). Meta-analysis of consumers’ willingness to pay for sustainable food products. *Appetite* 163, 105239.
- Li, X., K. L. Jensen, C. D. Clark, and D. M. Lambert (2016). Consumer willingness to pay for beef grown using climate friendly production practices. *Food Policy* 64, 93–106.
- Li, X., K. L. Jensen, D. M. Lambert, and C. D. Clark (2018). Consequentiality beliefs and consumer valuation of extrinsic attributes in beef. *Journal of Agricultural and Applied Economics* 50(1), 1–26.
- List, J. A. and C. A. Gallet (2001). What experimental protocol influence disparities between actual and hypothetical stated values? *Environmental and Resource Economics* 20(3), 241–254.
- Lombardi, G. V., R. Berni, and B. Rocchi (2017). Environmental friendly food. Choice experiment to assess consumer’s attitude toward “climate neutral” milk: The role of communication. *Journal of Cleaner Production* 142, 257–262.
- Macdiarmid, J. I., S. Cerroni, D. Kalentakis, and C. Reynolds (2021). How important is healthiness, carbon footprint and meat content when purchasing a ready meal? Evidence from a non-hypothetical discrete choice experiment. *Journal of Cleaner Production* 282, 124510.

- Magistris, T. d., A. Gracia, et al. (2014). Do consumers care about organic and distance labels? An empirical analysis in Spain. *International Journal of Consumer Studies* 38(6), 660–669.
- Mattmann, M., I. Logar, and R. Brouwer (2016). Hydropower externalities: A meta-analysis. *Energy Economics* 57, 66–77.
- Menapace, L. and R. Raffaelli (2017). Preferences for locally grown products: Evidence from a natural field experiment. *European Review of Agricultural Economics* 44(2), 255–284.
- Meyerding, S. G., A.-L. Schaffmann, and M. Lehberger (2019). Consumer preferences for different designs of carbon footprint labelling on tomatoes in Germany - does design matter? *Sustainability* 11(6), 1587.
- Michaud, C., D. Llerena, and I. Joly (2013). Willingness to pay for environmental attributes of non-food agricultural products: A real choice experiment. *European Review of Agricultural Economics* 40(2), 313–329.
- Möerbeck, S. (2022). Communicating sustainability in the food industry: The impact of carbon labels on the willingness to pay for food products. Master's thesis. Universidade Católica Portuguesa.
- Moon, D., K. Lee, and C. Kim (2015). Measuring willingness to pay for CO₂ information on consumption goods in Korea. SIER Working Paper No. 2015-13, Sungkyun Institute of Economic Research, Sungkyunkwan University.
- Mostafa, M. M. (2016). Egyptian consumers' willingness to pay for carbon-labeled products: A contingent valuation analysis of socio-economic factors. *Journal of Cleaner Production* 135, 821–828.
- My Emissions. Food carbon accounting platform. <https://myemissions.co/>. Accessed: 01 March 2023.

- Nassivera, F., G. Gallenti, S. Troiano, F. Marangon, M. Cosmina, P. Bogoni, B. Campisi, and M. Carzedda (2020). Italian millennials' preferences for wine: An exploratory study. *British Food Journal* 122(8), 2403–2423.
- National Academies of Sciences, Engineering, and Medicine (2017). *Valuing Climate Damages: Updating Estimation of the Social Cost of Carbon Dioxide*. The National Academies Press, Washington, DC.
- Nelson, J. P. and P. E. Kennedy (2009a). The use (and abuse) of meta-analysis in environmental and natural resource economics: An assessment. *Environmental and Resource Economics* 42(3), 345–377.
- Nelson, J. P. and P. E. Kennedy (2009b). The use (and abuse) of meta-analysis in environmental and natural resource economics: An assessment. *Environmental and Resource Economics* 42(3), 345–377.
- Nesselhauf, L., R. Fleuchaus, and L. Theuvsen (2020). What about the environment? A choice-based conjoint study about wine from fungus-resistant grape varieties. *International Journal of Wine Business Research* 32(1), 96–121.
- Nyberg, E. (2018). Meat production preferences among Swedish consumers: A choice experiment with lasagna. Master's thesis, University of Gothenburg, School of Business, Economics and Law.
- 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.
- Osawe, O. W., G. Grilli, and J. Curtis (2023). Examining food preferences in the face of environmental pressures. *Journal of Agriculture and Food Research* 11, 100476.
- Owusu-Sekyere, E., Y. Mahlathi, and H. Jordaan (2019). Understanding South African consumers' preferences and market potential for products with low water and carbon footprints. *Agrekon* 58(3), 354–368.

- Ozdemir-Oluk, B. (2025). Transparency of carbon-neutral labels: evidence from a choice experiment. Working paper.
- Ozkan, M. (2011). Purchaser preferences on carbon labels: Conventional vs. organic milk. Master's thesis, University of Arkansas.
- Penn, J. and W. Hu (2019). Cheap talk efficacy under potential and actual hypothetical bias: A meta-analysis. *Journal of Environmental Economics and Management* 96, 22–35.
- Peschel, A. O., C. Grebitus, B. Steiner, and M. Veeman (2016). How does consumer knowledge affect environmentally sustainable choices? Evidence from a cross-country latent class analysis of food labels. *Appetite* 106, 78–91.
- Pezzey, J. C. V. (2019). Why the social cost of carbon will always be disputed. *Wiley Interdisciplinary Reviews: Climate Change* 10(1), e558.
- Pindyck, R. S. (2013). Climate change policy: What do the models tell us? *Journal of Economic Literature* 51(3), 860–872.
- Piracci, G., E. Lamonaca, F. G. Santeramo, F. Boncinelli, and L. Casini (2024). On the willingness to pay for food sustainability labelling: A meta-analysis. *Agricultural Economics* 55(2), 329–345.
- Plate Up for the Planet. Carbon food calculator. <https://assets.plateupfortheplanet.org/carbon-calculator/>. Accessed: 01 March 2023.
- Rahmani, D., Z. Kallas, M. Pappa, and J. M. Gil (2019). Are consumers' egg preferences influenced by animal-welfare conditions and environmental impacts? *Sustainability* 11(22), 6218.
- Ratliff, E. L. (2021). *Consumer Preferences for Value Added Products: A Case Study for South Carolina*. Ph. D. thesis, Clemson University.

- Rennert, K., F. Errickson, B. C. Prest, L. Rennels, R. G. Newell, W. Pizer, C. Kingdon, J. Wingenroth, R. Cooke, B. Parthum, et al. (2022). Comprehensive evidence implies a higher social cost of CO₂. *Nature* 610(7933), 687–692.
- Rogelj, J., O. Geden, A. Cowie, and A. Reisinger (2021). Net-zero emissions targets are vague: Three ways to fix. *Nature* 591(7850).
- Rondoni, A. and S. Grasso (2021). Consumers behaviour towards carbon footprint labels on food: A review of the literature and discussion of industry implications. *Journal of Cleaner Production* 301, 127031.
- Selina Wamucii. Food & agriculture B2B marketplace. <https://www.selinawamucii.com/>. Accessed: 01 March 2023.
- Severens, F. (2021). Choice experiment amongst Dutch consumers: Willingness to pay for sustainability attributes of pork products. Master's thesis, Institute for Environmental Studies (IVM), Vrije Universiteit Amsterdam.
- Smith, V. K. and S. K. Pattanayak (2002). Is meta-analysis a Noah's ark for non-market valuation? *Environmental and Resource Economics* 22(1), 271–296.
- Sonntag, W. I., D. Lemken, A. Spiller, and M. Schulze (2023). Welcome to the (label) jungle? Analyzing how consumers deal with intra-sustainability label trade-offs on food. *Food Quality and Preference* 104, 104746.
- Sporleder, E. M., M. Kayser, N. Friedrich, and L. Theuvsen (2014). Consumer preferences for sustainably produced bananas: A discrete choice experiment. *International Food and Agribusiness Management Review* 17(1030-2016-82973), 59–82.
- Stanley, T. D. and H. Doucouliagos (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods* 5(1), 60–78.

- Staples, A. J., C. J. Reeling, N. J. O. Widmar, and J. L. Lusk (2020). Consumer willingness to pay for sustainability attributes in beer: A choice experiment using eco-labels. *Agribusiness* 36(4), 591–612.
- Steiner, B. E., A. Peschel, and C. Grebitus (2017). Multi-product category choices labeled for ecological footprints: Exploring psychographics and evolved psychological biases for characterizing latent consumer classes. *Ecological Economics* 140, 251–264.
- Stiglitz, J. E., N. Stern, M. Duan, O. Edenhofer, G. Giraud, G. Heal, E. L. la Rovere, A. Morris, E. Moyer, M. Pangestu, Y. Sokona, P. R. Shukla, and Winkler (2017). Report of the High-Level Commission on Carbon Prices. Technical report, Carbon Pricing Leadership Coalition.
- Tait, P., C. Saunders, M. Guenther, P. Rutherford, and S. Miller (2016). Exploring the impacts of food label format on consumer willingness to pay for environmental sustainability: A choice experiment approach in the United Kingdom and Japan. *International Food Research Journal* 23(4).
- Tappwater. Water filters and filtration systems. <https://tappwater.co/>. Accessed: 01 March 2023.
- Thøgersen, J. and K. S. Nielsen (2016). A better carbon footprint label. *Journal of Cleaner Production* 125, 86–94.
- Tol, R. S. (2011). The social cost of carbon. *Annual Review of Resource Economics* 3(1), 419–443.
- Tu, V. H., S. W. Kopp, N. T. Trang, A. Kontoleon, and M. Yabe (2021). UK consumers’ preferences for ethical attributes of floating rice: Implications for environmentally friendly agriculture in Vietnam. *Sustainability* 13(15), 8354.
- Tully, S. M. and R. S. Winer (2014). The role of the beneficiary in willingness to pay for socially responsible products: A meta-analysis. *Journal of Retailing* 90(2), 255–274.

- Tunçel, T. and J. K. Hammitt (2014). A new meta-analysis on the WTP/WTAs disparity. *Journal of Environmental Economics and Management* 68(1), 175–187.
- United States Department of Agriculture. Economic Research Service, Historical monthly price spread data for beef, pork, and broilers. <https://www.ers.usda.gov/data-products/meat-price-spreads/>, Accessed: 01 March 2023.
- Van Loo, E. J., V. Caputo, R. M. Nayga Jr, H.-S. Seo, B. Zhang, and W. Verbeke (2015). Sustainability labels on coffee: Consumer preferences, willingness-to-pay and visual attention to attributes. *Ecological Economics* 118, 215–225.
- Van Loo, E. J., V. Caputo, R. M. Nayga Jr, and W. Verbeke (2014). Consumers' valuation of sustainability labels on meat. *Food Policy* 49, 137–150.
- Vecchio, R. (2013). Determinants of willingness-to-pay for sustainable wine: Evidence from experimental auctions. *Wine Economics and Policy* 2(2), 85–92.
- Vecchio, R. and A. Annunziata (2015). Willingness-to-pay for sustainability-labelled chocolate: An experimental auction approach. *Journal of Cleaner Production* 86, 335–342.
- World Bank (2025). *State and Trends of Carbon Pricing 2025*. Washington, DC: World Bank.
- Xu, Y., B. Xian, Y. Ren, Y. Wang, L. Lang, and B. Wang (2024). Do carbon labels increase chinese consumers' willingness to pay for carbon-labeled agricultural products? *Journal of Cleaner Production* 434, 140299.
- Yang, X., Q. Chen, Z. Xu, Q. Zheng, R. Zhao, H. Yang, C. Ruan, F. Han, and Q. Chen (2021). Consumers' preferences for health-related and low-carbon attributes of rice: A choice experiment. *Journal of Cleaner Production* 295, 126443.

- Zhao, R., Y. Geng, Y. Liu, X. Tao, and B. Xue (2018). Consumers' perception, purchase intention, and willingness to pay for carbon-labeled products: A case study of Chengdu in China. *Journal of Cleaner Production* 171, 1664–1671.
- Zhao, R., M. Yang, J. Liu, L. Yang, Z. Bao, and X. Ren (2020). University students' purchase intention and willingness to pay for carbon-labeled food products: A purchase decision-making experiment. *International Journal of Environmental Research and Public Health* 17(19), 7026.
- Zheng, Y. (2014). *Consumer preference and willingness-to-pay for locally produced, organic food: a stated choice approach*. Ph. D. thesis. University of Guelph.

Appendix

A.1 Data collection

A.1.1 *Excluded studies and exclusion criteria*

In this section, we provide a more detailed discussion of the excluded studies and the reasons for their exclusion, complementing the study selection procedure described in Section 2.1.

Our initial sample includes 83 studies. First, we exclude 21 studies that lack information that would allow us to derive or make assumptions about the amount of CO₂ reductions. Second, we exclude 12 studies categorized as “carbon transparent,” which value carbon footprint labels without providing information that would allow us to derive the amount of CO₂ reduction associated with the label. In addition, we exclude 6 studies that focus on reduced transportation distance unless the study values the carbon footprint of the product (from its production, distribution, or overall lifecycle) and just “frames” it in terms of distance traveled by car, for example, to make it easier for consumers to gauge the amount of CO₂ emissions. Furthermore, we exclude 5 studies that discretely code cost levels (prices), and one study that reports the WTP as a percentage premium on unspecified product price, not allowing a derivation of WTP in currency units. Additionally, we exclude a study that does not specify the type and amount of a product. Lastly, we exclude a study that reports WTP for a sustainability label, which refers to organic, fair trade, and carbon-neutral attributes, not allowing derivation of WTP of the carbon-neutral label alone. In Table A.1, we list the excluded studies by exclusion reason.

After these exclusions, our final dataset consists of 37 studies.

Exclusion Reason	Excluded Studies
Unknown carbon reduction	Michaud et al. (2013), Van Loo et al. (2015), Vecchio and Annunziata (2015), Tait et al. (2016), De Marchi et al. (2016), Feucht and Zander (2017), Lombardi et al. (2017), Menapace and Raffaelli (2017), Janßen and Langen (2017), Feucht and Zander (2018), Asioli et al. (2018), Boehm et al. (2019), Staples et al. (2020), Dudinskaya et al. (2020), Broeckhoven et al. (2021), Cubero Dudinskaya et al. (2021), Ratliff (2021), Asioli et al. (2022), Cuong et al. (2022), Asioli et al. (2023), Asioli et al. (2023), Sonntag et al. (2023)
Carbon transparent	Ozkan (2011), Caputo et al. (2013), Echeverría et al. (2014), Colantuoni et al. (2016), Moon et al. (2015), Kim et al. (2016), Erraach et al. (2017), Zhao et al. (2018), Nassivera et al. (2020), Zhao et al. (2020), Asioli et al. (2022), Chang et al. (2023)
Transportation distance reduction	Kovalsky and Lusk (2013), De-Magistris et al. (2013), Magistris et al. (2014), Zheng (2014), Adalja et al. (2015), Carroll (2018)
Discretely coded cost	Boesch and Weber (2012), Thøgersen and Nielsen (2016), Peschel et al. (2016), Steiner et al. (2017), Meyerding et al. (2019)
Percentage premium WTP on unspecified product price	Xu et al. (2024)
Unspecified product	Mostafa (2016)
Multiple sustainability labels	Sporleder et al. (2014)

Table A.1: Excluded studies and rationales

A.1.2 List of studies and characteristics

In Table A.2, we summarize the studies included in the meta-analysis and report the product, country, method, and the information provided to participants for each study.

Study	No. obs. ^a	Product	Country	Method ^b	Information provided to participants ^c
Ajayi and Reiner (2020)	2	Plastic bottle	United Kingdom	DCE	Percentage of carbon capture
Akaichi et al. (2016)	6	Banana	France, Netherlands, United Kingdom	DCE	Carbon footprint from transportation
Akaichi et al. (2017)	1	Rice	United States	AFE	Carbon footprint
Akaichi et al. (2020)	4	Ground beef	Spain, United Kingdom	DCE	GHG emissions
Aoki and Akai (2022)	3	Mandarin	Japan	DCE	Carbon footprint
Bek (2022)	1	Coffee	Germany	DCE	Carbon neutrality
Birkenberg et al. (2021)	1	Coffee	Germany	DCE	Carbon neutrality
Breustedt (2014)	2	Juice, milk	Germany	DCE	Carbon footprint
Caputo et al. (2018)	4	Chicken	Belgium	DCE	Carbon footprint from transportation
Carlsson et al. (2022)	6	Lasagne	Sweden	DCE	GHG emission categories
Cerroni et al. (2019a) & Cerroni et al. (2019b)	4	Lasagne	United Kingdom	DCE, AFE	Carbon footprint emission categories

Table A.2: Literature covered in the meta-analysis

^a“No. obs.” refers to the number of WTP observations included from each respective study.

^bIn our database, stated preference studies include DCE (Discrete Choice Experiments), and CV (Contingent Valuation Method). Based on Harrison and List (2004), we classified revealed preference studies as AFE (Artificial Field Experiments) or CLE (Conventional Lab Experiments).

^cNote that unless otherwise stated, the terms such as “carbon footprint,” “carbon neutrality,” and “carbon reduction” refer to emissions from either production or the entire life cycle of the product.

Study	No. obs.	Product	Country	Method	Information provided to participants
Chaimesh (2022)	2	Rice	Thailand	DCE	Percentage of carbon reduction
Chen et al. (2018)	2	Water	China	CLE	Carbon footprint
Drichoutis et al. (2016)	2	Egg, olive oil	Greece	CV	Carbon neutrality
Gassler (2015)	1	Milk	Austria	DCE	Carbon neutrality
Greibitus et al. (2013)	1	Ground Beef	Canada	DCE	Carbon footprint
Greibitus et al. (2015)	1	Potatoes	Germany	DCE	Carbon footprint
Greibitus et al. (2016)	8	Ground beef, potatoes, toilet paper, yogurt	Canada, Germany	DCE	Carbon footprint
Hassen (2016)	1	Flower	Ethiopia	DCE	Percentage of carbon reduction
Henderson (2018)	3	Chicken	United States	DCE	Carbon footprint
Kimura et al. (2010)	16	Candy, chips, chocolate, juice	Japan	CV	Carbon footprint
Lami et al. (2022)	2	Beef	Spain	DCE	Carbon footprint
Li et al. (2016)	1	Beef	United States	CV	Climate-friendly label and annual GHG emission reduction in percentages
Li et al. (2018)	2	Beef, ground beef	United States	DCE	Climate-friendly label
Macdiarmid et al. (2021)	2	Lasagne	United Kingdom	DCE	Carbon footprint

Table A.2: Literature covered in the meta-analysis (continued)

Study	No. obs.	Product	Country	Method	Information provided to participants
Mørbeck (2022)	10	Apples, beef, butter, cheese, chicken, eggs, flour, milk, rice, tomatoes	Germany	CV	Carbon neutrality
Nesselhauf et al. (2020)	3	Wine	Germany	DCE	Percentage of carbon reduction
Nyberg (2018)	4	Lasagne	Sweden	DCE	GHG emissions, expressed in terms of carbon dioxide equivalents
Onozaka and McFadden (2011)	2	Apples, tomatoes	United States	DCE	Carbon footprint
Osawe et al. (2023)	6	Beef, chicken, vegetables	Ireland	DCE	GHG emissions, expressed in terms of carbon dioxide equivalents
Owusu-Sekyere et al. (2019)	2	Beef	South Africa	DCE	GHG emissions, expressed in terms of carbon dioxide equivalents
Rahmani et al. (2019)	12	Egg	Spain	DCE	Percentage of carbon reduction
Severens (2021)	2	Pork	Netherlands	DCE	Carbon reduction categories, expressed as equivalent kilometers driven by a car
Tu et al. (2021)	3	Rice	United Kingdom	DCE	Percentage of carbon reduction
Van Loo et al. (2014)	2	Chicken	Belgium	DCE	Carbon footprint reduction
Vecchio (2013)	1	Wine	Italy	CLE	Carbon neutrality
Yang et al. (2021)	1	Rice	China	DCE	Percentage of carbon reduction

Table A.2: Literature covered in the meta-analysis (continued)

A.1.3 Data collection, covariate construction, and aggregation rules

In this section, we outline our strategies for data collection, covariate construction, and data aggregation. First, we provide detailed information on how we derive the respective WTP estimates and amounts of CO₂ reduction from each study. Second, we present the rules that we follow for aggregating WTP estimates for meta-analysis. Third, we define the independent variables used in our analysis, discuss the general approaches we use to construct them, and note any exceptional cases.

First, we start by providing paper-by-paper details on the WTP derivation in Table A.3. Note that we describe our general WTP derivation approach in Section 2.1.

Second, we describe our data aggregation rule for the main analysis and explain an exception in which we apply an alternative aggregation approach. We initially collected 37 studies and 224 observations. We then aggregate observations that remained constant across studies, product categories, and the independent variables that are used in the last column of Table 3.2 in Section 3.2. Consequently, our final dataset comprises 37 studies and 126 observations.

There is an instance, as described below, in which we use a different aggregation rule. Before we test publication bias, we aggregate observations at the sample level (identified by equal n and the same country within each study), yielding 52 observations from 37 studies, as multiple estimates can originate from the same underlying sample. We do so because the publication-bias analysis applies weights proportional to the square root of the sample size; therefore, we aggregate estimates at the sample level to ensure that each underlying sample receives a consistent weight.

Study	Details
Ajayi and Reiner (2020)	The WTP for 50% and 100% carbon capture relative to 1% capture is reported in the study (Tables 3 and 4). Assuming 1 kg of PET plastic in Europe leads to 2.15 kg carbon emissions (sourced from Tappwater.co) and that the study values a 100 ml PET-type plastic weighing 0.25 kg, its emissions are equal to 0.54 kg of carbon. Therefore, 49% carbon capture corresponds to a reduction of 0.25 kg of carbon emissions, while a 99% emission capture corresponds to a reduction of 0.50 kg carbon emissions. Following rule 6 from the Section A.1.3, we average the WTP values from different choice models as well as from preference and WTP spaces.
Akaichi et al. (2016)	The WTP estimates (from WTP space) for reducing carbon emissions by 1 kg are reported in the study (Table 4). The WTP estimates (from preference space) are derived from the study (Table 3).
Akaichi et al. (2017)	The WTP differences of four types of rice – local hybrid, non-local hybrid, local conventional, and non-local conventional – are reported in the study (Tables 5 and 6). Conventional rice has approximately 0.05 kg (1.76 oz) higher greenhouse gas (GHG) emissions from production, expressed as carbon dioxide equivalents than hybrid rice. The difference between Round 2 (WTP after GHG emissions information) and Round 1 (WTP based on appearance) is used to derive the WTP for 0.05 kg carbon reduction from Table 5, lines 2 and 3. Similarly, the difference between Round 3 (WTP after GHG information and food miles information) and Round 1 (WTP based on appearance) is used to derive the WTP for 0.05 kg carbon reduction from Table 6, lines 2 and 3. Note that our approach focuses on differences between hybrid and conventional rice while keeping the locality attribute constant (lines 2 & 3 only). Therefore, the focus is only on the derivation of the WTP for carbon reductions, not on the WTP for reduction in food miles (the distance at which food is transported from the place of production to the store).

Table A.3: WTP derivation strategy

Study	Details
Akaichi et al. (2020)	The WTP estimates for low (5.9 kg) and moderate (19.1 kg) relative to high (32.2 kg) GHG emissions are reported in the study (Table 3). As this is a common practice, we assume that GHG emissions are expressed in terms of carbon dioxide equivalents in this study. Therefore, the WTP for low relative to high carbon emissions corresponds to 26.3 kg (32.2 kg - 5.9 kg), while the WTP for moderate relative to high emissions corresponds to 13.1 kg (32.2 kg - 19.1 kg) kg of reduction in carbon emissions.
Aoki and Akai (2022)	The WTP estimates for a 0.001 kg increase in carbon emissions are reported in the study (Table 5). The WTP for decreasing carbon emissions by the same amount is derived by taking the negative of these values. Following rule 6 from Section A.1.3, the WTP estimates from hypothetical online surveys with and without cheap talk are averaged.
Bek (2022)	The WTP estimates for offsetting and reducing a product's full supply chain emissions are reported in the study (Table 7). We assume 0.5 kg of coffee leads to 2.5 kg of carbon emissions based on MyEmissions carbon calculator. Following rule 6 from Section A.1.3, the WTP estimates for offsetting and reducing product emissions are averaged.
Birkenberg et al. (2021)	The WTP estimates for carbon-neutral product are reported in the study (Table 4). WTP estimates for carbon neutrality are calculated by subtracting the WTP for the product with a carbon-neutral label from the WTP for the unlabeled product. Based on the study, carbon emissions of 1 kg of green coffee equals 4.82 kg. We use a weight conversion rate of 1.176:1 (as given in the study) between green and roasted coffee to calculate the respective emissions of 1 kg of roasted coffee, which is equivalent to 5.67 kg of carbon dioxide. Therefore, we assume 1.42 kg of carbon emissions for 0.25 kg of roasted coffee. Following rule 6 from Section A.1.3, the WTP estimates from Models 2 and 3 are averaged.
Breustedt (2014)	The WTP estimates for reducing carbon emissions by 1 kg are reported in the study (Tables 5 and 7). Following rule 6 from Section A.1.3, WTP estimates from MNL and RPL models are averaged.
Caputo et al. (2018)	The WTP estimates for 20% (1.4 kg) and 30% (2.1 kg) reduction in carbon emissions are reported in the study (Table 4).

Table A.3: WTP derivation strategy (continued)

Study	Details
Carlsson et al. (2022)	The WTP for large and small relative to medium GHG emissions are provided in the study (Table 2). The value for large emissions (relative to medium emissions) was multiplied by -1 to calculate the WTP for medium emissions (relative to large emissions). In the study, GHG emissions of 4 kg are classified as large, levels between 3 kg and 4 kg as medium, and levels less than 3 kg as small emissions. We assume that GHG emission levels are expressed in terms of carbon dioxide equivalents, as this is commonly the practice. A high carbon-footprint lasagna, weighing 0.4 kg, contains approximately 0.08 kg of ground beef, as given in Macdiarmid et al. (2021). Its carbon emissions are equal to 1.88 kg (according to the Plate up for the Planet calculator), which falls within the “small emissions” category defined in this study. In the meta-analysis, as we treat offsets and reductions in the same way, we assume that the carbon emissions of products can be reduced to 0 kg. Therefore, we assume an average emission of 1.5 kg for small levels (averaging 0 kg and 3 kg of carbon emissions), 3.5 kg for medium levels (averaging 3 kg and 4 kg of carbon emissions), and 4 kg for large levels. Respective carbon reduction levels for WTP estimates are calculated by determining the differences between small and medium ($2\text{ kg} = 3.5\text{ kg} - 1.5\text{ kg}$), and between medium and large ($0.5\text{ kg} = 4\text{ kg} - 3.5\text{ kg}$) carbon emissions.

Table A.3: WTP derivation strategy (continued)

Study	Details
Cerroni et al. (2019a,b)	The WTP estimates for low and medium carbon emissions, relative to large emissions, are provided in the studies (Tables 3, 4, D2, E2, 6, and F3 in Cerroni et al., 2019b) (and Tables 4 and 9 in Cerroni et al., 2019a). Carbon emissions are categorized as small for emissions of 0.26 kg or less, medium for emissions between 0.26 kg and 0.4 kg, and large for emissions of more than 0.4 kg per 100 g of lasagna. These are multiplied by 4 for a portion (0.4 kg) of lasagna. A high carbon-footprint lasagna, weighing 0.4 kg, contains approximately 0.08 kg of ground beef, as provided in Macdiarmid et al. (2021). Its carbon emissions amount to 1.88 kg (according to the Plate Up for the Planet calculator). In the meta-analysis, as we treat offsets and reductions in the same way, we assume that the carbon emissions of products can be reduced to 0 kg. Therefore, we assume an average emission of 0.52 kg for low levels (averaging 0 kg and 1.04 kg of carbon emissions), 1.32 kg for medium levels (averaging 1.04 kg and 1.60 kg of carbon emissions), and 1.74 kg for large levels (averaging 1.60 kg and 1.88 kg of carbon emissions). Respective carbon reduction levels for WTP estimates are calculated by determining the differences between low and large (1.22 kg = 1.74 kg - 0.52 kg), and between medium and large (0.42 kg = 1.74 kg - 1.32 kg) carbon emissions. Following rule 6 from Section A.1.3, the WTP estimates obtained from colored, grey, and plain-text labels, from WTP space and preference space estimations from different models, have been averaged.
Chaiyesh (2022)	The WTP estimates for 20% (1.35 kg) and 40% (2.71 kg) carbon reductions are reported in the study (Table 4).
Chen et al. (2018)	The WTP estimates for carbon-labeled products are reported in the study (Tables 3 and 4). We subtract the WTP for a product with 0.10 kg of carbon emissions from that for a product with 0.15 kg, the WTP for a product with 0.15 kg from that for a product with 0.20 kg, and the WTP for a product with 0.10 kg from that for a product with 0.20 kg. This yields two observations for the WTP for a 0.05 kg carbon reduction and one observation for a 0.10 kg carbon reduction.
Drichoutis et al. (2016)	The WTP for carbon-neutral claims are provided by the author through direct correspondence. The carbon emissions for 1 liter of olive oil (2.53 kg of carbon) are sourced from the myEmissions carbon calculator, while the emissions for 0.38 kg of eggs (1.81 kg of carbon), assumed to be equivalent to a six pack, are sourced from the Plate up for the Planet carbon calculator. Following rule 6 from Section A.1.3, we average the WTP observations obtained from both inferred and contingent valuation methods, as well as from dichotomous choice and payment card formats.

Table A.3: WTP derivation strategy (continued)

Study	Details
Gassler (2015)	The WTP estimate for the carbon-neutral label is reported in the study (Section 4.2). The carbon emissions for 0.75 liters of wine (2.9 kg of carbon) are obtained from the Plate up for the Planet carbon calculator.
Grebitus et al. (2013)	The WTP estimates for a 1 kg increase in carbon emissions are derived from the study (Table 3). To obtain the WTP for a 1 kg reduction in carbon emissions, the negative of these estimates is taken. Following rule 6 from Section A.1.3, the WTP estimates from models 1-5 are then averaged.
Grebitus et al. (2015)	The WTP estimate for a 1 kg reduction in carbon emissions is derived from the study (Table 4).
Grebitus et al. (2016)	The WTP estimates for a 1 kg reduction in carbon emissions are reported in the study (Figure 2).
Hassen (2016)	The WTP estimates for percentage carbon reductions are derived from the study (Tables 4 and 6). The carbon reduction attribute has three levels: 25%, 50%, and high (which for simplification we assume to represent a 0% reduction). Because the carbon attribute is discretely coded, the average of 25% and 50% is taken to determine the overall percentage of carbon reduction (37.5%). The carbon emissions of 2.44 kg for the flower (assuming a Dutch rose) is obtained from the not-for-profit organization “Flowers from the Farm.” Therefore, the amount of carbon reduction valued is assumed to be equal to 1.83 kg. Following rule 6 from Section A.1.3, the WTP estimates from the MNL and RPL models have been averaged.
Henderson (2018)	The WTP estimates for low (79 oz $\approx$ 2.23 kg), medium (90 oz $\approx$ 2.55 kg), and high (112 oz $\approx$ 3.18 kg) carbon footprints are derived from the study (Tables 4, 5, and 8). We subtract the WTP for low (2.23 kg) carbon emissions from that for medium (2.55 kg) carbon emissions, the WTP for low (2.23 kg) from that for high (3.18 kg), and the WTP for medium (2.55 kg) from that for high (3.18 kg). This yields observations for the WTP for carbon reductions of 0.32 kg, 0.95 kg, and 0.63 kg. Following rule 6 from Section A.1.3, the WTP estimates from MNL and LC (Latent Class) models are averaged.

Table A.3: WTP derivation strategy (continued)

Study	Details
Kimura et al. (2010)	The WTP estimates for low carbon products (0.06 kg for chocolate, 0.07 kg for chips, 0.065 kg for candy, 0.075 kg for juice), medium carbon products (0.07 kg for chocolate, 0.08 kg for chips, 0.075 kg for candy, 0.085 kg for juice), and high carbon products (0.08 kg for chocolate, 0.09 kg for chips, 0.085 kg for candy, 0.095 kg for juice) are provided by the authors through direct correspondence. We computed the WTP estimates for carbon reductions by subtracting the WTP estimates for low emission products from those of medium and high emission products, as well as the estimates for medium emission products from high emission products. This yields two observations for the WTP for a 0.01-kg carbon reduction and one observation for a 0.02-kg carbon reduction for each product.
Lami et al. (2022)	The WTP estimates for high carbon emissions (28 kg) and medium carbon emissions (18 kg) with respect to low carbon emissions (8 kg) are reported in the study (Table 7). Therefore, the WTP for medium with respect to high carbon emissions corresponds to a 10 kg (28 kg - 18 kg) carbon reduction, and the WTP for low carbon emissions with respect to high carbon emissions corresponds to a 20 kg (28 kg - 8 kg) carbon reduction.
Li et al. (2016)	The WTP estimates for annual beef consumption certified as “raised carbon friendly” are reported in the study (Section 5.2.6). To convert these values to per person and per kg of beef, we divide by 2.8 (average household size based on the study) and by 25.45 kg (annual beef consumption per person sourced from USDA). Since beef production represents 2.2% of total U.S. greenhouse gas emissions and these emissions could be reduced up to 2% if beef production was climate friendly (Li et al., 2016), we assume that the carbon reduction for beef is equivalent to 91% ($2\%/2.2\%$). Carbon emissions for beef are assumed to be 43.33 kg per kg (sourced from myEmissions), yielding a carbon reduction of 39.42 kg per kg of beef.
Li et al. (2018)	The WTP for a climate friendly label is reported in the study (Table 3). We calculate the carbon reduction as the same as Li et al. (2016) except for the fact that the amount of beef valued is 1 pound (≈ 0.45 kg), which leads to 17.74 kg of carbon emissions. Note that we average the WTP estimates for sub-sample groups.

Table A.3: WTP derivation strategy (continued)

Study	Details
Macdiarmid et al. (2021)	The WTP estimates for low-level carbon (green label) and moderate-level carbon (amber label), relative to high-level carbon (red label), are reported in the study (Table 2). Carbon emissions are categorized as low for emissions of 0.26 kg or less, moderate for emissions between 0.26 kg and 0.4 kg, and high for emissions more than 0.4 kg per 100g of lasagna. These are multiplied by 4 for a 0.4 kg lasagna. A high carbon-footprint lasagna, weighing 0.4 kg, contains approximately 0.08 kg of ground beef, as described in the study. Its carbon emissions amount to 1.88 kg (according to the Plate Up for the Planet calculator). In the meta-analysis, as we treat offsets and reductions in the same way, we assume that the carbon emissions of products can be reduced to 0 kg. Therefore, we assume an average emission of 0.52 kg for low levels (averaging 0 kg and 1.04 kg of carbon emissions), 1.32 kg for moderate levels (averaging 1.04 kg and 1.60 kg of carbon emissions), and 1.74 kg for high levels (averaging 1.60 kg and 1.88 kg of carbon emissions). Respective carbon reduction levels for WTP estimates are calculated by determining the differences between low and high ($1.22 \text{ kg} = 1.74 \text{ kg} - 0.52 \text{ kg}$) and between moderate and high ($0.42 \text{ kg} = 1.74 \text{ kg} - 1.32 \text{ kg}$) carbon emissions.
Möerbeck (2022)	The WTP for products labeled as carbon-neutral (group 2), both carbon-footprint and carbon-neutral (group 3), and those without any label (group 4) are reported in the study (Table 2). WTP estimates for carbon reductions are calculated by subtracting the WTP estimates for unlabeled products (group 4) from the WTP estimates for the other groups (2 and 3). Note that we average the WTP values obtained by subtracting group 4 from group 2 and group 4 from group 3 for each product.
Nesselhauf et al. (2020)	The WTP for 30% carbon reduction relative to 0% reduction, 30% carbon reduction relative to 50% carbon reduction, and 50% carbon reduction relative to 0% reduction are reported in the study (Table 7). Therefore, three WTP estimates are derived for 30%, 20%, and 50% carbon reduction, respectively. The corresponding carbon reduction amounts (7.03 kg, 4.69 kg, 11.73 kg) are calculated based on the 23.45 kg of emissions per 0.75 liters of wine, as sourced from the Plate up for the Planet calculator.

Table A.3: WTP derivation strategy (continued)

Study	Details
Nyberg (2018)	The WTP for low carbon emissions and medium carbon emissions, relative to large carbon emissions, are reported in the study (Tables 10, 12, A2, and A4). Carbon emissions are categorized as low for emissions of 7 kg or less, medium for emissions between 7 kg and 11 kg, and large for emissions of more than 11 kg per portion of lasagna (0.4 kg). A high carbon-footprint lasagna, weighing 0.4 kg, contains approximately 0.08 kg of ground beef, as provided in Macdiarmid et al. (2021). Its carbon emissions amount to 1.88 kg (according to the Plate Up for the Planet calculator), which falls within the “low emissions” category defined in this study. In the meta-analysis, as we treat offsets and reductions in the same way, we assume that the carbon emissions of products can be reduced to 0 kg. Therefore, we assume average emissions of 3.5 kg for low levels (averaging 0 kg and 7 kg of carbon emissions), 9 kg for medium levels (averaging 7 kg and 11 kg of carbon emissions), and 11 kg for large levels. Respective carbon reduction levels for WTP estimates are calculated by determining the differences between low and large (7.5 kg = 11 kg - 3.5 kg) and between medium and large (2 kg = 11 kg - 9 kg) carbon emissions. We average WTP estimates from the survey results from colored labels and text-only labels.
Onozaka and McFadden (2011)	The WTP estimates for an increase of 10% in carbon emissions are reported in the study (Table 4). We take the negative of the reported estimates to get WTP for a decrease of 10% in carbon emissions. The carbon emission reductions for apple (0.004 kg) and tomato (0.013 kg) are calculated based on information from the myEmissions calculator.

Table A.3: WTP derivation strategy (continued)

Study	Details
Osawe et al. (2023)	The WTP estimates for moderate and low carbon emissions, relative to high emissions, are derived from the study (Table 5). The WTP estimates from latent classes are multiplied by their class probabilities. Carbon emissions for beef are classified as low for emissions below 20 kg, moderate for emissions between 20 kg and 30 kg, and high for emissions exceeding 30 kg. For chicken, the categories are low for emissions below 5 kg, moderate for emissions between 5 kg and 7.5 kg, and high for emissions above 7.5 kg. For vegetables, emissions below 0.22 kg are considered low, those between 0.22 kg and 0.4 kg as moderate, and those exceeding 0.4 kg as high. In the meta-analysis, as we treat offsets and reductions in the same way, we assume that the carbon emissions of products can be reduced to 0 kg. We have checked the emissions of beef, chicken, and vegetables using carbon calculators. For beef and chicken, the emissions fall below the high emissions category. For vegetables, the carbon emissions amount to 2 kg based on the MyEmissions calculator. Therefore, we assume average carbon emissions of 10 kg for low levels for beef (averaging 0 kg and 20 kg), 25 kg for moderate levels (averaging 20 kg and 30 kg), and 30 kg for high levels. For chicken, we assume average carbon emissions of 2.5 kg for low levels (averaging 0 kg and 5 kg), 6.25 kg for moderate levels (averaging 5 kg and 7.5 kg), and 7.5 kg for high levels. For vegetables, we assume average carbon emissions of 0.11 kg for low levels (averaging 0 kg and 0.22 kg), 0.31 kg for moderate levels (averaging 0.22 kg and 0.40 kg), and 1.22 kg for high levels (averaging 0.4 kg and 2.05 kg). Respective carbon reduction levels for WTP estimates are calculated by determining the differences between low and high, as well as between moderate and high carbon emissions for each product type. For beef, chicken, and vegetables, the amount of carbon reductions are 20 kg, 5 kg, and 11.11 kg for the difference between low and high categories, respectively; and 5 kg, 1.25 kg, and 9 kg for the difference between moderate and high categories.
Owusu-Sekyere et al. (2019)	The WTP estimates for high GHG emissions (27.50 kg) and medium GHG emissions (26.37 kg) with respect to low GHG emissions (22.90 kg), measured in carbon equivalents, are reported in the study (Table 8). Hence, the WTP for medium relative to high GHG emissions corresponds to a 1.13 kg GHG reduction (27.50 - 26.37) in carbon equivalents, and the WTP for low relative to high GHG emissions corresponds to a 4.6 kg GHG reduction (27.50 - 22.90) in carbon equivalents.

Table A.3: WTP derivation strategy (continued)

Study	Details
Rahmani et al. (2019)	The WTP for 10%, 20%, and 30% GHG reduction, expressed in terms of carbon equivalents, are reported in the study (Table 6). The emissions of each type of egg are provided in the study. Therefore, the respective amount of carbon emission reduction calculated for caged eggs are 0.15 kg, 0.30 kg, and 0.44 kg; for barn eggs 0.17 kg, 0.35 kg, 0.52 kg, for free range eggs, 0.17 kg, 0.34 kg, 0.51 kg, and for organic eggs 0.17 kg, 0.34 kg, and 0.51 kg. Note that we average the WTP estimates for four types of eggs.
Severens (2021)	The WTP estimates for low, average emissions with respect to high emissions are reported in the study (Table 4). Carbon emissions levels of 4.3 kg or less are classified as low, levels between 4.4 and 6.6 kg as average, and levels more than 6.6 kg as high. Carbon emissions of 1 kg of pork equals 9.3 kg, which is sourced from the Plate up for Planet calculator. In the meta-analysis, as we treat offsets and reductions in the same way, we assume that the carbon emissions of products can be reduced to 0 kg. We assume an average emission of 2.15 kg for low levels (averaging 0 kg and 4.3 kg), 5.50 kg for average levels (averaging 4.4 kg and 6.6 kg), and 7.95 kg (averaging 9.3 kg and 6.6 kg) for high levels. We subtract high and average (2.45 kg = 7.95 kg - 5.50 kg), and high and low (5.80 kg = 7.95 kg - 2.15 kg) carbon emissions to calculate the respective amount of reductions.
Tu et al. (2021)	The WTP estimates for 34%, 25%, and 17% carbon reduction, relative to a 12% carbon reduction, are reported in the paper (Table 8). Carbon emissions of 1 kg of rice equal to 1.35 kg, which is sourced from the myEmissions calculator. We use this information to calculate respective carbon emission reductions (0.08 kg, 0.18 kg, and 0.30 kg).
Van Loo et al. (2014)	The WTP estimates for 20% (1.4 kg) and 30% (2.1 kg) carbon reduction are reported in the study (Table 7).
Vecchio (2013)	The WTP for the carbon-neutral product is reported in the study (Figure 2). The WTP for carbon neutrality is calculated by subtracting the WTP for conventional product from the WTP for carbon-neutral product. The carbon emissions of 0.75 liters of wine (1.03 kg) is obtained from the myEmissions calculator.
Yang et al. (2021)	WTP for a 38% carbon reduction is derived from the study (Table 4). The amount of carbon emissions of 1 kg of rice (0.68 kg) is obtained from the myEmissions calculator. We use this information to calculate the respective carbon emission reduction (0.26 kg).

Table A.3: WTP derivation strategy (continued)

Third, we detail the independent variables included in the regression analysis.

The first variable is the “amount of CO₂ reduction” in kilograms. In cases where the study does not specify the baseline CO₂ emissions of products, which is necessary to calculate the corresponding amount of CO₂ emission reduction, as it is sometimes the case for food and drink products, we, as mentioned, use online food/drink carbon calculators, specifically, My Emissions and Plate Up for the Planet. A few instances also involve non-food products, specifically flowers and plastic bottles. For flowers, we refer to Flowers from the Farm, an association supporting cut flower growers in the United Kingdom, and for plastic bottles, Tappwater, a company that makes water-filtration products. Note that if the study does not specify the amount of CO₂ emission reduction, we have a control (dummy) variable for such observations, which is described later in this section.

The “product price” variable is measured in 2020 USD. If a study does not specify a product’s price, we use the WTP for the unlabeled product, as reported in the study. If this information is also unavailable and if the study confirms that these levels are aligned with observed market prices, we use the average of the price levels specified in the study as a proxy for price. We have three cases for which price information is unavailable: rice in Thailand, apple juice in Germany, and beef in the United States. We obtained rice price data from GlobalProductPrices.com for Chaiyesh (2022), apple juice prices from Selina Wamucii, an agricultural produce marketplace, for Breustedt (2014), and beef prices from the United States Department of Agriculture for Li et al. (2016).

“Stated preference method” is a dummy variable that takes on a value of 1 for observations derived from CV and DCE methods, and 0 for those obtained from revealed preference methods.

“In-person” is another dummy variable, taking value 1 for studies conducted face-to-face and 0 for online or computer-assisted surveys.

“Sample size” is a variable indicating the number of participants, which is generally available in all studies. However, there are two exceptions involving sub-samples. Van Loo et al. (2014) does not specify the sample sizes for income

clusters. In this case, we assume an even distribution between high and low-income groups. In Kimura et al. (2010), the sample size varies between 18, 19, 20, and 21 for different treatment groups. A fixed sample size of 19 is assumed for all WTP observations to facilitate aggregation over fixed covariates (including sample size).

The dummy variable “publication” takes the value of 1 for published studies, and 0 for working papers, conference proceedings, or theses.

The “study year” variable refers to the year in which a study was conducted. For studies that span two consecutive years, we use the first year, while for those covering three years, the middle year is used. If a study is a conference paper and does not specify the year, as in the case of Gassler (2015), we refer to the year in which the respective conference took place.

“GDP per capita” refers to the per capita Gross Domestic Product of the country where the study was conducted, measured in 2020 USD.

“Europe” is a dummy variable that takes the value of 1 for studies conducted in Europe and 0 for those conducted in Africa, the Americas, or Asia.

“CO₂ reduction assumptions” is a dummy variable that takes the value of 1 if a study lacks specific information on the amount of carbon reduction, requiring us to make assumptions, as described in detail for each study in Table A.3.

“WTP derivation” variable takes the value of 1 if we had to derive the WTP estimates ourselves, and 0 if these are directly reported in the original study. Note that in some cases, this variable can take the values of 0 and 1 for different observations originating from the same study.

We also include two additional variables in the robustness tests: “colored labels” and “carbon-neutral labels.” “Colored label” variable takes the value of 1 if colors are used to differentiate the CO₂ intensity of the valued product. “Carbon-neutral label” variable takes the value of 1 for carbon-neutral certifications and 0 for carbon reduction and carbon footprint certifications.

Finally, the “product CO₂ emissions” variable, which is measured in kilograms, indicates the CO₂ emissions produced during the production or life-cycle of the product. This variable is not included in the meta-regressions but used

for calculating WTP_{CN} and $WTP_{CN\%}$ and plotting their distributions.

A.2 Descriptive statistics and supplementary material

This section includes the main descriptive statistics for the sample used for the meta-analysis and supplementary material. Table A.4 shows the summary statistics of the (unweighted) sample of 126 observations, which includes one or more observations from each study. Table A.5 shows the mean of study-specific mean WTP estimates for each product category.

Figures A.1 show WTP_R estimates by product category. Figures A.2 shows the relationship of WTP_R with CO_2 emission reduction levels, while Figure A.3 shows residuals versus fitted values from the main model (OLS III) shown in Table 5 in Section 3.2. The results of the Breusch-Pagan test for heteroskedasticity are reported in Table A.6.

	N	Mean	Std. Dev.	Min	Max
WTP _R (USD)	126	1.18	1.59	-0.09	9.06
WTP _{kg} (USD)	126	4.34	9.26	-1.38	45.28
WTP _{CN} (USD)	126	10.66	38.84	-0.13	311.56
WTP _{CN%} (%)	126	168.72	360.28	-10.94	1874.74
CO ₂ reduction (kg)	126	2.56	5.88	0.00	39.43
Product CO ₂ emissions (kg)	126	5.34	9.63	0.02	43.33
Carbon neutral label	126	0.13	0.33	0.00	1.00
Colored label	126	0.14	0.34	0.00	1.00
Price (USD)	126	4.28	4.92	0.09	22.15
Stated pref. method	126	0.95	0.21	0.00	1.00
In-person	126	0.21	0.41	0.00	1.00
Sample size	126	549	609	19	3085
Publication	126	0.81	0.39	0.00	1.00
Study year	126	2015	4.22	2008	2021
GDP per capita (100 USD)	126	427.37	166.69	5.84	935.47
Europe	126	0.68	0.47	0.00	1.00
CO ₂ reduction assump.	126	0.56	0.50	0.00	1.00
WTP derivation	126	0.42	0.50	0.00	1.00

Table A.4: Summary statistics: unweighted sample

This table displays the number of observations (N), and the summary statistics of the variables. Standard deviations are provided in parentheses. The WTP_R denotes (un-standardized) WTP for CO₂ reductions. WTP_{kg} is the WTP for 1 kg CO₂ reduction, WTP_{CN} is the WTP for carbon-neutrality, and WTP_{CN%} is the proportion of a product's price that consumers would be willing to pay extra for carbon neutrality. For detailed variable descriptions, see Section A.1.3.

Product Category	N	CO ₂ (kg)	WTP _R (USD)	WTP _{kg} (USD)	WTP _{CN} (USD)	WTP _{CN%} (%)
Dairy & eggs	6	1.39 (0.46)	0.71 (0.63)	1.07 (1.34)	0.93 (0.75)	51.73 (30.86)
Fruits & vegetables	7	2.77 (2.98)	0.43 (0.49)	1.33 (1.76)	2.24 (3.24)	200.75 (297.28)
Meat	17	15.18 (15.04)	2.09 (1.97)	1.27 (1.60)	31.10 (70.26)	359.79 (549.37)
Non-food	3	1.34 (0.99)	0.77 (0.37)	1.34 (1.60)	0.88 (0.53)	543.45 (893.59)
Oil & grain	6	2.10 (2.41)	0.57 (0.49)	1.61 (2.11)	2.14 (3.77)	60.62 (97.33)
Snacks	1	0.07 (0.00)	0.33 (0.00)	21.94 (0.00)	0.02 (0.00)	2.03 (0.00)
Water & drinks	7	0.92 (0.77)	1.13 (0.84)	5.33 (10.51)	1.55 (2.07)	40.11 (34.74)

Table A.5: Means of study means: WTP estimates by product category

This table displays the product categories, their respective number of studies (N), and the means of study specific means of the outcome variables. Standard deviations are provided in parentheses. The third column presents the CO₂ emissions associated with the products, which vary according to type and amount of product valued in studies. The non-standardized measure WTP_R denotes non-standardized WTP for carbon reductions. WTP_{kg} is WTP per 1 kg carbon reduction, WTP_{CN} is WTP for carbon-neutrality, and WTP_{CN%} is the proportion of a product's price that consumers would be willing to pay extra for carbon neutrality. For detailed variable descriptions, see Section A.1.3

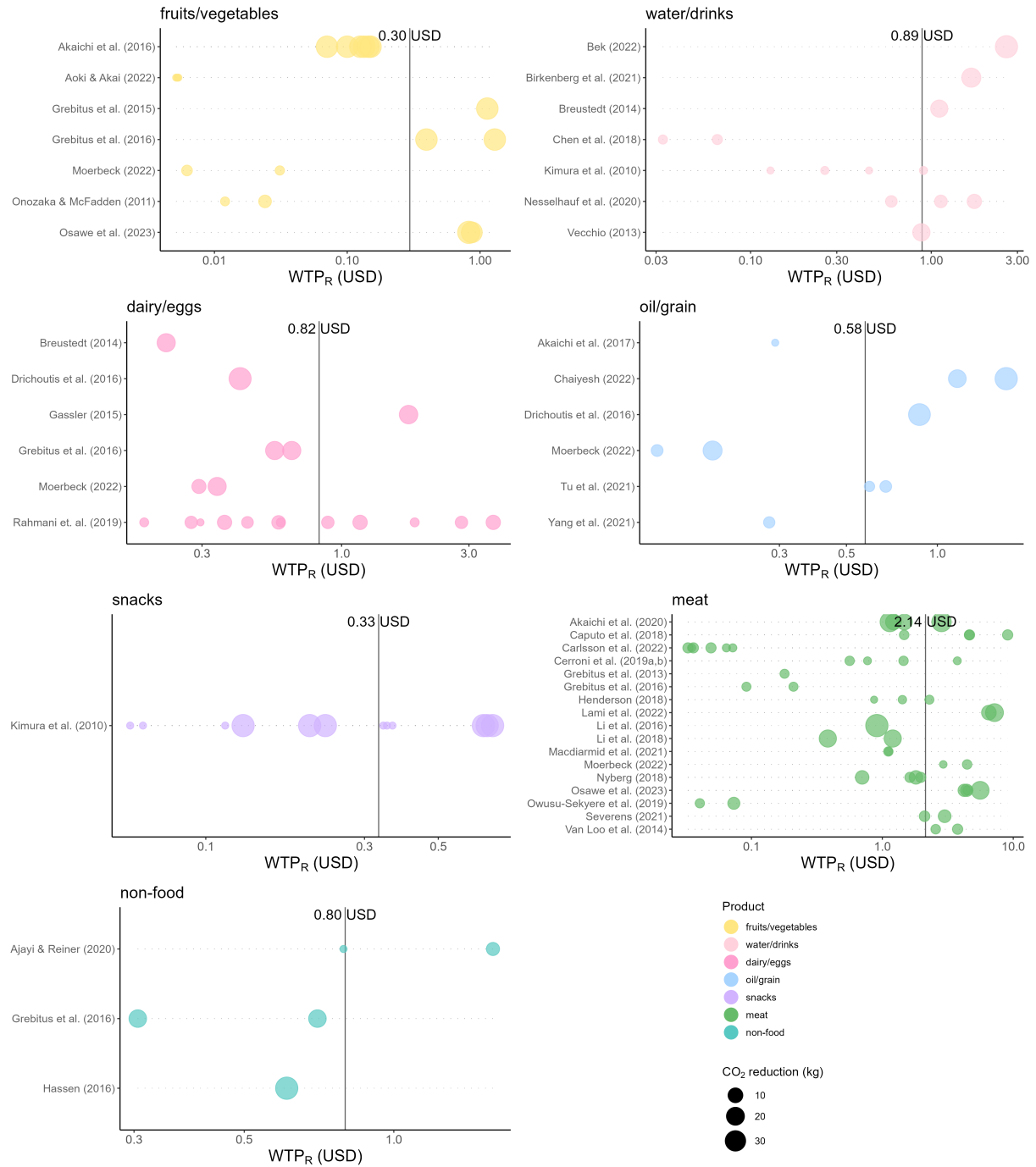


Figure A.1: WTP for carbon reductions (WTP_R) by product category

A logarithmic axis (base 10) is used. Vertical lines show the mean of the study means. Panels by product category display the non-standardized measure; WTP_R denotes the non-standardized willingness to pay for carbon reductions. Circle size represents the amount of CO₂ emission reduction (kg).

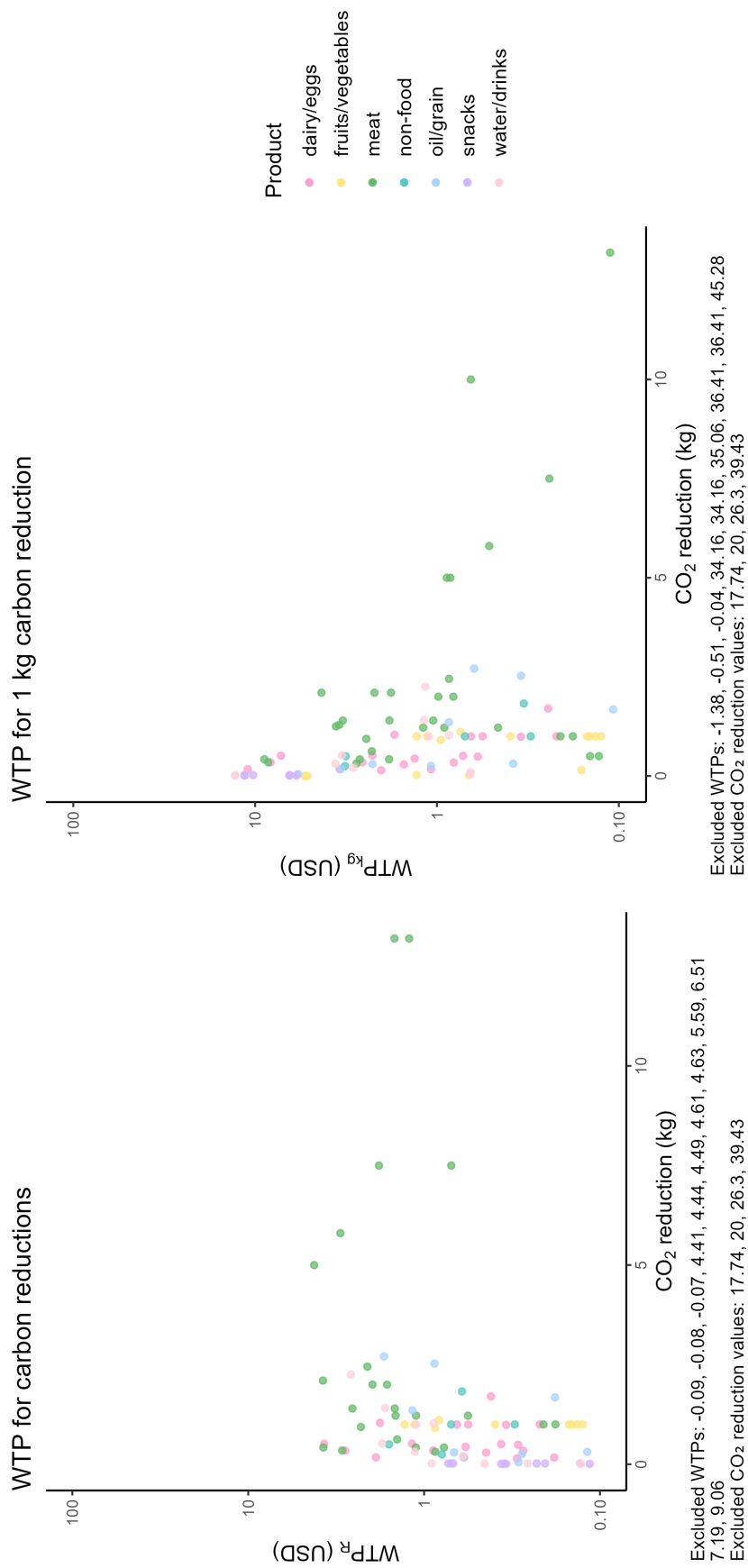


Figure A.2: WTP_R and WTP_{kg} by the amount of CO_2 emission reduction

A logarithmic y-axis (base 10) is used to create this figure. The left graph displays WTP_R (non-standardized WTP for CO_2 reductions). The right-hand graph shows WTP_{kg} (WTP for 1 kg CO_2 reduction), which is calculated by dividing WTP_R by the amount of CO_2 reduction. Both WTP_R and WTP_{kg} are measured in 2020 USD. Outliers, defined as values exceeding two standard deviations from the mean, and negative values are excluded from the figure.

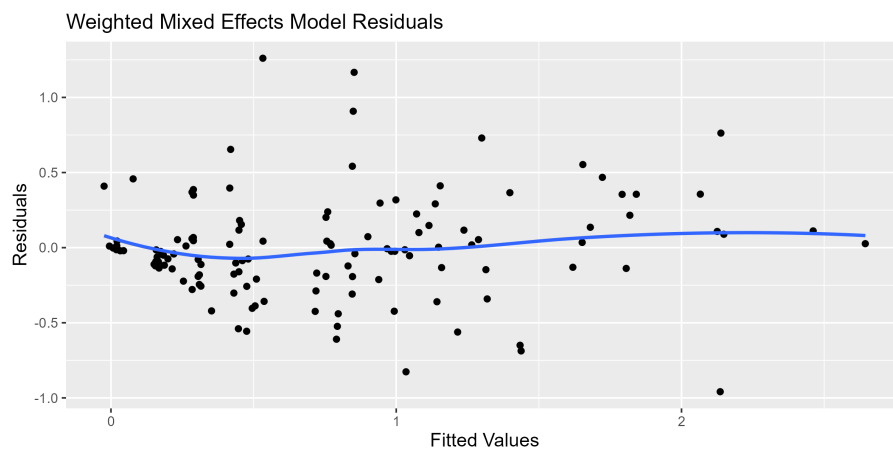
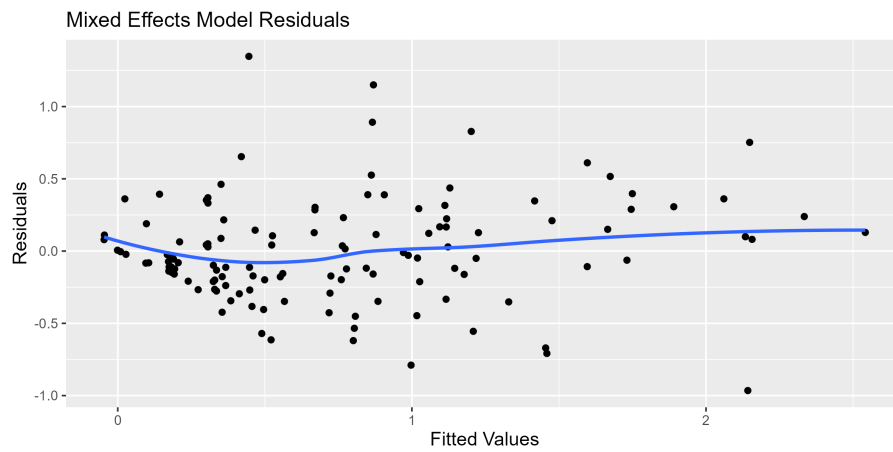
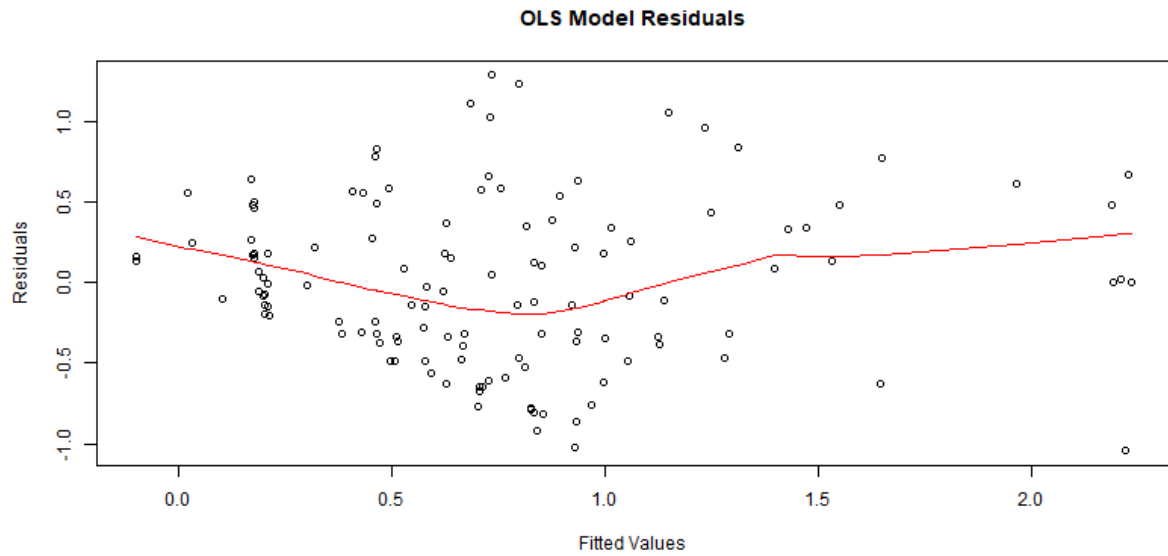


Figure A.3: Residuals versus fitted values

The top, middle, and bottom panels display residuals versus fitted values from the OLS model (Table 5, Section 3.2, column 3), the mixed-effects model (Table A.7, Section A.3, column 2), and the weighted mixed-effects model (Table A.7, column 3), respectively.

Model	Breusch-Pagan Stat.	p-value
OLS I	5.02	0.08
OLS II	22.69	0.01***
OLS III	21.91	0.03**

*p<0.05 **p<0.01 ***p<0.001

Table A.6: Breusch-Pagan test for OLS models presented in Table 5

The table displays the Breusch–Pagan heteroskedasticity test from the OLS models in Table 5, Section 3.2.

A.3 Robustness tests

This section presents and discusses the robustness tests conducted to test the sensitivity of our main meta-analytical results.

First, we run regressions with different models. The first column of Table A.7 provides baseline OLS III model estimations with standard errors clustered by study. The second column shows mixed-effects model estimations with random effects for studies and product categories. The third column includes the weighted mixed-effects model where weights are based on the inverse number of estimates obtained from each study to equally weigh studies in the meta-analysis.

Second, Tables A.8 and A.9 include two additional variables, carbon-neutral certification and colored labels, analyzed using OLS and mixed-effects models, compared to the baseline regression results presented in Table 5 in Section 3.2. In the first column, we include only the CO₂ reduction, price, colored label, and carbon-neutral label variables. In the second column, we add variables for stated preference studies, in-person studies, sample size, published studies, study year, GDP per capita, and studies conducted in Europe. In the last column, we further add controls for studies requiring making assumptions about the amount of CO₂ reduction, and for which WTP estimates had to be derived.

Third, we report WLS and weighted mixed-effects model estimations in Tables A.10 and A.11, respectively. For comparison purposes, we use the baseline unweighted models in the first column. In the second column, we apply weights based on the inverse number of estimates from each study to equalize each study’s contribution. In the third column, we use the sample size of the

study as weights.

Fourth, we report OLS and mixed-effects regression results using different transformations of the dependent variable in Tables A.12 and A.13, respectively. For comparison purposes, the first column shows baseline estimations with the untransformed dependent variable. The second column displays the results based on the inverse hyperbolic sine transformation, and the last column uses the logarithmic transformation.

Fifth, we show the OLS model results with two-way clustered errors in Table A.14. For comparison, the first column includes the baseline OLS model with clustered standard errors for studies. The second column presents the OLS model with two-way clustered standard errors for studies and product categories. The third column includes two-way clustered errors for studies and countries.

Sixth, we run the mixed-effects regressions with alternative random effects as shown in Table A.15. For comparison, the first column includes the baseline mixed-effects model with random effects for studies only. The second column incorporates random effects for both studies and product categories. The third column includes random effects for both studies and study countries.

Seventh, we include the square of the z-scored CO₂ emission reduction variable in OLS and mixed-effects regressions. The results are shown in Tables A.16 and A.17, respectively. The first column includes the CO₂ reduction, CO₂ reduction squared, and product price variables. The second column includes additional variables capturing study characteristics and contextual factors. In the last column, we add control variables for CO₂ reduction assumptions and WTP derivations.

Eighth, we present the results from OLS and mixed-effects regressions with different subsets of observations in Tables A.18 and A.19 respectively. The first column is based on the complete set of observations. The second column displays results without observations that require assumptions about the amount of CO₂ reductions. The third column includes only observations where WTP values are sourced directly from the studies, omitting those requiring further calculations or derivations. The final column shows model results that excludes

both types of observations: those requiring CO₂ reduction assumptions and those with derived WTP estimates.

Across the robustness checks in Tables A.7-A.19, the coefficient on CO₂ reduction is positive and statistically significant in most specifications, including models with alternative transformations, clustering choices, and functional forms. Including indicators for observations where WTP_R is derived by us or where assumptions about CO₂ reductions are required does not change the results (Table A.7). However, in three restrictive subsample columns that exclude the former or both of such observations, the CO₂ reduction coefficient loses statistical significance (the last columns of Tables A.18 and A.19, and the third column of Table A.19). These cases rely on much smaller samples (e.g., 28 observations, 8 studies) and therefore have larger standard errors; importantly, the estimates remain positive and in most cases large. Overall, the evidence points to a robust positive association between WTP_R and CO₂ reduction; the few non-significant cases seem mostly consistent with lower precision rather than a change in the underlying relationship.

The product price is positive and significant across all robustness checks, always at the 1% level.

Even while controlling for the GDP per capita, studies conducted in Europe are robustly positively significant, mainly at the 1% level.

We do not find significant results for colored labels, carbon-neutral labels, WTP derivation, or CO₂ reduction assumptions.

The remaining variables are mostly insignificant, or become significant in only a few of the regressions, as described in what follows. Sample size becomes significant with a negative sign in the second column of Table A.18. The coefficient for stated preference studies becomes significant at the 5% level with a positive sign in the third column of Table A.7, the third column of Table A.9, and the second column of Table A.11, and at the 1% level with a negative sign in the second column of Table A.18. The coefficient for published studies becomes negative and significant in the last column of Table A.10, the second, third, and fourth columns of Table A.18, and the second and fourth columns of Table A.19

at significance levels ranging from 1% to 10%. The coefficient for the dummy variable indicating in-person studies is significant and negative in the second and fourth columns of Tables A.18 and A.19. GDP per capita becomes positively significant in the last column of Table A.18 at the 1% level. The study year becomes positively significant in the second and fourth columns of Table A.19 and in the fourth column of Table A.18.

	OLS	Mixed Eff.	Weighted Mixed Eff.
Intercept	0.58* (0.32)	0.20 (0.39)	0.11 (0.36)
CO ₂ reduction	0.11*** (0.04)	0.10** (0.05)	0.09** (0.04)
Price	0.35*** (0.05)	0.33*** (0.07)	0.30*** (0.07)
Stated pref. method	-0.02 (0.25)	0.35 (0.28)	0.48** (0.23)
In-person	-0.10 (0.23)	0.01 (0.21)	0.02 (0.18)
Sample size	-0.06 (0.08)	-0.06 (0.09)	-0.08 (0.09)
Publication	0.00 (0.22)	-0.13 (0.21)	-0.15 (0.21)
Study year	0.02 (0.10)	0.04 (0.10)	0.04 (0.09)
GDP per capita	0.07 (0.08)	0.02 (0.07)	0.01 (0.07)
Europe	0.33* (0.18)	0.44*** (0.17)	0.48*** (0.17)
CO ₂ reduction assump.	0.07 (0.15)	0.09 (0.19)	0.06 (0.18)
WTP derivation	-0.09 (0.12)	0.02 (0.15)	0.03 (0.15)
Number of obs.	126	126	126
Var (study random effect)		0.15	0.18
Var (product random eff.)		0.01	0.01
Adjusted-R ²	0.45	0.45	0.45
AIC	212.91	230.16	230.80
BIC	249.78	272.70	273.34
Log Likelihood	-93.45	-100.08	-100.40

***p<0.01; **p<0.05; *p<0.1

Table A.7: Factors associated with WTP for CO₂ reductions: robustness tests with alternative models

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. In the first column, we use an OLS model with clustered standard errors by study. In the second and third columns, we use mixed-effects models, including studies and product categories as random effects. In the third column, we use weights, which correspond to the inverse of the number of WTP_R estimates obtained from each study. The dependent variable is the unstandardized WTP for carbon emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

	OLS I	OLS II	OLS III
Intercept	0.70*** (0.09)	0.51* (0.32)	0.56** (0.28)
CO ₂ reduction	0.12** (0.06)	0.11*** (0.04)	0.08*** (0.03)
Price	0.38*** (0.05)	0.35*** (0.05)	0.33*** (0.07)
Colored label	0.16 (0.19)	0.03 (0.20)	0.08 (0.20)
Carbon neutral label	0.13 (0.16)	0.19 (0.22)	0.11 (0.25)
Stated pref. method		-0.01 (0.24)	-0.00 (0.21)
In-person		-0.05 (0.24)	-0.07 (0.25)
Sample size		-0.08 (0.08)	-0.08 (0.09)
Publication		-0.03 (0.23)	-0.17 (0.20)
Study year		0.01 (0.10)	0.06 (0.09)
GDP per capita		0.08 (0.08)	0.01 (0.07)
Europe		0.37*** (0.16)	0.54*** (0.18)
CO ₂ reduction assump.			0.03 (0.17)
WTP derivation			0.01 (0.15)
Number of obs.	126	126	126
Adjusted-R ²	0.40	0.46	0.53
AIC	218.84	211.85	233.36
BIC	235.86	248.73	275.90
Log Likelihood	-103.42	-92.93	-101.68

***p<0.01; **p<0.05; *p<0.1

Table A.8: Factors associated with WTP for CO₂ reductions: OLS model, robustness tests with additional variables

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use an OLS model with clustered standard errors by study. The dependent variable is the (unstandardized) WTP for CO₂ emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

	Mixed Eff. I	Mixed Eff. II	Mixed Eff. III
Intercept	0.74*** (0.11)	0.23 (0.35)	0.09 (0.36)
CO ₂ reduction	0.09* (0.05)	0.10* (0.05)	0.09* (0.04)
Price	0.35*** (0.06)	0.32*** (0.06)	0.29*** (0.07)
Colored label	0.08 (0.16)	0.04 (0.16)	0.11 (0.17)
Carbon neutral label	0.01 (0.19)	0.12 (0.21)	0.09 (0.27)
Stated pref. method		0.34 (0.27)	0.50** (0.24)
In-person		0.02 (0.21)	0.03 (0.19)
Sample size		-0.08 (0.10)	-0.09 (0.10)
Publication		-0.14 (0.21)	-0.15 (0.22)
Study year		0.05 (0.09)	0.04 (0.10)
GDP per capita		0.03 (0.07)	0.03 (0.08)
Europe		0.45*** (0.16)	0.50*** (0.17)
CO ₂ reduction assump.			-0.00 (0.24)
WTP derivation			0.03 (0.15)
Number of obs.	126	126	126
Var (study random effect)	0.16	0.14	0.19
Var (product random eff.)	0.01	0.01	0.03
AIC	216.88	229.64	236.84
BIC	239.57	272.19	285.05
Log Likelihood	-100.44	-99.82	-101.42

***p<0.01; **p<0.05; *p<0.1

Table A.9: Factors associated with WTP for CO₂ reductions: mixed effects model, robustness tests with additional variables

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use mixed-effects models, including studies and product categories as random effects. The dependent variable is the unstandardized WTP for carbon emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

	OLS	WLS I (w = inv. number obs.)	WLS II (w = sample size)
Intercept	0.58* (0.32)	0.57** (0.28)	1.21*** (0.29)
CO ₂ reduction	0.11*** (0.04)	0.09*** (0.03)	0.10*** (0.02)
Price	0.35*** (0.05)	0.33*** (0.07)	0.29*** (0.08)
Stated pref. method	-0.02 (0.25)	-0.00 (0.21)	-0.27 (0.22)
In-person	-0.10 (0.23)	-0.09 (0.26)	-0.33 (0.25)
Sample size	-0.06 (0.08)	-0.07 (0.08)	-0.06 (0.04)
Publication	0.00 (0.22)	-0.16 (0.19)	-0.44** (0.20)
Study year	0.02 (0.10)	0.06 (0.09)	0.11 (0.12)
GDP per capita	0.07 (0.08)	-0.00 (0.07)	-0.00 (0.12)
Europe	0.33* (0.18)	0.52*** (0.17)	0.46*** (0.11)
CO ₂ reduction assump.	0.07 (0.15)	0.10 (0.13)	-0.13 (0.10)
WTP derivation	-0.09 (0.12)	0.01 (0.14)	-0.02 (0.09)
Number of obs.	126	126	126
Adjusted R ²	0.50	0.57	0.41
AIC	212.91	230.12	306.73
BIC	249.78	266.99	343.61
Log Likelihood	0.45	0.53	0.36

***p<0.01; **p<0.05; *p<0.1

Table A.10: Factors associated with WTP for CO₂ reductions: OLS and WLS models, robustness tests with weights

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use an OLS and WLS models with clustered standard errors by study. This table shows coefficient estimates, and associated standard errors, which are indicated within parentheses. The standard errors are clustered across studies. In the first column, we do not weigh the outcome variable. In the second column, we weigh based on the inverse of the number of estimates derived or obtained from each study. In the last column, we use the sample size as the weight. The dependent variable is the (unstandardized) WTP for CO₂ emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

	Mixed Eff.	Weighted Mixed Eff. I (w = inv. number obs.)	Weighted Mixed Eff. II (w = sample size)
Intercept	0.20 (0.39)	0.11 (0.36)	0.97 (0.63)
CO ₂ reduction	0.10** (0.05)	0.09** (0.04)	0.11*** (0.04)
Price	0.33*** (0.07)	0.30*** (0.07)	0.28*** (0.08)
Stated pref. method	0.35 (0.28)	0.48** (0.23)	-0.10 (0.55)
In-person	0.01 (0.21)	0.02 (0.18)	-0.26 (0.27)
Sample size	-0.06 (0.09)	-0.08 (0.09)	-0.08 (0.09)
Publication	-0.13 (0.21)	-0.15 (0.21)	-0.29 (0.22)
Study year	0.04 (0.10)	0.04 (0.09)	0.13 (0.10)
GDP per capita	0.02 (0.07)	0.01 (0.07)	0.03 (0.07)
Europe	0.44*** (0.17)	0.48*** (0.17)	0.44*** (0.13)
CO ₂ reduction assumpt.	0.09 (0.19)	0.06 (0.18)	-0.11 (0.20)
WTP derivation	0.02 (0.15)	0.03 (0.15)	-0.01 (0.15)
Number of obs.	126	126	126
Var (study random effect)	0.15	0.18	0.10
Var (country random effect)	0.01	0.02	0.01
AIC	230.16	230.80	324.34
BIC	272.70	273.34	366.88
Log Likelihood	-100.08	-100.40	-147.17

*** p<0.01; ** p<0.05; * p<0.1

Table A.11: Factors associated with WTP for CO₂ reductions: mixed effects and weighted mixed effects models, robustness tests with weights

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use mixed-effects models, including studies and product categories as random effects. In the first column, we do not weigh the outcome variable. In the second column, we weigh based on the inverse of the number of estimates derived or obtained from each study. In the last column, we use the sample size as the weight. The dependent variable is the (unstandardized) WTP for CO₂ emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

	OLS I (not transformed)	OLS II (inv. hyperbolic sine trans.)	OLS III (log. trans.)
Intercept	0.78 (0.64)	0.58* (0.32)	-1.50* (0.89)
CO ₂ reduction	0.20* (0.10)	0.11*** (0.04)	0.29*** (0.09)
Price	0.82*** (0.11)	0.35*** (0.05)	0.64*** (0.16)
Stated pref. method	-0.05 (0.54)	-0.02 (0.25)	-0.18 (0.47)
In-person	-0.31 (0.45)	-0.10 (0.23)	-0.36 (0.53)
Sample size	-0.26 (0.19)	-0.06 (0.08)	0.01 (0.19)
Publication	0.31 (0.40)	0.00 (0.22)	-0.30 (0.64)
Study year	0.21 (0.21)	0.02 (0.10)	-0.28 (0.32)
GDP per capita	0.14 (0.16)	0.07 (0.08)	-0.04 (0.24)
Europe	0.71** (0.34)	0.33* (0.18)	1.19** (0.57)
CO ₂ reduction assumpt.	-0.27 (0.35)	0.07 (0.15)	0.58 (0.51)
WTP derivation	-0.14 (0.20)	-0.09 (0.12)	0.20 (0.49)
Number of obs.	126	126	123
Adjusted R ²	0.55	0.50	0.38
AIC	399.33	212.91	437.14
BIC	436.20	249.78	473.70
Log Likelihood	-186.67	-93.45	-205.57

***p<0.01; **p<0.05; *p<0.1

Table A.12: Factors associated with WTP for CO₂ reductions: OLS model, robustness tests with transformations of the dependent variable

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use an OLS model with clustered standard errors by study. For the first column, we do not transform the outcome variable. In the second column, we transform it using the inverse hyperbolic sine function. In the third column, we use logarithmic transformation, resulting in the loss of three negative observations. The dependent variable is the (unstandardized) WTP for CO₂ emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

	Mixed Eff. I (not transformed)	Mixed Eff. II (inv. hyperbolic sine trans.)	Mixed Eff. III (log. trans.)
Intercept	0.33 (0.81)	0.20 (0.39)	-2.19** (0.92)
CO ₂ reduction	0.25** (0.10)	0.10** (0.05)	0.19* (0.11)
Price	0.82*** (0.13)	0.33*** (0.07)	0.58*** (0.16)
Stated pref. method	0.44 (0.59)	0.35 (0.28)	0.69 (0.61)
In-person	-0.09 (0.44)	0.01 (0.21)	0.10 (0.50)
Sample size	-0.22 (0.19)	-0.06 (0.09)	0.02 (0.24)
Publication	0.05 (0.42)	-0.13 (0.21)	-0.43 (0.56)
Study year	0.23 (0.20)	0.04 (0.10)	-0.07 (0.25)
GDP per capita	0.05 (0.15)	0.02 (0.07)	0.01 (0.18)
Europe	0.85** (0.35)	0.44*** (0.17)	1.14*** (0.38)
CO ₂ reduction assumpt.	-0.24 (0.38)	0.09 (0.19)	0.51 (0.48)
WTP derivation	0.01 (0.32)	0.02 (0.15)	0.23 (0.34)
Number of obs.	126	126	123
Var (study random effect)	0.50	0.15	1.18
Var (product random effect)	0.00	0.01	0.19
AIC	411.77	230.16	394.93
BIC	454.32	272.70	437.12
Log Likelihood	-190.89	-100.08	-182.47

***p<0.01; **p<0.05; *p<0.1

Table A.13: Factors associated with WTP for CO₂ reductions: mixed effects model, robustness tests with transformations of the dependent variable

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use mixed-effects models, including studies and product categories as random effects. For the first column, we do not transform the outcome variable. In the second column, we transform it using the inverse hyperbolic sine function. In the third column, we use logarithmic transformation, resulting in the loss of three negative observations. The dependent variable is the (unstandardized) WTP for CO₂ emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

	OLS I (clst. by study)	OLS II (clst. by study & product)	OLS III (clst. by study & country)
Intercept	0.58* (0.32)	0.56*** (0.07)	0.58* (0.30)
CO ₂ reduction	0.11*** (0.04)	0.11*** (0.02)	0.11** (0.05)
Price	0.35*** (0.05)	0.35*** (0.04)	0.35*** (0.04)
Stated pref. method	-0.02 (0.25)	-0.03 (0.13)	-0.02 (0.23)
In-person	-0.10 (0.23)	-0.11 (0.14)	-0.10 (0.26)
Sample size	-0.06 (0.08)	-0.06 (0.15)	-0.06 (0.09)
Publication	0.00 (0.22)	-0.01 (0.16)	0.00 (0.25)
Study year	0.02 (0.10)	0.02 (0.04)	0.02 (0.12)
GDP per capita	0.07 (0.08)	0.06 (0.06)	0.07 (0.08)
Europe	0.33* (0.18)	0.35* (0.21)	0.33* (0.18)
CO ₂ reduction assump.	0.07 (0.15)	0.09 (0.08)	0.07 (0.21)
WTP derivation	-0.09 (0.12)	-0.08 (0.09)	-0.09 (0.11)
Number of obs.	126	126	126
Adjusted R ²	0.45	0.45	0.45
AIC	212.91	212.91	212.91
BIC	249.78	249.78	249.78
Log Likelihood	-93.45	-93.45	-93.45

***p<0.01; **p<0.05; *p<0.1

Table A.14: Factors associated with WTP for CO₂ reductions: OLS model, robustness tests with alternative cluster variables

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. In the first column, we use an OLS model with standard errors clustered by studies. In the second column, the standard errors of the OLS model are clustered by both study and product category. In the second column, the study year variable is not included due to the insufficient variation within the study and product clusters. In the third column, the standard errors are clustered by study and country. The dependent variable is the (unstandardized) WTP for CO₂ emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

	Mixed Eff. I (study r.e.)	Mixed Eff. II (study & product r.e.)	Mixed Eff. III (study & country r.e.)
Intercept	0.24 (0.38)	0.20 (0.39)	0.24 (0.38)
CO ₂ reduction	0.11** (0.05)	0.10** (0.05)	0.11** (0.05)
Price	0.36*** (0.06)	0.33*** (0.07)	0.36*** (0.06)
Stated pref. method	0.31 (0.28)	0.35 (0.28)	0.31 (0.28)
In-person	0.02 (0.21)	0.01 (0.21)	0.02 (0.21)
Sample size	-0.06 (0.09)	-0.06 (0.09)	-0.06 (0.09)
Publication	-0.14 (0.21)	-0.13 (0.21)	-0.14 (0.21)
Study year	0.04 (0.10)	0.04 (0.10)	0.04 (0.10)
GDP per capita	0.02 (0.07)	0.02 (0.07)	0.02 (0.07)
Europe	0.44*** (0.17)	0.44*** (0.17)	0.44*** (0.17)
CO ₂ reduction assump.	0.10 (0.19)	0.09 (0.19)	0.10 (0.19)
WTP derivation	0.03 (0.15)	0.02 (0.15)	0.03 (0.15)
Number of obs.	126	126	126
Var (study random effect)	0.14	0.15	0.14
Var (product random eff.)		0.01	
Var (country random effect)			0.00
AIC	228.40	230.16	230.40
BIC	268.11	272.70	272.95
Log Likelihood	-100.20	-100.08	-100.20

***p<0.01; **p<0.05; *p<0.1

Table A.15: Factors associated with WTP for CO₂ reductions: mixed effects model, robustness tests with alternative random effects

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use mixed-effects models. For the first column, we include studies as random effects. In the second column, we include both studies and product categories as random effects. In the third column, we include studies and countries as random effects. The dependent variable is the (unstandardized) WTP for CO₂ emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

	OLS I	OLS II	OLS III
Intercept	0.80*** (0.09)	0.57* (0.32)	0.55* (0.31)
CO ₂ reduction	0.33** (0.16)	0.25** (0.12)	0.26* (0.15)
CO ₂ reduction ²	-0.04* (0.02)	-0.03 (0.02)	-0.03 (0.02)
Price	0.35*** (0.05)	0.33*** (0.05)	0.33*** (0.05)
Stated pref. method		-0.02 (0.24)	-0.02 (0.24)
In-person		-0.10 (0.24)	-0.10 (0.23)
Sample size		-0.05 (0.08)	-0.05 (0.08)
Publication		0.01 (0.22)	0.00 (0.22)
Study year		0.03 (0.09)	-0.01 (0.10)
GDP per capita		0.06 (0.08)	0.06 (0.08)
Europe		0.35* (0.18)	0.32* (0.19)
CO ₂ reduction assump.			0.11 (0.18)
WTP derivation			-0.03 (0.12)
Number of obs.	126	126	126
Adjusted R ²	0.42	0.46	0.46
AIC	213.68	209.45	212.45
BIC	227.86	243.49	252.16
Log Likelihood	-101.84	-92.73	-92.22

***p<0.01; **p<0.05; *p<0.1

Table A.16: Factors associated with WTP for CO₂ reductions: OLS model, robustness tests with alternative functional form of CO₂ reduction

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use an OLS model with clustered standard errors by study. The dependent variable is the (unstandardized) WTP for CO₂ emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored. CO₂ reduction² variable is obtained by squaring the z-scored CO₂ reduction variable.

	Mixed Eff. I	Mixed Eff. II	Mixed Eff. III
Intercept	0.80*** (0.09)	0.32 (0.35)	0.20 (0.39)
CO ₂ reduction	0.29** (0.12)	0.22* (0.12)	0.25** (0.13)
CO ₂ reduction ²	-0.04* (0.02)	-0.02 (0.02)	-0.03 (0.02)
Price	0.34*** (0.06)	0.31*** (0.06)	0.33*** (0.06)
Stated pref. method		0.32 (0.27)	0.34 (0.27)
In-person		0.01 (0.21)	0.03 (0.21)
Sample size		-0.06 (0.09)	-0.05 (0.09)
Publication		-0.12 (0.21)	-0.14 (0.21)
Study year		0.05 (0.09)	0.02 (0.10)
GDP per capita		0.03 (0.07)	0.01 (0.07)
Europe		0.42*** (0.16)	0.42** (0.17)
CO ₂ reduction assump.			0.14 (0.19)
WTP derivation			0.08 (0.16)
Number of obs.	126	126	126
Var (study random effect)	0.15	0.14	0.15
Var (product random eff.)	0.00	0.01	0.00
AIC	214.83	229.69	236.43
BIC	234.68	269.40	281.81
Log Likelihood	-100.41	-100.85	-102.21

***p<0.01; **p<0.05; *p<0.1

Table A.17: Factors associated with WTP for CO₂ reductions: mixed effects model, robustness tests with alternative functional form of CO₂ reduction

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use mixed-effects models, including studies and product categories as random effects. The dependent variable is the unstandardized WTP for carbon emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored. CO₂ reduction² variable is obtained by squaring the z-scored CO₂ reduction variable.

	OLS I (original)	OLS II (no CO ₂ assump.)	OLS III (no WTP derivation)	OLS IV (neither)
Intercept	0.51 (0.33)	1.57*** (0.23)	0.78* (0.43)	1.70*** (0.15)
CO ₂ reduction	0.11*** (0.04)	0.08* (0.04)	0.10*** (0.03)	0.03 (0.02)
Price	0.35*** (0.05)	0.24*** (0.03)	0.33*** (0.07)	0.29*** (0.05)
Stated pref. method	0.00 (0.24)	-0.60*** (0.11)	0.01 (0.28)	
In-person	-0.08 (0.24)	-0.83*** (0.19)	-0.27 (0.32)	-0.65*** (0.17)
Sample size	-0.05 (0.08)	-0.17*** (0.04)	-0.09 (0.07)	-0.05 (0.06)
Publication	0.01 (0.23)	-0.35** (0.15)	-0.45** (0.20)	-0.85*** (0.06)
Study year	0.04 (0.09)	0.15 (0.10)	0.02 (0.11)	0.48*** (0.07)
GDP per capita	0.06 (0.08)	0.04 (0.05)	-0.13 (0.12)	0.32*** (0.05)
Europe	0.37** (0.17)	0.62*** (0.11)	0.56*** (0.20)	0.47*** (0.03)
Number of obs.	126	55	73	28
Number of studies	37	14	24	8
Adjusted R ²	0.46	0.78	0.51	0.79
AIC	209.92	51.26	120.70	39.72
BIC	241.12	73.34	145.90	53.05
Log Likelihood	-93.96	-14.63	-49.35	-9.86

***p<0.01; **p<0.05; *p<0.1

Table A.18: Factors associated with WTP for CO₂ reductions: OLS model, robustness tests with subsets of the data

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use an OLS model with clustered standard errors by study. Each column presents OLS outputs based on different subsets of data. The first column is based on the complete set of observations. The second column excludes observations for which we need to make assumptions regarding the amount of CO₂ reduction through external calculators or other third-party sources. The third column is based on data where WTP_R values are obtained directly from the studies, without any additional calculations. The fourth column combines these criteria, excluding both observations with CO₂ reduction assumptions and derived WTP values. The dependent variable is the (unstandardized) WTP for CO₂ emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

	Mixed Eff. I (original)	Mixed Eff. II (no CO ₂ assump.)	Mixed Eff. III (no WTP derivation)	Mixed Eff. IV (neither)
Intercept	0.26 (0.35)	1.62*** (0.37)	0.19 (0.52)	1.58*** (0.54)
CO ₂ reduction	0.10** (0.05)	0.15** (0.07)	0.09 (0.07)	0.10 (0.09)
Price	0.32*** (0.06)	0.46*** (0.07)	0.31*** (0.08)	0.43*** (0.09)
Stated pref. method	0.34 (0.27)	-0.36 (0.26)	0.64 (0.40)	
In-person	0.00 (0.21)	-0.62*** (0.19)	-0.19 (0.25)	-0.59** (0.29)
Sample size	-0.07 (0.09)	0.06 (0.09)	-0.10 (0.11)	0.08 (0.15)
Publication	-0.12 (0.21)	-0.44** (0.21)	-0.41 (0.27)	-0.88* (0.46)
Study year	0.06 (0.09)	0.31** (0.12)	0.08 (0.12)	0.45** (0.20)
GDP per capita	0.02 (0.07)	0.13 (0.09)	-0.08 (0.11)	0.23 (0.17)
Europe	0.44*** (0.16)	0.49*** (0.13)	0.50** (0.20)	0.48** (0.20)
Number of obs.	126	55	73	28
Number of studies	37	14	24	8
Var (study random effect)	0.14	0.00	0.14	0.00
Var (product random eff.)	0.01	0.24	0.00	0.13
AIC	222.89	83.06	147.34	65.21
BIC	259.76	109.16	177.12	81.20
Log Likelihood	-98.44	-28.53	-60.67	-20.60

***p<0.01; **p<0.05; *p<0.1

Table A.19: Factors associated with WTP for CO₂ reductions: mixed effects model, robustness tests with subsets of the data

This table shows coefficient estimates and associated standard errors, which are indicated within parentheses. We use mixed effects models, including studies and product categories as random effects. Each column presents mixed effects model outputs based on different subsets of data. The first column is based on the complete set of observations. The second column excludes observations for which we need to make assumptions regarding the amount of CO₂ reduction through external calculators or other third-party sources. The third column is based on data where WTP_R values are obtained directly from the studies, without any additional calculations. The fourth column combines these criteria, excluding both observations with CO₂ reduction assumptions and derived WTP values. The dependent variable is the (unstandardized) WTP for CO₂ emission reductions (WTP_R), which is transformed using the inverse hyperbolic sine function. CO₂ reduction, price, sample size, study year, and GDP per capita variables are z-scored.

Chapter 2: Demand for carbon-neutral products: evidence from Amazon's online marketplace

coauthored with Stefano Carattini, Fabian Dvorak, Ivana Logar

Abstract

A key policy question concerns the market's capability to supply (global) public goods, such as climate change mitigation. Markets are evolving rapidly, given the widespread transition toward carbon neutrality in companies' operations and products. Therefore, it is an empirical question how consumers value carbon-neutral products. We address this question by leveraging information on prices and product characteristics from one of the largest online marketplaces, Amazon. Using a hedonic approach, we infer consumers' valuation of carbon-neutral products from revealed preferences. The staggered process of carbon-neutral certification leads to a series of quasi-natural experiments, and we use these experiments for identification. We find no evidence that consumers value carbon neutrality in actual market data. This result contrasts with the evidence from the existing literature, which relies primarily on stated-preference methods.

Keywords: carbon-neutral labels; corporate social responsibility; revealed preferences; hedonic pricing

JEL codes: D12; H41; Q51; Q54

1 Introduction

As the transition to a cleaner economy becomes increasingly urgent, reliance on voluntary private action to provide global public goods has grown. Reflecting this trend, many firms have recently declared their intentions to make their operations or products carbon neutral (Rogelj et al., 2021), primarily in response to pressure from company boards and investors preparing for a low-carbon future (Kim and Lyon, 2011). Whether consumers value these efforts, however, remains an open empirical question. Existing evidence relies largely on stated-preference methods, which may be prone to hypothetical bias. Consequently, revealed-preference evidence on consumers’ willingness to pay a premium for carbon-neutral products remains very limited, and this paper fills this gap.

Using actual market data, we implement a hedonic difference-in-differences design to estimate consumers’ valuation of carbon neutral certifications. In particular, we use publicly available information on prices, product characteristics, and reviews from the Amazon marketplace, covering a wide range of products, tracked over several months in different countries, the United States, the United Kingdom, and Germany, to assess in a hedonic framework whether there is any demand for the carbon neutrality of products.

Amazon is by far the most important online marketplace in many countries around the world.¹ In line with other marketplaces, many products on Amazon have recently been certified carbon neutral. Consumers buying these products would in principle access a climate-friendly product, without the need for any action on their part, unlike in the case of offsetting. Carbon-neutral products remain niche products, but their recent expansion allows the application of a hedonic approach to investigate whether consumers value their climate-friendly nature. From an empirical standpoint, the emergence of these certifications creates a multitude of quasi-natural experiments, which we leverage to causally

¹Of the 129 million households in the United States (US Census Bureau, 2025), more than half had an Amazon Prime subscription already in 2019 (Statista, 2022b). Amazon plays a dominant role in European markets as well. 86% of shoppers in the United Kingdom use Amazon. More than 25% of British adults had an Amazon Prime membership in 2018 (Mintel, 2019). More than 75% of online German consumers made at least one purchase per month on Amazon in 2022 (Statista, 2022a)

identify consumers' interest in the carbon neutrality of products. We do so with an event-study analysis, which also allows us to test pre-trends and ensure the suitability of control groups.

We collected data on over 120,000 products across the three Amazon marketplaces over many months on a weekly basis. In the United States, we identify 207 treated products that received carbon-neutral certifications on Amazon during the observation period and 7,526 category-matched control products without such certification. In the United Kingdom, we identify 50 treated products and 5,876 control products. In Germany, we identify 84 treated products and 5,451 controls. Given the staggered nature of these certifications, we rely on recent advances in the difference-in-difference literature, and in particular on Callaway and Sant'Anna (2021). This estimator allows us to examine causally the effect of treatment in the presence of dynamic effects and heterogeneity across cohorts. This estimator also allows leveraging propensity scores to maximize comparability between treatment and control groups, which we also inspect visually based on pre-trends (or absence thereof).

With the hedonic difference-in-difference approach, we find no evidence of a relationship between carbon-neutral labeling and product prices. Treatment effects derived from the hedonic approach are statistically indistinguishable from zero across all three markets, implying that consumers do not value carbon neutrality. Crucially, keyword searches applied to customer reviews of treated products reveal that mentions of carbon neutrality and related terms are negligible in all three markets ($<0.14\%$). This observation is consistent with the absence of a positive relationship between carbon-neutral labeling and product prices. Hence, the substantial WTP for the carbon neutrality of products reported in the literature, which is mainly consists of stated preference studies, is not reflected in how consumers value carbon neutrality on the main online marketplace, Amazon, across three countries. We also do not find consistent evidence for an effect of carbon-neutral labeling on the quantities sold – proxied by customer ratings.

This paper contributes to multiple strands of the literature. First, we add

to the body of work examining corporate social responsibility (e.g. Fehr et al., 1993; Shleifer, 2004; Besley and Ghatak, 2007; Falk and Szech, 2013; Bartling et al., 2015; see also Bénabou and Tirole, 2010, and Kitzmueller and Shimshack, 2012) and emissions reduction (e.g. Kim and Lyon, 2011; Doda et al., 2016) by quantifying the consumers' value of climate change mitigation efforts. If consumers reward carbon-neutral products with a price premium, firms have an incentive to reduce and offset emissions. The meta-analysis of the existing literature indicates a WTP strong enough to drive mass adoption of such labels (Carattini et al., 2025). However, our hedonic estimates paint a different picture: we find no evidence that consumers value carbon neutrality.

Second, we speak to a broad literature on the adoption of pro-social behavior (e.g. Dawes and Thaler, 1988; Fehr et al., 1993; Fehr and Schmidt, 1999; Bénabou and Tirole, 2006; Ellingsen and Johannesson, 2008; Andreoni and Bernheim, 2009; Ariely et al., 2009), including a recent focus on the adoption of non-normative pro-social behavior (e.g. Sparkman and Walton, 2017; Kraft-Todd et al., 2018; Mortensen et al., 2019; Spencer et al., 2019; Andreoni et al., 2021; Bicchieri and Dimant, 2022; Carattini and Blasch, 2024). Our focus is on carbon-neutral products, whose purchase represents a relatively recent form of climate-friendly behavior. Our findings from the hedonic model indicate that carbon-neutral products would remain niche in the absence of a change in consumer preferences, unless corporate social responsibility motives continue to support their expansion. Future research may investigate whether making carbon-neutral purchases more visible generates social rewards and increases attention, as suggested by randomized evidence from peer-to-peer solar, which allows households without rooftop solar to access solar electricity from nearby households and finds higher interest when adoption can be shared publicly via “green reports” (Carattini et al., 2024).

Third, we contribute to analyses of people's cooperativeness in a global social dilemma, such as climate change mitigation (see Carattini et al., 2019, for a review), including private demand for carbon offsets (Kotchen, 2009; Jacobsen, 2011; Kesternich et al., 2016; Rodemeier, 2023). We extend this literature by

examining demand for carbon-neutral products, rather than optional carbon offsets. Carbon offsets are typically an add-on at checkout: consumers can choose whether to pay an extra amount to offset emissions. By contrast, carbon-neutral products embed the mitigation cost in the product itself. Producers incur the cost of reducing or offsetting emissions upfront, and consumers choose between purchasing the carbon-neutral product or an alternative product.

Fourth, we build upon the diverse scholarship estimating WTP for labeled products, particularly carbon-neutral or low-carbon labels (e.g. Akaichi et al., 2017; Birkenberg et al., 2021; Muller et al., 2019). We also provide evidence relevant to the theoretical literature on environmental labels (Fischer and Lyon, 2014; Brécard, 2017; Heyes and Martin, 2018; Poret, 2019; Fischer and Lyon, 2019), which examines how certifications can shape market outcomes by shifting demand toward green products when consumers value the labeled attribute. Using a hedonic design based on revealed preferences, we test whether carbon-neutral certification is associated with a price premium in actual market settings.

Fifth, we connect to a strand of literature comparing various methods, in particular in light of their ability to uncover actual preferences, including WTP (e.g. Arrow et al., 1993; Adamowicz et al., 1994; Bateman et al., 2002; Johnston et al., 2017). We do so by comparing estimates from the existing literature and the hedonic model.

Sixth, we contribute to an established literature applying hedonic methods to a wide range of questions in environmental economics and beyond (e.g. Rosen, 1974; Smith and Desvousges, 1986; Chay and Greenstone, 2005; Muehlenbachs et al., 2015; Banzhaf, 2020, 2021). In this new application, we use a hedonic approach to examine demand for carbon-neutral products, also leveraging recent advances in econometrics allowing for the staggered timing of treatment.

In terms of policy implications, assessing the demand for carbon-neutral products contributes to understanding the potential for expanding the market for carbon-neutral products beyond niche, thus achieving additional voluntary carbon reductions in the private sector, while ambitious climate policy gradually ramps up. Although large publicly-traded firms pledged to become carbon

neutral due to pressure from investors, including at the time in expectation of future policy tightening, this paper examines whether there is a rationale for many other firms to pursue carbon neutrality.

Unlike the meta-analysis of the underlying literature (Carattini et al., 2025), which relies heavily on stated preferences and controlled environments, our hedonic approach finds no evidence that consumers value carbon neutrality, based on prices across three large markets. Overall, people's private motivations to contribute to (global) public goods do not seem a sufficiently strong force to spur the adoption of carbon-neutral products.

The remainder of the paper proceeds as follows. Section 2 introduces our data and empirical approach. In turn, Section 3 provides empirical evidence. Section 4 concludes.

2 Data and empirical approach

This section describes the data used for the hedonic analyses as well as the corresponding empirical approach, while also pointing the reader to additional information in the Appendix. The goal of the hedonic model is to provide empirical evidence from revealed preferences, to be compared with the evidence covered by the meta-analysis of the existing literature presented in Carattini et al. (2025).

To ensure comparability with the data from the meta-analysis, and for reasons of external validity, we cover a wide range of products and several geographies (United States, United Kingdom, Germany) from a large online marketplace with global coverage, Amazon. Amazon's marketplace provides detailed information about product characteristics, including prices, as well as customer reviews, which may point to carbon neutrality as a feature that consumers value. Further, over the last few years, Amazon gave high importance to carbon neutrality, among other environmental aspects, by collaborating with several organizations to provide labels for carbon-neutral products.

More than 50 third-party sustainability certifications are currently displayed

under Amazon’s “Climate Pledge Friendly” program. Products that meet the relevant standards can display a Climate Pledge Friendly label, if requested by the vendor. These certifications cover a range of sustainability dimensions (e.g., USDA Organic, Fairtrade International, SCS Carbon Neutral Certification). For the purpose of this paper, we focus exclusively on carbon-neutral certifications, which are offered by independent organizations, including Climate Impact Partners (previously named Natural Capital Partners), SCS Global Services, ClimatePartner, Carbonfund, and Carbon Trust.

On Amazon, carbon-neutral certifications are displayed at different points of the platform: through the “Sustainably recognized” badge on product search results pages, via the Climate Pledge Friendly page where products can be filtered by a specific carbon-neutral certification, and through certification information displayed on individual product pages. These features allow consumers to see whether a product is carbon-neutral and enable us to identify certified products and their certifying organizations. Figure 1 shows an example of a carbon-neutral product on Amazon.com. Appendix A.1 provides additional details on Amazon’s Climate Pledge Friendly program.

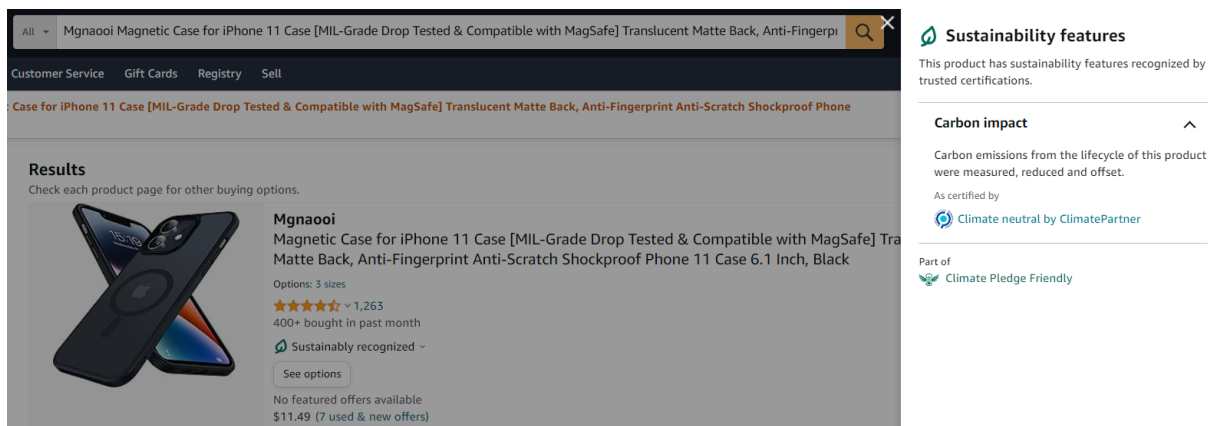


Figure 1: A product certified carbon neutral by Climate Partner on Amazon.com

Our hedonic analysis is based on a weekly panel of products sold on Amazon’s marketplace. We employ the following strategy to construct the panel. First, we identify a list of several thousand products with carbon-neutral labels using built-in filters available on Amazon.com. Next, we identify the Amazon

category nodes assigned to the carbon-neutral products, i.e., Amazon’s category identifiers that group products within the same defined product category. For each category node that we identify, we scrape many untreated products without carbon-neutral labels. This process ensures that for each already-treated product, we obtain many control units from the same product category. We proceed similarly with Amazon.co.uk and Amazon.de.

The benefits of this product selection strategy are threefold. First, selecting products to be monitored from categories that already contain treated products ensures that, in principle, it is possible to make such products carbon neutral and label them accordingly. Second, it also implies that there are some motivations among producers to make these products carbon-neutral in the near future, thereby increasing the likelihood of treatment within the time horizon of our study. Third, it allows us to estimate the dynamic effect of treatment by controlling for the category-specific price trend of untreated products in the same category.

Our product selection strategy results in a set of 38,974 products from 35 categories on Amazon’s American marketplace, 41,384 products from 134 categories on Amazon’s marketplace in the United Kingdom, and 40,031 products from 48 categories on Amazon’s German marketplace. We scrape the same set of product information each week for all these products, as displayed on Amazon’s website, for IP addresses from the United States, the United Kingdom, and Germany. The data collection started in March 2023 and ended in November 2024. Data from Amazon.com, the American marketplace, was scraped for this entire period. For the United Kingdom and Germany, we have data from May 2024 through November 2024. Across markets, the gradual adoption of carbon-neutral labels during the sample period generates many quasi-natural experiments that we exploit for identification.

From this broad set of products, we identify treated products that obtain new carbon-neutral labels, and, in turn, control products. Most importantly, we retrieve information about the price of the product and the treatment status of the product, which allows us to perform a staggered difference-in-difference

analysis. The staggered adoption of carbon-neutral labels by products sold on Amazon provides the ideal features of a quasi-natural experiment. Here, a product changes its treatment status when it receives one of the abovementioned carbon-neutral labels. Control units are represented by products that are arguably comparable within the same product category assigned by Amazon and pre-trends are examined accordingly. We exclude Amazon's own products as well as books from the analysis, since these categories might have different market dynamics than the other products on Amazon.

The underlying assumption of this exercise is that the prices observed on Amazon's marketplace are equilibrium prices. It is well known that Amazon uses dynamic pricing for its own products and offers its in-house dynamic pricing engine to all sellers. As a result, the prices displayed on the website tend to automatically adjust to changes in demand, in principle quickly approaching equilibrium prices. Still, we also examine effects on quantities, which are proxied by the number of ratings a product receives, to provide a fuller picture.

For the difference-in-differences analysis, we restrict attention to products whose information, including labels and prices, can be scraped consistently and that are available for purchase in most weeks. We exclude product pages that redirect to a different item (e.g., when an item is replaced or merged into another listing). When weekly prices are missing, we fill gaps using linear interpolation (and extrapolate at the beginning or end of a product's series when needed), while, as a robustness check, we also estimate the main model using an unbalanced panel without any interpolation or extrapolation. We further exclude always-treated products from the hedonic analysis (2,759 for the United States, 1,881 for the United Kingdom, and 1,943 for Germany).

We identify 207 treated and 7,526 control products in the United States, 50 treated and 5,876 control products in the United Kingdom, and 84 treated and 5,451 control products in Germany. When selecting suitable control products among untreated candidate products, i.e. products without a carbon-neutral label at the beginning of the timeframe of reference (23,935 products in the United States, 32,486 in the United Kingdom, and 33,224 in Germany), we

first consider products that, at the beginning of the panel, do not have another “Climate Pledge Friendly” label (e.g. organic, fair trade) and then match each treated product’s category node to find control products that share at least one of these nodes. Each product on Amazon is assigned up to 10 category nodes, ranging from broader to more specific classifications. If fewer than 10 suitable controls are available within the most precise node, we move to a higher-level node. Only control products matched to the category of at least one treated product are retained in the final dataset. Tables 1, 2, and 3 present the number of treated and control products by highest level category across the three markets.

Additional details are provided in Appendix A.2. Table A.1 summarizes the restrictions used to define treated and control products. Table A.2 shows that treated and control products are balanced in terms of their initial prices at the start of the panel. Tables A.3-A.5 list treated products for each country, including their categories and prices at the start and end of the panel.

We estimate the treatment effect of carbon neutrality in non-canonical difference in differences. We assume that the timing at which products get treated is plausibly random and assess pre-trends accordingly. We also take treatment as irreversible (at least within our time frame), leading to a difference-in-difference setup with staggered treatment assignment and treatment being absorbed. Accordingly, we proceed with Callaway and Sant’Anna (2021), as now standard in these circumstances. Our setting is also such that only a small share of products receives the treatment, making economically meaningful general equilibrium effects unlikely. In our main estimations, we also control for 57 different finest level product categories from Amazon United States, 15 from Amazon United Kingdom, and 21 from Amazon Germany, as well as initial product prices.

Category	Treated	Control
Beauty & Personal Care	12	304
Cell Phones & Accessories	53	2864
Electronics	91	1716
Grocery & Gourmet Food	1	225
Health & Household	14	932
Musical Instruments	1	37
Office Products	1	356
Safety & Security	2	25
Tools & Home Improvement	9	42
Toys & Games	3	49
Video Games	20	976
Total	207	7526

Table 1: Number of treated and control products by category (United States)

This table presents the number of treated products (products receiving carbon-neutral labels) and control products (without a label) by highest level category for the U.S. market.

Category	Treated	Control
Accessories	2	665
Arts & Crafts	1	277
Gardening	2	77
Headphones, Earphones & Accessories	4	833
Hi-Fi & Home Audio	1	490
Microphones	2	96
Mobile Phones & Communication	28	2654
Sports	2	423
Wearable Technology	8	361
Total	50	5876

Table 2: Number of treated and control products by category (United Kingdom)

This table presents the number of treatment products (products receiving carbon-neutral labels) and control products (without a label) by highest level category for the U.K. market.

Category	Treatment	Control
Computer & Accessories	45	2660
Electronics & Photo	30	1814
Games	5	437
Health & Personal Care	2	358
Stationery & Office Supplies	1	84
Toys	1	98
Total	84	5451

Table 3: Number of treated and control products by category (Germany)

This table presents the number of treated products (products receiving carbon-neutral labels) and control products (without a label) by highest level category for the German market.

Event analyses are used to verify that the (conditional) parallel trend assumption holds (when conditioning on the finest-level product category available and initial price) and to examine dynamic treatment effects. We focus on products that receive the treatment (label) during the period of interest and, as control group, use not-yet-treated products and never-treated products. Event-time effects are obtained by aggregating cohort-time effects, as in Callaway and Sant’Anna (2021):

$$ATT^{\text{Event Study}}(e) = \sum_{c \in C_e} ATT(c, c + e) \Pr(C = c \mid C \in C_e)$$

where $e = t - c$ is the event time for cohort c , $C_e = \{c \in C : c + e \leq T\}$ is the set of cohorts that we observe to be treated for at least e periods up to period T , $\Pr(C = c \mid C \in C_e)$ is the share of cohort c among other cohorts observed at event time e . Each Average Treatment Effect on the Treated ($ATT_{c,t}$) compares cohort c at time t with not-yet-treated and never-treated units at time t .

We express the outcome variable in terms of the percentage change in a product’s price, relative to its baseline (initial) price, in our main specifications, but also provide estimates based on absolute prices. We define a product’s baseline price as its price in the first month of the panel (March 2023 for the United States and May 2024 for the United Kingdom and Germany). Since we

aggregate weekly prices to the monthly level by taking their average, the baseline price coincides with the average weekly price in the initial month. Additionally, we translate the effect into USD per kilogram (ton) of CO₂ by accounting for the average carbon footprint of the treated products. Appendix A.3 provides more details on the carbon footprint calculations. As mentioned, we also consider the number of ratings a product receives as a proxy for quantities (sales).

We conduct a variety of robustness tests: alternative sample restrictions; using the unbalanced-data option instead of filling missing data by interpolation or extrapolation; excluding control variables; including additional controls (e.g., the number of initial product ratings as a proxy for sales); and using the absolute price level as the outcome. We also estimate treatment effects on products' price change relative to the baseline price, accounting for product heterogeneity across four dimensions: price ranges within and above the range of prices observed in the meta-analysis of the existing literature provided in Carattini et al. (2025), product categories, carbon intensity, and certifiers.

To explore mechanisms, we analyze the content of customer reviews, paying particular attention to references to carbon neutrality or climate friendliness more generally. We scrape the first page of product reviews weekly from Amazon's marketplace in the United States (May 2024-November 2024), the United Kingdom (March 2024-November 2024), and Germany (March 2024-November 2024). We then select keywords pointing to carbon neutrality or climate friendliness in general and count their occurrence. Because reviews reflect what customers choose to comment on, we interpret keyword mentions as a proxy for the salience of carbon neutrality in consumers' post-purchase narratives. We use this evidence to complement the hedonic results.

We analyze keyword occurrences in the always-treated group, which includes products with a carbon-neutral label at the start of the panel, in comparison with the never-treated group, which includes products that never receive such a label. We restrict the always-treated sample to reviews posted after the panel start date, to ensure we only analyze reviews of products at a time when they are treated. In total, we have 127,268 reviews for always-treated products and 450,448 reviews

for never-treated products, written by individuals making purchases on Amazon.

3 Empirical Evidence

In this section, we describe the findings from the hedonic analyses using Amazon's data from the United States, the United Kingdom, and Germany.

3.1 Results from Amazon's U.S. marketplace

We start by showing the standard event analysis, in Figure 2, for the United States, based on Callaway and Sant'Anna (2021). Figure 2 serves two purposes. First, it allows to test how comparable the control group is to the group that receives the treatment, by inspecting pre-trends. Second, Figure 2 shows the effect of receiving a carbon-neutral label on the price of the product in the months after treatment, allowing us to examine the dynamic effect of carbon neutrality. The y-axis indicates the relative price difference in percentages to the product's price in March 2023. Table A.6 in Appendix A.4 contains the coefficients describing the dynamic effects.

The pre-treatment coefficients generally fluctuate around zero and are not statistically significant, with standard errors that are quite tight. In terms of the causal effects of carbon-neutral labels, Figure 2 shows that carbon neutrality does not have a positive effect on the price of a product after treatment. The dynamic effects of a carbon-neutral label are never statistically significant. 95% confidence intervals are a bit wider than in the pre-period, as may be expected, but are tight enough long enough to rule out any substantial positive effects following the treatment. Standard errors widen over time, mechanically, due to the fact that the number of observations thins out as one moves towards the right-hand side of the graph due to the staggered nature of treatment assignment. Overall, the average effect of a carbon-neutral label over the 19-month observation period is -1.91% (95% CI: -3.88, 0.06) of the initial product price, and not significantly different from zero. In short, consumers do not seem to value carbon neutrality, given the absence of a relationship between prices and labels.

While the effect that we document is clearly indistinguishable from zero, there may be readers who still wonder about how to interpret the point estimate. The average effect of a carbon-neutral label on product price translates to USD -0.60 in absolute terms, given the average initial price of treated products of USD 31.2. This corresponds to USD -0.13 per kilogram of reduced CO₂ emissions, based on the average CO₂ emissions (4.6 kg) of treated products (from 191 out of 207 products with available carbon footprint information). Once again, -0.13 per kilogram of reduced CO₂ emissions is statistically indistinguishable from zero.

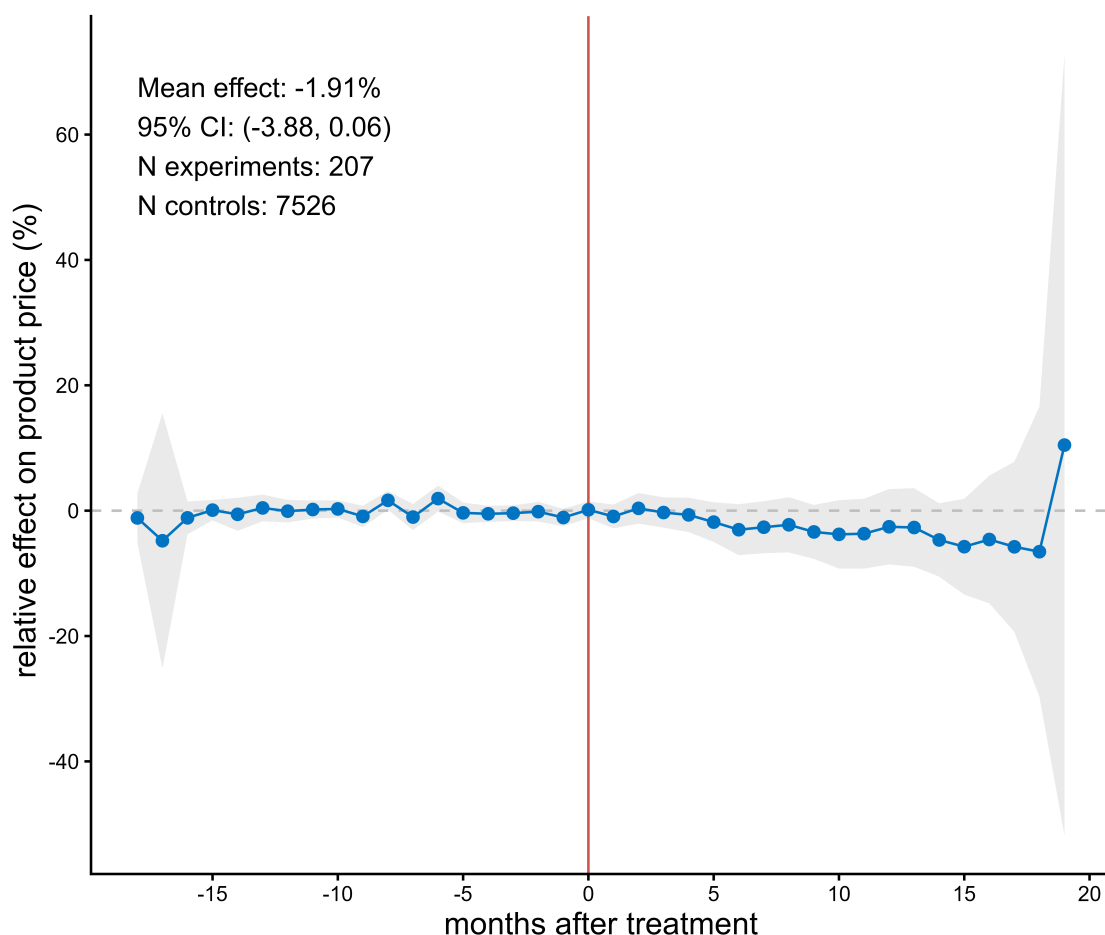


Figure 2: Effect of carbon-neutral label on products' price change (United States)

This plot shows the dynamic treatment effect in percentages relative to the product's price at baseline. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of the treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories and the product's initial price at the beginning of the panel.

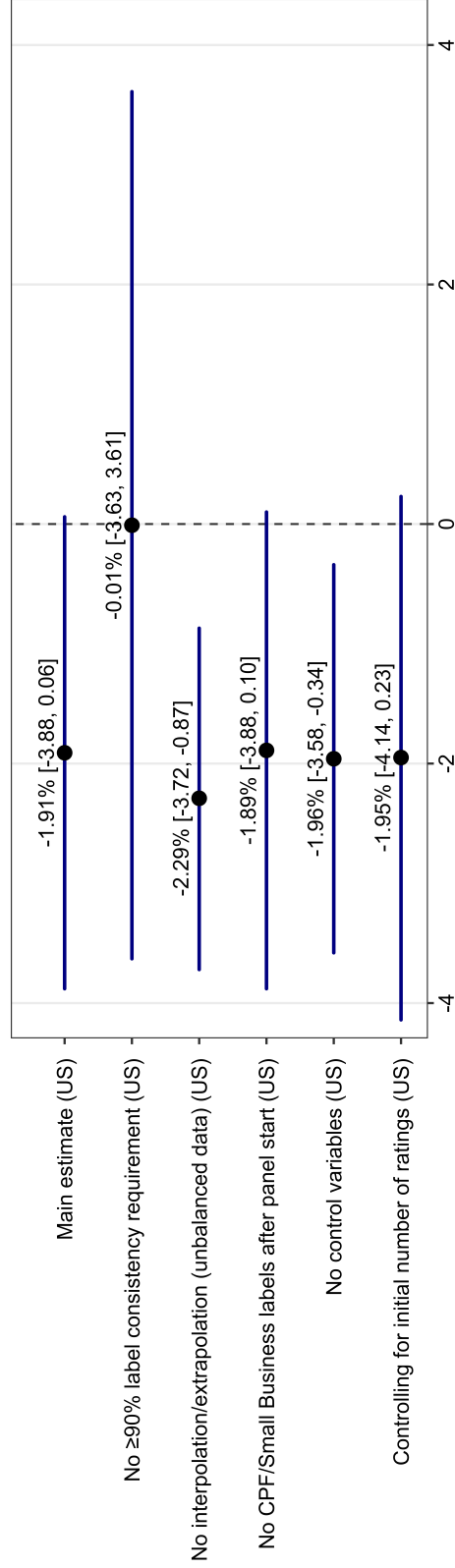


Figure 3: Summary of results for the average effect of the carbon-neutral label on product price changes (United States)

This figure summarizes (a) the main estimate and (b) robustness checks, detailed in Appendix A.6: (i) removing the 90% post-treatment consistency requirement for treated products (i.e., allowing treated products to be classified as treated even if the carbon-neutral label is not observed in at least 90% of their post-treatment period observations); (ii) not filling missing points by interpolation and extrapolation and allowing an unbalanced panel; (iii) excluding the Climate Pledge Friendly labels (introduced after the panel started) and Small Business labels; and (iv) excluding control variables; (v) controlling for the initial number of ratings.

Evidence from the hedonic approach suggests that consumers do not value carbon neutrality. To provide additional evidence on the relationship – or absence thereof – between carbon neutrality and prices, we report the occurrence of carbon neutrality and related keywords in customer reviews of both always-treated and never-treated products. Our headline finding is that the share of reviews mentioning carbon neutrality and related, less technical, terms is negligible for treated products. This statement is true both at the level of each keyword and when summing occurrences over multiple relevant keywords (0.1% of reviews). In short, consumers do not seem to care about carbon neutrality. Appendix A.5 provides more details about our approach and Table A.8 provides the corresponding summary statistics for the United States.

Our estimates from the hedonic model are robust to a variety of tests, summarized in Figure 3 and the event-study plots provided in Appendix A.6. Overall, we do not find evidence that consumers value carbon neutrality. On occasion, the negative point estimate turns significant. Magnitudes, however, remain moderate and are generally consistent across robustness tests. That is, even as we deviate from our preferred specification, the main conclusions are unchanged. We describe our robustness tests in what follows.

First, we relax the restriction that the scraped product information needs to contain the carbon-neutral label in more than 90% of the observations after treatment. We identify 395 treated products; for 188 of these products (48%), the carbon-neutral label is displayed in fewer than 90% of post-treatment weeks. Relaxing this condition does not affect our conclusions, as seen in Figure A.1 in Appendix A.6. The average treatment remains statistically insignificant.

Second, we estimate treatment effects using an unbalanced version of the panel without filling gaps in the scraping process of price data by interpolating (or extrapolating). The corresponding robustness test is presented in Figure A.2, Appendix A.6. While this approach restricts the inclusion of control variables in the estimation, our conclusions remain again unchanged. The average treatment effects are, if anything, negative, also using unbalanced data and in the absence of control variables.

Third, since for our main analysis we only exclude the products with other Climate Pledge Friendly labels (e.g, organic, fair trade), which we detected at the beginning of our panels, we further detect and exclude from the analysis any products that received another label after March 2023, as well as those with a Small Business badge.² This approach results in 207 (the same number as the main analysis) treated products and 7,476 control products. Appendix A.6, Figure A.3 shows that once more, the effect of a carbon-neutral label is overall statistically non-significant.

Fourth, we examine the impact of removing control variables over the balanced panel. While the upper bound of the confidence interval is still close to zero, the effect of the carbon neutral label in this case is -1.96% of the baseline price (95% CI: [-3.58, -0.34]) and significant. This pattern is consistent with the possibility that treated products are concentrated in categories with flatter underlying price growth, in which case omitting category controls could mechanically shift the estimate toward more negative values. The overall effect is close to that of the main estimation, as shown in Figure A.4 in Appendix A.6.

Fifth, we assess the effect of controlling for the initial number of product ratings, as a proxy for products' initial popularity (sales), alongside product categories and products' initial prices. Here too the treatment effect is undistinguishable from zero. The estimate is at -1.95% of products' baseline price [-4.14, 0.23], as shown in Figure A.5 in Appendix A.6.

Lastly, using the absolute price levels as the dependent variable in the hedonic analysis instead of the relative price change confirms that consumers tend not to value carbon neutrality. The average effect of a carbon-neutral label on the absolute product price over the 19-month observation period is -0.81 USD (95% CI: -1.81, 0.19), as shown in Figure A.6, Appendix A.6. This effect is also indistinguishable from zero. Overall, these robustness tests confirm that our main findings for the United States are not sensitive to the use of different specifications or sample restrictions.

We also examine variation in treatment timing. Given the staggered nature

²Amazon provides products with a Small Business badge if the producer has fewer than 100 employees and less than 50 million USD in annual revenue.

of treatment assignment, calendar effects (effects in specific calendar months) differ from dynamic effects (effects observed after certain months of treatment) and are estimated separately. As shown in Table A.7 in Appendix A.4, no specific pattern emerges from the calendar effects.

To provide a fuller picture, we also consider treatment effects on product sales, which we proxy by the number of ratings. When doing so, we also find a non-significant effect, although the point estimate is positive, at 1560 ratings (95% CI: -5, 3126), as shown in Figure A.7 in Appendix A.6. That is, there is a hint of increased popularity of carbon-neutral products on Amazon, at the same time that we provide consistent evidence that consumers do not care about their climate-friendly properties and do not seem to value them. The fact that based on our proxy there is some indication of movement around sales in this market may suggest that relying on prices only, as is usually done in the non-market valuation literature, could potentially mask some effects that may contribute to explain the adoption of labels.

3.2 Heterogeneity analyses

In this section, we assess whether consumers may value carbon neutrality in specific subsamples by estimating heterogeneous treatment effects using the same outcome as in the main U.S. analysis (percentage price change relative to the initial price) while accounting for product heterogeneity. Specifically, we distinguish products along four dimensions: price ranges within and above those covered in the meta-analysis by Carattini et al. (2025), product categories, carbon-intensity groups, and certifiers. Overall, the results indicate that the main conclusion is unchanged: consumers generally do not value carbon-neutral labels, including within these subgroups.

Figure 4 distinguishes between products falling within the price range observed in the meta-analysis of the existing literature conducted by Carattini et al. (2025) (USD 0.09–22.15) and those above it. Consumers may value the same amount of carbon reductions more for expensive goods, relative to cheaper ones, a point that is discussed in the meta-analysis by Carattini et al. (2025). Based

on Figure 4, our conclusions are unchanged. Either way, consumers do not value carbon-neutral labels. The estimates for the two subgroups are not statistically different. The estimates are -1.56% [-5.59, 2.46] for products within the meta-analysis price range and -2.49% [-4.85, -0.13] for those above, respectively.

Figure 5 provides estimates for the different product categories provided by Amazon: Electronics -5.33% [-7.22, -3.44], Cell Phones & Accessories 2.28% [-2.39, 6.94], Video Games -4.92% [-11.36, 1.53], Health & Household -5.96% [-9.76, -2.16], and Beauty & Personal Care -1.58% [-6.17, 3.02]. Once more, our main conclusions are confirmed. In no case there is evidence that consumers value carbon-neutral labels. Confidence intervals across categories overlap and include the overall estimate (-1.91%). Note that in the case of Health & Household products, the treated sample is small ($N = 14$).

Figure 6 groups products by carbon intensity. We define product carbon intensity as a product's carbon footprint (kg CO₂) divided by its weight (kg). We rank treated products by this measure and split them at the median into 'Low' and 'High' carbon-intensity groups. A category is labeled Low (High) only if all experiments in that category fall in that group; categories with both levels are marked 'Mixed' and excluded, and all control products (as well as 7 treated products with missing carbon intensity information) get their category's Low/High label. The carbon-intensity ranges for the Low and High groups are, respectively, 0.40–4.52 kg CO₂/kg (Low) and 4.82–723 kg CO₂/kg (High). Here too, we find that consumers do not value carbon-neutral labels. Point estimates are -2.98% [-5.92, -0.05] for products with high carbon intensity and -1.49% [-6.14, 3.17] for products with low carbon intensity.

Finally, Figure 7 compares products grouped according to the third-party certifiers certifying them. Also in this case, we conclude that consumers do not value carbon-neutral labels. The effect for CarbonFree Certified (Carbonfund) is -3.23% [-5.86, -0.60], for ClimatePartner -1.34% [-3.98, 1.30], and for all other certifiers -4.05% [-10.46, 2.36]. Confidence intervals overlap substantially across certifiers.

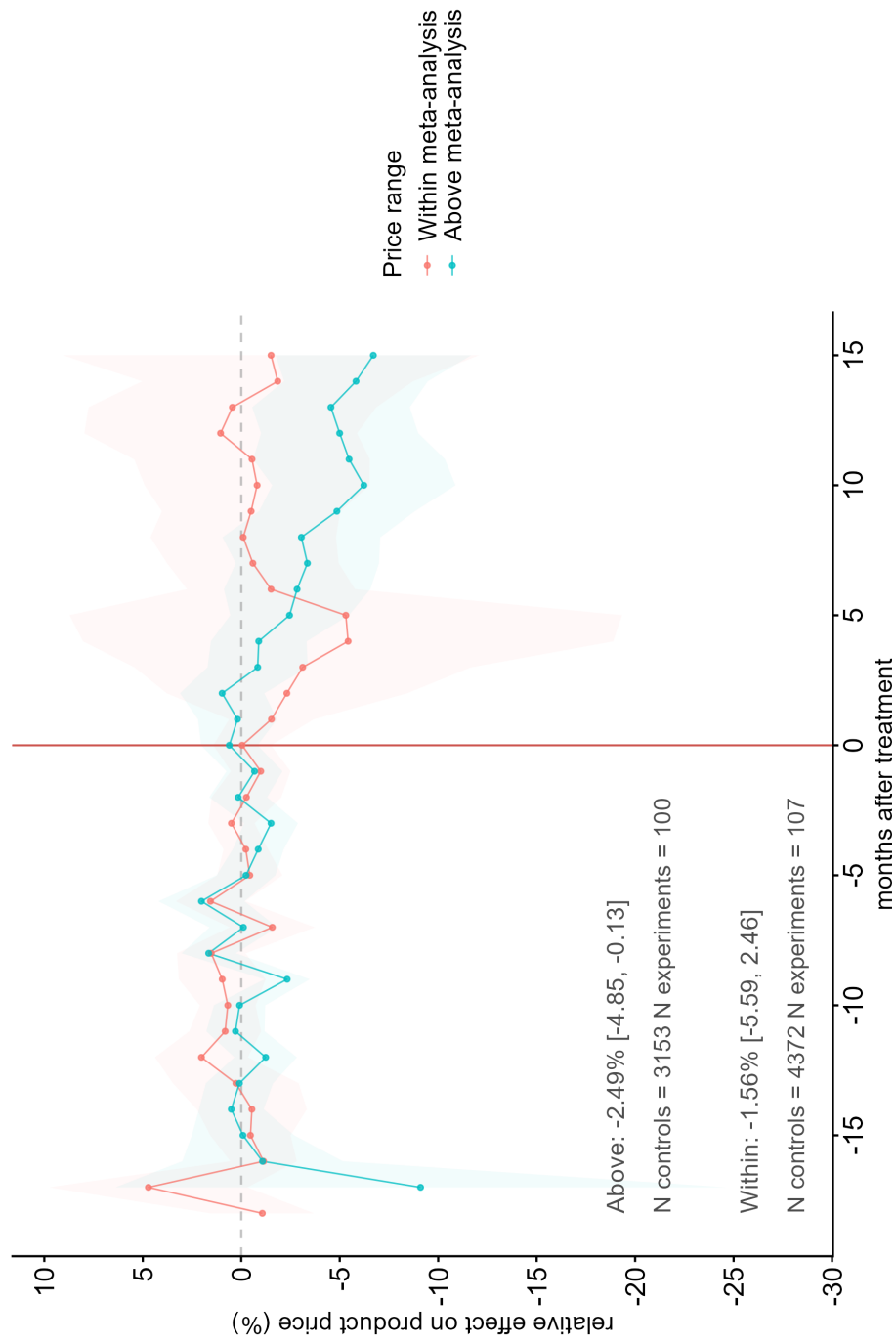


Figure 4: Heterogeneous effect of the carbon-neutral label on product price changes: prices within and above the meta-analysis range (US)

This plot shows the dynamic treatment effect of the carbon-neutral label on product prices relative to March 2023. Following Carattini et al. (2025), prices are restricted to the meta-analysis range (USD 0.09-22.15) and above; no observations fall below this range. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of the treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories and the initial price of products. In the heterogeneity analysis, the event-study plot is trimmed at period 15 because subsequent months in the panel have very small sample sizes by construction.

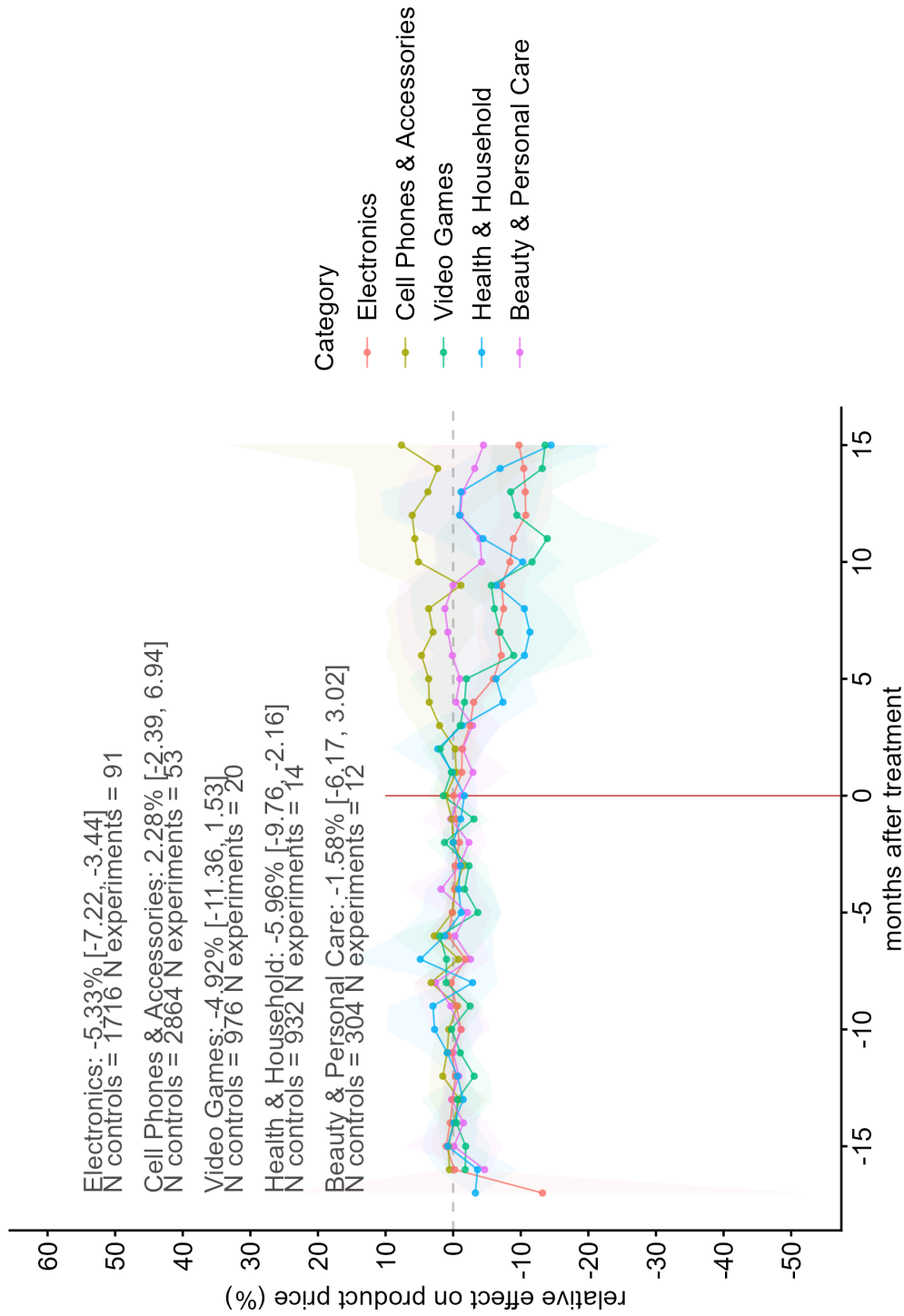


Figure 5: Heterogeneous effect of carbon-neutral label on product's price change: product categories (US)

This plot shows the dynamic treatment effect of the carbon-neutral label on product prices relative to March 2023 by product categories with 10 or more treated products. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of the treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories and the initial price of products. In the heterogeneity analysis, the event-study plot is trimmed at period 15 because subsequent months in the panel have very small sample sizes by construction, limiting the ability to study heterogeneity.

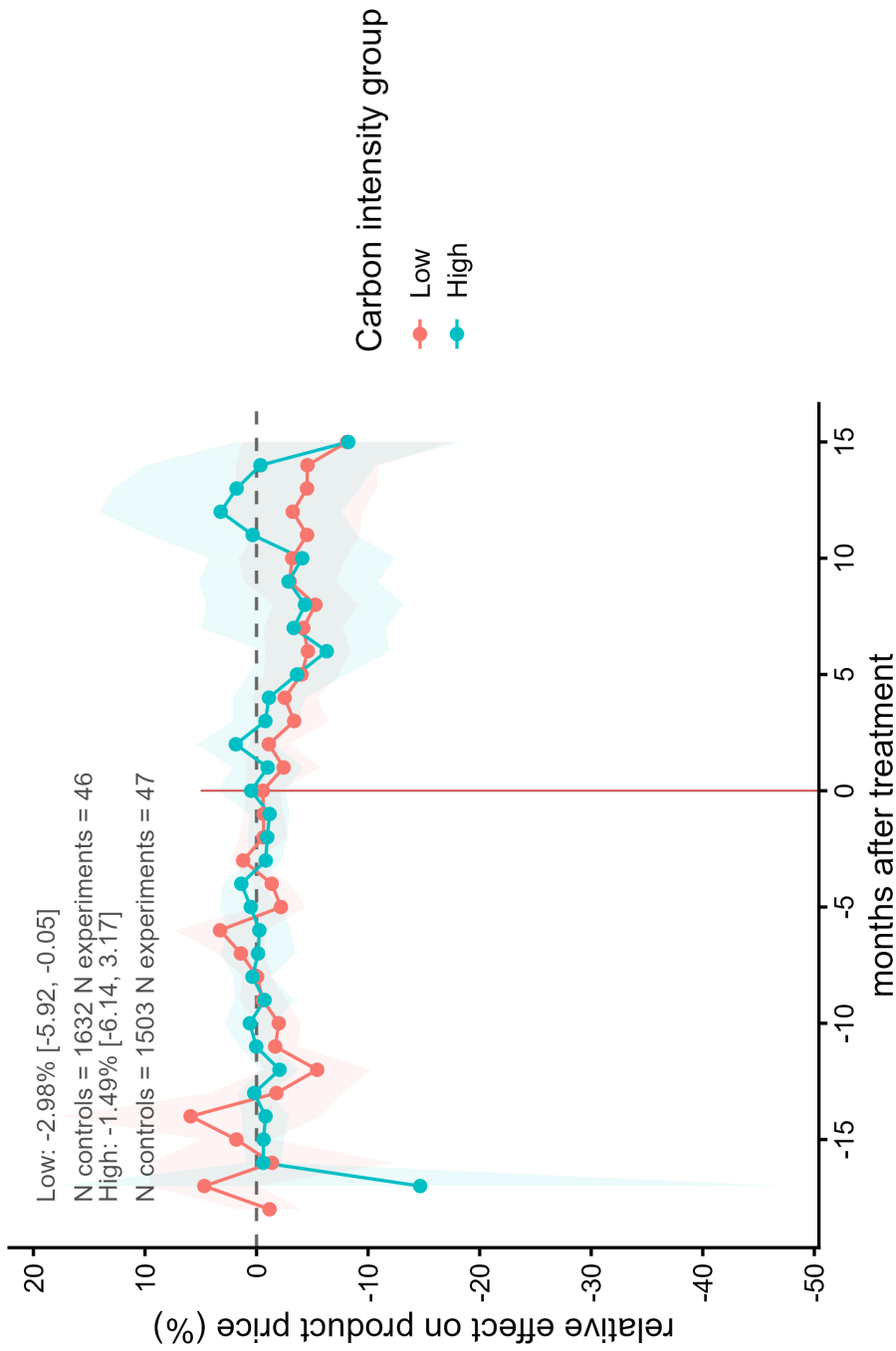


Figure 6: Heterogeneous effect carbon-neutral label on product's price change: high and low carbon-intensive product categories (US)

This plot displays the dynamic treatment effect of the carbon-neutral label on products' price change relative to its price in March 2023, with different carbon intensity categories. Carbon-intensity heterogeneity is constructed by ranking experiment-level carbon intensities (carbon footprint per kilogram of product) and splitting the distribution into two groups (Low, High). A category is labeled Low (High) only if all experiments in that category fall in that group; categories with mixed groups are excluded. Control group products get the Low/High label of their category. Low categories: Accessories, Adapters, Backpacks, Black, Cloths & Towellees, Covers, Dolls, Gloves, Headsets, Hubs, Keyboard Cases, Non-Sterile Gloves, Screen Protectors, Sleeves, USB Cables, Wrist Rests. High categories: Chargers & Power Adapters, Charging Stations, Cradles, Creams & Moisturizers, Elderberry, Eyes, Face Mists, FM Transmitters, Gaming Keyboards, Laptop Accessories, Menstrual Cups, Mice, Omega-3, Pore Cleansing Strips, Portable Bluetooth Speakers, Sports & Outdoor Play, Stylus Pens, Underwater Photography. The shaded area indicates a 95% confidence interval of the treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories and the initial price of products. In the heterogeneity analysis, the event-study plot is trimmed at period 15 because subsequent months in the panel have very small sample sizes by construction, limiting the ability to study heterogeneity

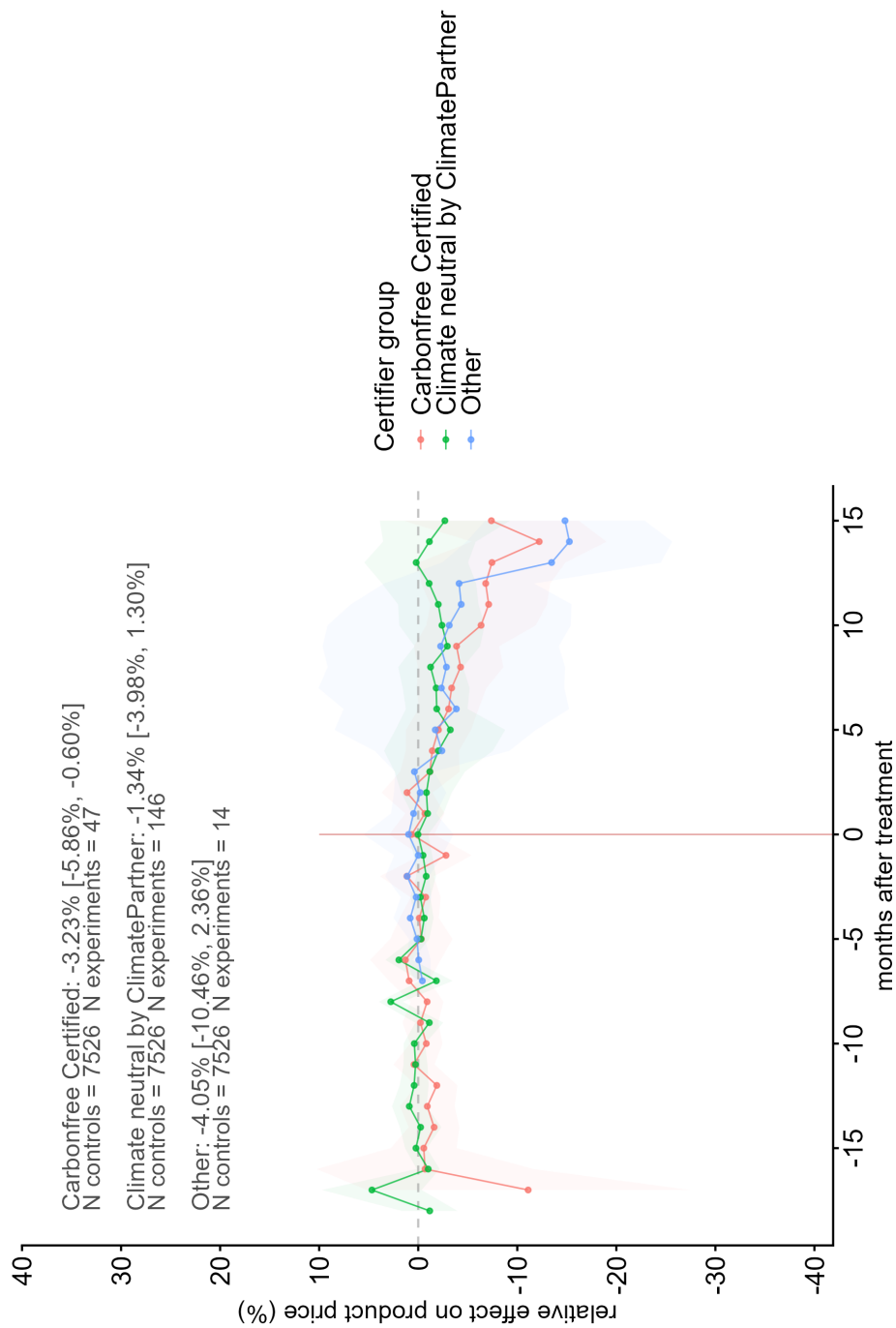


Figure 7: Heterogeneous effect of carbon-neutral label on product's price change: label certifiers (US)

This plot displays the dynamic treatment effect of the carbon-neutral label on products' price change relative to its price in March 2023, with different carbon-neutral label certifiers. Certifiers with at least 10 treated products are displayed individually; all remaining certifiers are grouped together as 'Other', which includes Carbon Neutral by Carbon Trust and Carbon Neutral Certified by SCS Global Services. Control variables include product categories. The shaded area indicates a 95% confidence interval of the treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories and the initial price of products. In the heterogeneity analysis, the event-study plot is trimmed at period 15 because subsequent months in the panel have very small sample sizes by construction, limiting the ability to study heterogeneity

3.3 Results from Amazon's U.K. and German marketplaces

To provide additional evidence concerning whether consumers value carbon neutrality, we move beyond American consumers shopping on Amazon in the United States. In particular, we examine evidence from two more countries, the United Kingdom and Germany, where we perform the same analyses using as outcome variables products' price change relative to their baseline price as well as the number of ratings. We provide a concise summary here and present the event-study plots in Appendix A.7.

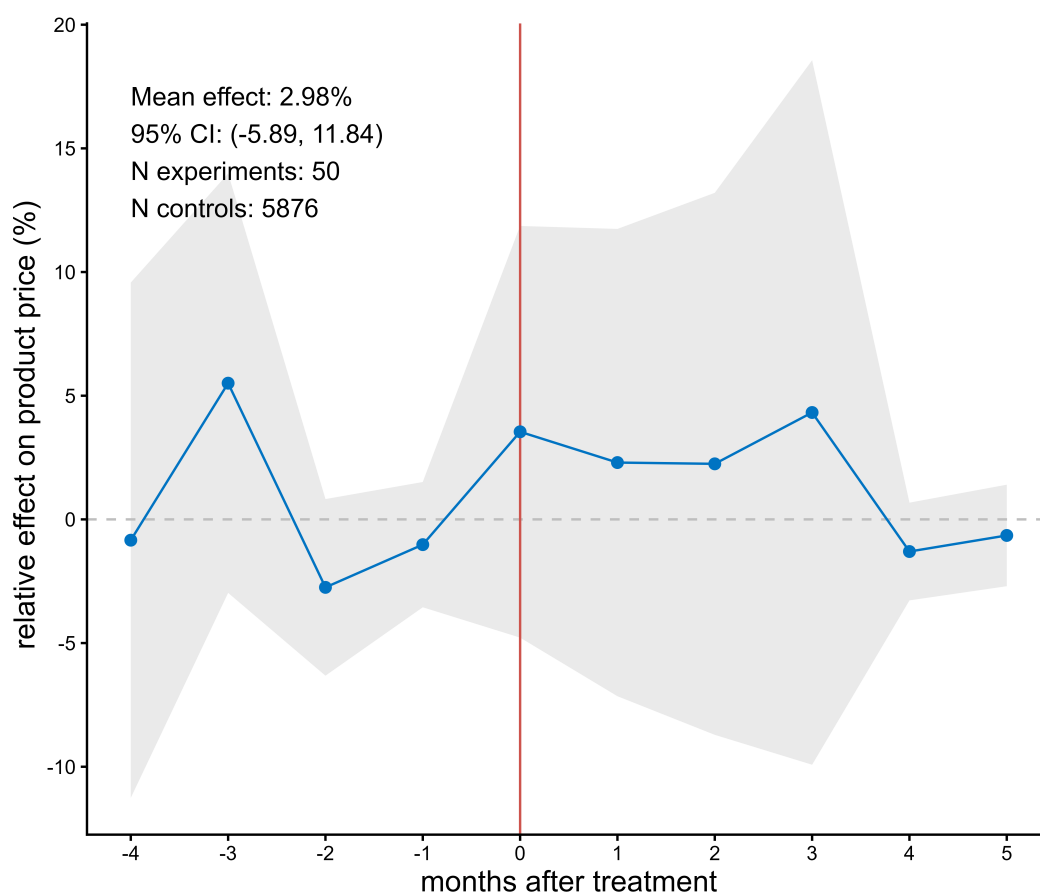


Figure 8: Effect of carbon-neutral label (UK)

This plot shows the dynamic treatment effect in percentages relative to the product's price at baseline. Vertical red line marks treatment onset. Shaded area indicates 95% confidence interval of treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories and the product's initial price at the beginning of the panel.

Figure 8 presents the main results for the United Kingdom. It shows the dynamic effects of receiving a carbon-neutral label on the price change of the

product relative to its initial price. Here too, pre-treatment estimates are not statistically significant, although with fewer observations from which to derive inference, the time series tends to be a bit more volatile than as observed with data for the United States. In terms of treatment effects, the average effect of a carbon-neutral label over the five-month observation period is again indistinguishable from zero. The point estimate is slightly positive, but small at 2.98% (95% CI: -5.89, 11.84). Here too, we observe that consumers do not tend to value carbon neutrality. Accordingly, we return to the analysis of customer reviews, this time analyzing reviews by customers shopping on Amazon in the United Kingdom. Once more, we find that mentions of carbon neutrality and similar keywords are negligible, in isolation as well as when summing over keywords, as shown in Table A.9 in Appendix A.5. Hence, it seems that also in the United Kingdom, consumers do not really value carbon neutrality.

As just described, the effect of carbon-neutral labels on prices is indistinguishable from zero also in the United Kingdom. However, there may be readers still interested in how to interpret this point estimate as well. The effect of a carbon-neutral label on product price translates to GBP 0.69, given that the average initial price of treated products is GBP 23.3. This corresponds to 0.33 GBP per kg of reduced CO₂ emissions, given that the average CO₂ emissions of treated products is 2.1 kg (based on 49 out of 50 products with available carbon footprint information). If taken at face value, this effect would be substantial, albeit still considerably lower than observed in the meta-analysis conducted by (Carattini et al., 2025). However, the effect is statistically indistinguishable from zero, and so is 0.33 GBP per 1 kg of reduced CO₂ emissions. Recall also that reviews do not provide any indication that consumers in the United Kingdom would value carbon neutrality.

We now turn to product sales for the United Kingdom. In this case, the average treatment effect on our proxy of product sales, the number of ratings a product receives, is negative and statistically insignificant. We find an average effect of -612 ratings (95% CI: -1346, 122), as shown in Figure A.8 in Appendix A.7.

Next, we turn to Amazon’s German marketplace. We start again with product prices and use an event analysis to test pre-trends and estimate treatment effects. We do so in Figure 9. Similarly to the case of the United States, pre-trends are flat and confidence intervals tight, confirming the validity of our approach. In terms of treatment effects, the average effect of a carbon-neutral label over the five-month observation period is undistinguishable from zero here too. The point estimate is -11.42% (95% CI: -32.25, 9.4), but standard errors are wide. Also for Germany, we examine customer reviews. Here too, we find that mentions of carbon neutrality and similar keywords are negligible, in isolation as well as when summing over keywords, as shown in Table A.10 in Appendix A.5. Hence, it seems that also in Germany, consumers do not really value carbon neutrality.

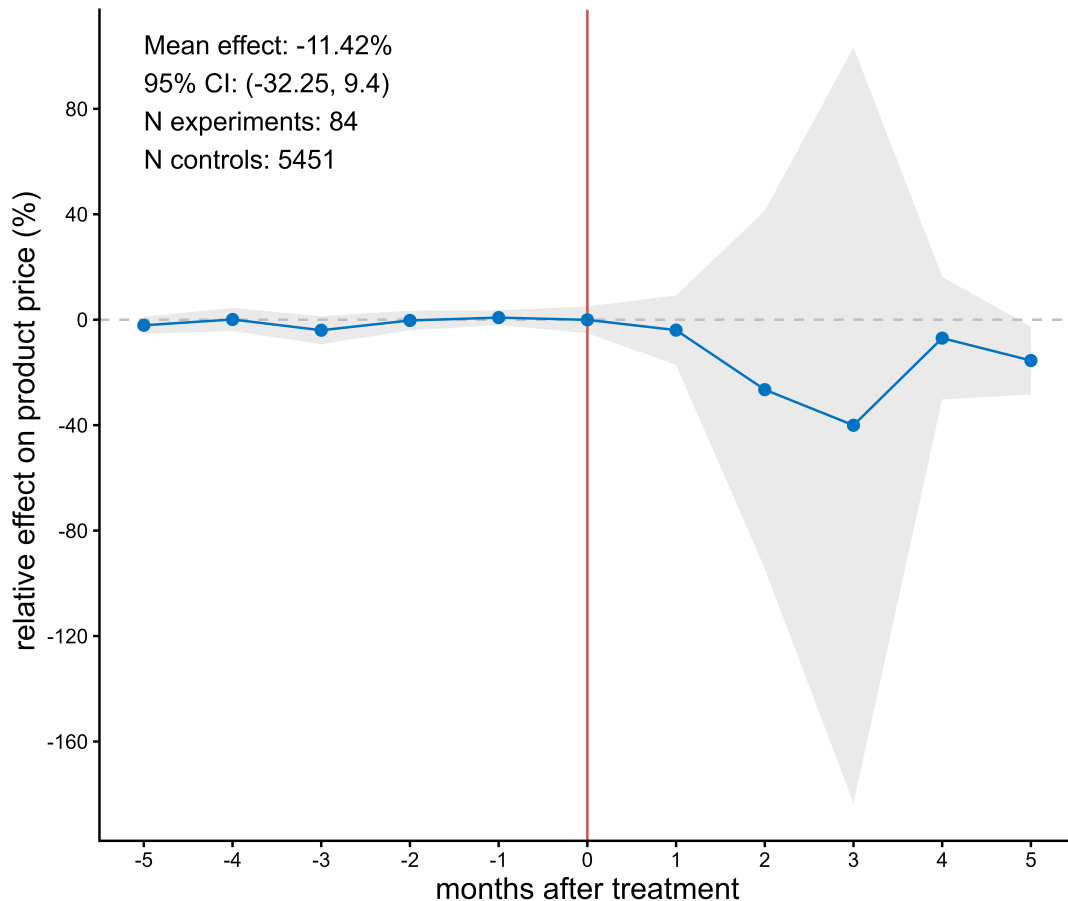


Figure 9: Effect of carbon-neutral label (Germany)

This plot shows the dynamic treatment effect in percentages relative to the product’s price at baseline. Vertical red line marks treatment onset. Shaded area indicates 95% confidence interval of the treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories and the product’s initial price at the beginning of the panel.

As just described, the effect of carbon-neutral labels on prices is indistinguishable from zero also in Germany. In terms of point estimate, though, the effect of a carbon-neutral label on product price translates to EUR -2.49 in absolute terms, given the average initial price of treated products of EUR 21.8. This corresponds to EUR -0.45 per kg of reduced CO₂ emissions, based on the average CO₂ emissions of 5.48 kg of treated products (based on 76 out of 84 products with available carbon footprint information).

Also for Germany, we estimate treatment effects on the number of ratings as a proxy for product sales. Once again, we find a non-statistically significant effect. The point estimate is 145 ratings (95% CI: -680, 970), as shown in Figure A.9 in Appendix A.7.

Based on the hedonic approach, we conclude that there is little evidence to support the notion that consumers value carbon neutrality. This statement follows from the analysis of product prices, product sales, and customer reviews across three major markets, the United States, the United Kingdom, and Germany. The substantial WTP found in the meta-analysis conducted by Carattini et al. (2025), even exceeding many estimates of the social cost of carbon, do not find support when looking at a popular marketplace such as Amazon's.

4 Conclusions

In this paper, we analyze the demand for carbon-neutral products empirically. We focus on online marketplaces in three different countries (United States, United Kingdom, and Germany) and their staggered introduction of carbon-neutral certified products to causally estimate the effect of carbon-neutral labels on product prices. To shed more light on whether consumers value carbon neutrality, we also examine the content of thousands of customer reviews.

We find no evidence of demand for carbon neutrality based on real-world market data from three markets across two continents. In other words, our analysis suggests consumers do not value carbon-neutral labels. Consistently, in customer reviews the fraction of customers that refers to the climate-friendly

nature of carbon-neutral products is minuscule.

Hence, the hedonic analysis of consumer demand for carbon-neutral products suggests that, based on actual market data, the potential of carbon-neutral labeling for climate change mitigation is to be considered with caution, in contrast with the current state of the literature valuing carbon-neutral and low-carbon labels, as summarized in the meta-analysis conducted by Carattini et al. (2025).

Future research could explore the conditions under which consumers notice and value carbon-neutral products. One avenue is to examine how prominently a product's carbon-neutral feature is communicated (i.e., the salience of the label) and how this visibility influences consumer attention and choices. Another avenue is to investigate whether consumers are more inclined to choose carbon-neutral products when the act of purchasing them is socially observable, in other words, when others can see and recognize this pro-environmental choice. Additionally, further studies might gather evidence using observational data or field experiments in a variety of real-world contexts. Expanding the range of research methods and settings in this way would diversify the literature and clarify whether, and under what circumstances, consumers value carbon-neutrality of products.

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.
- Andreoni, J. and B. D. Bernheim (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica* 77(5), 1607–1636.
- Andreoni, J., N. Nikiforakis, and S. Siegenthaler (2021). Predicting social tipping and norm change in controlled experiments. *Proceedings of the National Academy of Sciences* 118(16), e2014893118.
- Ariely, D., A. Bracha, and S. Meier (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review* 99(1), 544–555.
- Arrow, K., R. Solow, P. R. Portney, E. E. Leamer, R. Radner, and H. Schuman (1993). Report of the NOAA Panel on Contingent Valuation. Technical report, National Oceanic and Atmospheric Administration.
- Banzhaf, H. S. (2020). Panel data hedonics: Rosen’s first stage as a ‘sufficient statistic’. *International Economic Review* 61(2), 973–1000.
- Banzhaf, S. (2021). Difference-in-differences hedonics. *Journal of Political Economy* 129(8), 2385–2414.
- Bartling, B., R. A. Weber, and L. Yao (2015). Do markets erode social responsibility? *Quarterly Journal of Economics* 130(1), 219–266.

- Bateman, I., R. Carson, B. Day, M. Hanemann, N. Hanley, T. Hett, M. Jones-Lee, and G. Loomes (2002). *Economic Valuation with Stated Preference Techniques*. Edward Elgar Publishing.
- Bénabou, R. and J. Tirole (2006). Incentives and prosocial behavior. *American Economic Review* 96(5), 1652–1678.
- Bénabou, R. and J. Tirole (2010). Individual and corporate social responsibility. *Economica* 77(305), 1–19.
- Besley, T. and M. Ghatak (2007). Retailing public goods: The economics of corporate social responsibility. *Journal of Public Economics* 91(9), 1645–1663.
- Bicchieri, C. and E. Dimant (2022). Nudging with care: The risks and benefits of social information. *Public Choice* 191(3), 443–464.
- 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.
- Brécard, D. (2017). Consumer misperception of eco-labels, green market structure and welfare. *Journal of Regulatory Economics* 51, 340–364.
- Callaway, B. and P. H. C. Sant’Anna (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics* 225(2), 200–230.
- Carattini, S. and J. Blasch (2024). Nudging when the descriptive norm is low: Evidence from a carbon offsetting field experiment. *Journal of Behavioral and Experimental Economics* 110, 102194.
- Carattini, S., F. Dvorak, I. Logar, and B. Ozdemir-Oluk (2025). Demand for carbon-neutral products. CESifo Working Paper 12232.
- Carattini, S., K. Gillingham, X. Meng, and E. Yoeli (2024). Peer-to-peer solar and social rewards: Evidence from a field experiment. *Journal of Economic Behavior & Organization* 219, 340–370.

- Carattini, S., S. Levin, and A. Tavoni (2019). Cooperation in the climate commons. *Review of Environmental Economics and Policy* 13(2), 227–247.
- Chay, K. and M. Greenstone (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy* 113(2), 376–424.
- Dawes, R. M. and R. H. Thaler (1988). Anomalies: Cooperation. *Journal of Economic Perspectives* 2(3), 187–197.
- Doda, B., C. Gennaioli, A. Gouldson, D. Grover, and R. Sullivan (2016). Are corporate carbon management practices reducing corporate carbon emissions? *Corporate Social Responsibility and Environmental Management* 23(5), 257–270.
- Ellingsen, T. and M. Johannesson (2008). Pride and prejudice: The human side of incentive theory. *American Economic Review* 98(3), 990–1008.
- Falk, A. and N. Szech (2013). Morals and markets. *Science* 340(6133), 707–711.
- Fehr, E., G. Kirchsteiger, and A. Riedl (1993). Does fairness prevent market clearing? An experimental investigation. *Quarterly Journal of Economics* 108(2), 437–459.
- Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics* 114(3), 817–868.
- Fischer, C. and T. P. Lyon (2014). Competing environmental labels. *Journal of Economics & Management Strategy* 23(3), 692–716.
- Fischer, C. and T. P. Lyon (2019). A theory of multitier ecolabel competition. *Journal of the Association of Environmental and Resource Economists* 6(3), 461–501.
- 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.

- Jacobsen, G. D. (2011). The Al Gore effect: An Inconvenient Truth and voluntary carbon offsets. *Journal of Environmental Economics and Management* 61(1), 67–78.
- Johnston, R. J., K. J. Boyle, W. V. Adamowicz, J. Bennett, R. Brouwer, T. A. Cameron, W. M. Hanemann, N. Hanley, M. Ryan, R. Scarpa, R. Tourangeau, and C. A. Vossler (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists* 4(2), 319–405.
- Kesternich, M., A. Löschel, and D. Römer (2016). The long-term impact of matching and rebate subsidies when public goods are impure: Field experimental evidence from the carbon offsetting market. *Journal of Public Economics* 137(C), 70–78.
- Kim, E.-H. and T. Lyon (2011). When does institutional investor activism increase shareholder value?: The Carbon Disclosure Project. *The B.E. Journal of Economic Analysis & Policy* 11(1).
- Kitzmueller, M. and J. Shimshack (2012). Economic perspectives on corporate social responsibility. *Journal of Economic Literature* 50(1), 51–84.
- Kotchen, M. J. (2009). Voluntary provision of public goods for bads: A theory of environmental offsets. *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.
- Meinrenken, C. J., D. Chen, R. A. Esparza, V. Iyer, S. P. Paridis, A. Prasad, and E. Whillas (2022). The carbon catalogue, carbon footprints of 866 commercial products from 8 industry sectors and 5 continents. *Scientific Data* 9(1), 87.
- Mintel (2019). Mintel online retailing report. <https://www.mintel.com/press-centre/the-amazon-effect-nine-in-ten-brits-shop-on-amazon/>. Accessed: 20 March 2025.

- 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.
- Muehlenbachs, L., E. Spiller, and C. Timmins (2015). The housing market impacts of shale gas development. *American Economic Review* 105(12), 3633–3659.
- Muller, L., A. Lacroix, and B. Ruffieux (2019). Environmental labelling and consumption changes: A food choice experiment. *Environmental and Resource Economics* 73(3), 871–897.
- Poret, S. (2019). Label wars: Competition among NGOs as sustainability standard setters. *Journal of Economic Behavior & Organization* 160, 1–18.
- Rodemeier, M. (2023). Willingness to pay for carbon mitigation: Field evidence from the market for carbon offsets. SSRN Scholarly Paper 4360822, Rochester, NY.
- Rogelj, J., O. Geden, A. Cowie, and A. Reisinger (2021). Net-zero emissions targets are vague: Three ways to fix. *Nature* 591(7850).
- Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy* 82(1), 34–55.
- Shleifer, A. (2004). Does competition destroy ethical behavior? *American Economic Review* 94(2), 414–418.
- Smith, V. K. and W. H. Desvousges (1986). The value of avoiding a Lulu: Hazardous waste disposal sites. *The Review of Economics and Statistics* 68(2), 293–299.
- Sparkman, G. and G. M. Walton (2017). Dynamic norms promote sustainable behavior, even if it is counternormative. *Psychological Science* 28(11), 1663–1674.

- Spencer, G., S. Carattini, and R. B. Howarth (2019). Short-term interventions for long-term change: Spreading stable green norms in networks. *Review of Behavioral Economics* 6(1), 53–93.
- Statista (2022a). Share of online shoppers purchasing on Amazon in Germany in 2022, by frequency. <https://www.statista.com/statistics/1368653/amazon-shopping-by-frequency-germany/>. Accessed: 11 March 2025.
- Statista (2022b). Total number of households with Amazon Prime subscription in the USA. <https://www.statista.com/statistics/861060/total-number-of-households-amazon-prime-subscription-usa/>. Accessed: 21 July 2022.
- Stichting Sustainability Impact Metrics (2024). Idemat Database. <https://www.ecocostsvalue.com/data-tools-books/>. Accessed: 01 May 2025.
- US Census Bureau (2025). Total households. FRED, Federal Reserve Bank of St. Louis, <https://fred.stlouisfed.org/data/TTLHH>, Accessed: 8 October 2025.

Appendix

A.1 Carbon neutrality on Amazon's marketplace

Amazon, in collaboration with Global Optimism, an organization focused on environmental and social change, initiated the Climate Pledge in 2019. Amazon is a co-founder and participant, starting this initiative to promote the sale of more sustainable products among its vendors.

Amazon uses the term “Climate Pledge Friendly labels” to refer to various aspects of sustainability, not just those related to carbon neutrality. These other sustainability labels cover various aspects, including energy efficiency, recycling, organic certification, fair trade, animal welfare, and CO₂ reduction. We focus on carbon-neutral labels, in particular excluding a related label implying much less commitment, the “Reducing CO₂” label.

The list of Climate Pledge Friendly labels evolves over time. As of March 2023, the labels focusing on sustainability aspects other than carbon neutrality on Amazon.com included: BIFMA LEVEL, Blue Angel, Bluesign, Certified Animal Welfare Approved, Compact by Design, Cradle to Cradle Certified, ECOLOGO, ENERGY STAR Most Efficient, EPEAT, EWG Verified, Fair Trade Certified, Fairtrade International, Fair for Life, The Forest Stewardship Council, Global Organic Textile Standard, Global Recycled Standard, GreenCircle Certified, Green Seal, Higg Index Materials, Made in Green by Oeko-Tex, MADE SAFE, Natrue, Nordic Swan Ecolabel, Organic Content Standard 100, Organic Content Standard Blended, Pre-owned Certified, Rainforest Alliance, Recycled Claim Standard 100, Recycled Claim Standard Blended, Reducing CO₂, Regenerative Organic Certified, Responsible Wool Standard, Soil Association, STANDARD 100 by OEKO-TEX, TCO Certified, USDA Organic, U.S. EPA Safer Choice, and WaterSense.

A.2 Changes in treatment status

We successfully scraped data for the United States from March 2023 through November 2024, and from May 2024 through November 2024 for the United Kingdom and Germany.

Table A.1 summarizes the restrictions used to identify changes in treatment status and the corresponding control products, as explained in Section 2. Table A.2 shows that treated and control products are balanced on the initial price at the start of the panel.

Tables A.3, A.4, and A.5 provide detailed information on the timing of changes in products' treatment status for the United States, United Kingdom, and Germany, including ASINs,³ product categories (as defined by Amazon), the dates the products were first identified as carbon-neutral, and the product prices at the start and at the end of the panel.

³ASIN numbers can be added after "amazon.com/dp/," "amazon.co.uk/dp/," "amazon.de/dp/" to find a specific product on the respective sites for the United States, United Kingdom, and Germany.

Category	Restrictions
All products	Sufficient price data: More than 25% of price observations are available. Sufficient availability: ‘Currently unavailable’ less than 50% of the time.
Control products	Never treated: No carbon-neutral label for the entire time series. No other labels: Cannot have additional Climate Pledge Friendly labels that we identify and track at the start of the panel (e.g., organic, fairtrade). Category matching: Matched to treated product categories. Higher-level categories used if fewer than 10 suitable controls are available.
Treated products	Initial untreated sequence: At least 3 weekly untreated observations before treatment. Treatment: Status changes from 0 to 1 (receives a carbon-neutral label). Consistent treatment: Treated at least 90% of the time. Frequent availability: Meets the 50% availability threshold after receiving the treatment. No other labels: No additional Climate Pledge Friendly labels that we identify and track at the start of the panel (e.g., organic, fairtrade).

Table A.1: Restrictions to define treated and control products

This table shows the sample restrictions used to identify changes in treatment status and suitable controls from our product panel across Amazon’s three marketplaces.

	Treated	Control	T-stat (Welch p)
Initial Price	31.2	31.54	-0.167 (p=0.867)

Table A.2: Balance

The table reports the mean initial prices for treated products (those receiving carbon-neutral labels) and control products (without a label). The two-sample t-test is used to test covariate balance between the two groups.

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in Mar 2023 (USD)	Price in Nov 2024 (USD)
B00M48YNOU	Grocery & Gourmet Food	2023-04-03	12.48	21.99
B0771VVJRW	Beauty & Personal Care	2023-04-10	19.99	19.99
B0B6V9D89F	Cell Phones & Accessories	2023-04-10	13.06	6.99
B0BF9N2RP2	Cell Phones & Accessories	2023-04-24	14.66	14.24
B095BZT4SD	Electronics	2023-05-08	11.99	8.26
B09NKJ5MCV	Electronics	2023-05-08	15.32	8.34
B09NNJYGB4	Video Games	2023-05-08	229.62	156.35
B08M5L57KT	Electronics	2023-05-08	23.99	22.24
B07M91R8PN	Electronics	2023-05-08	17.99	16.99
B0B5KDNTWS	Health & Household	2023-05-15	58.00	50.16
B09W36YKY7	Health & Household	2023-05-15	42.99	42.99
B09W35DHLH	Health & Household	2023-05-15	36.99	31.81
B0948ZFQFR	Electronics	2023-05-22	25.99	24.04
B09W363MVD	Health & Household	2023-05-22	42.99	36.79
B09CTLNCFG	Electronics	2023-06-12	19.99	18.99
B0B1BSLRGT	Cell Phones & Accessories	2023-06-19	33.29	22.79
B08J3K4N15	Electronics	2023-06-19	23.99	26.49
B007N3H26M	Electronics	2023-06-19	56.99	46.66
B0073UBRP2	Electronics	2023-06-19	23.99	25.64
B014G1G10Q	Beauty & Personal Care	2023-06-19	27.95	25.46
B0BHMKVFI P	Video Games	2023-06-26	31.66	23.90
B0BL66ZW9H	Video Games	2023-06-26	59.89	56.33
B0BL67RHS6	Video Games	2023-06-26	183.76	148.98
B0BL6G3NRJ	Video Games	2023-06-26	136.36	107.84
B0BHMMH9KM	Video Games	2023-06-26	34.99	28.16
B07GSLHXXQ	Electronics	2023-07-03	9.98	10.23
B0B1TVD3HK	Electronics	2023-07-03	20.99	19.99
B0B1TQTNMC	Electronics	2023-07-03	20.99	19.59
B0BLRGQF3M	Cell Phones & Accessories	2023-07-03	9.98	15.98
B093PT44N1	Electronics	2023-07-10	30.99	24.29
B0BJFFGLHM	Electronics	2023-07-10	27.66	29.99
B07GZFJ4G5	Cell Phones & Accessories	2023-07-24	36.37	37.32

Table A.3: Timing of the changes in treatment status (United States)

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in Mar 2023 (USD)	Price in Nov 2024 (USD)
B08L3WX26S	Electronics	2023-07-24	24.99	19.94
B09FDJFJ6Z	Electronics	2023-07-24	7.99	6.40
B07RQRMGKB	Electronics	2023-07-24	11.32	9.74
B08GJ3F11N	Beauty & Personal Care	2023-07-24	11.99	9.99
B085ZXC2HS	Cell Phones & Accessories	2023-07-24	15.99	14.99
B09YFH1C8X	Beauty & Personal Care	2023-07-24	15.09	16.99
B086JBZW48	Health & Household	2023-07-31	12.97	14.99
B08PHY1PJF	Health & Household	2023-07-31	12.99	8.66
B089LDX88M	Cell Phones & Accessories	2023-07-31	15.03	11.99
B09L9RKN7W	Health & Household	2023-07-31	10.99	9.99
B08P27Y27M	Health & Household	2023-07-31	12.99	10.79
B0BLK79BZ2	Electronics	2023-08-14	28.95	23.99
B0BGHM8SY4	Electronics	2023-08-14	11.90	8.88
B09HKX6HRB	Electronics	2023-08-28	10.99	8.52
B08BR4V18G	Electronics	2023-08-28	14.99	14.99
B086QW23YD	Electronics	2023-08-28	12.99	12.48
B0BBSP2JNQ	Beauty & Personal Care	2023-08-28	19.92	21.75
B09J1DFTTV	Electronics	2023-09-04	19.91	17.99
B0BGHRM5DV	Electronics	2023-09-04	18.99	15.99
B093T7GQWB	Electronics	2023-09-04	18.79	9.95
B074KV9TT4	Electronics	2023-09-04	37.32	27.39
B088RHCSG3	Electronics	2023-09-04	14.99	12.96
B09J1FYF9V	Electronics	2023-09-04	20.95	22.47
B0BLTDYG2B	Cell Phones & Accessories	2023-09-11	14.99	14.71
B08XQQ5XTZ	Cell Phones & Accessories	2023-09-18	27.66	16.99
B0831BF1FH	Cell Phones & Accessories	2023-09-18	28.49	23.99
B08883JK8Y	Electronics	2023-09-18	33.99	29.86
B08K8S4ZDW	Electronics	2023-09-18	45.99	41.49
B095GJDXNG	Electronics	2023-09-18	31.99	27.32
B08RDF9B3F	Cell Phones & Accessories	2023-09-18	12.38	16.99
B09PR1BTM7	Tools & Home Improvement	2023-09-18	15.99	13.66
B08GG42WXY	Tools & Home Improvement	2023-09-18	11.19	8.79

Table A.3: Timing of the changes in treatment status (United States)

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in Mar 2023 (USD)	Price in Nov 2024 (USD)
B0BGN9R72N	Tools & Home Improvement	2023-09-18	10.99	8.09
B0BHH7M4YJ	Cell Phones & Accessories	2023-09-18	9.99	8.62
B09PV827TS	Electronics	2023-09-18	31.99	27.99
B0BHHVN541	Cell Phones & Accessories	2023-09-18	9.99	8.62
B0BM4QL882	Cell Phones & Accessories	2023-09-18	9.99	8.99
B07Y9G18V7	Electronics	2023-09-18	36.32	29.99
B0BK99PT9K	Electronics	2023-09-18	32.99	28.49
B0BC21JFYH	Cell Phones & Accessories	2023-09-18	10.99	8.62
B0BM4LPT4Y	Cell Phones & Accessories	2023-09-25	9.99	8.99
B08K8DNVB4	Cell Phones & Accessories	2023-10-02	43.88	36.25
B0B96PKNVL	Video Games	2023-10-16	20.32	18.66
B0B66RHD7B	Video Games	2023-10-16	28.49	27.99
B01M11FLUJ	Office Products	2023-10-30	14.17	9.97
B016XTADG2	Electronics	2023-11-06	25.99	25.99
B01MTB55WH	Electronics	2023-11-06	36.66	33.99
B07Z4RF1D3	Electronics	2023-11-06	16.73	16.15
B0B1MCHS14	Electronics	2023-11-13	26.99	15.49
B096BCMK8N	Electronics	2023-11-13	27.32	16.39
B087LRK3H4	Electronics	2023-11-13	17.99	12.99
B093C2B4K3	Electronics	2023-11-13	19.32	16.99
B082Y6YDZZ	Electronics	2023-11-20	64.98	58.45
B01BT02Q88	Beauty & Personal Care	2023-11-20	14.99	14.99
B0BLBQ9G2C	Cell Phones & Accessories	2023-11-20	26.99	26.99
B0BRC415HH	Cell Phones & Accessories	2023-11-20	21.99	20.99
B0BTRTFK4S	Cell Phones & Accessories	2023-11-20	26.99	26.99
B07J4TNYV8	Electronics	2023-11-20	139.99	128.32
B0B2BSQQL7	Electronics	2023-11-20	89.89	79.99
B0BQB8JNFB	Cell Phones & Accessories	2023-11-20	23.99	24.69
B0BP7HG18T	Electronics	2023-11-20	64.99	61.72
B07JR1XZ78	Electronics	2023-11-20	84.99	71.66
B07QXV6N1B	Cell Phones & Accessories	2023-11-27	21.99	22.32
B0BJZ5VMD6	Cell Phones & Accessories	2023-11-27	26.99	26.99

Table A.3: Timing of the changes in treatment status (United States)

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in Mar 2023 (USD)	Price in Nov 2024 (USD)
B0874M3KW4	Electronics	2023-11-27	69.99	46.99
B07ZCRWPPV	Electronics	2023-12-04	66.99	62.52
B08SJDXF73	Electronics	2023-12-11	10.99	9.97
B0BRKVSXG4	Electronics	2023-12-18	109.99	106.36
B07CSBYNWG	Beauty & Personal Care	2023-12-25	19.99	16.95
B0B7DP9CGN	Health & Household	2023-12-25	24.69	24.30
B09P37WCS4	Electronics	2024-01-22	15.99	15.79
B09X9JCLR7	Video Games	2024-01-29	29.99	26.46
B09P5BBPVY	Video Games	2024-01-29	29.99	27.22
B09YR1J35N	Electronics	2024-02-05	27.99	26.99
B09WDH6K1T	Toys & Games	2024-02-05	40.32	46.66
B0BJ2D5X2R	Cell Phones & Accessories	2024-02-05	74.10	59.20
B0BHZ6MWC1	Cell Phones & Accessories	2024-02-05	68.40	59.20
B0BHZ84Z9C	Cell Phones & Accessories	2024-02-05	68.40	59.20
B0BQ35Q43S	Cell Phones & Accessories	2024-02-05	74.10	49.60
B0BQ2L5KJJ	Cell Phones & Accessories	2024-02-05	94.73	56.86
B0BN1P1KKL	Cell Phones & Accessories	2024-02-05	68.40	53.65
B0BRY2FPMK	Cell Phones & Accessories	2024-02-05	74.10	59.20
B08FTDWPTX	Electronics	2024-02-12	11.99	11.69
B088NGVY4C	Electronics	2024-02-12	12.78	11.19
B0BRXY8RH7	Cell Phones & Accessories	2024-02-12	64.60	54.57
B089PYQQSQ	Electronics	2024-02-19	17.97	17.07
B09ZQQPNXD	Beauty & Personal Care	2024-02-19	11.99	11.54
B07HC4NBQ8	Video Games	2024-02-19	30.89	29.52
B08SJ5Z8JL	Video Games	2024-02-19	19.99	18.32
B01NAI2TXC	Video Games	2024-02-19	55.79	55.36
B0BHSVQXHG	Cell Phones & Accessories	2024-02-19	59.93	62.00
B0BQ2QJCQQ	Cell Phones & Accessories	2024-02-26	77.90	50.45
B07VCS8QTK	Electronics	2024-03-04	17.99	12.83
B09BN32Y86	Electronics	2024-03-04	11.32	7.89
B0B9SK4WD3	Cell Phones & Accessories	2024-03-04	10.98	16.98
B09JLLD5QH	Electronics	2024-03-11	15.95	9.61

Table A.3: Timing of the changes in treatment status (United States)

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in Mar 2023 (USD)	Price in Nov 2024 (USD)
B08YNKKB7M	Electronics	2024-03-11	18.79	9.95
B09R6JP7K5	Electronics	2024-03-11	18.99	15.99
B0BJV4888V	Electronics	2024-03-11	14.99	10.49
B0B3R73C5F	Electronics	2024-03-11	19.99	15.99
B09JBBPC9K	Electronics	2024-03-11	17.95	15.79
B0BGH6L5B6	Electronics	2024-03-11	18.99	14.99
B09KXSJZ6K	Electronics	2024-03-11	21.99	20.66
B09KZFH1JP	Electronics	2024-03-11	22.99	21.99
B09J4RQFK7	Electronics	2024-03-11	25.99	24.17
B084FSYC1B	Toys & Games	2024-03-11	19.99	29.99
B082ZYNNMC8	Health & Household	2024-03-18	9.49	9.66
B08FFFGHJF	Toys & Games	2024-03-18	22.24	27.99
B07H8TJMX7	Electronics	2024-06-03	10.40	9.98
B07SQP1GHC	Electronics	2024-06-03	35.99	35.99
B0BJ7GST13	Cell Phones & Accessories	2024-06-03	13.98	16.98
B0B9SP1CZ2	Cell Phones & Accessories	2024-06-03	11.98	16.98
B093L2Y8KQ	Electronics	2024-06-03	9.95	7.74
B09VPBF8NY	Video Games	2024-06-03	19.99	19.99
B07PMBCTSY	Electronics	2024-06-03	66.66	55.32
B09BYJLZ16	Electronics	2024-06-03	32.66	24.99
B0B63LNZBW	Electronics	2024-06-03	29.99	13.99
B09MHFSSFB	Tools & Home Improvement	2024-06-03	11.67	10.99
B07L2LS9SK	Electronics	2024-06-03	33.96	32.26
B0B73JCBRZ	Electronics	2024-06-03	19.99	20.99
B09HQFY88	Tools & Home Improvement	2024-06-03	11.32	10.83
B09YRVDWCP	Tools & Home Improvement	2024-06-03	15.99	13.49
B09YRVFVK	Tools & Home Improvement	2024-06-03	15.99	12.99
B09ZPCC1J3	Tools & Home Improvement	2024-06-03	15.99	12.24
B09ZPDMMRW	Tools & Home Improvement	2024-06-03	15.99	12.24
B0BMLFV2DJ	Electronics	2024-06-03	25.99	24.26
B09TDHLMXZ	Electronics	2024-06-03	28.49	27.99

Table A.3: Timing of the changes in treatment status (United States)

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in Mar 2023 (USD)	Price in Nov 2024 (USD)
B07P7MB88J	Electronics	2024-06-03	35.14	34.19
B09BN47YHV	Electronics	2024-06-03	22.99	35.64
B08R5CCRFD	Electronics	2024-06-03	43.98	37.99
B0BBG5RRXF	Electronics	2024-06-03	28.49	27.35
B0B2K2SMH7	Cell Phones & Accessories	2024-06-03	13.99	14.91
B09SW6L7H8	Cell Phones & Accessories	2024-06-03	15.99	12.24
B0BPYJWMP7	Cell Phones & Accessories	2024-06-03	16.98	17.98
B095VLRB2J	Electronics	2024-06-03	32.99	27.99
B0BK1T5PF4	Cell Phones & Accessories	2024-06-03	16.98	17.98
B08VVWRFLS	Electronics	2024-06-03	46.99	47.48
B0B8X44B4Y	Cell Phones & Accessories	2024-06-03	15.99	13.29
B07TS6R1SF	Electronics	2024-06-03	30.79	32.66
B0921JJMZT	Electronics	2024-06-03	15.99	13.06
B09ZTXVNVD	Cell Phones & Accessories	2024-06-03	28.99	16.19
B01H6GUCCQ	Video Games	2024-06-17	25.99	21.91
B0B16VD9RQ	Beauty & Personal Care	2024-06-17	19.99	19.98
B09Q7LTBTR	Beauty & Personal Care	2024-06-17	34.18	27.55
B081JP3MJK	Video Games	2024-06-17	33.99	28.86
B0BG21S94B	Electronics	2024-06-24	52.49	48.66
B0BNPBTJDP	Video Games	2024-06-24	42.99	37.85
B07KCRTN9Q	Video Games	2024-07-08	44.99	37.96
B098S48QWM	Electronics	2024-07-22	10.58	9.98
B01MTVC775	Electronics	2024-07-29	56.99	49.10
B0BPCHQBS7	Electronics	2024-07-29	119.99	99.99
B076Q6442Z	Beauty & Personal Care	2024-08-05	7.98	6.99
B0811RYZ2J	Cell Phones & Accessories	2024-08-12	17.99	15.99
B0811RH5MF	Cell Phones & Accessories	2024-08-12	16.99	13.19
B08F79BQD3	Cell Phones & Accessories	2024-08-12	16.99	15.74
B07L793HPW	Cell Phones & Accessories	2024-08-12	17.99	13.29
B09D941CFQ	Cell Phones & Accessories	2024-08-12	16.99	16.64
B07GBXVX7W	Cell Phones & Accessories	2024-08-12	15.99	16.99

Table A.3: Timing of the changes in treatment status (United States)

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in Mar 2023 (USD)	Price in Nov 2024 (USD)
B07THZ6MVP	Cell Phones & Accessories	2024-08-12	16.99	11.42
B09D952JQQ	Cell Phones & Accessories	2024-08-12	17.99	18.24
B0BCJ8Q1QN	Cell Phones & Accessories	2024-08-12	18.99	17.99
B08M636GG3	Cell Phones & Accessories	2024-08-12	16.42	16.99
B0BBKW53L3	Cell Phones & Accessories	2024-08-12	22.99	8.99
B01K2UMMI0	Beauty & Personal Care	2024-08-19	24.99	22.39
B08L3K9LC6	Health & Household	2024-08-19	19.99	32.80
B081S71B77	Health & Household	2024-08-19	19.97	23.95
B0932BCM2T	Musical Instruments	2024-08-26	31.30	30.75
B07XVCP7F5	Video Games	2024-08-26	66.99	55.59
B08JCV3J5P	Video Games	2024-08-26	79.99	75.97
B07R1R1MKW	Health & Household	2024-08-26	17.97	23.39
B08FDPW8KR	Electronics	2024-09-09	35.32	35.99
B0B979GR4Z	Electronics	2024-09-16	31.32	27.49
B00N1YTJRC	Health & Household	2024-09-23	38.37	41.74
B0B6BW56FF	Safety & Security	2024-10-07	8.99	8.35
B0B6BVHS4W	Safety & Security	2024-10-07	9.99	8.99

Table A.3: Timing of the changes in treatment status (United States)

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in May 2024 (GBP)	Price in Nov 2024 (GBP)
B0773F8S74	Mobile Phones & Communication	2024-06-18	15.99	15.99
B0C3C8M5X6	Accessories	2024-07-16	9.99	9.99
B09BW1T7X2	Headphones, Earphones & Accessories	2024-07-16	129.95	119.71
B09BW1QVVT	Headphones, Earphones & Accessories	2024-07-16	129.95	119.71
B0CD226VG1	Arts & Crafts	2024-07-30	6.29	6.73
B0BZCMKWGV	Mobile Phones & Communication	2024-08-06	12.99	12.99
B0CENVGM3R	Headphones, Earphones & Accessories	2024-08-06	32.99	30.99
B0CENVHKVVL	Headphones, Earphones & Accessories	2024-08-06	29.99	40.49
B07D6N526S	Hi-Fi & Home Audio	2024-08-06	10.99	19.99
B09996WC24	Mobile Phones & Communication	2024-08-06	13.99	13.46
B09D952JQQ	Mobile Phones & Communication	2024-08-06	11.89	14.96
B07GBXVX7W	Mobile Phones & Communication	2024-08-06	12.99	13.24
B08F7Q8Y5W	Mobile Phones & Communication	2024-08-06	13.99	13.08
B081H1N3PJ	Mobile Phones & Communication	2024-08-06	13.99	13.29
B07Z1CVWD2	Mobile Phones & Communication	2024-08-06	19.99	19.99
B099968ZQK	Mobile Phones & Communication	2024-08-06	13.99	13.21
B07TFD7KR3	Mobile Phones & Communication	2024-08-06	13.99	12.99
B09D94T83V	Mobile Phones & Communication	2024-08-06	14.99	14.43
B0BCHZFSFD	Mobile Phones & Communication	2024-08-06	13.99	10.92
B0CHRFQPCC	Mobile Phones & Communication	2024-08-06	13.59	15.99
B0BCJ4KZWG	Mobile Phones & Communication	2024-08-06	12.99	12.99
B0BN54PK3X	Mobile Phones & Communication	2024-08-06	13.99	11.24
B0BN5HTP3K	Mobile Phones & Communication	2024-08-06	15.99	15.19
B08R78YY9M	Mobile Phones & Communication	2024-08-06	13.99	13.99
B08DHT9KJZ	Wearable Technology	2024-08-06	7.99	6.80
B0BZH6ZF7M	Sports	2024-08-06	15.99	16.99
B0C8N8D6GF	Sports	2024-08-06	17.99	16.14
B0CB5XH9J5	Mobile Phones & Communication	2024-08-06	23.99	24.53
B09L4GBWDV	Wearable Technology	2024-08-13	7.99	6.80
B0CCN9Q27D	Gardening	2024-08-13	65.13	32.28
B0C8N7MKF6	Mobile Phones & Communication	2024-08-13	12.99	12.12
B08ZXQLJN9	Accessories	2024-08-13	19.99	21.49

Table A.4: Timing of the changes in treatment status (United Kingdom)

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in MaY 2024 (GBP)	Price in Nov 2024 (GBP)
B0CHLXSBND	Mobile Phones & Communication	2024-08-13	29.40	21.16
B0BHYKBQLP	Gardening	2024-08-13	29.22	28.94
B0BK84XP9K	Microphones	2024-08-27	49.99	46.24
B09B3LJVfJ	Microphones	2024-08-27	36.99	35.60
B0CHNSY4S8	Mobile Phones & Communication	2024-09-17	43.99	37.49
B0CHYL54H8	Mobile Phones & Communication	2024-09-17	33.43	41.79
B0CHYKQQ6S	Mobile Phones & Communication	2024-09-17	43.99	41.79
B0BDFDRCKV	Wearable Technology	2024-09-24	6.58	7.29
B09H2JD8F8	Wearable Technology	2024-09-24	6.49	7.58
B08GCNLQT6	Wearable Technology	2024-09-24	6.49	7.58
B0BPNYZ8NN	Wearable Technology	2024-09-24	6.49	7.62
B0BPNT6XYY	Wearable Technology	2024-09-24	6.49	7.62
B09KN7CNRZ	Wearable Technology	2024-09-24	6.49	7.58
B0CHYBKQPM	Mobile Phones & Communication	2024-10-15	16.99	16.34
B0C273FR2T	Mobile Phones & Communication	2024-10-15	17.99	17.84
B082SNXJ4G	Mobile Phones & Communication	2024-10-15	21.99	20.89
B0915H9JD8	Mobile Phones & Communication	2024-10-15	19.99	19.17
B0CB1C69BY	Mobile Phones & Communication	2024-10-15	16.98	17.73

Table A.4: Timing of the changes in treatment status (United Kingdom)

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in May 2024 (EUR)	Price in Nov 2024 (EUR)
B096FX9226	Electronics & Photo	2024-06-19	21.99	20.52
B09KLYT52T	Electronics & Photo	2024-06-19	9.99	9.99
B0C9ZVN154	Computer & Accessories	2024-06-26	49.99	45.96
B0C748DZRH	Electronics & Photo	2024-06-26	22.99	32.99
B0C7Q55ZM5	Computer & Accessories	2024-06-26	26.99	19.44
B0C6KF5HKT	Computer & Accessories	2024-06-26	16.59	14.79
B0CD7LYR4C	Computer & Accessories	2024-06-26	24.64	26.99
B09VB3WXQS	Computer & Accessories	2024-06-26	13.99	12.73
B0CTQ2346R	Computer & Accessories	2024-06-26	30.99	26.05
B0B31FVQPQ	Computer & Accessories	2024-06-26	16.99	15.52
B0C68N2BH1	Computer & Accessories	2024-06-26	16.99	15.91
B0BW95FDB6	Computer & Accessories	2024-06-26	19.99	18.97
B09X1H9VNZ	Computer & Accessories	2024-06-26	15.99	15.99
B09X1FZQPX	Computer & Accessories	2024-06-26	23.99	24.99
B0C9DJK1QX	Computer & Accessories	2024-06-26	14.09	13.50
B0CMCGLT21	Computer & Accessories	2024-06-26	39.67	45.99
B0CTQBG9YP	Computer & Accessories	2024-06-26	28.99	25.54
B0BVL3JNJ	Computer & Accessories	2024-06-26	17.99	15.99
B0CTQ94TV3	Computer & Accessories	2024-06-26	39.99	35.09
B0C3LB86PN	Computer & Accessories	2024-06-26	54.99	44.99
B0BVLX7BXW	Computer & Accessories	2024-06-26	16.09	14.72
B0CGZVDMKG	Computer & Accessories	2024-07-10	55.99	46.63
B0BM5XSKDR	Computer & Accessories	2024-07-17	9.99	25.77
B075V27G2R	Computer & Accessories	2024-07-17	11.19	12.59
B07F2YJRN2	Computer & Accessories	2024-07-17	55.38	43.63
B0C3L93F2Q	Stationery & Office Supplies	2024-07-24	25.49	27.74
B08M5PSFWF	Computer & Accessories	2024-07-24	7.89	9.46
B0CSYMR7GZ	Electronics & Photo	2024-07-24	32.38	37.80
B0CH7VQJKH	Games	2024-08-07	23.95	19.32
B07JNJGM1G	Electronics & Photo	2024-08-07	13.98	12.50
B07GBXVX7W	Electronics & Photo	2024-08-07	13.95	13.45
B09996WC24	Electronics & Photo	2024-08-07	13.56	15.95

Table A.5: Timing of the changes in treatment status (Germany)

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in May 2024 (EUR)	Price in Nov 2024 (EUR)
B0CH346J32	Games	2024-08-07	23.98	19.99
B0CN8LRLH	Games	2024-08-07	39.99	39.99
B0CN8GG24V	Games	2024-08-07	39.99	37.19
B0BR5JG23N	Games	2024-08-07	19.99	19.99
B0CL6LL5SF	Electronics & Photo	2024-08-14	12.74	14.99
B0CPDZT72H	Electronics & Photo	2024-08-14	12.74	13.87
B08ZXQLJN9	Computer & Accessories	2024-08-14	29.99	29.99
B01MY4L8BV	Health & Personal Care	2024-08-21	9.95	9.95
B0BLND9W1C	Electronics & Photo	2024-09-04	17.68	14.09
B0873358VL	Computer & Accessories	2024-10-02	15.99	20.47
B07D5QDZTY	Computer & Accessories	2024-10-02	15.99	21.74
B0CN8G45K1	Electronics & Photo	2024-10-02	59.99	55.93
B0CGM1DJZ8	Computer & Accessories	2024-10-02	36.97	43.24
B0CGLXNPHL	Computer & Accessories	2024-10-02	47.99	58.13
B073RY7XD7	Computer & Accessories	2024-10-02	14.99	19.49
B0CGR7H7BT	Computer & Accessories	2024-10-02	41.62	55.89
B0CGHRZD4Q	Computer & Accessories	2024-10-02	45.96	48.52
B0CGHF5G95	Computer & Accessories	2024-10-02	47.97	61.49
B09XR315M4	Computer & Accessories	2024-10-02	59.01	42.59
B0CPPDCGBT	Computer & Accessories	2024-10-02	47.95	62.78
B0CB3DRLCT	Electronics & Photo	2024-10-16	15.99	15.93
B0B3MKD39C	Electronics & Photo	2024-10-16	23.99	24.19
B0C273FR2T	Electronics & Photo	2024-10-16	19.99	17.99
B0CB1C69BY	Electronics & Photo	2024-10-16	19.99	18.99
B0C3C9GCGY	Computer & Accessories	2024-10-16	10.99	12.99
B0C3C8CWYQ	Computer & Accessories	2024-10-16	10.99	12.99
B08V8NPY3Y	Electronics & Photo	2024-10-16	16.99	16.99
B0BV6NT6CH	Computer & Accessories	2024-10-23	10.99	25.95
B08NSJR3TN	Computer & Accessories	2024-10-23	55.91	60.57
B09GG64C6G	Toys	2024-10-30	17.99	17.09
B0CL94BQ56	Electronics & Photo	2024-10-30	17.97	11.64
B0749MNW3N	Health & Personal Care	2024-11-06	8.49	9.99

Table A.5: Timing of the changes in treatment status (Germany)

Product ASIN	Category	First Treated (yyyy-mm-dd)	Price in May 2024 (EUR)	Price in Nov 2024 (EUR)
B0BVRD97MT	Electronics & Photo	2024-11-06	6.95	6.43
B09VH599VG	Electronics & Photo	2024-11-06	6.89	6.69
B09X777HXL	Electronics & Photo	2024-11-06	6.89	6.68
B09F3P3DQD	Electronics & Photo	2024-11-06	6.95	6.84
B0CJMJS3DS	Electronics & Photo	2024-11-06	6.89	6.67
B0BVVN31X8	Computer & Accessories	2024-11-06	10.95	8.99
B09ZYJB6RB	Electronics & Photo	2024-11-06	6.89	6.95
B09F3RCJJR	Electronics & Photo	2024-11-06	6.95	6.69
B0BMQPWBK6	Electronics & Photo	2024-11-06	6.89	6.43
B0C4LRY247	Electronics & Photo	2024-11-06	6.95	6.95
B0BD5J9M98	Electronics & Photo	2024-11-06	6.95	6.68
B0C4LTQZ62	Electronics & Photo	2024-11-06	6.89	6.95
B09FYBKN69	Computer & Accessories	2024-11-06	9.89	9.31
B09FYCQDTK	Computer & Accessories	2024-11-06	9.89	9.11
B09QT5H713	Computer & Accessories	2024-11-06	9.89	9.23
B09QT6M1WS	Computer & Accessories	2024-11-06	9.89	8.99
B0CG1R72SM	Computer & Accessories	2024-11-06	17.99	15.99
B0C84Q8TFV	Computer & Accessories	2024-11-06	17.95	17.95
B09X768KKV	Electronics & Photo	2024-11-06	6.89	6.95
B0BD5D5JHQ	Electronics & Photo	2024-11-06	6.89	6.69

Table A.5: Timing of the changes in treatment status (Germany)

A.3 Carbon footprint estimation of quasi-natural experiments

Information about a product’s mitigated carbon footprint (that is, the amount reduced or offset) may be readily available on Amazon or can be estimated via external sources. Overall, we rely on four sources for carbon footprint estimation: the product’s Amazon webpage, ClimatePartner’s webpage, Carbon Catalogue (Meinrenken et al., 2022) and the Idemat database (Stichting Sustainability Impact Metrics, 2024). In practice, to estimate a product’s carbon footprint we follow these steps. First, we check whether carbon footprint information is available on the product’s Amazon page or ClimatePartner’s page. ClimatePartner lists some of its product footprint information on the product webpage and allows consumers to search for a product’s carbon footprint using its identification number on the company’s website. If not, we verify whether the exact product or a similar one is listed in the Carbon Catalogue. If it is, we rely on that figure. Otherwise, we identify the product’s main material (often specified first in online descriptions) and consult the Idemat database for relevant data, combining these two databases. Because carbon footprint data is typically provided in terms of carbon dioxide (or equivalent) emissions per kilogram of product, we multiply the reported figure by the product’s actual weight to determine its total carbon footprint. Caveats apply. In particular, we could not estimate the carbon footprint for 8 out of 207 quasi-natural experiments in the US, 1 out of 50 experiments in the UK, and 8 out of 84 experiments in Germany.

A.4 Difference-in-difference analysis

This section presents the output of the difference-in-difference model (Callaway and Sant’Anna, 2021) described in Section 3.1. Tables A.6 and A.7 report the dynamic and calendar effects for the United States, respectively.

Months -18 to 0 (Pre-treatment)				Months 1 to 19 (Post-treatment)			
Months	Estimate (%)	Lower 95% CI	Upper 95% CI	Months	Estimate (%)	Lower 95% CI	Upper 95% CI
-18	-1.16	-5.05	2.73	1	-0.93	-2.82	0.97
-17	-4.80	-25.12	15.52	2	0.37	-2.07	2.81
-16	-1.15	-3.76	1.47	3	-0.30	-2.71	2.12
-15	0.09	-1.54	1.71	4	-0.67	-3.41	2.06
-14	-0.58	-3.22	2.05	5	-1.82	-4.98	1.33
-13	0.44	-1.69	2.56	6	-3.04	-7.10	1.02
-12	-0.08	-1.87	1.71	7	-2.64	-6.78	1.50
-11	0.18	-1.26	1.63	8	-2.25	-6.66	2.15
-10	0.28	-1.08	1.64	9	-3.39	-7.69	0.90
-9	-0.91	-2.61	0.79	10	-3.78	-9.22	1.67
-8	1.65	0.15	3.15	11	-3.67	-9.25	1.90
-7	-1.02	-3.09	1.05	12	-2.57	-8.58	3.44
-6	1.93	-0.14	4.00	13	-2.69	-8.96	3.58
-5	-0.35	-1.99	1.29	14	-4.67	-10.52	1.17
-4	-0.50	-1.79	0.79	15	-5.75	-13.39	1.90
-3	-0.40	-1.65	0.85	16	-4.60	-14.78	5.58
-2	-0.16	-1.72	1.41	17	-5.77	-19.34	7.80
-1	-1.10	-2.49	0.30	18	-6.55	-29.68	16.58
0	0.14	-1.20	1.47	19	10.46	-51.74	72.67

Table A.6: Dynamic effects (US)

This table presents the effect of carbon-neutral labels on product price changes, expressed in percentages, relative to the product’s price in March 2023 on Amazon’s U.S. marketplace. Furthermore, the plot displays the upper and lower 95% confidence intervals based on 1,000 bootstrap samples. The control group is defined as products without a carbon-neutral label. The control variables include the product categories and the product’s initial price at the beginning of the panel.

Time	Estimate (%)	Lower 95% CI	Upper 95% CI
April 2023	3.17	-3.12	9.46
May 2023	-0.53	-4.31	3.24
June 2023	1.00	-2.39	4.38
July 2023	-1.06	-4.14	2.03
August 2023	-1.88	-6.59	2.82
September 2023	-0.58	-3.54	2.37
October 2023	-0.72	-4.90	3.47
November 2023	-0.85	-5.16	3.46
December 2023	-0.12	-4.71	4.46
January 2024	-2.39	-6.48	1.71
February 2024	-1.56	-5.49	2.36
March 2024	-2.71	-7.04	1.63
April 2024	-2.29	-6.94	2.37
May 2024	-2.87	-7.08	1.35
June 2024	-1.73	-4.94	1.48
July 2024	-1.14	-4.28	2.00
August 2024	-1.17	-4.09	1.75
September 2024	-2.43	-5.51	0.65
October 2024	-2.23	-5.55	1.08
November 2024	-4.59*	-8.00	-1.19

Table A.7: Calendar Effects (US)

This table presents the effect of carbon-neutral labels on product price changes, expressed in percentages, relative to the product's price in March 2023 on Amazon's U.K. marketplace. Furthermore, the plot displays the upper and lower 95% confidence intervals based on 1,000 bootstrap samples. The control group is defined as products without a carbon-neutral label. The control variables include the product categories and the product's initial price at the beginning of the panel.

A.5 Customer reviews

In this section, we examine the content of customer reviews. We are interested in assessing whether consumers talk about carbon neutrality. Since carbon neutrality is a rather technical term, we consider additional keywords and then proceed by searching keywords in the corpus of reviews. The first page of reviews were scraped every week from May to November 2024 from Amazon's marketplace in the United States and from March to November 2024 in the United Kingdom and in Germany. We use a list of 40 keywords, as provided in this section. We collected 201,673 unique reviews for the United States, 90,617 reviews for the United Kingdom, and 285,426 unique reviews for Germany.

Tables A.8, A.9, and A.10 show occurrences of keywords for always-treated products, which already had a carbon-neutral label at the beginning of the panel, and never-treated products without a carbon-neutral label. For always-treated products, we further exclude reviews posted before the panel start date in each country to ensure we focus on reviews written when the products had the label. We proceed in this way to capture customers talking about products that we can confidently consider carbon neutral at the time of the review.

We report the occurrence of each keyword, as well as the total occurrence of the 40 keywords that we selected, providing overall shares for always-treated and never-treated units at the bottom of the table. Keyword frequencies are very low for both always-treated and never-treated products: United States, 0.114% and 0.028%; United Kingdom, 0.040% and 0.016%; Germany, 0.134% and 0.082%, respectively. That is, while treated products attract marginally more talk of climate-friendly features, the overall level is negligible. Further, the presence of similar values in control and treated products, as in the case of 'sustainable' features, points to product characteristics other than the labels that we are interested in driving the marginal coverage that we document. Overall, the summary statistics provided in this section confirm our main conclusions: consumers do not really value carbon-neutrality as a product feature.

Keyword	Always treated products (US)	Never treated products (US)
CO ₂ neutral [†]	0/63276	0/138397
Carbon Neutral Certified by SCS Global Services	0/63276	0/138397
Carbon Neutral by Carbon Trust	0/63276	0/138397
Carbon Trust	0/63276	0/138397
CarbonNeutral product by Climate Impact Partners	0/63276	0/138397
CarbonNeutral product by Natural Capital Partners	0/63276	0/138397
Carbonfree	0/63276	0/138397
Carbonfree Certified	0/63276	0/138397
Carbonfund	0/63276	0/138397
Carbonfund.org	0/63276	0/138397
Climate Impact Partners	0/63276	0/138397
Climate neutral by ClimatePartner	0/63276	0/138397
ClimatePartner	0/63276	0/138397
Natural Capital Partners	0/63276	0/138397
SCS Global Services	0/63276	0/138397
carbon credits	0/63276	0/138397
carbon footprint	1/63276	2/138397
carbon offsetting	0/63276	0/138397
carbon dioxide [†]	1/63276	0/138397
carbon neutral [†]	15/63276	2/138397
carbon reducing [†]	0/63276	0/138397
climate change	0/63276	0/138397
climate friendl*	0/63276	0/138397
climate pledge friendl*	0/63276	0/138397
corporate social responsibility	0/63276	0/138397
eco-label*	0/63276	0/138397
emissions	0/63276	0/138397
environment friendly [†]	0/63276	1/138397
environmentally friendl*	26/63276	3/138397
global warming	0/63276	0/138397
greenhouse gas	0/63276	0/138397
low carbon [†]	0/63276	0/138397
nature friendly [†]	0/63276	0/138397
net zero [†]	0/63276	0/138397
offsetting	0/63276	0/138397
reducing CO ₂ [†]	0/63276	0/138397
renewable energy	0/63276	0/138397
sustainabl*	29/63276	31/138397
zero carbon	0/63276	0/138397
zero emissions	0/63276	0/138397
Total (%)	0.114%	0.028%

Table A.8: Keyword occurrences (US)

This table shows keyword occurrences in always-treated and never-treated product reviews in Amazon’s US marketplace, scraped between May 2024 and November 2024. Words marked with “*” denote wildcards, allowing any continuation of the stem (e.g., “sustainabl*” matches “sustainability” and “sustainable”). Terms marked with [†] include both hyphenated and non-hyphenated forms. For “reducing CO₂,” we also cover the reversed order, i.e., “CO₂ reducing.”

Keyword	Always treated products (UK)	Never treated products (UK)
CO ₂ neutral [†]	0/17617	0/73000
Carbon Neutral Certified by SCS Global Services	0/17617	0/73000
Carbon Neutral by Carbon Trust	0/17617	0/73000
Carbon Trust	0/17617	0/73000
CarbonNeutral product by Climate Impact Partners	0/17617	0/73000
CarbonNeutral product by Natural Capital Partners	0/17617	0/73000
Carbonfree	0/17617	0/73000
Carbonfree Certified	0/17617	0/73000
Carbonfund	0/17617	0/73000
Carbonfund.org	0/17617	0/73000
Climate Impact Partners	0/17617	0/73000
Climate neutral by ClimatePartner	0/17617	0/73000
ClimatePartner	0/17617	0/73000
Natural Capital Partners	0/17617	0/73000
SCS Global Services	0/17617	0/73000
carbon credits	0/17617	0/73000
carbon footprint	0/17617	0/73000
carbon offsetting	0/17617	0/73000
carbon dioxide [†]	0/17617	5/73000
carbon neutral [†]	0/17617	0/73000
carbon reducing [†]	0/17617	0/73000
climate change	0/17617	0/73000
climate friendl*	0/17617	0/73000
climate pledge friendl*	0/17617	0/73000
corporate social responsibility	0/17617	0/73000
eco-label*	0/17617	0/73000
emissions	0/17617	0/73000
environment friendly [†]	0/17617	0/73000
environmentally friendl*	4/17617	3/73000
global warming	0/17617	0/73000
greenhouse gas	0/17617	0/73000
low carbon [†]	0/17617	0/73000
nature friendly [†]	0/17617	0/73000
net zero [†]	0/17617	0/73000
offsetting	0/17617	1/73000
reducing CO ₂ [†]	0/17617	0/73000
renewable energy	0/17617	0/73000
sustainabl*	3/17617	3/73000
zero carbon	0/17617	0/73000
zero emissions	0/17617	0/73000
Total (%)	0.040%	0.016%

Table A.9: Keyword occurrences (UK)

This table shows keyword occurrences in always-treated and never-treated product reviews on Amazon’s UK marketplace, scraped between March 2024 and November 2024. Words marked with “*” denote wildcards, allowing any continuation of the stem (e.g., “sustainabl*” matches “sustainability” and “sustainable”). Terms marked with [†] include both hyphenated and non-hyphenated forms. For “reducing CO₂,” we also cover the reversed order, i.e., “CO₂ reducing.”

Keyword	Always treated products (Germany)	Never treated products (Germany)
CO ₂ neutral [†]	0/46375	0/239051
Carbon Neutral Certified by SCS Global Services	0/46375	0/239051
Carbon Neutral by Carbon Trust	0/46375	0/239051
Carbon Trust	0/46375	0/239051
CarbonNeutral product by Climate Impact Partners	0/46375	0/239051
CarbonNeutral product by Natural Capital Partners	0/46375	0/239051
Carbonfree	0/46375	0/239051
Carbonfree Certified	0/46375	0/239051
Carbonfund	0/46375	0/239051
Carbonfund.org	0/46375	0/239051
Climate Impact Partners	0/46375	0/239051
Climate neutral by ClimatePartner	0/46375	0/239051
ClimatePartner	0/46375	0/239051
Natural Capital Partners	0/46375	0/239051
SCS Global Services	0/46375	0/239051
carbon credits	0/46375	0/239051
carbon footprint	1/46375	0/239051
carbon offsetting	0/46375	0/239051
carbon dioxide [†]	0/46375	5/239051
carbon neutral [†]	0/46375	0/239051
carbon reducing [†]	0/46375	0/239051
climate change	0/46375	6/239051
climate friendl*	0/46375	6/239051
climate pledge friendl*	0/46375	6/239051
corporate social responsibility	0/46375	0/239051
eco-label*	0/46375	0/239051
emissions	0/46375	2/239051
environment friendly [†]	9/46375	35/239051
environmentally friendl*	16/46375	64/239051
global warming	0/46375	0/239051
greenhouse gas	0/46375	0/239051
low carbon [†]	0/46375	0/239051
nature friendly [†]	0/46375	0/239051
net zero [†]	0/46375	0/239051
offsetting	0/46375	0/239051
reducing CO ₂ [†]	0/46375	0/239051
renewable energy	0/46375	0/239051
sustainabl*	36/46375	73/239051
zero carbon	0/46375	0/239051
zero emissions	0/46375	0/239051
Total (%)	0.134%	0.082%

Table A.10: Keyword occurrences (Germany)

This table shows keyword occurrences in always-treated and never-treated product reviews in Amazon’s Germany marketplace, scraped between March 2024 and November 2024. Words marked with “*” denote wildcards, allowing any continuation of the stem (e.g., “sustainabl*” matches “sustainability” and “sustainable”). Terms marked with [†] include both hyphenated and non-hyphenated forms. For “reducing CO₂,” we also cover the reversed order, i.e., “CO₂ reducing.”

A.6 Robustness tests

This section presents a list of robustness tests for the main estimation using data from Amazon's United States marketplace.

First, Figure A.1 illustrates the results of relaxing the condition that the carbon-neutral label should appear in more than 90% of observations after treatment. Second, Figure A.2 shows the estimation results using unbalanced data for the United States. Third, Figure A.3 shows the results excluding products with Climate Pledge Friendly or Small Business labels introduced after March 2023. Fourth, Figure A.4 shows the results excluding control variables. Fifth, Figure A.5 shows results controlling for the initial number of ratings, in addition to the product's initial price and product categories.

We also list two additional robustness estimations using alternative outcome variables. First, Figure A.6 shows the results using absolute price level as the outcome variable. Second, A.7 shows the results using the number of ratings as the outcome variable.

A discussion of these additional findings is provided in Section 3.1 in the main body of the text.

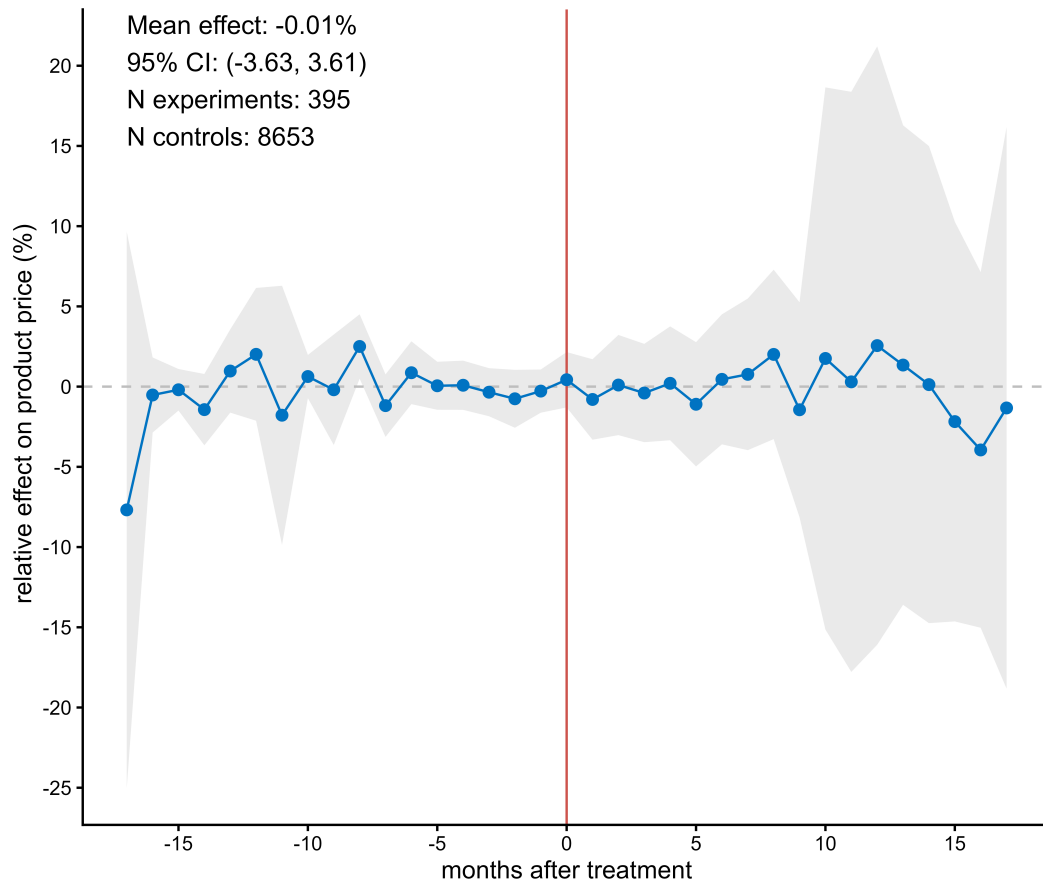


Figure A.1: The effect of carbon-neutral label when removing label consistency constraint (United States)

This plot shows the dynamic treatment effect in percentages relative to the product's price at baseline. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories and the product's initial price at the beginning of the panel.

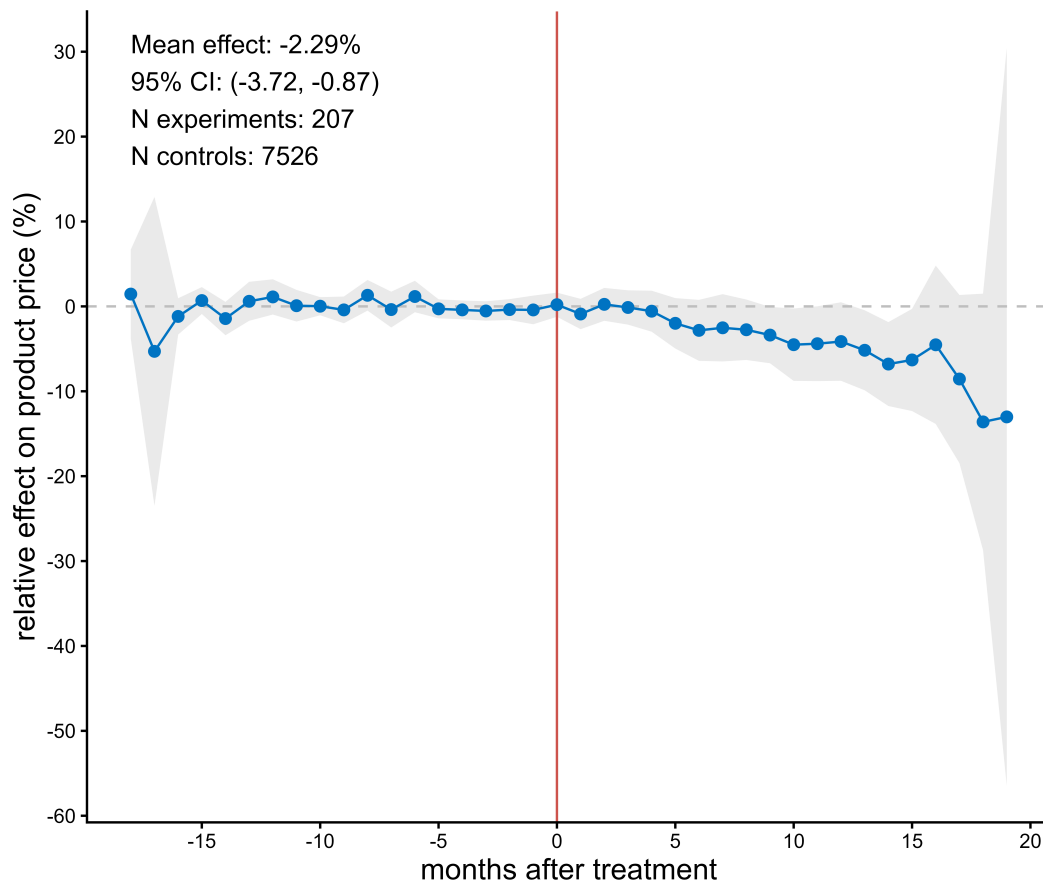


Figure A.2: Effect of carbon-neutral label using unbalanced data option for estimation (United States)

This plot shows the dynamic treatment effect in percentages relative to the product's price at baseline. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing.

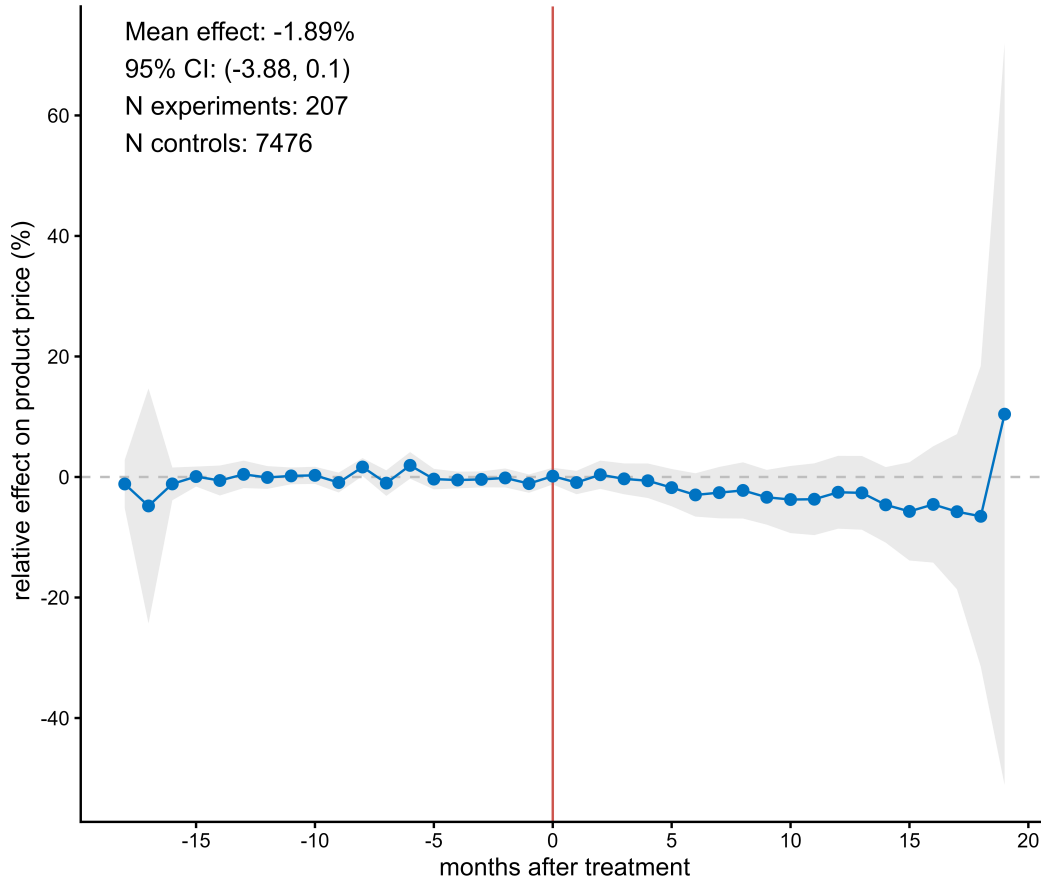


Figure A.3: Effect of carbon-neutral label excluding new Climate Pledge Friendly labels and Small Business labels (United States)

This plot shows the dynamic treatment effect in percentages relative to the product's price at baseline. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories and the product's initial price at the beginning of the panel.

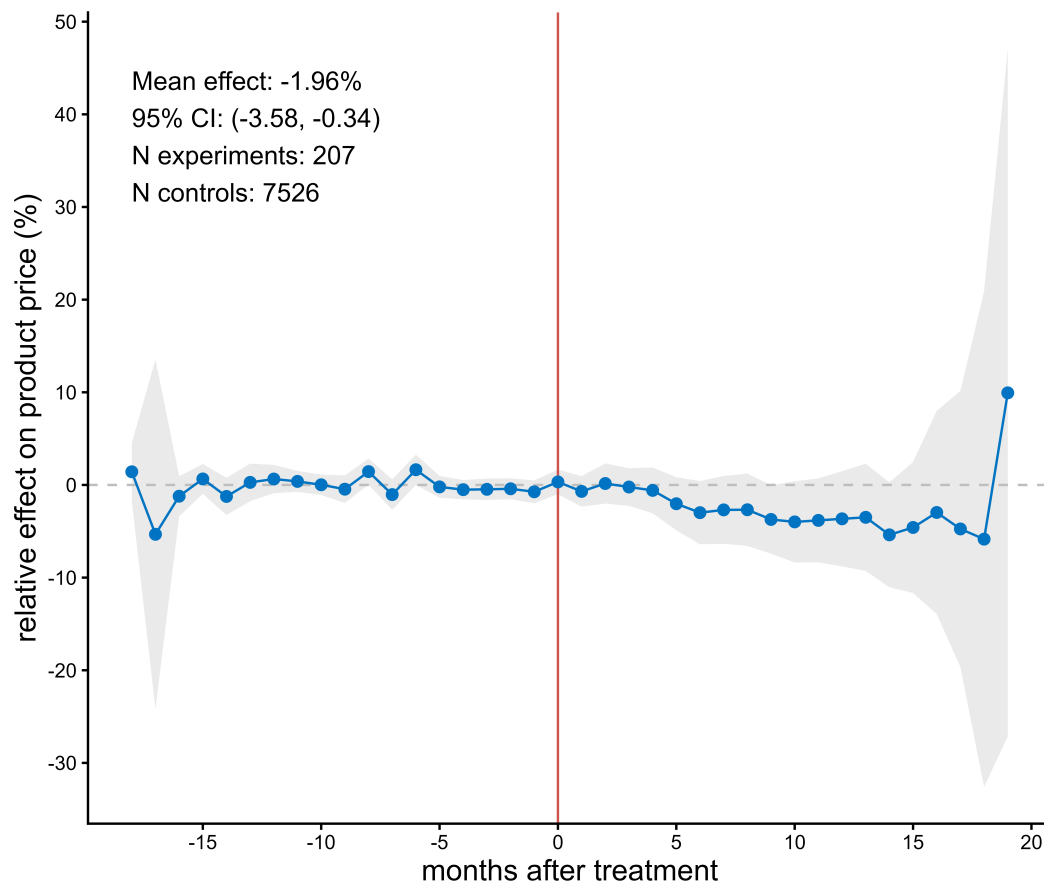


Figure A.4: Effect of carbon-neutral label without control variables (United States)

This plot shows the dynamic treatment effect in percentages relative to the product's price at baseline. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of the treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing.

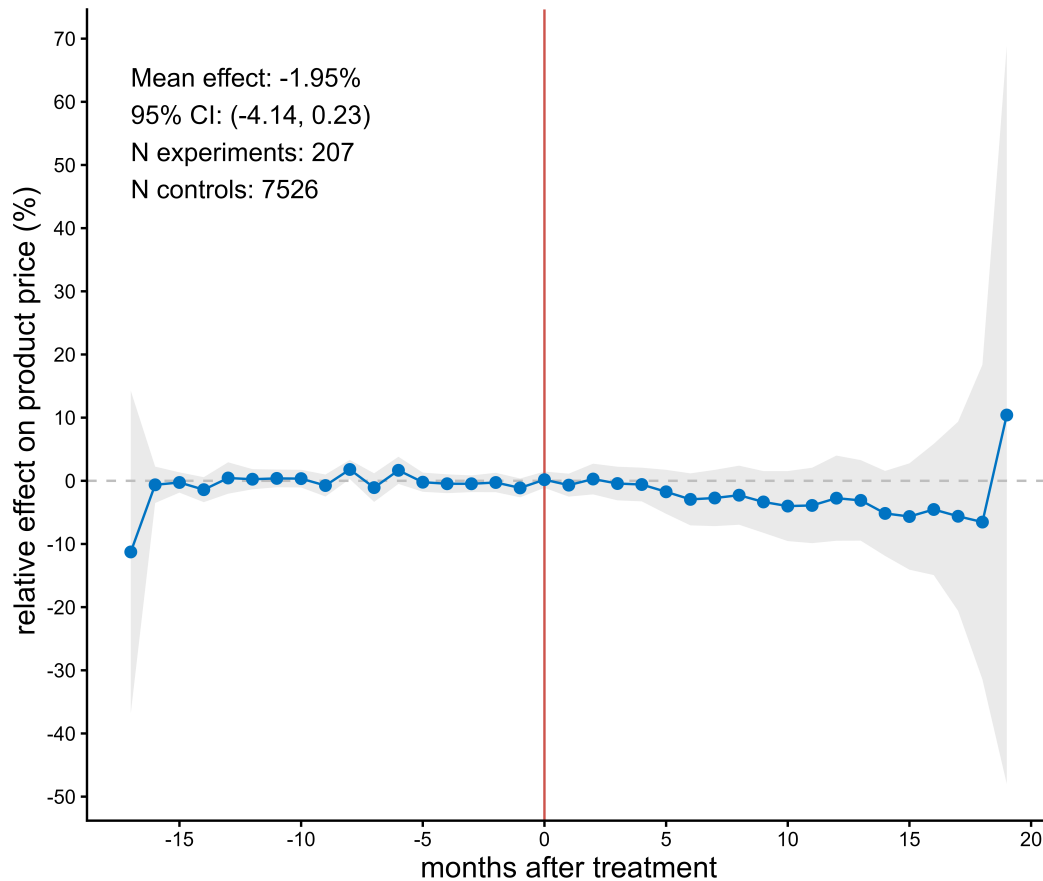


Figure A.5: Effect of carbon-neutral label with the initial number of ratings as an additional control variable (United States)

This plot shows the dynamic treatment effect in percentages relative to the product's price at baseline. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of the treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories, the product's initial price, and the initial number of ratings at the beginning of the panel.

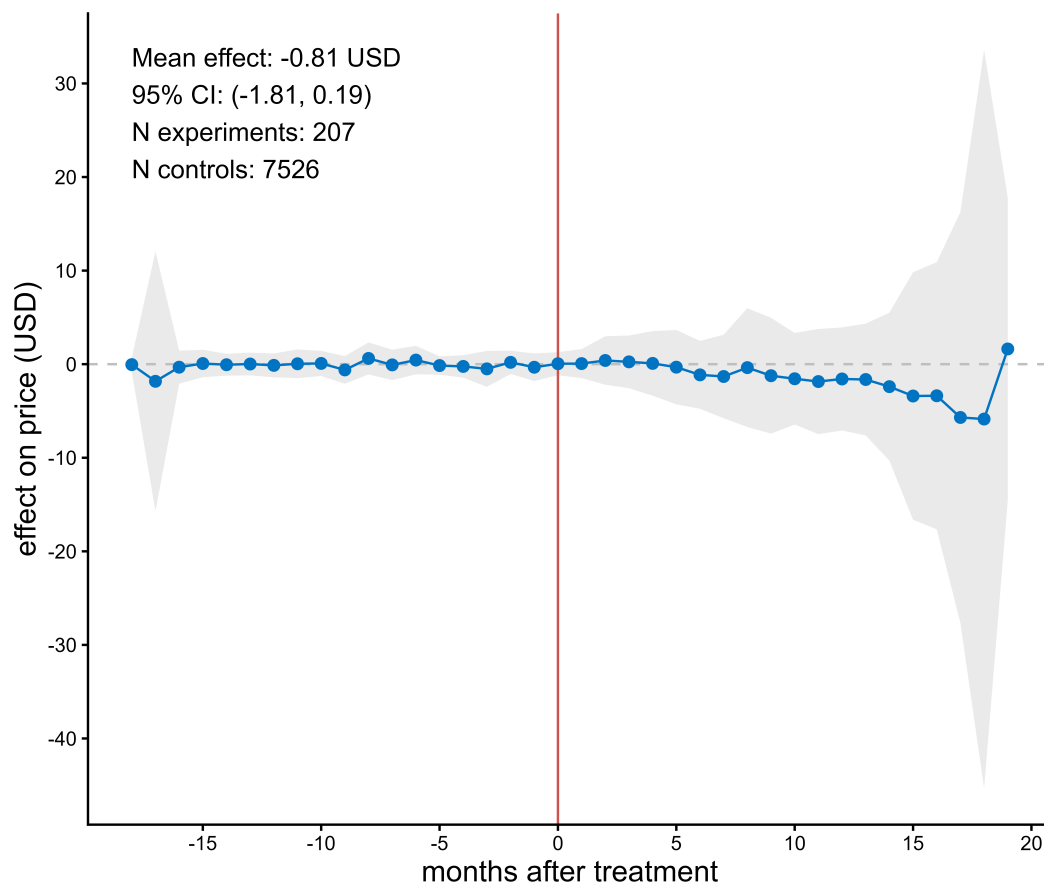


Figure A.6: Effect of carbon-neutral label using absolute price level as outcome variable (United States)

This plot shows the dynamic treatment effect in USD. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories and the product's initial price at the beginning of the panel.

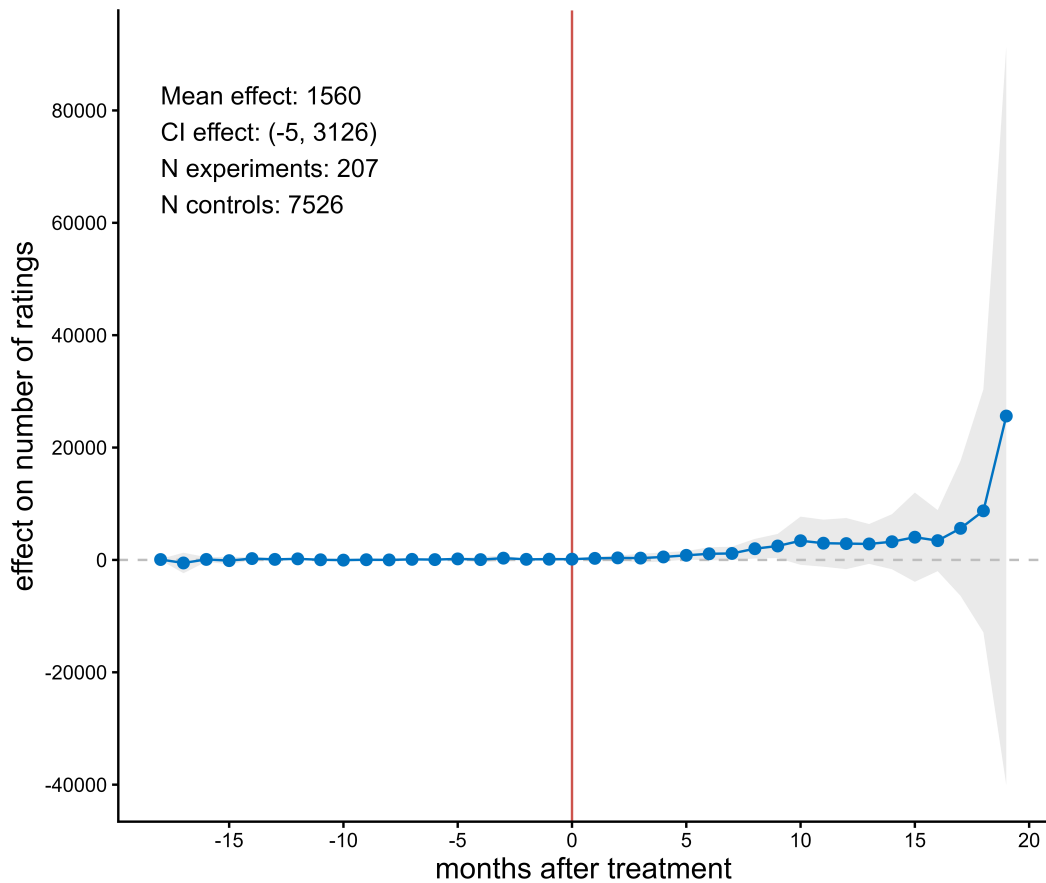


Figure A.7: Effect of carbon-neutral label using the number of ratings as the outcome variable (US)

This plot shows the dynamic treatment effect of carbon-neutral label on the rating count in March 2023. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories.

A.7 Additional results from Amazon's U.K. and Germany marketplaces

This section presents event-study plots using the number of product ratings for the United Kingdom and Germany, shown in Figures A.8 and A.9, respectively.

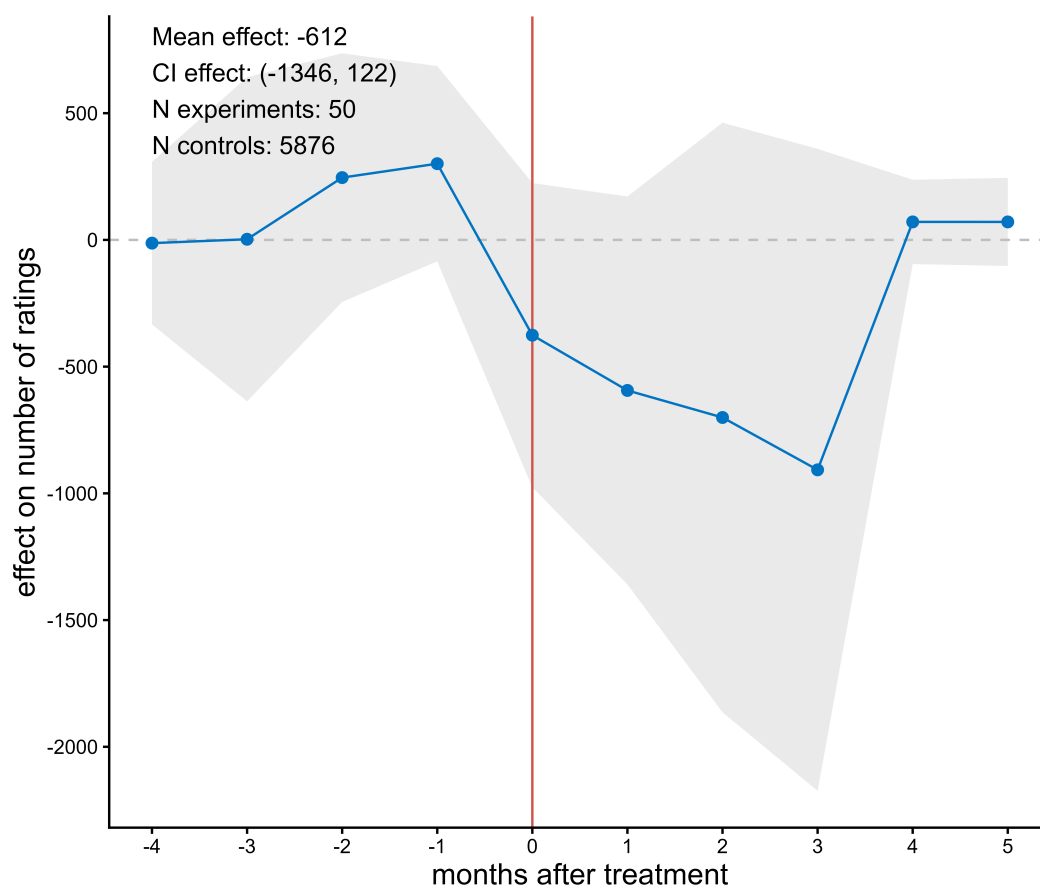


Figure A.8: Effect of carbon-neutral label using the number of ratings as the outcome variable (UK)

This plot shows the dynamic treatment effect of the carbon-neutral label on the rating count in May 2024. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of the treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories.

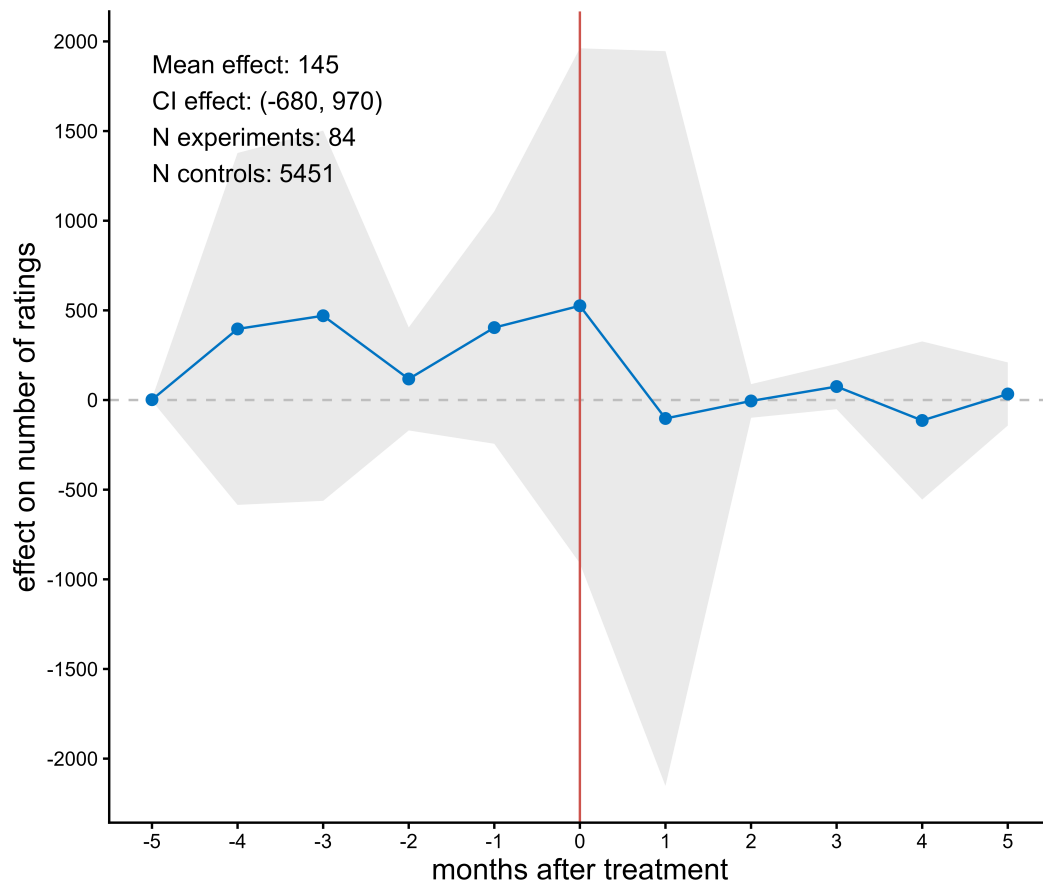


Figure A.9: Effect of carbon-neutral label using the number of ratings as the outcome variable (Germany)

This plot shows the dynamic treatment effect of carbon-neutral label on the rating count in May 2024. Vertical red line marks treatment onset. The shaded area indicates a 95% confidence interval of treatment effect based on 1,000 bootstrap samples that control for multiple hypothesis testing. The control variables include product categories.

Chapter 3: Transparency of carbon-neutral labels: evidence from a choice experiment

single-authored

Abstract

This paper examines the effect of transparency in carbon-neutral labeling on consumers' willingness to pay. Carbon-neutral labels indicate that a product's greenhouse gas emissions have been offset (compensated) outside the company and/or directly reduced within it. Although expert assessments view CO₂ offsets as less effective than CO₂ reductions in mitigating emissions, most labels on the market lack transparency regarding the proportion of CO₂ offset and CO₂ reduction. This study empirically investigates whether consumers are willing to pay for transparency in carbon-neutral labels that disclose the shares of CO₂ offset and CO₂ reduction. Using a discrete choice experiment among UK tea consumers, I compare the willingness to pay for standard versus transparent carbon-neutral labels. The control group saw a standard carbon-neutral label, while the treatment groups saw transparent labels indicating varying shares of CO₂ offsetting and CO₂ reduction. I find no evidence of consumers' willingness to pay for transparency on carbon-neutral labels or a preference for CO₂ reductions over CO₂ offsets, indicating a mismatch between expert and consumer valuations. I further find that the presence of organic and ethical trade labels crowds out the marginal value of the carbon-neutral claim, yet does not change the null effect of transparency.

Keywords: carbon-neutral labeling; competing sustainability labels; discrete choice experiment; greenwashing; information asymmetry; transparency

JEL codes: C25; C83; C90; D12; Q51; Q58

1 Introduction

As the effects of climate change become more severe, many companies, responding to investor and board pressure, have committed to carbon neutrality (Kim and Lyon, 2011; Rogelj et al., 2021). Nearly two-thirds (63%) of the Forbes Global 2000 companies have announced carbon-neutral or net-zero targets (Net Zero Tracker, 2025). This paper asks whether consumers value transparency in carbon-neutral labels that disclose the shares of CO₂ offset and CO₂ reduction, and whether they prefer CO₂ reductions over offsets. It finds no evidence that any of these effects are present.

Companies increasingly label their products as carbon neutral; however, the labels lack transparency regarding how carbon neutrality is achieved. Carbon neutrality requires measuring greenhouse-gas (GHG) emissions in carbon-dioxide equivalents (CO₂e) and then reducing emissions internally (e.g., through energy-efficiency improvements) and/or offsetting remaining emissions. CO₂ offsetting involves financing external projects, such as reforestation, to compensate for residual emissions. These offsets are traded on voluntary carbon markets, with transaction values totaling USD 723 million in 2023 (Forest Trends' Ecosystem Marketplace, 2024).

Carbon-neutral labels denote CO₂ offsets, direct CO₂ reductions, or a combination thereof. However, these activities are not equivalent, and how neutrality is achieved matters. Established frameworks and principles for climate change mitigation prioritize direct CO₂ reductions and treat offsets as a last resort (IPCC, 2022; Axelsson et al., 2024). Voluntary offset credits may not deliver real, additional abatement (Calel et al., 2025), and one ton of offset is worth only a fraction of a ton of permanent CO₂ reduction (Groom and Venmans, 2023). Concerns arise from both practical and ethical considerations regarding CO₂ offsets. From a practical perspective, offsets are seen as less effective than direct reductions in lowering atmospheric CO₂, due to issues of credibility (Bumpus and Liverman, 2008; Hooper et al., 2008), additionality (Hyams and Fawcett, 2013; Schneider and Kollmuss, 2015), double counting (Schneider et al., 2015;

Calel et al., 2025; Trencher et al., 2024), and leakage (Murray et al., 2004). From an ethical perspective, concerns include moral licensing (Dorner, 2019) and the commodification of nature through pricing (Aldred, 2012).

Transparency of carbon-neutral labels is becoming increasingly important for policymakers. The European Parliament proposed limits on unverified generic claims such as “climate neutral” (European Parliament, 2023); German courts require firms to clarify the meaning of “climate neutral” in advertising (The Wall Street Journal, 2024); and lawsuits have challenged airline neutrality claims (The Guardian, 2023). In response, some certifiers have withdrawn carbon-neutral labels (Carbon Trust, 2023; ClimatePartner, 2023). Yet, most carbon-neutral certifiers still do not disclose the offset–reduction shares on the label.¹ Therefore, the empirical questions arise: *Do consumers value transparency on carbon-neutral labels? Do they prefer CO₂ reductions over offsets?*

Beyond these main research questions, I investigate an additional question. In practice, it is common for products to carry multiple sustainability claims. To situate the transparency question in this broader choice environment, I additionally explore: *How does the coexistence of other sustainability labels change the marginal willingness to pay (MWTP) for carbon-neutral labels and transparency?* This extension clarifies whether consumer valuations reflect absolute preferences for carbon-neutral claim (and its transparency) or relative trade-offs within a broader labeling environment.

This pre-registered study² leverages a discrete choice experiment (DCE), a stated preference method, to elicit consumers’ preferences and MWTP for carbon-neutral labels. An online survey was conducted among 1,339 tea consumers in the UK. The experimental design consists of one control group and two treatment arms. The control group saw the standard carbon-neutral label without any information about the shares of CO₂ offsetting and CO₂ reduction. The treatment groups saw a transparent version of the carbon-neutral label, which

¹An exception is Swiss International Air Lines, which offers offsetting or a combination of offsetting and reductions after flight purchases (Swiss International Air Lines, 2024).

²Ozdemir 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>

includes information on the shares of CO₂ offsetting and CO₂ reduction. The first treatment group saw a carbon-neutral label indicating 95% CO₂ offsetting and 5% CO₂ reduction, while the second treatment group saw a carbon-neutral label with equal shares, i.e., 50% CO₂ offsetting and 50% CO₂ reduction. All participants were provided with a short, neutral description of the CO₂ offsetting and CO₂ reduction activities that the label entails, with or without information on the exact share of each activity, depending on the experimental group³. Across all arms, the choice tasks also included additional sustainability labels, namely organic and ethical trade, to study coexistence effects (that is, how other sustainability labels on the same product affect the valuation of carbon-neutral labels and their transparency); these labels were displayed identically across groups and were not manipulated by the transparency treatments.

Consumer demand for carbon-neutral labels can be driven by psychological factors such as feelings of guilt about environmental harm (Kotchen, 2009) and the warm glow from environmentally friendly choices (Andreoni, 1990; Kahneman and Knetsch, 1992). However, media scepticism and concerns about corporate greenwashing can negatively affect trust in labels involving CO₂ offsets. While many studies find that consumers are willing to pay a premium for carbon-neutral labels (Gassler, 2015; Vecchio and Annunziata, 2015; Driouchis et al., 2016; Birkenberg et al., 2021; Bek, 2022), other evidence is mixed; for instance, a hedonic analysis by Carattini et al. (2025) finds no evidence of willingness to pay (WTP).

Transparency, by disclosing the shares of CO₂ offsetting versus CO₂ reduction, can mitigate negative perceptions and enhance trust by clarifying sellers' incentives (Darby and Karni, 1973). It may also reduce demand when the offsetting share is high relative to reductions, given practical and ethical concerns (Carattini and Tavoni, 2016). In markets without standardized labels, additional details can confuse consumers and enable adverse selection (Akerlof, 1970; Brécard, 2017). Thus, the net effect of transparency likely hinges on whether

³While the experiment provides participants with information on CO₂ offsetting and reduction, the descriptions are neutral and non-evaluative. Thus, the design is primarily intended to reveal consumers' preferences rather than to form new ones.

consumers distinguish offsets from reductions and on how they process added information. Although this study does not elicit prior beliefs or normative preferences about typical shares to avoid priming, it provides reduced-form evidence on how disclosure affects MWTP for carbon-neutral labels.

I find that UK consumers are willing to pay a premium for carbon-neutral labels on tea products (with an estimated MWTP of £0.55 in the control group, £0.52 in Treatment 1 with a higher share of offsetting (95%), and £0.32 in Treatment 2 with equal shares of offsetting and reduction). However, there is no evidence that transparency affects consumers' MWTP for carbon-neutral label. If anything, transparency slightly reduces MWTP: £-0.03 [£-0.28, £0.22] when the label shows a higher share of CO₂ offsetting, and £-0.23 [£-0.47, £0.01] when it shows an equal share of CO₂ offsetting and reduction. To put these effects into perspective, they correspond to reductions of about 5.5% and 41.8% of the MWTP for the standard label, respectively.

Furthermore, there is no evidence that consumers prefer CO₂ reductions over offsets: moving from a 5% to a 50% reduction share lowers MWTP by about £0.20 [£-0.46, £0.06], i.e., roughly 38% relative to the 5% group.

I also explore a range of potential mechanisms—psychological and contextual factors such as confusion, trust, concern, climate worry, guilt, social approval, and beliefs about the polluter-pays principle—together with resource constraints (financial and time), but none of these variables are found to systematically explain bilateral differences in MWTP across experimental groups, which are statistically insignificant in any case.

These results also hold for conditional MWTP for carbon-neutral label when other sustainability claims are present. The findings show that organic and ethical-trade labels compete with the carbon-neutral claim and substantially reduce—and, when combined, can even reverse—the premium for carbon neutrality. However, neither the transparency treatment nor the shift from a 5% to a 50% reduction share has a statistically significant effect, regardless of whether other sustainability labels are present concurrently with the carbon-neutral label.

These findings provide three policy-relevant insights.⁴ First, the market cannot sort out the quality of carbon-neutral claims by rewarding transparent labels more than non-transparent ones: consumers do not reward transparency per se, and the estimated transparency effects are small, negative, and statistically insignificant.

Second, a policy intervention that ensures transparency regarding the composition of carbon-neutral labels does not lead consumers to demand more CO₂ reductions rather than offsets. The findings show that consumers value labels dominated by offsets at least as highly as those with an equal-share disclosure.

Third, the presence of organic and ethical-trade labels substantially reduces—and, when combined, can even reverse—the WTP premium for the carbon-neutral claim. While the effects of transparency and a higher reduction share remain statistically insignificant, other sustainability claims can erode the marginal value of carbon-neutral certifications. Because several mechanisms could drive this crowd-out (e.g., salience competition or substitution toward other valued sustainability attributes), these findings motivate the exploration of alternative label architecture policies (e.g., standardized placement and prominence) to preserve the visibility of carbon-neutral claims.

Regulating labels remains important⁵, and is consistent with the intentions of policymakers. Taken together, the results indicate that, given current information and beliefs, consumers neither generate stronger demand for CO₂ reductions than for CO₂ offsets nor show a higher WTP for transparency itself. The mismatch between expert and consumer valuations provides a rationale for further information regarding the inferior quality of low quality offsets, as well as the need for additional regulation. The challenge is to design policies that enhance the composition and credibility of carbon-neutral (or related) claims without undermining the demand for climate claims altogether.

⁴Throughout this study, I examine the disclosure effect in a between-subjects setting to avoid contrast and experimenter demand effects. The policy implications are based on the assumption that consumers do not make side-by-side comparisons between standard and transparently labeled versions of the same product of focus.

⁵Information asymmetry can distort market mechanisms, leading to adverse selection and the selection of lower-quality goods (Akerlof, 1970). A body of theoretical literature suggests that, under certain assumptions and in the presence of information asymmetry, competition can favour labels with lower environmental quality (Brécard, 2014; Heyes and Martin, 2017; Brécard, 2017; Heyes and Martin, 2018; Poret, 2019).

This study contributes to several lines of the existing literature. First, 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 examined consumers' MWTP for climate labels, to my knowledge, none of them has focused on the effect of transparency in carbon-neutral labels. This paper fills that gap by studying how making the share of CO₂ offsetting explicit affects consumers' valuation of the carbon-neutral labels.

Second, this paper contributes to the extensive literature on the economics of CO₂ 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), which examines how consumers value CO₂ offsets. I contribute by focusing on the offsetting share indicated on carbon-neutral labels relative to direct CO₂ reductions. Two recent studies have examined how German consumers value CO₂ offsets relative to direct CO₂ reductions (Bek, 2022; Roemer et al., 2023). Bek (2022) uses a hypothetical DCE on coffee and find that consumers are willing to pay 66% more for lifecycle CO₂ reductions than for offsets, whereas Roemer et al. (2023) use an incentivized DCE on shipping mode choice and report roughly equal valuations for internal reductions and offsets. However, both studies rely on within-subjects designs, where participants are shown multiple label types, either as attribute levels or as distinct attributes, making it difficult to disentangle absolute preferences from contrast or salience effects. In this study, I instead adopt a between-subjects design, randomly assigning each participant to only one of three label conditions. This approach minimizes contrast and experimenter demand effects and allows me to identify the effect of transparency on MWTP for carbon-neutral labels and to quantify the effect of transparency in offsetting claims.

Third, this study connects to the literature on how non-normative pro-social behaviors become widely accepted (Sparkman and Walton, 2017; Carattini et al., 2024; Mortensen et al., 2019; Kraft-Todd et al., 2018) by investigating whether

increased transparency in environmental labeling can foster the diffusion of green consumption. Carbon-neutral products currently remain a niche market, and I examine whether making the offsetting share explicit in carbon-neutral labels can enhance the adoption of such products.

Finally, the study contributes to the empirical literature on the competition or complementarity between climate labels and other sustainability labels (Onozaka and McFadden, 2011; Akaichi et al., 2020). Onozaka and McFadden (2011) find that negative perceptions of carbon footprint or imported products can be mitigated when paired with organic or fair trade labels, while Akaichi et al. (2020) show that Spanish consumers perceive health and sustainability labels as complements, whereas UK consumers show weaker complementarity and even substitutability. By contrast, my findings point to competition across sustainability claims: organic and ethical-trade labels crowd out the marginal value of the carbon-neutral claim.

The remainder of the paper is outlined as follows: Section 2 details the methodology, including the survey and DCE design (2.1), data (2.2), and the empirical approach (2.3). Section 3 presents the results, which discusses the effect of transparency on MWTP for carbon neutral labels (3.1), the underlying mechanisms of consumers' MWTP (3.3), and competing labels (3.4). Section 4 concludes.

2 Methodology

2.1 Survey and choice experiment design

This section provides an overview of the survey and experimental design. It begins by introducing the tea product that is the focus of the survey and outlining the sample characteristics. Next, it presents the experimental design. It then describes the DCE, including the attributes and levels, as well as the choice tasks presented to participants. Finally, it summarizes the additional survey questions that followed the DCE, which were designed to elicit participants' understanding, perceptions, and motivations related to climate-friendly products and carbon-neutral labels.

Participants are randomly assigned to one of three experimental groups. All groups receive identical survey content and DCE designs. The only difference is in the type of carbon-neutral label that participants in each group see, as shown in Table 1. The control group is shown a standard carbon-neutral label, stating “CO₂ neutral” only. The information on the shares of CO₂ offsetting and CO₂ reduction are not revealed to the participants ⁶. The two treatment groups are shown a “transparent” carbon-neutral label with additional text indicating the shares of CO₂ offsetting and CO₂ reduction that make the life cycle of the tea product carbon neutral. 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.

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 CO₂ emissions (Crippa et al., 2021). Although a box of tea itself may not be considered CO₂-intensive,

⁶Participants in the control group are provided the following information: “The 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.”

its frequent consumption can 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).

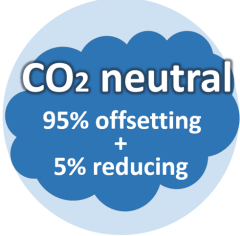
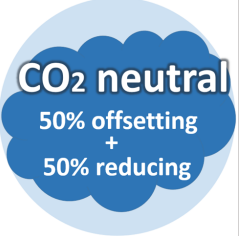
	Control	Treatment 1	Treatment 2
Label explanation	Yes	Yes	Yes
Carbon-neutral label			
CO ₂ offset share	No information	%95	%50
CO ₂ reduction share	No information	%5	%50
No. of participants	448	447	444

Table 1: Experimental design

In the survey, which is provided in Appendix E, respondents are asked to consider 80 teabags or 200 grams of tea in their preferred form. The pre-registered online survey⁷ is pretested on 157 respondents, and the main survey included 1,339 tea drinkers in the UK⁸.

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.

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

Table 2: Choice attributes and levels

The DCE focuses on a tea box with the following attributes: carbon-neutral label, organic label, ethical trade label, and price. Table 2 shows the attributes

⁷Ozdemir 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>

⁸Any deviations from the pre-registration, as well as exploratory and confirmatory hypotheses, are listed in the Appendix D.

and attribute levels included in the DCE: three sustainability labels with two possible levels each, the presence or absence of the label, and the price that can take 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.

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 situations. Organic and ethical trade labels do not vary across sub-samples, unlike carbon-neutral labels.⁹ 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. Respondents were also informed about the hypothetical nature of the choice scenarios.

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

DCEs and contingent valuation methods (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

⁹The 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.

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. In the budget reminder, I remind respondents that any amount they choose to spend on tea would reduce the money available for other purchases.

Ngene software is used to generate the D-efficient DCE design, which consists of 16 different choice tasks divided into 2 blocks containing 8 choice tasks each. These blocks are two different versions of the questionnaire: each respondent is randomly assigned to one block and therefore answers only 8 choice tasks, which reduces respondent fatigue while using all 16 tasks in the estimation. Figure 1 shows one of the choice cards shown to the control group. There are two tea products, and “none of the two” choice opinions. 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) and the highest price level (6.9) tea alternatives. If the price is 0.9 pounds, it should not have any labels; similarly, if it is 6.9, it must have all labels. The estimated parameters from the pre-test were used to create the final DCE design. Please refer to Table A.10 in Appendix A.2 for the respective DCE design used for the main survey.

	Tea product A	Tea product B	None of the two
			
CO ₂ neutral label		X No label	
Organic label	X No label	X No label	
Ethical trade label	X No label	X No label	
	Price of tea product A £3.9	Price of tea product B £2.9	Price £0
I choose...			

Figure 1: Example choice card presented to participants in the control group

Each participant was shown one of eight different cards. This example shows one of the cards presented to participants in the control group, block 2.

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 CVM questions. Additionally, the survey explores participants' attitudes toward climate change and sustainability labels using Likert-scale statements, such as trust in carbon-neutral labels, confusion about carbon-neutral labels, concerns about CO₂ 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 pur-

chases)¹⁰. Then, participants are shown all three versions of carbon-neutral labels and asked which label they trust the most and find the most confusing. Finally, it collects information on participants' tea consumption habits and sociodemographic characteristics.

2.2 Data

The survey data were collected online in September 2024 in collaboration with a professional survey company (Bilendi). 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 out or dropped out. Individuals who indicated that they never purchase or never drink tea, constituting 16% of the total sample, were disqualified from continuing the survey¹¹.

A total of 1,339 individuals successfully completed the main survey¹². After excluding protest responses (1% of the full sample), the analyzed sample consists of 1,321 individuals. Protest responses were identified as cases where respondents consistently chose the status quo option (no tea purchase), with reasons including opposition to one or more labels, insufficient information about the products, or disagreement with the question itself.

The sample represents the UK population aged 18 and over, with some variations with respect to general population statistics due to the focus on adult tea consumers older than 18 years in this study. The median age of the sample is 46 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). 50% of the sample holds a post-secondary certificate (NQF Level 4) level or above, very close to the national

¹⁰In the survey instrument, the statements capturing respondents' concerns about CO₂ offsetting are designed based on the ethical and practical considerations discussed in (Carattini and Blasch, 2024). The protest-response screeners and the statement assessing beliefs about the polluter-pays principle are based on Brouwer (2011). Finally, the items measuring social approval and guilt are adapted from the questionnaire used by Theotokis and Manganari (2015) and Gruchmann et al. (2025).

¹¹Please 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.

¹²For the pre-test, 157 participants were recruited. The pre-test data is not included in the main survey data.

figure of 49% in the population (Gov.uk, 2021a). Additionally, the median household income of the sample¹³ is 35,000, which is close to the national median of approximately 34,500 (Office for National Statistics, 2024a). Finally, the sample’s employment rate of 65% is slightly below the national average of 75% (Office for National Statistics, 2024b). A summary of all socioeconomic characteristics of tea consumers/drinkers is provided in Table A.1. Table A.9 in Appendix A.1 reports covariate balance. Standardized mean differences are generally small, with most covariates showing values below 0.2 in absolute terms. The only deviation is the attention-check question on the definition of CO₂ offsetting, where standardized mean differences are around 0.3.

Among tea consumers, approximately 78% drink tea daily, and 78% purchase tea at least once a month or more frequently. Detailed information on tea consumption and purchasing habits is provided in Table A.3 in Appendix A.1. Over 73% of respondents reported consuming black tea, while the remainder consume green tea, herbal tea, or other blends.

2.3 Empirical approach

To estimate consumers’ MWTP for the carbon-neutral label, I leverage a mixed logit (MXL) model estimated directly in WTP space.¹⁴ A summary of deviations from the pre-analysis plan, along with the exploratory and confirmatory hypotheses, is provided in Appendix D.

The foundation of my approach is based on the Random Utility Model (RUM) (Lancaster, 1966; McFadden, 1973), where individuals choose the alternative that maximizes their expected utility. Each choice situation n consists of a deterministic component (sustainability labels and tea price) and a random component capturing unobserved factors. In a standard “preference space” specification, the utility U_{inj} that individual i derives from alternative j in

¹³Income was estimated by assigning the midpoint of each reported income bracket, with the lowest and highest categories approximated using a Pareto distribution.

¹⁴Although the preregistration specified a preference-space specification, the main analysis relies on an MXL model estimated in WTP space because of methodological advantages of WTP-space models, including the direct estimation of MWTP and greater interpretability. All preference-space model results are reported in the robustness checks.

choice situation n is defined as:

$$U_{inj} = \alpha_i p_{inj} + \beta_i^\top \mathbf{X}_{inj} + s_{in} + \varepsilon_{inj}, \quad (1)$$

where p_{inj} denotes price, and $\mathbf{X}_{inj}$ is a vector of non-price attributes (carbon-neutral, organic, and ethical trade labels). The parameters α_i and β_i represent the individual-specific marginal utility of price and non-price attributes, respectively, while s_{in} captures the status quo (no purchase) alternative. The error term ε_{inj} is assumed to be i.i.d. Type I extreme value.

In this framework, the MWTP for an attribute is derived by calculating the ratio of the attribute coefficient to the price coefficient ($-\beta_i/\alpha_i$). However, when both coefficients are random, this ratio can yield a heavy-tailed distribution with undefined moments, leading to unstable estimates (Hensher et al., 2005b; Daly et al., 2012). To resolve this, I follow Train and Weeks (2005) and estimate the model directly in WTP space. This involves reparameterizing the utility function so that the coefficients represent monetary values directly:

$$U_{inj} = -\alpha_i (p_{inj} - \mathbf{w}_i^\top \mathbf{X}_{inj}) + s_{in} + \varepsilon_{inj}. \quad (2)$$

In this specification, $\mathbf{w}_i$ is the vector of random coefficients interpreted directly as the MWTP for the attribute levels. The price coefficient α_i becomes a scaling parameter representing the variance of unobserved utility, which I assume to be log-normally distributed to ensure it is strictly positive. The MWTP parameters $\mathbf{w}_i$ and the status quo parameter s_{in} are assumed to be normally distributed.

I use the MXL model to estimate the choice parameters, allowing coefficients to vary randomly across respondents to capture more realistic substitution patterns among alternatives and better reflect the distribution of MWTP. I use MXL model approach instead of the standard multinomial logit (MNL) model, with fixed coefficients because of two reasons: (i) MXL model accounts for unobserved preference heterogeneity across individuals and (ii) relaxes the independence of irrelevant alternatives (IIA) assumption, which implies that the relative odds of choosing between any two alternatives are unaffected by the

presence or characteristics of other options.

I estimate the model using Simulated Maximum Likelihood. To approximate the high-dimensional integrals in the log-likelihood function, I use 1,000 Halton draws, which provide greater precision than pseudo-random draws (Train, 2009). Following Mariel et al. (2021), I initialize the standard deviations of the random parameters at 0.5 to facilitate convergence as suggested by Mariel et al. (2021), while mean starting values are taken from an initial MNL estimation.

To compare MWTP estimates across experimental groups, I follow the combinatorial approach proposed by Poe et al. (2005). This non-parametric test is particularly suitable here as it allows for the comparison of full distributions of simulated MWTP measures rather than relying solely on point estimates. Although the Poe test is most relevant for MWTP estimates derived from preference-space models, where the ratio of two random coefficients can result in an analytically undefined distribution, I apply it to all specifications, including WTP-space estimations, to ensure full comparability between the main analysis (WTP space) and the robustness checks (which also include preference space estimations).

To address the issue of multiple hypothesis testing—standard in experimental economics (List et al., 2019)—I adjust p-values using the Holm-Bonferroni procedure. This correction is applied to the means of the random parameters (carbon-neutral, organic, ethical trade labels) and pairwise comparisons across experimental groups.

To understand why transparency may or may not increase MWTP, or consumers may prefer CO₂ reductions over offset, I examine several mechanisms suggested by the literature on pro-environmental behavior. Specifically, I interact the MWTP parameter for the carbon-neutral label with the variables¹⁵ listed in Table A.7 in Appendix A.1¹⁶. These variables capture mechanisms

¹⁵Since the warm glow variable is highly correlated with the guilt and social approval variables, it is excluded from the main estimation and included in a separate robustness test that omits the other two variables.

¹⁶Pre-registered binary indicators identifying participants who are confused only by transparent labels but not by standard labels, and those who trust only transparent but not standard labels, based on survey questions 18 and 19 in Section E of the Appendix. The preregistered psychological and contextual variables—confusion, trust, concern, climate worry, guilt, social approval, and beliefs about the polluter-pays principle—along with resource constraints (financial and time), are measured using the statements in the first ten rows of Table A.6 in the Appendix.

such as label trust and confusion, climate-related worry, guilt, perceived social approval, beliefs about who should bear the cost of mitigation, and constraints related to time and financial resources. I then compare the resulting coefficients across experimental groups using Wald tests. Because the 7-point Likert scale variables are mean-centered, their coefficients represent the change in MWTP associated with a one-unit increase; for dummy variables, they indicate a discrete shift compared to the reference group.

Because products frequently carry multiple sustainability claims, the marginal value of a carbon-neutral label is likely determined not only by absolute preferences for carbon-neutral claims but also by trade-offs and salience in a broader labeling environment. To explore potential crowd-out or complementarity between sustainability labels, I therefore include interaction terms between the MWTP parameter for the carbon-neutral label and the organic and ethical-trade labels. This specification enables estimation of the MWTP for the carbon-neutral label when no other label is present, when either organic or ethical-trade is present, and when both labels appear simultaneously, and it allows testing whether these differ across experimental groups in these different situations.

3 Results

3.1 Transparency and willingness to pay

In this section, I begin by presenting MWTP estimates for the carbon-neutral label across the three experimental groups: a control group which saw a standard (non-transparent) label; Treatment 1, which saw a transparent label with 95% offsetting and 5% reduction; and Treatment 2, which saw a transparent label with a 50%–50% split. I then show two key results: (i) transparency regarding the shares of CO₂ offsetting and reduction does not increase consumers' MWTP; and (ii) consumers do not show a preference for CO₂ reductions over CO₂ offsets.

I first present the MWTP estimates for the carbon-neutral label across the three experimental groups. Table 3 reports the results from a MXL model es-

estimated in WTP space. In this specification, the estimated coefficients for the attributes, including the carbon-neutral labels, are expressed directly in monetary terms, while the corresponding standard deviations capture unobserved taste heterogeneity across respondents. The coefficients for the carbon-neutral label are positive and statistically significant at the 1% level in all three groups, indicating that respondents are willing to pay a premium for carbon neutrality. The implied MWTP is £0.55 in the control group, £0.52 in Treatment 1, and £0.32 in Treatment 2. Relative to the average product price used in the experiment (£3.90), these values correspond to 14%, 13%, and 8% of the average price, respectively.

Building on the baseline MWTP estimates, I first examine whether transparency affects consumers' MWTP, corresponding to the first main research question. Poe tests with Holm–Bonferroni correction reported in Table 4 show no statistically significant differences between the control group and either Treatment 1 or Treatment 2.¹⁷ Relative to the control group, MWTP is lower by £0.03 [£-0.28, £0.22] in Treatment 1 and by £0.23 [£-0.47, £0.01] in Treatment 2. These correspond to reductions of about 5.5% and 41.8% of the control-group MWTP, respectively, but neither is statistically significant. Overall, there is no evidence that transparency increases MWTP for carbon-neutral labels; if anything, it slightly reduces it.

The results indicate that consumers do not value CO₂ reductions more than CO₂ offsets. If reductions were preferred, MWTP should be higher in Treatment 2 than in Treatment 1. The results indicate the opposite: MWTP in Treatment 2 is about £0.20 [£-0.46, £0.06] lower than in Treatment 1, which corresponds to a reduction of about 38% of MWTP. This difference is not statistically significant after correcting for multiple comparisons (Holm–Bonferroni). Overall, an equal share of reduction and offsetting does not lead to a higher MWTP premium; if anything, it leads to valuations similar to or lower than those for the label with a higher share of offsetting, indicating no preference for carbon reductions over offsets, and vice versa.

¹⁷Holm–Bonferroni adjusted Wald tests comparing coefficients across groups yield the same conclusion.

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of parameters</i>				
MWTP _{Carbon neutral}	0.47*** (0.11)	0.55*** (0.08)	0.52*** (0.10)	0.32*** (0.09)
MWTP _{Organic}	1.10*** (0.07)	1.04*** (0.09)	1.15*** (0.11)	1.13*** (0.10)
MWTP _{Ethical trade}	1.17*** (0.07)	1.15*** (0.11)	1.14*** (0.12)	1.23*** (0.11)
μ_{Price}	-0.17*** (0.04)	-0.11 (0.07)	-0.18*** (0.06)	-0.21*** (0.07)
$\mu_{\text{Status quo}}$	-4.26*** (0.19)	-4.73*** (0.33)	-3.97*** (0.27)	-4.28*** (0.29)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	1.24*** (0.46)	1.31*** (0.12)	1.36*** (0.24)	1.25*** (0.12)
σ_{Organic}	1.64*** (0.10)	1.41*** (0.14)	1.57*** (0.14)	1.69*** (0.10)
$\sigma_{\text{Ethical trade}}$	1.48*** (0.12)	1.40*** (0.19)	1.43*** (0.15)	1.47*** (0.15)
σ_{Price}	0.77*** (0.07)	0.78*** (0.07)	0.74*** (0.09)	0.82*** (0.07)
$\sigma_{\text{Status quo}}$	2.27*** (0.13)	2.49*** (0.24)	2.12*** (0.18)	2.22*** (0.19)
Log Likelihood	-8540.44	-2808.57	-2898.27	-2822.83
AIC	17100.88	5637.13	5816.54	5665.67
BIC	17173.54	5698.88	5878.24	5727.21
Pseudo-R ²	0.26	0.28	0.25	0.26
Number of observations	10568	3552	3536	3480
Number of participants	1321	444	442	435

Table 3: Mixed logit model in WTP space, all parameters randomized

This table reports the results from a Mixed logit model estimated in WTP space. Robust standard errors are shown in parentheses. Holm-Bonferroni adjusted p -values (four tests for MWTP estimates and status quo parameter) are used to determine statistical significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$. Holm-Bonferroni adjusted Wald tests for MWTP estimates (for pairwise comparisons across the three experimental groups) indicate no statistically significant differences in MWTP coefficients for carbon-neutral, organic, and ethical trade labels.

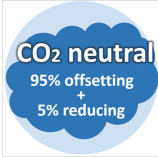
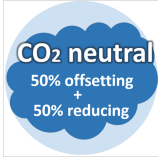
			
	Control	Treatment 1	Treatment 2
MWTP_{Carbon neutral}	0.55*** (0.08)	0.52*** (0.10)	0.32*** (0.09)
<i>Holm-Bonferroni adjusted Poe test</i>			
Control vs Treatment 1	0.40 (p = 0.793)		
Control vs Treatment 2	0.03 (p = 0.152)		
Treatment 1 vs Treatment 2		0.06 (p = 0.251)	
<i>Unadjusted Poe test</i>			
Control vs Treatment 1	0.40 (p = 0.793)		
Control vs Treatment 2	0.03* (p = 0.051)		
Treatment 1 vs Treatment 2		0.06 (p = 0.125)	

Table 4: Poe test results comparing MWTP estimates for the carbon-neutral label between experimental groups

MWTP estimates are reported with robust standard errors in parentheses. Significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$. Poe test statistics and the corresponding Holm-Bonferroni adjusted p -values (for three pairwise comparisons across experimental groups) are provided in parentheses. Unadjusted p -values are also reported separately.

3.2 Robustness tests

The robustness tests reported in Appendix B show that the conclusions about the insignificant differences between experimental groups do not change. Because the MWTP estimates could, in principle, change based on specific utility specifications, distributional assumptions, or sample restrictions, I conduct a comprehensive set of robustness checks to assess their stability, in line with best practice in the discrete choice literature (Hensher et al., 2005a; Train, 2009).

The first set of robustness tests varies (i) whether models are estimated in preference space or WTP space and (ii) whether coefficients are treated as fixed or random. In the preference-space models, Table A.12 presents an MNL specification without random parameters, while Tables A.13 and A.14 report MXL specifications with non-monetary attributes and all parameters treated as random. In the WTP-space models,¹⁸ Tables A.18 and A.19 report an MNL specification and an MXL specification with non-monetary attributes treated as random, respectively.

The second set of robustness tests with WTP-space models varies the simulation precision and model specification. Table A.20 doubles the number of Halton draws to 2,000; Table A.21 introduces alternative-specific error components; and Table A.22 allows for correlations among the error terms of non-monetary attributes.

The third set of robustness checks involves relaxing or tightening sample restrictions. Tables A.23, A.24, and A.25 estimate the WTP-space models by (i) including protest responses, (ii) excluding survey speeders, and (iii) excluding participants who failed attention or manipulation checks, respectively.

The corresponding Poe test results for the robustness checks, reported for preference-space estimates in Tables A.16 and for WTP-space estimates in Tables A.31, A.32, and A.33, show no significant differences in MWTP between groups. Across all these robustness checks, there are only two marginal exceptions: the specification with twice as many Halton draws (Table A.20) and

¹⁸Sensitivity checks with cost parameter starting values of 0.1 or 1 yield identical results (Table 3) and are not reported separately.

the model that includes protest responses (Table A.23), both of which indicate marginal 10% significance for Treatment 2 versus the control group, with the corresponding Poe test results reported in Tables A.31 and A.33, respectively.

A complementary check using an open-ended CVM approach also does not change the main conclusion of no statistically significant differences in MWTP across treatments. The CVM estimates in Table A.34, based on 10,000 bootstrapped samples, yield higher WTP in Treatment 2 than in the control group and Treatment 1, in contrast to the DCE results. However, none of these differences are statistically significant after applying the Holm adjustment, again indicating no statistically robust differences in MWTP across treatments.

Finally, Appendix A.2 examines whether the effects of transparency and preferences for CO₂ reductions vary by socioeconomic characteristics (income, education, age, gender, employment). Table A.35 indicates that only the high-income interaction with MWTP for the carbon-neutral label differs marginally between Treatment 1 and Treatment 2 (Holm-adjusted $p = 0.091$); all other pairwise comparisons of interaction coefficients are not significant. Overall, transparency effects appear broadly similar across socioeconomic groups, with weak evidence of a larger effect among higher-income respondents when the label indicates an equal share of reduction and offsetting.

3.3 Mechanisms

In this section, I examine mechanisms that may underlie the null findings, in particular the statistically insignificant effect of transparency on MWTP for carbon-neutral labels and the absence of a preference for CO₂ reductions over offsets.

I focus on three sets of channels. First, I evaluate preregistered *psychological and contextual factors*—confusion, trust, concern, climate worry, guilt, social approval, and beliefs about the polluter-pays principle—together with *resource constraints* (financial and time), based on statements listed in the first ten rows of Table A.6 in the Appendix. I also consider preregistered binary variables that identify a subgroup of participants who are confused only by transparent labels

but not by standard labels, and who trust only transparent but not standard labels, based on survey questions 18 and 19 in Section E of the Appendix. Second, I examine correct *understanding* of the term “offsetting.” Third, I consider *additional factors*, such as participants’ familiarity with carbon-neutral labels, the perceived consequentiality of the survey (i.e., whether participants think that the survey results affect labeling or pricing policies), and attribute non-attendance for the carbon-neutral label (participants ignoring the carbon-neutral label while making choices).¹⁹

To motivate this mechanism analysis, I first present descriptive evidence on trust and confusion for the transparent and standard labels. Table 5 reports these patterns based on a follow-up question shown after the choice experiment, where participants were presented with all three labels. In this question, the transparent 50/50 label is reported as the most trusted (39.29%), whereas the 95% offset/5% reduction label is reported as the most confusing (34.07%). These descriptive differences provide a natural starting point for the analysis: do *confusion* about the additional information or *trust* in the labels relate to MWTP differences between the control and treatment groups?

Turning to the results, I first find no evidence that trust, confusion, other psychological and contextual factors, or resource constraints are associated with differences in MWTP across experimental groups. Table 6 estimates associations between MWTP and the listed preregistered factors.²⁰ Holm-adjusted Wald tests are used to check whether these associations differ by treatment status. Despite variation in signs and magnitudes, I find no statistically significant between-group differences. These results are robust to alternative specifications shown in Tables A.26—A.28 in Appendix B.²¹

¹⁹Further factors tested include beliefs about the effectiveness of offsetting; views on putting a price on nature; moral licensing; perceptions of greenwashing; survey time, which might capture attention to the survey; a manipulation check; additional attention checks; membership in an environmental organization; and response certainty (how certain respondents are about their choices).

²⁰Seven-point Likert covariates are mean-centred and treated as continuous; a one-unit change corresponds to a change in MWTP (GBP). Binary indicators for respondents who are only (i) confused by or (ii) trusting of transparent labels (but not standard labels) reflect average MWTP differences relative to the remainder of the sample.

²¹While the “trust” variable is positively associated with MWTP on average, its association does not vary with transparency or a higher reduction share, indicating that it captures general attitudes rather than responses to specific disclosure content.

Second, the descriptive statistics in Table A.4 shows limited understanding of (and attention to) the meaning of “CO₂ offsetting” among nearly half of participants. At the same time, the main results remain the same when restricting the sample to those who pass this attention check,²² as reported in Table A.29. This suggests that correct understanding of the term “offsetting” does not translate into a higher WTP for transparency, noting that only neutral definitions were provided and potential criticisms of offsetting were not communicated.

Third, I examine an additional set of factors which may be associated with differences in MWTP between experimental groups. Familiarity with carbon-neutral labels, measured by whether respondents have purchased such labeled products while grocery shopping, is consistently associated with higher WTP, indicating the importance of prior exposure in shaping valuation. Table A.30 indicates an average increase of about £1.57 associated with familiarity in the full sample. Holm-adjusted Wald tests for the coefficients across experimental groups also reject the equality of coefficients across groups, with the association being stronger when (i) the label is transparent and (ii) the CO₂ reduction share is higher. Other coefficients, including perceptions of survey consequentiality and attribute non-attendance for the carbon-neutral label, are smaller and statistically insignificant.

Overall, the evidence suggests that differences (and the lack thereof) in MWTP across experimental groups are driven less by a wide range of psychological and contextual factors or resource and time constraints, and more by prior familiarity with carbon-neutral certifications. At the same time, I do not identify a clear mechanism that explains why transparent labels sometimes yield lower (or higher) MWTP than standard labels. Future research could explore alternative explanations, such as order effects in the presentation of offsetting and reduction on the labels or biases related to large numerical shares of offsetting versus reduction.

²²Holm-adjusted Wald statistics for the interacted coefficient of participants who correctly understand “offsetting” across groups are not significant: control vs. treatment 1 = 0.878 ($p = 1.000$), control vs. treatment 2 = 0.450 ($p = 1.000$), treatment 1 vs. treatment 2 = 0.383 ($p = 1.000$).

		Full sample		Control		Treatment 1		Treatment 2	
		N	Share	N	Share	N	Share	N	Share
Trust for label 1		309	23.39%	109	24.55%	90	20.36%	110	25.29%
		330	24.98%	114	25.68%	124	28.05%	92	21.15%
		519	39.29%	179	40.32%	164	37.10%	176	40.46%
Trust (none)		264	19.98%	80	18.02%	97	21.95%	87	20.00%
Confusion about label 1		341	25.81%	111	25.00%	130	29.41%	100	22.99%
Confusion about label 2		450	34.07%	162	36.49%	124	28.05%	164	37.70%
Confusion about label 3		336	25.44%	106	23.87%	116	26.24%	114	26.21%
Confusion (none)		393	29.75%	131	29.50%	144	32.58%	118	27.13%

Table 5: Trust in and confusion about standard and transparent carbon-neutral labels

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 options (Label 1, Label 2, Label 3) and a 'none' option. This question was asked after the choice experiment. Multiple selections were allowed.

Label 1 represents the standard carbon-neutral label. Label 2 is the transparent carbon-neutral label with 95% CO₂ offsetting and 5% CO₂ reduction. Label 3 is the transparent carbon-neutral label with an equal split between CO₂ reduction and CO₂ offsetting (50%-50%).

	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral}	-0.26 (0.16)	-0.43 (0.29)	-0.25 (0.26)	-0.19 (0.29)
Main interactions				
MWTP _{Carbon neutral} x Only confused with transparent labels	0.03 (0.15)	0.21 (0.20)	0.24 (0.19)	-0.32 (0.24)
MWTP _{Carbon neutral} x Only trust transparent labels	0.65*** (0.14)	0.80*** (0.23)	0.39 (0.20)	0.72** (0.23)
MWTP _{Carbon neutral} x Concern level	0.18*** (0.05)	0.08 (0.10)	0.30*** (0.08)	0.15 (0.18)
MWTP _{Carbon neutral} x Trust level	0.09 (0.05)	0.10 (0.08)	0.12 (0.07)	-0.05 (0.09)
MWTP _{Carbon neutral} x Confusion level	-0.03 (0.06)	0.01 (0.06)	-0.00 (0.06)	-0.16 (0.08)
MWTP _{Carbon neutral} x Climate worry	0.06 (0.05)	0.01 (0.10)	-0.05 (0.08)	0.25 (0.22)
MWTP _{Carbon neutral} x Guilt	0.12 (0.05)	0.15 (0.07)	0.11 (0.08)	0.13 (0.13)
MWTP _{Carbon neutral} x Social approval	0.18*** (0.05)	0.29* (0.11)	0.21 (0.09)	0.00 (0.09)
MWTP _{Carbon neutral} x Polluter pays	-0.02 (0.04)	-0.01 (0.09)	-0.12 (0.06)	0.03 (0.09)
MWTP _{Carbon neutral} x Financial constraints	-0.18*** (0.04)	-0.16 (0.11)	-0.15 (0.07)	-0.22 (0.09)
MWTP _{Carbon neutral} x Time restrictions	-0.02 (0.04)	-0.08 (0.07)	-0.03 (0.06)	0.06 (0.06)
Socioeconomic interactions	Yes	Yes	Yes	Yes
Sustainability label attributes, price, status quo	Yes	Yes	Yes	Yes
Standard deviations of random parameters	Yes	Yes	Yes	Yes
Log Likelihood	-8371.60	-2746.65	-2835.84	-2756.99
AIC	16797.20	5547.30	5725.68	5567.97
BIC	16993.37	5714.04	5892.29	5734.15
Pseudo R ²	0.28	0.30	0.27	0.28
Number of observations	1056	3552	3536	3480
Number of participants	1321	444	442	435

Table 6: Mixed logit model in WTP space with interaction variables

This table shows the choice model output from the mixed logit model estimated in WTP space with interaction variables. Robust standard errors are reported in parentheses. All parameters, except for the interaction terms, are randomized. Holm-Bonferroni adjusted p-values (for 21 comparisons) and significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$. Pre-registered binary indicators identifying participants who are confused only by transparent labels but not by standard labels, and those who trust only transparent but not standard labels, based on survey questions 18 and 19 in Section E of the Appendix. The preregistered psychological and contextual variables—confusion, trust, concern, climate worry, guilt, social approval, and beliefs about the polluter-pays principle—along with resource constraints (financial and time), are measured using the statements in the first ten rows of Table A.6 in the Appendix. Bilateral Wald tests were conducted to assess the equivalence of interaction coefficients across experimental groups. Using coefficients reported in the table, adjusted Wald test results show that none of the above interactions are statistically significant after adjusting for multiple testing using Holm-Bonferroni method for 21 comparisons, which include all of the main interactions, socioeconomic variables, WTP parameters, and the status quo parameter. Unadjusted Wald test results show that the interaction with concern level (about carbon offsets) is significant at $p = 0.079$ when comparing the control group and treatment group 1. The interaction with the binary variable indicating individuals confused only by transparent labels (but not by standard labels) is significant at $p = 0.065$ for the comparison between treatment group 1 and treatment group 2. For the comparison between the control group and treatment group 2, the interaction with the social approval variable is significant at $p = 0.038$, and the interaction with confusion level (about carbon-neutral labels) is significant at $p = 0.095$. Wald test results for coefficients not reported in the table (e.g., socioeconomic variables) are also not included in the table notes.

3.4 Competing labels

This section analyzes consumers' MWTP for two further sustainability labels on tea, organic and ethical trade, and examines how their presence changes the MWTP for transparent and standard (intransparent) versions of the carbon-neutral label.

First, I present the estimated MWTPs for each sustainability label across scenarios and show how these labels compete with each other (rather than complement each other). Second, I assess whether MWTP for the carbon-neutral label differs across experimental groups to test whether the presence of competing labels changes the main conclusions (the null effects of transparency and of a higher CO₂ reduction share).

Table 3 reports an average MWTP of 0.47 for the carbon-neutral label, 1.10 for the organic label, and 1.17 for the ethical-trade label. Relative to the average product price of 3.90 in the choice experiment, these values correspond to approximately 12%, 28%, and 30% of the product price. Similar values are obtained in all robustness checks, including the preference-space results summarised in Table A.17 and the additional WTP-space estimates in Tables A.18-A.25.

To understand whether other sustainability claims complement or compete with the carbon-neutral claim, two-way and three-way interaction terms between the sustainability labels, as well as correlated error terms, are introduced into the MXL model. Figure 2 plots the estimated MWTP for the carbon-neutral label under four scenarios. When no other labels are present, the MWTP for the carbon-neutral label equals £1.01 in the control group, £0.89 in treatment 1, and £0.99 in treatment 2. Adding only the organic label lowers the corresponding MWTP estimates to £0.69, £1.00, and £0.46. When only the ethical trade label is present, the MWTP estimates fall to £0.70, £0.90, and £0.51. Finally, when all labels are displayed simultaneously, the WTP for the carbon-neutral claim turns negative, reaching -£0.80 in the control group, -£0.57 in treatment 1, and -£0.80 in treatment 2, and is statistically significant.

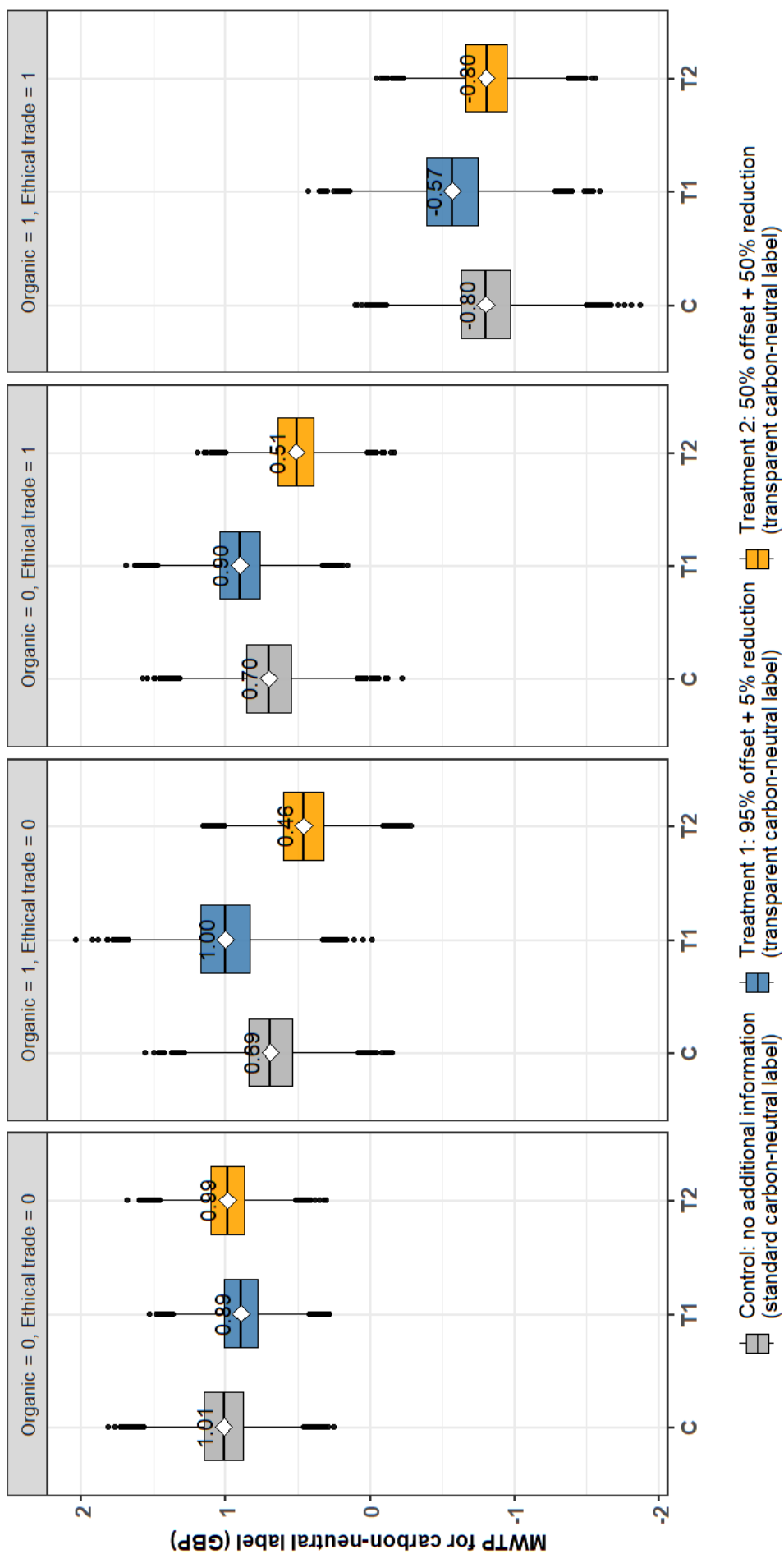


Figure 2: MWTP for the carbon-neutral label in the presence or absence of other sustainability labels

This figure shows the MWTP estimates derived from a mixed logit model with both two-way and three-way interaction terms for sustainability labels, and correlated error terms. It displays the WTP estimates for the carbon-neutral label depending on whether the other labels are present (coded as 1) or not present (coded as 0). The corresponding choice model output is provided in Table A.36 in Appendix C.

The negative interaction coefficients in Table A.37 confirm competition among the labels for the control group and treatment group 1: the three-way interaction equals $-\pounds 1.18$ in the control group, $-\pounds 1.58$ in treatment 1, and is statistically significant, and $-\pounds 0.78$ in treatment 2 but is statistically insignificant. While the interaction effects are smaller for the label implying a higher offsetting share, the results indicate that the two competing labels substantially erode the WTP premium otherwise paid for carbon neutrality.

Finally, I demonstrate that the presence of competing labels changes the main conclusions (i.e., the null effects of transparency and a higher CO₂ reduction share). Holm-adjusted Poe tests (Table A.38) show that, despite competition among labels, the differences between the experimental groups remain statistically insignificant. This result is robust when only two two-way interactions between the carbon-neutral label with organic and ethical-trade labels are included, and no correlations between error terms are introduced. See Table A.37 for the MXL model and Table A.39 for the Poe test results.

To sum up, although consumers value each sustainability certification, the organic and ethical-trade labels reduce, and when combined, even reverse, the additional amount they are willing to pay for a carbon-neutral label. However, in either case, there is no evidence that transparency or a higher CO₂ reduction share affects consumers' MWTP for the carbon-neutral label.

4 Conclusion

The motivation of this paper is based on the growing criticism of CO₂ offsets and increasing regulatory scrutiny of environmental claims. While carbon-neutral labels are becoming more common, many lack information on how carbon neutrality is achieved. Given that CO₂ offsets and reductions may differ substantially in their environmental effectiveness, transparency may play a key role in shaping consumers' WTP.

I find that consumers value carbon-neutral labels in general, but they do not value transparency per se, nor do they value a higher reduction share once

transparency is provided. Transparent versions are valued no more than the standard label, and an equal offset-reduction split is not valued more than a label with a larger share of offsets.

I further show that organic and ethical-trade claims compete with the carbon-neutral claim and substantially erode, and when combined can even reverse, its marginal value. However, the labels' coexistence does not change the core results: neither transparency nor a higher share of CO₂ reduction leads to a statistically significant increase in MWTP.

The policy implications are as follows. First, if policymakers expect markets to sort out the quality of carbon-neutral claims because consumers reward transparency, the evidence does not support this expectation: transparency alone is not valued. Second, if the goal is to shift demand from offsets toward reductions, simply requiring disclosure of offset and reduction shares is unlikely to achieve it. Transparency mandates are nonetheless consistent with anti-greenwashing objectives, but they could be paired with communication that distinguishes offsets from reductions in terms of their effectiveness in mitigating climate change. Furthermore, because multiple labels can crowd out the value of carbon-neutral claims, policy could also regulate label architecture.

Three avenues for future work emerge. First, rather than testing disclosure per se, experiments could randomize clarity interventions (brief explanations and simple visuals) to assess whether improved comprehension and an emphasis on the distinction between offsets and reductions change valuations of reduction-versus offset-based neutrality. Second, experiments could vary the order in which reductions and offsets are presented and the size of the reported reduction share to test for ordering and large-number effects. Third, research could study label architecture in realistic shopping environments, examining how order, salience, and the joint display of multiple sustainability claims shape attention, crowd-out, and the marginal value of carbon-neutral claims.

References

- 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.
- Akaichi, F., C. Revoredo Giha, K. Glenk, and J. M. Gil (2020). How consumers in the UK and Spain value the coexistence of the claims low fat, local, organic and low greenhouse gas emissions. *Nutrients* 12(1), 120.
- 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.
- Axelsson, K., A. Wagner, I. Johnstone, M. Allen, B. Caldecott, N. Eyre, S. Fankhauser, T. Hale, C. Hepburn, C. Hickey, R. Khosla, S. Lezak, E. Mitchell-Larson, Y. Malhi, N. Seddon, A. Smith, and S. M. Smith (2024). Oxford principles for net zero aligned carbon offsetting (revised 2024).
- Bek, D. (2022). Pricing sustainable shipping of coffee: Consumers' 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.
- Brouwer, R. (2011). A mixed approach to payment certainty calibration in discrete choice welfare estimation. *Applied Economics* 43(17), 2129–2142.
- Brouwer, R., L. Brander, and P. Van Beukering (2008). “A convenient truth”: air travel passengers’ willingness to pay to offset their CO₂ emissions. *Climatic Change* 90, 299–313.
- 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 (2025). Do carbon offsets offset carbon? *American Economic Journal: Applied Economics* 17(1), 1–40.
- Carattini, S. and J. Blasch (2024). Nudging when the descriptive norm is low: Evidence from a carbon offsetting field experiment. *Journal of Behavioral and Experimental Economics* 110, 102194.
- Carattini, S., F. Dvorak, I. Logar, and B. Ozdemir-Oluk (2025). Demand for carbon-neutral products. CESifo Working Paper 12232.
- Carattini, S., K. Gillingham, X. Meng, and E. Yoeli (2024). Peer-to-peer solar and social rewards: Evidence from a field experiment. *Journal of Economic Behavior & Organization* 219, 340–370.
- Carattini, S. and A. Tavoni (2016). How green are green economists? *Economics Bulletin* 36(4), 2311–2323.
- Carbon Trust (2023). Carbon neutral verification. Accessed: 2024-12-19.

- 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.
- 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.
- Daly, A., S. Hess, and K. Train (2012). Assuring finite moments for willingness to pay in random coefficient models. *Transportation* 39(1), 19–31.
- Darby, M. R. and E. Karni (1973). Free competition and the optimal amount of fraud. *The Journal of Law and Economics* 16(1), 67–88.
- 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.
- Forest Trends' Ecosystem Marketplace (2024). State of the voluntary carbon market 2024. Technical report, Forest Trends Association, Washington, DC.
- Gassler, B. (2015). How green is your “Grüner”? Millennial wine consumers' preferences and willingness-to-pay for eco-labeled wine. *Jahrbuch der Österreichischen Gesellschaft für Agrarökonomie* 24, 131–140.
- 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.
- Groom, B. and F. Venmans (2023). The social value of offsets. *Nature* 619(7971), 768–773.

- Gruchmann, T., G. Maugeri, and R. Wagner (2025). Do you feel guilty? a consumer-centric perspective on green nudging in last-mile deliveries. *International Journal of Physical Distribution & Logistics Management* 55(5), 540–566.
- Hensher, D. A., J. M. Rose, and W. H. Greene (2005a). *Applied Choice Analysis* (2nd ed.). Cambridge University Press.
- Hensher, D. A., J. M. Rose, and W. H. Greene (2005b). *Applied Choice Analysis: a Primer*. 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>.

- IPCC (2022). *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK and New York, NY, USA: Cambridge University Press.
- Kahneman, D. and J. L. Knetsch (1992). Valuing public goods: the purchase of moral satisfaction. *Journal of Environmental Economics and Management* 22(1), 57–70.
- Kim, E.-H. and T. Lyon (2011). When does institutional investor activism increase shareholder value?: The Carbon Disclosure Project. *The B.E. Journal of Economic Analysis & Policy* 11(1).
- 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.
- List, J. A., A. M. Shaikh, and Y. Xu (2019). Multiple hypothesis testing in experimental economics. *National Bureau of Economic Research Working Paper No. 25848*.
- 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.
- Mariel, P., D. Hoyos, J. Meyerhoff, M. Czajkowski, T. Dekker, K. Glenk, J. B. Jacobsen, U. Liebe, S. B. Olsen, J. Sagebiel, et al. (2021). *Environmental valuation with discrete choice experiments: Guidance on design, implementation and data analysis*. Springer Nature.

- 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.
- 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.
- Murray, B. C., B. A. McCarl, and H.-C. Lee (2004). Estimating leakage from forest carbon sequestration programs. *Land Economics* 80(1), 109–124.
- Net Zero Tracker (2025). Net zero stocktake 2025. Accessed 2025-11-03.
- 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.
- 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.
- Rogelj, J., O. Geden, A. Cowie, and A. Reisinger (2021). Net-zero emissions targets are vague: Three ways to fix. *Nature* 591(7850).
- 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.

- Swait, J. and J. Louviere (1993). The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of marketing research* 30(3), 305–314.
- 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 (2023). 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 Wall Street Journal (2024). Companies must explain what “climate neutral” means—or not use it, german court says. Accessed: March 25, 2025.
- Theotokis, A. and E. Manganari (2015). The impact of choice architecture on sustainable consumer behavior: The role of guilt. *Journal of Business Ethics* 131, 423–437.
- Train, K. and M. Weeks (2005). Discrete choice models in preference space and willingness-to-pay space. In R. Scarpa and A. Alberini (Eds.), *Applications of Simulation Methods in Environmental and Resource Economics*, Volume 6 of *The Economics of Non-Market Goods and Resources*, pp. 1–17. Springer, Dordrecht.
- 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.

- Vecchio, R. and A. Annunziata (2015). Willingness-to-pay for sustainability-labelled chocolate: An experimental auction approach. *Journal of Cleaner Production* 86, 335–342.
- Ziegler, A., J. Schwarzkopf, and V. H. Hoffmann (2012). Stated versus revealed knowledge: Determinants of offsetting CO₂ 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, attendance, consequentiality, and protest responses of participants. Table A.6 presents the level of agreement with various statements using a Likert scale. Table A.7 lists the variables interacted with the carbon-neutral label attribute and their definitions. Finally, Table A.9 shows covariate balance.

Overall, the descriptive tables document three main facts. First, the three treatment groups are very similar in terms of age, gender, education, employment, and income, even when using detailed category breakdowns, with differences across groups typically within a few percentage points, which supports the credibility of the randomized design. Second, the sample largely consists of regular tea consumers for whom the choice scenario is realistic: about 78% report drinking tea daily and roughly 88% at least once a week, and most purchase tea at least monthly. Third, survey-quality indicators are reassuring: 99% of respondents report clear instructions, only about 7% report confusion with the survey, more than half perceive the choices as consequential, and protest responses remain below 2% of the sample, while attribute non-attendance and misperceptions about offsetting versus reduction are non-negligible but similar across experimental groups.

	Full sample		Control		Treatment 1		Treatment 2	
	N	Share	N	Share	N	Share	N	Share
Sample size	1321		444		442		435	
Age								
18 - 34 years	410	31.04%	135	30.41%	134	30.32%	141	32.41%
35 - 54 years	469	35.50%	157	35.36%	158	35.75%	154	35.40%
55+ years	442	33.46%	152	34.23%	150	33.94%	140	32.18%
Gender								
Male	645	48.83%	227	51.13%	206	46.61%	212	48.74%
Female	666	50.42%	214	48.20%	231	52.26%	221	50.80%
Non-binary	9	0.68%	3	0.68%	5	1.13%	1	0.23%
I prefer not to say	1	0.08%	-	-	-	-	1	0.23%
Education								
Primary school	3	0.23%	2	0.45%	1	0.23%	-	-
Secondary school: High school or equivalent	359	27.18%	130	29.28%	124	28.05%	105	24.14%
Post-secondary vocational training (up to 1 year)	55	4.16%	16	3.60%	27	6.11%	12	2.76%
Post-secondary vocational training (2 and more years)	153	11.58%	44	9.91%	55	12.44%	54	12.41%
Post-secondary academic below-degree level qualification (up to 1 year)	80	6.06%	31	6.98%	23	5.20%	26	5.98%
Post-secondary academic below-degree level qualification (2 and more years)	171	12.94%	49	11.04%	50	11.31%	72	16.55%
Bachelors or equivalent first degree qualification (e.g., BA, BSc, BEng)	311	23.54%	111	25.00%	95	21.49%	105	24.14%
Masters or equivalent higher degree level qualification (e.g., MA, MSc, MBA)	140	10.60%	49	11.04%	47	10.63%	44	10.11%
PhD or equivalent doctoral level qualification (e.g., PhD)	39	2.95%	8	1.80%	19	4.30%	12	2.76%
None of above	10	0.76%	4	0.90%	1	0.23%	5	1.15%

Table A.1: Summary statistics of sociodemographics

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

	Full sample		Control		Treatment 1		Treatment 2	
	N	Share	N	Share	N	Share	N	Share
Sample size	1321		444		442		435	
Employment								
Full or part time employment	760	57.53%	248	55.86%	257	58.14%	255	58.62%
Self-employed	95	7.19%	32	7.21%	29	6.56%	34	7.82%
Unemployed	82	6.21%	24	5.41%	28	6.33%	30	6.90%
Retired	263	19.91%	90	20.27%	87	19.68%	86	19.77%
Looking after family or home	77	5.83%	30	6.76%	26	5.88%	21	4.83%
Full-time student	32	2.42%	14	3.15%	11	2.49%	7	1.61%
None of above	12	0.91%	6	1.35%	4	0.90%	2	0.46%
Annual household income after taxes								
Under 10,000	75	5.68%	23	5.18%	22	4.98%	30	6.90%
10,000 - 19,999	177	13.40%	50	11.26%	62	14.03%	65	14.94%
20,000 - 29,999	250	18.93%	85	19.14%	88	19.91%	77	17.70%
30,000 - 39,999	174	13.17%	71	15.99%	48	10.86%	55	12.64%
40,000 - 49,999	136	10.30%	43	9.68%	53	11.99%	40	9.20%
50,000 - 59,999	103	7.80%	32	7.21%	34	7.69%	37	8.51%
60,000 - 69,999	68	5.15%	24	5.41%	16	3.62%	28	6.44%
70,000 - 79,999	49	3.71%	15	3.38%	18	4.07%	16	3.68%
80,000 - 89,999	43	3.26%	10	2.25%	21	4.75%	12	2.76%
90,000 - 99,999	31	2.35%	12	2.70%	9	2.04%	10	2.30%
100,000 - 129,999	70	5.30%	22	4.95%	27	6.11%	21	4.83%
130,000 or more	49	3.71%	16	3.60%	15	3.39%	18	4.14%
No answer	96	7.27%	41	9.23%	29	6.56%	26	5.98%
Environmental organization membership								
Member	149	11.28%	46	10.36%	55	12.44%	48	11.03%

Table A.1: Summary statistics of sociodemographics (continued)

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

	Non-tea consumers (N = 222)	Tea consumers (N = 1,339)	SMD
Descriptive statistics			
Age, (mean, SD)	50.54 (16.78)	47.54 (16.52)	0.180
Male, n (%)	111 (50.0%)	653 (48.8%)	0.025
Higher than post-secondary education (up to 1 year), n (%)	52 (23.4%)	491 (36.7%)	-0.292

Table A.2: Comparison of tea consumers and non-tea consumers

This table displays descriptive statistics and standardized mean differences (SMDs) between tea consumers and non-tea consumers across age, gender, and education characteristics. Note that the tea consumer sample includes protest responses.

	Full sample		Control		Treatment 1		Treatment 2	
	N	Mean	N	Mean	N	Mean	N	Mean
Sample size	1321		444		442		435	
Tea drinking frequency								
Daily	1027	77.74%	338	76.13%	343	77.60%	346	79.54%
Once a week	138	10.45%	52	11.71%	50	11.31%	36	8.28%
Once every two weeks	60	4.54%	18	4.05%	21	4.75%	21	4.83%
Once a month	35	2.65%	10	2.25%	10	2.26%	15	3.45%
Several times a year	61	4.62%	26	5.86%	18	4.07%	17	3.91%
Tea purchase frequency								
Once a week	252	19.08%	94	21.17%	83	18.78%	75	17.24%
Once every two weeks	271	20.51%	80	18.02%	98	22.17%	93	21.38%
Once a month	505	38.23%	173	38.96%	160	36.20%	172	39.54%
Several times a year	253	19.15%	83	18.69%	85	19.23%	85	19.54%
Once a year	40	3.03%	14	3.15%	16	3.62%	10	2.30%
Preferred tea blend								
Black tea	955	72.29%	317	71.40%	324	73.30%	314	72.18%
Green tea	212	16.05%	75	16.89%	62	14.03%	75	17.24%
Herbal tea	105	7.95%	36	8.11%	39	8.82%	30	6.90%
Other	49	3.71%	16	3.60%	17	3.85%	16	3.68%
Regular grocery shopping								
Carbon-neutral label	168	12.72%	54	12.16%	61	13.80%	53	12.18%
Organic label	321	24.30%	109	24.55%	111	25.11%	101	23.22%
Fair trade label	485	36.71%	169	38.06%	161	36.43%	155	35.63%
No sustainability label	249	18.85%	80	18.02%	81	18.33%	88	20.23%
Other label	26	1.97%	15	3.38%	4	0.90%	7	1.61%
No knowledge about label	391	29.60%	127	28.60%	131	29.64%	133	30.57%

Table A.3: Summary statistics: tea consumption

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

	Full		Control		Treatment 1		Treatment 2	
	N	Mean	N	Mean	N	Mean	N	Mean
Sample size	1321		444		442		435	
Survey completion time								
Choice experiment time (min.)	1321	3.09	444	4.69	442	1.86	435	2.70
Full survey time (min.)	1321	14.61	444	16.14	442	11.41	435	16.31
Device								
Smartphone	741	56.09%	239	53.83%	260	58.82%	242	55.63%
Tablet	42	3.18%	12	2.70%	13	2.94%	17	3.91%
Desktop	538	40.73%	193	43.47%	169	38.24%	176	40.46%
Attention and manipulation								
<i>The number of labels</i>								
One label	65	4.92%	28	6.31%	20	4.52%	17	3.91%
Two labels	334	25.28%	103	23.20%	119	26.92%	112	25.75%
Three labels (correct resp.)	821	62.15%	281	63.29%	266	60.18%	274	62.99%
Not remember.	101	7.65%	32	7.21%	37	8.37%	32	7.36%
<i>Percentage info. on label</i>								
Yes (correct for treat. groups)	567	42.41%	95	21.30%	242	54.14%	230	51.80%
No (correct for control group)	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%
<i>Definition of CO₂ offsetting</i>								
Correct resp.	650	49.21%	223	50.23%	225	50.90%	202	46.44%
False resp. (CO ₂ reduction def.)	475	35.96%	156	35.14%	149	33.71%	170	39.08%
Not remember.	196	14.84%	65	14.64%	68	15.38%	63	14.48%
Survey clarity								
<i>Clear instructions</i>								
No.	9	0.68%	-	-	5	1.13%	4	0.92%
Yes.	1312	99.32%	444	100.00%	437	98.87%	431	99.08%
<i>Confusion with survey</i>								
No.	1232	93.26%	422	95.05%	415	93.89%	395	90.80%
Yes.	89	6.74%	22	4.95%	27	6.11%	40	9.20%

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

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.

	Full sample		Control		Treatment 1		Treatment 2	
	N	Share	N	Share	N	Share	N	Share (%)
Certainty level (0-10)								
0-2 (very uncertain)	10	0.76%	3	0.68%	5	1.13%	2	0.46%
3-5	119	9.01%	39	8.78%	42	9.50%	38	8.74%
6-8	627	47.46%	214	48.20%	198	44.80%	215	49.43%
9-10 (very certain)	565	42.77%	188	42.34%	197	44.57%	180	41.38%
Attribute-non-attendance								
Not considered: CN label	440	33.31%	134	30.18%	154	34.84%	152	34.94%
Not considered: Organic label	450	34.07%	128	28.83%	154	34.84%	168	38.62%
Not considered: Ethical trade l.	315	23.85%	97	21.85%	108	24.43%	110	25.29%
Not considered: Price	266	20.14%	103	23.20%	73	16.52%	90	20.69%
Considered all attributes	400	30.28%	144	32.43%	138	31.22%	118	27.13%
Consequentiality								
Yes (policy and price impact).	706	53.44%	229	51.58%	253	57.24%	224	51.49%
No (no impact).	615	46.56%	215	48.42%	189	42.76%	211	48.51%
Consistent status-quo response								
Total protest responses	18	1.34%	4	0.89%	5	1.12%	9	2.03%
The products were too expensive.	10	0.75%	4	0.89%	3	0.67%	3	0.68%
I oppose one or more of the labels. [P]	3	0.22%	1	0.22%	-	-	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%	-	-
I disagree with the way the choice question was asked. [P]	2	0.15%	-	-	-	-	2	0.45%
Other reason* [P]	6	0.45%	1	0.22%	3	0.67%	2	0.45%

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

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

*This category includes protest responses based on various open-ended reasons. In addition to pre-specified categories, other responses were classified as protests if participants indicated that they consume only one brand, do not drink tea, or never purchase loose-leaf tea (possibly due to misinterpreting the question). One respondent mentioned that the options were expensive; however, since a pre-specified category already covered this reason, the response was coded accordingly and not classified as a protest.

Statement	SD	MD	Sltd	N	Slta	MA	SA
I worry about climate change.	5.98	5.45	7.57	13.02	28.08	22.79	17.11
Limited financial resources prevent me from buying climate-friendly products.	5.83	4.69	8.55	16.73	23.54	21.88	18.77
Lack of time prevents me from buying climate-friendly products.	16.35	13.02	19.53	24.30	12.94	8.71	5.15
My positive emotions increase when I choose climate-friendly products.	8.02	6.43	9.31	31.87	22.71	13.40	8.25
I feel guilty when I buy conventional products.	14.84	11.36	18.85	21.88	16.96	8.71	7.42
Most people approve of my choice of climate-friendly products.	5.53	3.71	8.10	47.16	15.97	12.34	7.19
Producers are responsible for climate change mitigation costs.	2.35	3.63	8.48	23.54	23.92	21.88	16.20
I trust carbon-neutral labels.	5.90	6.06	12.64	31.26	25.44	13.55	5.15
I am confused about carbon-neutral labels.	5.00	8.25	14.84	26.87	29.37	10.14	5.53
I am concerned about carbon offsets.	5.07	4.69	13.63	30.81	25.28	11.73	8.78
Carbon offsetting reduces carbon emissions.	5.60	4.92	12.49	31.11	27.93	12.41	5.53
Carbon offsetting allows producers to continue polluting.	1.89	2.95	8.02	30.20	30.96	15.37	10.60
Carbon offsetting is a misleading sense of relief.	2.57	1.67	6.28	27.18	32.10	19.15	11.05
Carbon offsetting is a form of greenwashing.	1.74	1.97	6.43	33.61	28.16	15.82	12.26

Table A.6: Agreement with the list of statements (full sample)

Values indicate the percentage of respondents who selected each agreement level. SD = Strongly Disagree, MD = Mostly Disagree, Sltd = Slightly Disagree, N = Neutral, Slta = Slightly Agree, MA = Mostly Agree, SA = Strongly Agree.

Variable	Description	Measurement
Only confused with transparent labels	Participants who are only confused with transparent labels but not with the standard labels	Binary variable (1 = yes, 0 = no)
Only trust transparent labels	Participants who only trust transparent labels but not the standard labels	Binary variable (1 = yes, 0 = no)
Concern level	Concern level about CO ₂ offsetting	7-point Likert scale (mean-centered)
Confusion level	Confusion about carbon neutral labels	7-point Likert scale (mean-centered)
Trust level	Trust in carbon-neutral labels	7-point Likert scale (mean-centered)
Climate worry	Worry level about climate change	7-point Likert scale (mean-centered)
Warm glow	Positive emotions from climate-friendly purchases	7-point Likert scale (mean-centered)
Guilt	Negative emotions when not making climate-friendly choices	7-point Likert scale (mean-centered)
Social approval	Perceived acceptance by others for climate-friendly choices	7-point Likert scale (mean-centered)
Polluter pays	Belief that producers should pay for climate mitigation	7-point Likert scale (mean-centered)
Financial constraints	Limited financial resources for climate-friendly purchases	7-point Likert scale (mean-centered)
Time restrictions	Limited time for climate-friendly choices	7-point Likert scale (mean-centered)
Gender	Male respondents	Binary variable (1 = male, 0 = female, non-binary, or not disclosed)
Age	Age of the respondent	Continuous variable (mean-centered)
High education	Highest educational qualification: Bachelor's degree or higher	Binary variable (1 = bachelor's degree or higher, 0 = lower than bachelor's degree)
Employed	Employment status	Binary variable (1 = employed, 0 = unemployed, retired, full-time student, looking after family or home, other)
High income	Yearly household income equal to or greater than 50,000 after taxes	Binary variable (1 = high income, 0 = low income or not disclosed)
Not disclosed income	Participants' income information disclosure in the survey	Binary variable (1 = not disclosed, 0 = disclosed)

Table A.7: Description of main interacted covariates

This table lists the covariates interacted with the carbon-neutral label to examine the factors associated with consumers' choices. Variables measured on 7-point Likert scales were converted to continuous, mean-centered variables, with higher values indicating stronger agreement and lower values indicating stronger disagreement.

Covariate	Control (Mean)	Treatment 1 (Mean)	Treatment 2 (Mean)	C - T1 (SMD)	C - T2 (SMD)	T1 - T2 (SMD)
Only confused with transparent labels	0.45	0.38	0.50	0.15	-0.09	-0.24
Only trust transparent labels	0.57	0.58	0.55	-0.01	0.05	0.06
Confusion level	4.25	4.16	4.18	0.06	0.05	-0.01
Trust level	4.28	4.21	4.28	0.05	0.00	-0.05
Concern level (for CO ₂ offsets)	4.45	4.32	4.34	0.09	0.07	-0.01
Climate worry	4.96	4.79	4.91	0.10	0.03	-0.08
Warm glow	4.43	4.20	4.21	0.15	0.14	-0.00
Guilt	3.80	3.67	3.64	0.08	0.09	0.02
Social approval	4.44	4.29	4.17	0.11	0.20	0.09
Polluter pays	4.97	4.86	4.97	0.08	0.00	-0.07
Financial constraints	4.97	4.74	4.94	0.14	0.02	-0.12
Time restrictions	3.62	3.36	3.55	0.16	0.04	-0.11
Gender (male)	0.51	0.47	0.49	0.09	0.05	-0.04
Age	47.72	47.88	46.73	-0.01	0.06	0.07
High education	0.38	0.36	0.37	0.03	0.02	-0.01
Employed	0.63	0.65	0.66	-0.03	-0.07	-0.04
High income	0.33	0.34	0.35	-0.03	-0.05	-0.02
Not disclosed income	0.09	0.07	0.06	0.10	0.12	0.02

Table A.8: Covariate balance

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

Covariate	Control (Mean)	Treatment 1 (Mean)	Treatment 2 (Mean)	C - T1 (SMD)	C - T2 (SMD)	T1 - T2 (SMD)
Effectiveness of offsetting	4.37	4.18	4.35	0.14	0.01	-0.12
Putting a price on nature	4.79	4.73	4.70	0.05	0.07	0.02
Moral licensing	4.91	4.87	4.81	0.03	0.07	0.04
Offsetting as greenwashing	4.88	4.71	4.84	0.13	0.03	-0.10
Survey time	16.14	11.41	16.31	0.11	-0.00	-0.10
Familiarity with carbon neutral label	0.12	0.14	0.12	-0.05	-0.00	0.05
Manipulation checker (percentage info.)	0.50	0.51	0.46	-0.01	0.08	0.09
Attention checker (number of labels)	0.63	0.60	0.63	0.06	0.01	-0.06
Attention checker (definition of CO ₂ offsetting)	0.37	0.55	0.52	-0.36	-0.32	0.05
Not consider carbon neutral label (ANA)	0.30	0.35	0.35	-0.10	-0.10	-0.00
Consequentiality	0.52	0.57	0.51	-0.11	0.00	0.12
Member of environmental organization	0.90	0.88	0.89	0.07	0.02	-0.04
Response certainty	8.03	8.00	8.05	0.02	-0.01	-0.03

Table A.9: Covariate balance (continued)

SMD refers to the standardized mean difference between experimental groups. c refers to the control group; t1 refers to treatment 1; t2 refers to treatment 2. The first part of the table includes key variables used to test mechanisms, including: only confused with transparent labels (participants confused by at least one transparent label but not the standard label), only trust transparent labels (participants who trust at least one transparent label but not the standard label), confusion level (degree of confusion about carbon-neutral labels), trust level (trust in carbon-neutral labels), concern level (for CO₂ offsets), climate worry (concern about climate change), warm glow (positive emotions from climate-friendly purchases), guilt (negative feelings when making non-climate-friendly choices), social approval (perceived social acceptance for climate-friendly behavior), polluter pays (belief that producers should bear climate mitigation costs), financial constraints (perceived financial constraints to climate-friendly purchases), and time restrictions (perceived lack of time to make climate-friendly choices). Among these, all except the two binary variables (only confused and only trust transparent labels) are measured using 7-point Likert scales. Demographic covariates include gender, age, high education (post-secondary or higher), employment status, high income (annual household income above 50,000), and undisclosed income. The part of the table reports additional covariates included as robustness checks. Among them, the following are measured using 7-point Likert scales: perceived effectiveness of offsetting (belief that offsetting reduces emissions), putting a price on nature (perception that offsetting is about putting monetary value to nature), moral licensing (belief that offsetting allows companies to pollute without real effort to reduce), and offsetting as greenwashing (belief that offsetting is a form of greenwashing). Other robustness covariates include: survey duration, familiarity with carbon-neutral labels, manipulation checker (percentage information), attention checkers (number of labels, definition of offsetting) and, -attendance (ignoring the carbon-neutral label), consequentiality perception (whether participants believe their responses affect tea prices or label policies), environmental organization membership, and response certainty.

A.2 Choice data

This section presents the choice experiment data. Table A.10 summarizes the overall choice design, including the different blocks, choice scenarios (cards), and respective attribute levels. Table A.11 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.10: Choice design

CN = Carbon-neutral label; ET = Ethical Trade label. Values in the table represent attribute levels shown on each choice card for Alternatives A and B. For the status quo option, all attribute levels take a value of 0 (i.e., the no-purchase option). Price values are in GBP. There were two blocks, shown in the first column.

Block	Choice card	N Alt A	N Alt B	N Alt SQ	Total	AltA	AltB	AltSQ
1	1	319	189	168	676	47.19%	27.96%	24.85%
1	2	363	166	147	676	53.70%	24.56%	21.75%
1	3	403	172	101	676	59.62%	25.44%	14.94%
1	4	228	203	245	676	33.73%	30.03%	36.24%
1	5	309	252	115	676	45.71%	37.28%	17.01%
1	6	174	333	169	676	25.74%	49.26%	25.00%
1	7	337	284	55	676	49.85%	42.01%	8.14%
1	8	110	449	117	676	16.27%	66.42%	17.31%
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.11: Choices: detailed information

This table reports the number of times each alternative was chosen in each choice card, as well as the share of respondents who selected each option. N Alt A, N Alt B, and N Alt SQ represent the number of times Alternatives A, B, and the status quo were chosen, respectively. Alt A, Alt B, and Alt SQ indicate their corresponding shares. Total refers to the total number of respondents per choice card. These results are based on the sample prior to the exclusion of protest responses.

B Robustness and heterogeneity

A.1 Robustness tests

This section presents the estimations for robustness checks using both preference-space and WTP-space specifications and shows estimations exploring the underlying mechanisms of main findings.

Preference-space estimations are reported in Section A.1.1, and begins with the MNL model estimates in Table A.12. Next, a MXL model where label attributes treated as random (Table A.13) and then with all coefficients, including price and the status-quo, treated as random (Table A.14). To check the mechanisms underlying consumer WTP, the model with interaction terms is estimated (Table A.15).

Holm-adjusted Poe statistics for WTP estimates are collected in Table A.16, and WTP point estimates derived from these preference-space models are collected in Table A.17.

Section A.1.2 repeats the same estimations directly in WTP space. It starts with a MNL model (Table A.18), proceeds to MXL (Table A.19), doubles the simulation precision with 2,000 Halton draws (Table A.20), adds alternative-specific error components (Table A.21), and introduces correlations for the error terms of non-monetary attributes (Table A.22).

Sensitivity analyses that change the starting value of the cost parameter to 0.1 or 1 (instead of 0.5) do not lead to any change in the main estimation output (Table 3 in Section 3); therefore, they are not reported separately.

Robustness to alternative samples is examined by estimating (i) the raw data of 1,339 responses without excluding protest responses (Table A.23), (ii) a dataset that excludes survey speeders (Table A.24), and (iii) a strictly-screened sample that also removes every participant who failed any manipulation or attention check (Table A.25).

Socio-economic interactions are added to the main MXL model (Table A.35); the interaction effect for individuals who correctly understand the term offsetting is incorporated (Table A.29); the main model without adjusting p-values for

multiple hypothesis testing is presented (Table A.26); the interaction set is reduced for parsimony (Table A.27); the warm-glow variable is introduced (Table A.28); and finally, additional interaction variables are incorporated (Table A.30).

Holm-adjusted Poe tests for the core WTP-space specifications are provided in Table A.31, and A.32. Associated Poe comparisons for estimations with alternative samples in Table A.33.

The results of the open-ended contingent-valuation question with 10,000 bootstrapped WTP samples and Holm-corrected Poe tests are summarized in Table A.34.

A.1.1 Preference space estimations

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of parameters</i>				
$\mu_{\text{Carbon neutral}}$	0.27*** (0.03)	0.31*** (0.05)	0.31*** (0.05)	0.20*** (0.05)
μ_{Organic}	0.63*** (0.03)	0.61*** (0.06)	0.65*** (0.06)	0.63*** (0.06)
$\mu_{\text{Ethical trade}}$	0.63*** (0.03)	0.64*** (0.06)	0.62*** (0.06)	0.62*** (0.06)
μ_{Price}	-0.48*** (0.02)	-0.50*** (0.03)	-0.48*** (0.03)	-0.47*** (0.03)
$\mu_{\text{Status quo}}$	-1.84*** (0.06)	-1.95*** (0.11)	-1.77*** (0.10)	-1.81*** (0.10)
Log Likelihood	-9986.11	-3312.74	-3367.50	-3302.05
AIC	19982.23	6635.47	6745.01	6614.09
BIC	20018.56	6666.35	6775.86	6644.87
Pseudo R ²	0.14	0.15	0.13	0.14
Number of observations	10568	3552	3536	3480
Number of participants	1321	444	442	435

Table A.12: Multinomial logit model in preference space estimates

This table reports the results from a multinomial logit model estimated in preference space. Robust standard errors are shown in parentheses.

P-values (Holm adjusted for 5 comparisons) and are used to determine statistical significance:

*** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of parameters</i>				
$\mu_{\text{Carbon neutral}}$	0.30*** (0.04)	0.37*** (0.07)	0.37*** (0.08)	0.18*** (0.07)
μ_{Organic}	0.76*** (0.05)	0.74*** (0.08)	0.80*** (0.08)	0.74*** (0.08)
$\mu_{\text{Ethical trade}}$	0.81*** (0.05)	0.84*** (0.08)	0.80*** (0.08)	0.79*** (0.08)
μ_{Price}	-0.67*** (0.02)	-0.69*** (0.04)	-0.67*** (0.04)	-0.65*** (0.04)
$\mu_{\text{Status quo}}$	-2.42*** (0.08)	-2.54*** (0.14)	-2.35*** (0.13)	-2.38*** (0.14)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	1.18*** (0.06)	1.18*** (0.10)	1.27*** (0.11)	1.05*** (0.12)
σ_{Organic}	1.22*** (0.06)	1.18*** (0.10)	1.24*** (0.09)	1.25*** (0.10)
$\sigma_{\text{Ethical trade}}$	1.13*** (0.06)	1.18*** (0.11)	1.11*** (0.10)	1.13*** (0.10)
Log Likelihood	-9218.19	-3063.44	-3098.03	-3051.75
AIC	18452.38	6142.88	6212.06	6119.50
BIC	18510.50	6192.28	6261.42	6168.74
Pseudo R ²	0.21	0.21	0.20	0.20
Number of observations	10568	3552	3536	3480
Number of participants	1321	444	442	435

Table A.13: Mixed logit model in preference space estimates, non-monetary parameters randomized

This table reports the results from a mixed logit model estimated in preference space. Robust standard errors are shown in parentheses.

P-values (Holm adjusted for 5 comparisons) and are used to determine statistical significance:

*** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of parameters</i>				
$\mu_{\text{Carbon neutral}}$	0.46*** (0.05)	0.59*** (0.08)	0.50*** (0.08)	0.34*** (0.08)
μ_{Organic}	0.97*** (0.05)	0.99*** (0.09)	0.97*** (0.09)	0.96*** (0.09)
$\mu_{\text{Ethical trade}}$	1.01*** (0.05)	1.10*** (0.10)	0.97*** (0.09)	1.01*** (0.09)
μ_{Price}	-0.23*** (0.05)	-0.13* (0.08)	-0.25*** (0.08)	-0.30*** (0.09)
$\mu_{\text{Status quo}}$	-5.19*** (0.25)	-5.93*** (0.56)	-4.66*** (0.37)	-5.41*** (0.49)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	0.83*** (0.08)	0.88*** (0.14)	0.88*** (0.12)	0.68*** (0.14)
σ_{Organic}	1.00*** (0.07)	1.00*** (0.11)	0.99*** (0.11)	1.04*** (0.12)
$\sigma_{\text{Ethical trade}}$	0.84*** (0.08)	0.93*** (0.14)	0.89*** (0.14)	0.78*** (0.13)
σ_{Price}	0.92*** (0.04)	0.93*** (0.08)	0.86*** (0.06)	0.98*** (0.07)
$\sigma_{\text{Status quo}}$	3.24*** (0.21)	3.78*** (0.41)	2.92*** (0.28)	3.37*** (0.40)
Log Likelihood	-8248.47	-2697.14	-2811.60	-2726.17
AIC	16516.94	5414.28	5643.19	5472.34
BIC	16589.60	5476.03	5704.90	5533.89
Pseudo R ²	0.29	0.31	0.28	0.29
Number of observations	10568	3552	3536	3480
Number of participants	1321	444	442	435

Table A.14: Mixed logit model in preference space estimates, all parameters randomized

This table reports the results from a mixed logit model estimated in preference space. Robust standard errors are shown in parentheses.

P-values (Holm adjusted for 5 comparisons) and are used to determine statistical significance:

*** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

	Full sample	Control	Treatment 1	Treatment 2
$\mu_{\text{Carbon neutral}}$	0.00 (0.12)	-0.03 (0.21)	0.01 (0.21)	-0.01 (0.23)
Main interactions				
$\mu_{\text{Carbon neutral}}$ * Only confused with transparent labels	-0.19 (0.09)	-0.13 (0.17)	-0.08 (0.17)	-0.31 (0.15)
$\mu_{\text{Carbon neutral}}$ * Only trust transparent labels	0.40*** (0.10)	0.51** (0.17)	0.19 (0.18)	0.60*** (0.17)
$\mu_{\text{Carbon neutral}}$ * Concern level	0.07 (0.04)	-0.06 (0.07)	0.18* (0.06)	0.09 (0.07)
$\mu_{\text{Carbon neutral}}$ * Trust level	0.04 (0.04)	0.10 (0.06)	0.07 (0.06)	-0.08 (0.07)
$\mu_{\text{Carbon neutral}}$ * Confusion level	-0.05 (0.03)	0.00 (0.06)	-0.05 (0.06)	-0.16* (0.06)
$\mu_{\text{Carbon neutral}}$ * Climate worry	0.12** (0.03)	0.13 (0.06)	0.02 (0.06)	0.25*** (0.06)
$\mu_{\text{Carbon neutral}}$ * Guilt	0.07 (0.04)	0.11 (0.06)	0.04 (0.07)	0.07 (0.06)
$\mu_{\text{Carbon neutral}}$ * Social approval	0.13** (0.04)	0.18 (0.08)	0.16 (0.07)	0.01 (0.07)
$\mu_{\text{Carbon neutral}}$ * Polluter pays	-0.04 (0.03)	-0.02 (0.06)	-0.09 (0.05)	0.03 (0.06)
$\mu_{\text{Carbon neutral}}$ * Financial constraints	-0.08* (0.03)	-0.10 (0.06)	-0.07 (0.05)	-0.13 (0.05)
$\mu_{\text{Carbon neutral}}$ * Time restrictions	-0.08 (0.03)	-0.10 (0.06)	-0.09 (0.06)	-0.05 (0.05)
Socioeconomic interactions	Yes	Yes	Yes	Yes
Sustainability label attributes, price, status quo	Yes	Yes	Yes	Yes
Standard deviations of random parameters	Yes	Yes	Yes	Yes
Log Likelihood	-8031.02	-2614.74	-2739.11	-2640.57
AIC	16116.05	5283.47	5532.23	5335.15
BIC	16312.22	5450.20	5698.84	5501.33
Pseudo R ²	0.31	0.33	0.29	0.31
Number of Observations	10568	3552	3536	3480
Number of Participants	1321	444	442	435

Table A.15: Mixed logit model in preference space estimates with interaction variables

P-values (Holm adjusted for 5 comparisons) and are used to determine statistical significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

All parameters, except for the interaction terms, are randomized. Holm-Bonferroni adjusted p-values (for 11 comparisons) and significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

Bilateral Wald tests were conducted to assess the equivalence of interaction coefficients across experimental groups. Holm-Bonferroni adjusted p-values (for 21 Wald test comparisons), covering all main interactions, socioeconomic variables, WTP parameters, and the status quo parameter, were used to account for multiple testing. Two of the interaction effects are statistically significant: concern level for carbon offsets (control vs. treatment 1, $p = 0.033$) and climate worry (treatment 1 vs. treatment 2, $p = 0.020$). Note that Wald test results for coefficients not reported in the table (e.g., socioeconomic variables) are also not included in the table notes.

WTP estimates derived from MNL model in preference space (Table A.12)			
	Full sample	Control	Treatment 1 Treatment 2
MWTP _{Carbon neutral}	0.56*** (0.06)	0.61*** (0.11)	0.66*** (0.11) 0.42*** (0.10)
Poe test: Control vs Treatment 1		0.61 (p = 0.78)	
Poe test: Control vs Treatment 2		0.09 (p = 0.37)	
Poe test: Treatment 1 vs Treatment 2			0.06 (p = 0.34)
WTP estimates derived from MXL model in preference space, non-monetary attributes as random par. (Table A.13)			
	Full sample	Control	Treatment 1 Treatment 2
MWTP _{Carbon neutral}	0.45*** (0.06)	0.54*** (0.10)	0.56*** (0.11) 0.27*** (0.10)
Poe test: Control vs Treatment 1		0.53 (p = 0.93)	
Poe test: Control vs Treatment 2		0.03 (p = 0.16)	
Poe test: Treatment 1 vs Treatment 2			0.03 (p = 0.16)
WTP estimates derived from MXL model in preference space, all parameters random (Table A.14)			
	Full sample	Control	Treatment 1 Treatment 2
MWTP _{Carbon neutral}	0.58*** (0.06)	0.67*** (0.10)	0.65*** (0.11) 0.45*** (0.11)
Poe test: Control vs Treatment 1		0.16 (p = 0.62)	
Poe test: Control vs Treatment 2			0.07 (p = 0.40)
Poe test: Treatment 1 vs Treatment 2			0.16 (p = 0.62)

Table A.16: Poe test results comparing MWTP estimates from preference-space models

WTP estimates are reported with standard errors in parentheses.

Significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

Poe test statistics are provided with p-values in parentheses, which are Holm-Bonferroni adjusted for three comparisons.

WTP estimates derived from MNL model in preference space (Table A.12)				
	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral}	0.56*** (0.06)	0.61*** (0.10)	0.65*** (0.11)	0.42*** (0.10)
MWTP _{Organic}	1.31*** (0.06)	1.22*** (0.11)	1.37*** (0.11)	1.33*** (0.12)
MWTP _{Ethical trade}	1.29*** (0.07)	1.28*** (0.11)	1.29*** (0.12)	1.31*** (0.12)
WTP estimates derived from MXL in preference space, non-monetary attributes as random parameters (Table A.13)				
	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral}	0.45*** (0.06)	0.54*** (0.10)	0.56*** (0.11)	0.27*** (0.10)
MWTP _{Organic}	1.13*** (0.06)	1.08*** (0.11)	1.19*** (0.11)	1.14*** (0.12)
MWTP _{Ethical trade}	1.20*** (0.06)	1.22*** (0.11)	1.19*** (0.11)	1.21*** (0.11)
WTP estimates derived from MXL model in preference space, all attributes and status quo as random p. (Table A.14)				
	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral}	0.58*** (0.06)	0.67*** (0.10)	0.65*** (0.11)	0.45*** (0.11)
MWTP _{Organic}	1.21*** (0.07)	1.13*** (0.10)	1.25*** (0.12)	1.29*** (0.13)
MWTP _{Ethical trade}	1.27*** (0.07)	1.26*** (0.11)	1.25*** (0.12)	1.35*** (0.13)
Number of participants	1321	444	442	435

Table A.17: MWTP estimates for carbon-neutral, organic, and ethical trade labels

This table shows the WTP estimated based on choice model outputs shown in Tables A.12, A.13, A.14.

Standard errors are reported in parentheses.

For the model shown in Tables A.12 and A.13, the Delta method is used to derive the standard errors, while for the model shown in Table A.14, the Krinsky-Robb simulation with 10,000 draws is used.

P-values (Holm adjusted for 3 comparisons) and are used to determine statistical significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

A.1.2 WTP space estimations

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of parameters</i>				
MWTP _{Carbon neutral}	0.56*** (0.06)	0.61*** (0.10)	0.65*** (0.11)	0.42*** (0.10)
MWTP _{Organic}	1.31*** (0.06)	1.22*** (0.11)	1.37*** (0.11)	1.33*** (0.12)
MWTP _{Ethical trade}	1.29*** (0.07)	1.28*** (0.11)	1.29*** (0.12)	1.31*** (0.12)
μ_{Price}	-0.48*** (0.02)	-0.50*** (0.03)	-0.48*** (0.03)	-0.47*** (0.03)
$\mu_{\text{Status quo}}$	-1.84*** (0.06)	-1.95*** (0.11)	-1.77*** (0.10)	-1.81*** (0.10)
Log Likelihood	-9986.11	-3312.74	-3367.50	-3302.05
AIC	19982.23	6635.47	6745.01	6614.09
BIC	20018.56	6666.35	6775.86	6644.87
Pseudo R ²	0.14	0.15	0.13	0.14
Number of observations	10568	3552	3536	3480
Number of participants	1321	444	442	435

Table A.18: Multinomial logit model in WTP space

This table shows the choice model output from the multinomial logit model estimated in WTP-space.

Robust standard errors are reported in parentheses.

P-values (Holm adjusted for 4 comparisons) and are used to determine statistical significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of parameters</i>				
MWTP _{Carbon neutral}	0.45*** (0.06)	0.54*** (0.10)	0.56*** (0.11)	0.27*** (0.10)
MWTP _{Organic}	1.13*** (0.06)	1.08*** (0.11)	1.19*** (0.11)	1.14*** (0.12)
MWTP _{Ethical trade}	1.20*** (0.06)	1.22*** (0.11)	1.19*** (0.11)	1.21*** (0.11)
μ_{Price}	-0.67*** (0.02)	-0.69*** (0.04)	-0.67*** (0.04)	-0.65*** (0.04)
$\mu_{\text{Status quo}}$	-2.42*** (0.08)	-2.54*** (0.14)	-2.35*** (0.13)	-2.38*** (0.14)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	1.75*** (0.10)	1.72*** (0.17)	1.89*** (0.18)	1.61*** (0.19)
σ_{Organic}	1.81*** (0.09)	1.70*** (0.16)	1.84*** (0.16)	1.92*** (0.17)
$\sigma_{\text{Ethical trade}}$	1.69*** (0.09)	1.71*** (0.16)	1.65*** (0.16)	1.73*** (0.16)
Log Likelihood	-9218.19	-3063.44	-3098.03	-3051.75
AIC	18452.38	6142.88	6212.06	6119.50
BIC	18510.50	6192.28	6261.42	6168.74
Pseudo R ²	0.21	0.21	0.20	0.20
Number of observations	10568	3552	3536	3480
Number of participants	1321	444	442	435

Table A.19: Mixed logit model in WTP space

This table shows the choice model output from the MXL model estimated in WTP-space.

Robust standard errors are reported in parentheses.

P-values (Holm adjusted for 4 comparisons) and are used to determine statistical significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of the parameters</i>				
MWTP _{Carbon neutral}	0.45*** (0.06)	0.53*** (0.09)	0.54*** (0.10)	0.25*** (0.08)
MWTP _{Organic}	1.09*** (0.05)	1.03*** (0.09)	1.16*** (0.10)	1.11*** (0.14)
MWTP _{Ethical trade}	1.16*** (0.06)	1.15*** (0.09)	1.17*** (0.12)	1.17*** (0.13)
μ_{Price}	-0.17*** (0.04)	-0.11 (0.07)	-0.20*** (0.07)	-0.21*** (0.08)
$\mu_{\text{Status quo}}$	-4.36*** (0.18)	-4.82*** (0.36)	-3.96*** (0.26)	-4.36*** (0.33)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	1.17*** (0.22)	1.13*** (0.26)	1.44*** (0.22)	0.98*** (0.09)
σ_{Organic}	1.58*** (0.08)	1.47*** (0.14)	1.58*** (0.13)	1.67*** (0.27)
$\sigma_{\text{Ethical trade}}$	1.47*** (0.09)	1.44*** (0.13)	1.45*** (0.15)	1.49*** (0.21)
σ_{Price}	0.82*** (0.06)	0.81*** (0.08)	0.72*** (0.07)	0.85*** (0.07)
$\sigma_{\text{Status quo}}$	2.30*** (0.13)	2.61*** (0.25)	2.08*** (0.20)	2.32*** (0.24)
Log Likelihood	-8536.84	-2808.48	-2898.52	-2823.37
AIC	17093.68	5636.96	5817.04	5666.73
BIC	17166.34	5698.71	5878.75	5728.28
Pseudo R ²	0.26	0.28	0.25	0.26
Number of Observations	10568	3552	3536	3480
Number of Participants	1321	444	442	435

Table A.20: Mixed logit model in WTP space with 2000 Halton draws

This table reports the results from a MXL model estimated in WTP space.

Robust standard errors are shown in parentheses.

P-values (Holm adjusted for 4 comparisons) and are used to determine statistical significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

2,000 Halton draws are used.

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of the parameters</i>				
MWTP _{Carbon neutral}	0.47*** (0.11)	0.55*** (0.08)	0.52*** (0.10)	0.32*** (0.08)
MWTP _{Organic}	1.10*** (0.08)	1.04*** (0.09)	1.15*** (0.11)	1.14*** (0.10)
MWTP _{Ethical trade}	1.17*** (0.07)	1.16*** (0.11)	1.14*** (0.12)	1.23*** (0.11)
μ_{Price}	-0.17*** (0.04)	-0.11 (0.07)	-0.18*** (0.07)	-0.21*** (0.07)
$\mu_{\text{Status quo}}$	-4.26*** (0.19)	-4.73*** (0.33)	-3.97*** (0.27)	-4.29*** (0.29)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	1.24*** (0.47)	1.31*** (0.12)	1.37*** (0.25)	1.25*** (0.12)
σ_{Organic}	1.64*** (0.10)	1.41*** (0.14)	1.57*** (0.14)	1.68*** (0.11)
$\sigma_{\text{Ethical trade}}$	1.48*** (0.12)	1.40*** (0.19)	1.43*** (0.15)	1.47*** (0.15)
σ_{Price}	0.77*** (0.07)	0.78*** (0.07)	0.74*** (0.09)	0.82*** (0.08)
$\sigma_{\text{Status quo}}$	2.27*** (0.13)	2.49*** (0.24)	2.13*** (0.19)	2.22*** (0.19)
<i>Alternative specific error components</i>				
$\sigma_{\text{Alternative 1}}$	0.02 (0.05)	-0.03 (0.05)	-0.12 (0.15)	0.13 (0.14)
$\sigma_{\text{Alternative 2}}$	-0.00 (0.10)	-0.05 (0.06)	-0.11 (0.18)	0.06 (0.18)
Log Likelihood	-8540.42	-2808.50	-2898.05	-2822.61
AIC	17104.84	5641.01	5820.10	5669.21
BIC	17192.03	5715.11	5894.15	5743.07
Pseudo R ²	0.26	0.28	0.25	0.26
Number of observations	10568	3552	3536	3480
Number of participants	1321	444	442	435

Table A.21: Mixed logit model in WTP space with alternative specific error components

This table reports the results from the mixed logit model estimated in WTP space. Robust standard errors are shown in parentheses.

P-values (Holm adjusted for 4 comparisons) and are used to determine statistical significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of the parameters</i>				
MWTP _{Carbon neutral}	0.44*** (0.07)	0.47*** (0.08)	0.48*** (0.10)	0.41*** (0.10)
MWTP _{Organic}	1.11*** (0.07)	1.00*** (0.08)	1.17*** (0.13)	1.16*** (0.10)
MWTP _{Ethical trade}	1.17*** (0.07)	1.10*** (0.09)	1.19*** (0.12)	1.20*** (0.10)
μ_{Price}	-0.20*** (0.04)	-0.12* (0.07)	-0.23*** (0.07)	-0.23*** (0.08)
$\mu_{\text{Status quo}}$	-4.32*** (0.18)	-4.88*** (0.36)	-4.00*** (0.27)	-4.40*** (0.32)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	0.62*** (0.13)	0.53*** (0.14)	0.59*** (0.18)	0.66*** (0.14)
σ_{Organic}	1.50*** (0.09)	1.35*** (0.11)	1.55*** (0.15)	1.54*** (0.15)
$\sigma_{\text{Ethical trade}}$	1.16*** (0.10)	1.18*** (0.12)	1.27*** (0.16)	1.17*** (0.19)
σ_{Price}	0.75*** (0.04)	0.81*** (0.07)	0.71*** (0.08)	0.80*** (0.06)
$\sigma_{\text{Status quo}}$	2.38*** (0.13)	2.72*** (0.29)	2.24*** (0.19)	2.42*** (0.23)
<i>Covariances</i>				
Cov (Carbon neutral, Organic)	0.57*** (0.09)	0.40*** (0.10)	0.74*** (0.16)	0.59*** (0.12)
Cov (Carbon neutral, Ethical trade)	0.78*** (0.07)	0.77*** (0.08)	0.71*** (0.12)	0.75*** (0.10)
Cov (Organic, Ethical trade)	0.28*** (0.07)	0.51*** (0.09)	0.27*** (0.09)	0.20** (0.09)
Log Likelihood	-8378.83	-2748.89	-2843.16	-2771.04
AIC	16783.65	5523.78	5712.31	5568.07
BIC	16878.11	5604.05	5792.53	5648.09
Pseudo R ²	0.28	0.30	0.27	0.28
Number of Observations	10568	3552	3536	3480
Number of Participants	1321	444	442	435

Table A.22: Mixed logit model in WTP space with correlated error terms

This table reports the results from the mixed logit model estimated in WTP space.

Robust standard errors are shown in parentheses.

P-values (Holm adjusted for 4 comparisons) and are used to determine statistical significance:

*** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of the parameters</i>				
MWTP _{Carbon neutral}	0.44*** (0.06)	0.53*** (0.09)	0.48*** (0.09)	0.24*** (0.08)
MWTP _{Organic}	1.07*** (0.06)	1.06*** (0.10)	1.04*** (0.09)	1.13*** (0.10)
MWTP _{Ethical trade}	1.13*** (0.06)	1.16*** (0.11)	1.09*** (0.10)	1.15*** (0.09)
μ_{Price}	-0.15*** (0.04)	-0.11 (0.07)	-0.18** (0.07)	-0.17** (0.08)
$\mu_{\text{Status quo}}$	-4.42*** (0.19)	-4.78*** (0.33)	-4.03*** (0.26)	-4.56*** (0.38)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	1.21*** (0.16)	1.25*** (0.19)	1.50*** (0.25)	-0.74*** (0.21)
σ_{Organic}	1.57*** (0.09)	1.50*** (0.16)	1.54*** (0.08)	1.71*** (0.09)
$\sigma_{\text{Ethical trade}}$	1.36*** (0.09)	1.41*** (0.16)	1.38*** (0.17)	1.55*** (0.09)
σ_{Price}	0.85*** (0.05)	0.80*** (0.08)	0.81*** (0.08)	0.97*** (0.11)
$\sigma_{\text{Status quo}}$	2.49*** (0.14)	2.64*** (0.24)	2.22*** (0.22)	2.74*** (0.26)
Log Likelihood	-8609.12	-2827.09	-2918.09	-2855.94
AIC	17238.24	5674.18	5856.17	5731.87
BIC	17311.03	5736.02	5917.99	5793.62
Pseudo R ²	0.27	0.28	0.26	0.27
Number of observations	10712	3584	3576	3552
Number of participants	1339	448	447	444

Table A.23: Mixed logit model in WTP space, with the sample including protest responses

This table reports the results from a MXL model estimated in WTP space.

Robust standard errors are shown in parentheses.

P-values (Holm adjusted for 4 comparisons) and are used to determine statistical significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of the parameters</i>				
MWTP _{Carbon neutral}	0.47*** (0.05)	0.56*** (0.09)	0.51*** (0.13)	0.31*** (0.11)
MWTP _{Organic}	1.10*** (0.07)	1.10*** (0.11)	1.12*** (0.11)	1.12*** (0.12)
MWTP _{Ethical trade}	1.14*** (0.07)	1.15*** (0.11)	1.09*** (0.11)	1.15*** (0.11)
μ_{Price}	-0.11*** (0.04)	-0.04 (0.07)	-0.15** (0.07)	-0.14* (0.08)
$\mu_{\text{Status quo}}$	-4.30*** (0.17)	-4.73*** (0.37)	-3.96*** (0.29)	-4.30*** (0.33)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	1.24*** (0.12)	1.18*** (0.22)	1.42*** (0.48)	1.17** (0.47)
σ_{Organic}	1.57*** (0.09)	1.40*** (0.14)	1.59*** (0.18)	1.64*** (0.18)
$\sigma_{\text{Ethical trade}}$	1.44*** (0.12)	1.39*** (0.16)	1.31*** (0.23)	1.33*** (0.18)
σ_{Price}	0.76*** (0.05)	0.72*** (0.10)	0.70*** (0.10)	0.80*** (0.08)
$\sigma_{\text{Status quo}}$	2.28*** (0.13)	2.62*** (0.25)	2.05*** (0.23)	2.30*** (0.24)
Log Likelihood	-8609.12	-2827.09	-2918.09	-2855.94
AIC	17238.24	5674.18	5856.17	5731.87
BIC	17311.03	5736.02	5917.99	5793.62
Pseudo R ²	0.27	0.28	0.26	0.27
Number of observations	10712	3584	3576	3552
Number of participants	1339	448	447	444

Table A.24: Mixed logit model in WTP space, all parameters randomized, without survey speeders

This table reports the results from a MXL model estimated in WTP space.

Robust standard errors are shown in parentheses.

P-values (Holm adjusted for 4 coefficient comparisons) and are used to determine statistical significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

Protest responses and survey speeders are excluded.

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of the parameters</i>				
MWTP _{Carbon neutral}	0.87*** (0.13)	0.97*** (0.27)	1.05*** (0.24)	0.67*** (0.20)
MWTP _{Organic}	1.14*** (0.14)	1.39*** (0.29)	0.92*** (0.22)	1.24*** (0.26)
MWTP _{Ethical trade}	1.17*** (0.14)	1.68*** (0.34)	0.75*** (0.23)	1.21*** (0.24)
μ_{Price}	0.05 (0.08)	0.22* (0.13)	-0.02 (0.12)	-0.00 (0.16)
$\mu_{\text{Status quo}}$	-4.19*** (0.40)	-4.65*** (0.92)	-3.80*** (0.56)	-4.74*** (0.87)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	1.59*** (0.22)	1.46*** (0.38)	1.56*** (0.31)	1.84*** (0.42)
σ_{Organic}	1.65*** (0.17)	1.76*** (0.31)	1.26*** (0.33)	1.68*** (0.18)
$\sigma_{\text{Ethical trade}}$	1.32*** (0.16)	1.74*** (0.46)	1.16*** (0.35)	0.95* (0.52)
σ_{Price}	0.53*** (0.11)	0.31 (0.20)	0.38** (0.18)	0.81*** (0.27)
$\sigma_{\text{Status quo}}$	2.17*** (0.31)	2.57*** (0.59)	1.79*** (0.45)	2.45*** (0.50)
Log Likelihood	-1423.90	-365.32	-524.64	-520.53
AIC	2867.79	750.64	1069.27	1061.07
BIC	2922.84	792.70	1114.01	1106.17
Pseudo R ²	0.29	0.33	0.26	0.29
Number of observations	1816	496	648	672
Number of participants	227	62	81	84

Table A.25: Mixed logit model in WTP space, excluding protest responses, survey speeders, and attention/manipulation test failers

This table reports the results from a MXL model estimated in WTP space.

Robust standard errors are shown in parentheses.

P-values (Holm adjusted for 4 comparisons) and are used to determine statistical significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

Protest responses, survey speeders and participants who fail at least one attention or manipulation tests are excluded.

	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral}	-0.26 (0.16)	-0.43 (0.29)	-0.25 (0.26)	-0.19 (0.29)
Main interactions				
MWTP _{Carbon neutral} x Only confused with transparent labels	0.03 (0.15)	0.21 (0.20)	0.24 (0.19)	-0.32 (0.24)
MWTP _{Carbon neutral} x Only trust transparent labels	0.65*** (0.14)	0.80*** (0.23)	0.39* (0.20)	0.72*** (0.23)
MWTP _{Carbon neutral} x Concern level	0.18*** (0.05)	0.08 (0.10)	0.30*** (0.08)	0.15 (0.18)
MWTP _{Carbon neutral} x Trust level	0.09** (0.05)	0.1 (0.08)	0.12** (0.07)	-0.05 (0.09)
MWTP _{Carbon neutral} x Confusion level	-0.03 (0.06)	0.01 (0.06)	-0.00 (0.06)	-0.16** (0.08)
MWTP _{Carbon neutral} x Climate worry	0.06 (0.05)	0.01 (0.10)	-0.05 (0.08)	0.25 (0.22)
MWTP _{Carbon neutral} x Guilt	0.12*** (0.05)	0.15*** (0.07)	0.11 (0.08)	0.13 (0.13)
MWTP _{Carbon neutral} x Social approval	0.18*** (0.05)	0.29*** (0.11)	0.21** (0.09)	0.00 (0.09)
MWTP _{Carbon neutral} x Polluter pays	-0.02 (0.04)	-0.01 (0.09)	-0.12* (0.06)	0.03 (0.09)
MWTP _{Carbon neutral} x Financial constraints	-0.18*** (0.04)	-0.16 (0.11)	-0.15** (0.07)	-0.22** (0.09)
MWTP _{Carbon neutral} x Time restrictions	-0.02 (0.04)	-0.08 (0.07)	-0.03 (0.06)	0.06 (0.06)
Socioeconomic interactions	Yes	Yes	Yes	Yes
WTP for other sustainability label attributes, price, status quo	Yes	Yes	Yes	Yes
Standard deviations of random parameters	Yes	Yes	Yes	Yes
Log Likelihood	-8371.60	-2746.65	-2835.84	-2756.99
AIC	16797.20	5547.30	5725.68	5567.97
BIC	16993.37	5714.04	5892.29	5734.15
Pseudo R ²	0.28	0.30	0.27	0.28
Number of Observations	1056	3552	3536	3480
Number of Participants	1321	444	442	435

Table A.26: Mixed logit model in WTP space, with interaction variables (unadjusted for multiple hypothesis testing)

This table shows the choice model output from the MXL model estimated in WTP-space.

Robust standard errors are reported in parentheses.

P-values and significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

	Full sample	Control	Treatment 1	Treatment 2
Interactions				
MWTP _{Carbon neutral} x Only confused with transparent labels	-0.01 (0.11)	0.24 (0.20)	0.21 (0.21)	-0.35 (0.17)
MWTP _{Carbon neutral} x Only trust transparent labels	0.77*** (0.12)	1.04*** (0.21)	0.57* (0.21)	0.77*** (0.19)
MWTP _{Carbon neutral} x Concern level	0.31*** (0.04)	0.25*** (0.07)	0.35*** (0.07)	0.30*** (0.07)
Means of the parameters				
MWTP _{Carbon neutral}	-0.37* (0.14)	-0.55 (0.29)	-0.45 (0.26)	-0.25 (0.25)
MWTP _{Organic}	1.09*** (0.06)	1.03*** (0.10)	1.17*** (0.11)	1.10*** (0.12)
MWTP _{Ethical trade}	1.17*** (0.06)	1.17*** (0.13)	1.19*** (0.11)	1.16*** (0.10)
μ_{Price}	-0.18*** (0.04)	-0.12* (0.07)	-0.18*** (0.06)	-0.21*** (0.08)
$\mu_{\text{Status quo}}$	-4.20*** (0.16)	-4.63*** (0.32)	-3.90*** (0.26)	-4.22*** (0.29)
Socioeconomic interactions	Yes	Yes	Yes	Yes
Standard deviations of random parameters	Yes	Yes	Yes	Yes
Log Likelihood	-8413.84	-2768.77	-2848.03	-2777.06
AIC	16865.68	5575.54	5734.06	5592.12
BIC	17003.73	5692.87	5851.30	5709.06
Pseudo R ²	0.28	0.29	0.27	0.27
Number of Observations	10568	3552	3536	3480
Number of Participants	1321	444	442	435

Table A.27: Mixed logit model in WTP space (fewer interactions)

This table shows the choice model output from the MXL model estimated in WTP-space. (with main variables Robust standard errors are reported in parentheses. P-values (Holm adjusted for 13 comparisons) and significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$ Bilateral Wald tests were conducted to assess the equivalence of interaction coefficients across experimental groups. Holm-Bonferroni adjusted p -values, covering all main interactions, socioeconomic variables, WTP parameters, and the status quo parameter, were used to account for multiple hypothesis testing. For the comparison between treatment group 1 and treatment group 2, the interaction with the binary indicator for respondents confused only by transparent labels is significant ($p=0.077$). In the comparison between the control group and treatment group 2, the same interaction is also significant ($p=0.084$).

	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral}	-0.27 (0.15)	-0.47 (0.25)	-0.17 (0.26)	-0.14 (0.25)
Main interactions				
MWTP _{Carbon neutral} x Only confused with transparent labels	0.02 (0.11)	0.21 (0.18)	0.15 (0.19)	-0.29 (0.20)
MWTP _{Carbon neutral} x Only trust transparent labels	0.66*** (0.12)	0.76*** (0.19)	0.33 (0.20)	0.70** (0.22)
MWTP _{Carbon neutral} x Trust level	0.08 (0.05)	0.14 (0.07)	0.07 (0.07)	-0.03 (0.09)
MWTP _{Carbon neutral} x Confusion level	-0.05 (0.04)	0.00 (0.07)	0.02 (0.06)	-0.16 (0.08)
MWTP _{Carbon neutral} x Concern level	0.21*** (0.05)	0.09 (0.08)	0.28*** (0.08)	0.16 (0.08)
MWTP _{Carbon neutral} x Climate worry	0.05 (0.04)	0.05 (0.07)	-0.07 (0.08)	0.25** (0.08)
MWTP _{Carbon neutral} x Warm glow	0.25*** (0.06)	0.31*** (0.08)	0.34*** (0.10)	0.14 (0.09)
MWTP _{Carbon neutral} x Polluter pays	-0.02 (0.04)	0.00 (0.07)	-0.11 (0.06)	0.02 (0.07)
MWTP _{Carbon neutral} x Financial constraints	-0.18*** (0.04)	-0.12 (0.08)	-0.16 (0.07)	-0.24*** (0.06)
MWTP _{Carbon neutral} x Time restrictions	-0.00 (0.04)	-0.05 (0.06)	-0.04 (0.06)	0.06 (0.06)
Socioeconomic interactions	Yes	Yes	Yes	Yes
WTP for other sustainability labels, price, status quo	Yes	Yes	Yes	Yes
Standard deviations of random parameters	Yes	Yes	Yes	Yes
Log Likelihood	-8372.29	-2750.23	-2834.09	-2757.73
AIC	16796.58	5552.46	5720.18	5567.45
BIC	16985.49	5713.01	5880.62	5727.48
Pseudo R ²	0.28	0.30	0.27	0.28
Number of Observations	10568.00	3552.00	3536.00	3480.00
Number of Participants	1321	444	442	435

Table A.28: Mixed logit model in WTP space with interaction variables (warm glow variable added)

This table shows the choice model output from the MXL model estimated in WTP-space. (with main variables Robust standard errors are reported in parentheses.

P-values (Holm-adjusted for 20 comparisons) and significance levels:

*** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral} x Understanding the term "offsetting"	0.46*** (0.13)	0.37* (0.20)	0.33 (0.21)	0.61** (0.25)
<i>Means of the parameters</i>				
MWTP _{Carbon neutral}	0.23*** (0.07)	0.36** (0.14)	0.35** (0.13)	0.03 (0.16)
MWTP _{Organic}	1.08*** (0.06)	1.04*** (0.09)	1.16*** (0.11)	1.12*** (0.11)
MWTP _{Ethical trade}	1.17*** (0.06)	1.14*** (0.11)	1.15*** (0.12)	1.18*** (0.11)
β_{Price}	-0.17*** (0.04)	-0.11 (0.07)	-0.18*** (0.06)	-0.21*** (0.07)
$\beta_{\text{Status quo}}$	-4.29*** (0.17)	-4.73*** (0.33)	-3.97*** (0.26)	-4.30*** (0.30)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	1.22*** (0.13)	1.27*** (0.13)	1.37*** (0.22)	1.17*** (0.16)
σ_{Organic}	1.59*** (0.10)	1.41*** (0.15)	1.58*** (0.15)	1.67*** (0.13)
$\sigma_{\text{Ethical trade}}$	1.47*** (0.06)	1.39*** (0.18)	1.43*** (0.14)	1.42*** (0.14)
σ_{Price}	0.79*** (0.04)	0.78*** (0.07)	0.73*** (0.09)	0.84*** (0.07)
$\sigma_{\text{Status quo}}$	2.28*** (0.13)	2.50*** (0.24)	2.13*** (0.18)	2.25*** (0.21)
Log Likelihood	-8532.81	-2806.72	-2897.00	-2818.81
AIC	17087.63	5635.45	5815.99	5659.62
BIC	17167.55	5703.37	5883.87	5727.32
Pseudo R ²	0.27	0.28	0.25	0.26
Number of observations	10568	3552	3536	3480
Number of participants	1321	444	442	435

Table A.29: Mixed logit model in WTP space, with interaction variable for participants who misunderstand the meaning of offsetting

This table shows the choice model output from the MXL model estimated in WTP-space. Robust standard errors are reported in parentheses. P-values and significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$. The Wald statistics (Holm adjusted for 4 coefficient comparisons) for comparisons of the interacted coefficient across experimental groups are $W_{C,T1} = 0.878$ ($p = 1.000$), $W_{C,T2} = 0.450$ ($p = 1.000$), and $W_{T1,T2} = 0.383$ ($p = 1.000$).

	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral}	0.51 (0.26)	0.36 (0.45)	1.05 (0.51)	0.09 (0.46)
Additional interactions				
MWTP _{Carbon neutral} x Effectiveness of offsetting	0.04 (0.04)	-0.02 (0.07)	0.16 (0.07)	-0.00 (0.07)
MWTP _{Carbon neutral} x Putting a price on nature	-0.02 (0.05)	-0.02 (0.09)	0.06 (0.09)	0.00 (0.07)
MWTP _{Carbon neutral} x Moral licensing	0.01 (0.06)	0.06 (0.10)	-0.05 (0.10)	0.01 (0.11)
MWTP _{Carbon neutral} x Offsetting as greenwashing	-0.07 (0.05)	-0.05 (0.08)	-0.17 (0.09)	-0.00 (0.10)
MWTP _{Carbon neutral} x Survey time	-0.00 (0.00)	-0.00 (0.00)	0.00 (0.01)	-0.00 (0.00)
MWTP _{Carbon neutral} x Familiarity with carbon neutral label	1.57*** (0.27)	1.14** (0.34)	1.29** (0.39)	2.71*** (0.55)
MWTP _{Carbon neutral} x Manipulation checker	0.16 (0.10)	-0.09 (0.16)	0.17 (0.18)	0.32 (0.17)
MWTP _{Carbon neutral} x Attention checker (number of labels)	0.28 (0.11)	0.34 (0.19)	0.14 (0.20)	0.37 (0.18)
MWTP _{Carbon neutral} x Understanding the term "offsetting"	0.11 (0.10)	0.18 (0.17)	0.10 (0.19)	0.03 (0.17)
MWTP _{Carbon neutral} x Not consider carbon neutral label (ANA)	-0.54*** (0.11)	-0.24 (0.17)	-0.64** (0.20)	-0.63** (0.18)
MWTP _{Carbon neutral} x Perceived consequentiality	-0.25 (0.10)	-0.11 (0.17)	-0.44 (0.19)	-0.20 (0.18)
MWTP _{Carbon neutral} x Environmental organization membership	-0.76*** (0.20)	-0.97 (0.35)	-0.97 (0.45)	-0.30 (0.34)
MWTP _{Carbon neutral} x Response certainty	-0.03 (0.03)	0.00 (0.05)	0.01 (0.06)	-0.14 (0.06)
Main interactions	Yes	Yes	Yes	Yes
Socioeconomic interactions	Yes	Yes	Yes	Yes
WTP for other sustainability labels, price, status quo	Yes	Yes	Yes	Yes
Standard deviations of random parameters	Yes	Yes	Yes	Yes
Log Likelihood	-8297.24	-2733.19	-2803.33	-2712.95
AIC	16674.49	5546.38	5686.66	5505.89
BIC	16965.11	5793.39	5933.49	5752.08
Pseudo-R ²	0.29	0.30	0.28	0.29
Number of Observations	10568	3552	3536	3480
Number of Participants	1321	444	442	435

Table A.30: Mixed logit model in WTP space, additional interactions

This table shows the choice model output from the MXL model estimated in WTP-space.

Robust standard errors are reported in parentheses.

P-values (Holm adjusted for 32 comparisons) and significance levels: *** $p \leq 0.01$, **

$0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

All parameters, except for the interaction terms, are randomized. Holm-Bonferroni adjusted p-values and significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$. Bilateral

Wald tests were conducted to assess the equivalence of interaction coefficients across experimental groups. Holm-Bonferroni adjusted p-values, covering all main interactions, socioeconomic variables, WTP parameters, and the status quo parameter, were used to account for multiple testing. Adjusted Wald test results indicate that the binary variable for participants who purchase carbon-neutral labeled products during grocery shopping (as captured by the label familiarity variable), the difference between the control group and treatment 1 is significant at the 5% level ($p = 0.043$), and the difference between treatment 1 and treatment 2 is marginally significant at the 10% level ($p = 0.080$).

WTP estimates: MNL model, WTP space (Table A.18)			
	Full sample	Control	Treatment 1 Treatment 2
MWTP _{Carbon neutral}	0.56*** (0.06)	0.61*** (0.10)	0.65*** (0.11) 0.41*** (0.10)
Poe test: Control vs Treatment 1		0.61 (p = 0.78)	
Poe test: Control vs Treatment 2		0.09 (p = 0.35)	
Poe test: Treatment 1 vs Treatment 2			0.06 (p = 0.34)
WTP estimates: MXL model, WTP space, label attributes as random parameters (Table A.19)			
	Full sample	Control	Treatment 1 Treatment 2
MWTP _{Carbon neutral}	0.45*** (0.06)	0.54*** (0.10)	0.56*** (0.11) 0.27*** (0.10)
Poe test: Control vs Treatment 1		0.54 (p = 0.925)	
Poe test: Control vs Treatment 2		0.03 (p = 0.163)	
Poe test: Treatment 1 vs Treatment 2			0.03 (p = 0.163)
WTP estimates: MXL model, WTP space, all parameters as random, using twice as many Halton draws (Table A.20)			
	Full sample	Control	Treatment 1 Treatment 2
MWTP _{Carbon neutral}	0.45*** (0.06)	0.53*** (0.10)	0.54*** (0.10) 0.25*** (0.08)
Poe test: Control vs Treatment 1		0.54 (p = 0.916)	
Poe test: Control vs Treatment 2		0.01* (p = 0.064)	
Poe test: Treatment 1 vs Treatment 2			0.01* (p = 0.064)

Table A.31: Poe test results, estimations with various model specifications estimated in WTP space

WTP estimates are reported with standard errors in parentheses.

Significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

Poe test statistics are provided with p-values in parentheses, which are Holm-Bonferroni adjusted for three comparisons.

WTP estimates: MXL model, WTP space, all parameters as random, error components for alternatives (Table A.21)				
	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral}	0.47*** (0.12)	0.55*** (0.08)	0.52*** (0.10)	0.32*** (0.08)
Poe test: Control vs Treatment 1		0.41 (p = 0.811)		
Poe test: Control vs Treatment 2		0.03 (p = 0.168)		
Poe test: Treatment 1 vs Treatment 2			0.06 (p = 0.255)	
WTP estimates: MXL model, WTP space, all parameters as random, correlated error terms (Table A.22)				
	Full sample	Control	Treatment 1	Treatment 2
MWTP _{carbon neutral}	0.44*** (0.07)	0.47*** (0.08)	0.48*** (0.10)	0.40*** (0.10)
Poe test: Control vs Treatment 1		0.54 (p = 1.000)		
Poe test: Control vs Treatment 2		0.29 (p = 1.000)		
Poe test: Treatment 1 vs Treatment 2			0.29 (p = 1.000)	

Table A.32: Poe test results, estimations with various model specifications estimated in WTP space

WTP estimates are reported with standard errors in parentheses.

Significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

Poe test statistics are provided with p-values in parentheses, which are Holm-Bonferroni adjusted for three comparisons.

WTP estimates: MXL model, WTP space, including protest responses (Table A.23)				
	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral}	0.44*** (0.06)	0.53*** (0.09)	0.48*** (0.09)	0.24*** (0.08)
Poe test: Control vs Treatment 1		0.35 (p = 0.705)		
Poe test: Control vs Treatment 2		0.01* (p = 0.054)		
Poe test: Treatment 1 vs Treatment 2			0.03 (p = 0.105)	
WTP estimates: MXL model, WTP space, excluding protest responses and survey speeders (Table A.24)				
	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral}	0.47*** (0.05)	0.56*** (0.09)	0.51*** (0.13)	0.32*** (0.11)
Poe test: Control vs Treatment 1		0.38 (p = 0.756)		
Poe test: Control vs Treatment 2		0.04 (p = 0.262)		
Poe test: Treatment 1 vs Treatment 2			0.13 (p = 0.526)	
WTP estimates: MXL model, WTP space, excluding protest responses, attention/manipulation checker failers (Table A.25)				
	Full sample	Control	Treatment 1	Treatment 2
MWTP _{Carbon neutral}	0.87*** (0.13)	0.96*** (0.27)	1.04*** (0.24)	0.67*** (0.20)
Poe test: Control vs Treatment 1		0.59 (p = 0.821)		
Poe test: Control vs Treatment 2		0.19 (p = 0.747)		
Poe test: Treatment 1 vs Treatment 2			0.11 (p = 0.673)	

Table A.33: Poe test results, estimations with different sample restrictions

WTP estimates are reported with standard errors in parentheses.

Significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

Poe test statistics are provided with p-values in parentheses, which are Holm-Bonferroni adjusted for three comparisons.

A.1.3 Open-ended contingent valuation method MWTP estimates

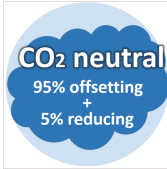
	 Control	 Treatment 1	 Treatment 2
MWTP_{Carbon neutral}	4.33*** (0.16)	4.45*** (0.18)	4.82 (0.16)
<i>Holm-Bonferroni adjusted Poe test</i>			
Control vs Treatment 1	0.70 (p=0.60)		
Control vs Treatment 2	0.98 (p=0.10)		
Treatment 1 vs Treatment 2		0.93 (p=0.26)	
<i>Unadjusted Poe test</i>			
Control vs Treatment 1	0.70 (p=0.60)		
Control vs Treatment 2	0.99** (p=0.03)		
Treatment 1 vs Treatment 2		0.94 (p=0.13)	

Table A.34: Poe test results, contingent valuation method

WTP estimates are based on 10,000 bootstrapped samples.

Significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

Poe test statistics are reported with Holm-Bonferroni adjusted p -values (adjusted for three pairwise comparisons).

A.2 Heterogeneity

This section examines socioeconomic heterogeneity—specifically income, education, age, gender, and employment. I first assess (i) how these variables are associated with MWTP for the carbon-neutral label, and then (ii) how they relate to the effect of transparency or a higher share of CO₂ reduction on MWTP.

Table A.35 reports results from an MXL specification that interacts MWTP with socioeconomic characteristics to explain variation across experimental groups. Some associations vary between the experimental groups. Income is positively associated with MWTP in the control group and in Treatment 2 (both significant at 1%), but not in Treatment 1, which saw a transparent label with a higher share of CO₂ offsetting. A preference for CO₂ reductions over offsets is stronger among higher-income participants: significant in the control group (5%) and Treatment 2 (1%), but not in Treatment 1. In Treatment 2, high-income participants are willing to pay about £1.24 more than low-income participants, compared with £0.50 in Treatment 1.

There is no consistent pattern for education. Participants with a Bachelor's degree or higher correlate with a higher MWTP in the control group and Treatment 1 (significant at 1% and 5%); however, it is insignificant in Treatment 2, where transparency indicates an equal split between CO₂ offsetting and reduction.

The age–MWTP interaction is negative and significant at the 5% level in Treatment 1, suggesting older individuals are more skeptical of offset-heavy labels; the effect is small, with each additional year lowering MWTP by about £0.02. Age is insignificant in the control group and Treatment 2. Therefore, the age effect is modest and specific to transparent labels with a greater offsetting share.

No consistent patterns emerge regarding gender or employment status. These characteristics do not systematically explain the differences in MWTP.

I now address the second question and examine how socioeconomic variables relate to the effect of transparency on MWTP. Bilateral Wald tests of interactions across groups suggest that transparency shapes some associations:

the difference in the high-income interaction between Treatment 1 and Treatment 2 is marginally significant after Holm adjustment ($p = 0.091$), indicating that higher-income participants reveal higher MWTP when transparency signals a 50% CO₂ reduction share as opposed to a lower share (5%).

Wald tests show no significant differences for the interactions involving age, education, gender, or employment.

	Full sample	Control	Treatment 1	Treatment 2
<i>Socioeconomic interactions</i>				
MWTP _{Carbon neutral} x Male	0.01 (0.12)	-0.09 (0.23)	0.41 (0.21)	-0.15 (0.23)
MWTP _{Carbon neutral} x Age	-0.01 (0.00)	-0.00 (0.01)	-0.02** (0.01)	-0.01 (0.01)
MWTP _{Carbon neutral} x High education	0.55*** (0.13)	0.63* (0.25)	0.76** (0.28)	0.40 (0.27)
MWTP _{Carbon neutral} x Employed	-0.09 (0.13)	-0.06 (0.27)	0.01 (0.25)	-0.08 (0.26)
MWTP _{Carbon neutral} x High income	0.75*** (0.13)	0.72** (0.25)	0.50 (0.24)	1.24*** (0.25)
MWTP _{Carbon neutral} x Not disclosed income	-0.29 (0.16)	-0.68* (0.27)	0.16 (0.40)	-0.47 (0.79)
<i>Means of parameters</i>				
MWTP _{Carbon neutral}	0.13 (0.12)	0.29 (0.20)	-0.09 (0.22)	-0.03 (0.24)
MWTP _{Organic}	1.07*** (0.06)	1.05*** (0.09)	1.16*** (0.10)	1.10*** (0.12)
MWTP _{Ethical trade}	1.16*** (0.06)	1.19*** (0.10)	1.16*** (0.12)	1.12*** (0.11)
$\mu_{\text{Status quo}}$	-4.23*** (0.17)	-4.63*** (0.31)	-3.93*** (0.26)	-4.22*** (0.29)
μ_{Price}	-0.17*** (0.04)	-0.11* (0.07)	-0.18*** (0.06)	-0.21*** (0.07)
Random coefficients included	Yes	Yes	Yes	Yes
Log Likelihood	-8480.40	-2790.12	-2869.27	-2800.80
AIC	16992.80	5612.25	5770.55	5633.59
BIC	17109.05	5711.05	5869.28	5732.07
Pseudo R ²	0.27	0.28	0.26	0.27
Number of observations	10568	3552	3536	3480
Number of participants	1321	444	442	435

Table A.35: Mixed logit model in WTP space

This table shows the choice model output from the MXL model estimated in WTP-space.

Robust standard errors are reported in parentheses. P-values (Holm adjusted for 10 comparisons) and are used to determine statistical significance: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$.

The equivalence of interaction coefficients across experimental groups was tested using bilateral Wald tests. For the comparison between treatment group 1 and treatment group 2, the interaction with *high income* remains marginally significant after applying a Holm-Bonferroni correction for six comparisons ($p = 0.091$).

C Competing labels

This section explores how the presence or absence of organic and ethical-trade labels affects the valuation of the carbon-neutral label. Two-way and three-way interaction estimates, along with correlated error terms, are presented in Table A.37, and the implied scenario-specific WTPs with their Poe comparisons are laid out in Table A.39. A more basic model, including only two of the two-way interactions and excluding correlated errors, is presented in Table A.36, with the corresponding WTPs and Poe comparisons shown in Table A.38.

	Full sample	Control	Treatment 1	Treatment 2
<i>Means of the parameters</i>				
MWTP _{Carbon neutral}	0.91*** (0.10)	1.01*** (0.20)	0.86*** (0.17)	0.84*** (0.17)
MWTP _{Organic}	1.38*** (0.08)	1.35*** (0.15)	1.36*** (0.15)	1.45*** (0.19)
MWTP _{Ethical trade}	1.43*** (0.09)	1.45*** (0.18)	1.38*** (0.21)	1.51*** (0.20)
μ_{Price}	-0.18*** (0.04)	-0.12 (0.07)	-0.19*** (0.07)	-0.22*** (0.08)
$\mu_{\text{Status quo}}$	-4.12*** (0.17)	-4.51*** (0.33)	-3.82*** (0.28)	-4.13*** (0.31)
<i>Interactions</i>				
MWTP _{Carbon neutral} x MWTP _{Organic}	-0.65*** (0.13)	-0.73** (0.29)	-0.46 (0.25)	-0.73*** (0.25)
MWTP _{Carbon neutral} x MWTP _{Ethical trade}	-0.52*** (0.12)	-0.47** (0.21)	-0.42 (0.23)	-0.54*** (0.18)
<i>Standard deviations of random parameters</i>				
$\sigma_{\text{Carbon neutral}}$	1.10*** (0.14)	1.32*** (0.19)	1.35*** (0.25)	1.20*** (0.20)
σ_{Organic}	1.71*** (0.06)	1.53*** (0.16)	1.62*** (0.13)	1.80*** (0.21)
$\sigma_{\text{Ethical trade}}$	1.55*** (0.06)	1.53*** (0.20)	1.52*** (0.18)	1.56*** (0.24)
σ_{Price}	0.89*** (0.06)	0.86*** (0.09)	0.78*** (0.10)	0.91*** (0.10)
$\sigma_{\text{Status quo}}$	2.25*** (0.13)	2.46*** (0.24)	2.10*** (0.18)	2.26*** (0.20)
Log Likelihood	-8517.55	-2800.77	-2893.98	-2813.97
AIC	17059.10	5625.53	5811.97	5651.95
BIC	17146.29	5699.64	5886.01	5725.80
Pseudo R ²	0.27	0.28	0.26	0.26
Number of observations	10568.00	3552.00	3536.00	3480.00
Number of participants	1321	444	442	435

Table A.36: Mixed logit model in WTP space, with sustainability label interactions, no correlated errors

This table presents the choice model results from the MXL model estimated in WTP space.

Robust standard errors are reported in parentheses.

Holm-Bonferroni adjusted p-values (for 6 comparisons) and significance levels:

*** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

	Full sample	Control	Treatment 1	Treatment 2
Means of the parameters				
MWTP _{Carbon neutral}	0.95*** (0.15)	1.01*** (0.20)	0.89*** (0.17)	0.99*** (0.18)
MWTP _{Organic}	1.70*** (0.20)	1.60*** (0.36)	1.80*** (0.18)	1.70*** (0.17)
MWTP _{Ethical trade}	1.75*** (0.21)	1.70*** (0.33)	1.80*** (0.20)	1.73*** (0.16)
β_{Price}	-0.17*** (0.05)	-0.10 (0.07)	-0.19*** (0.06)	-0.20*** (0.08)
$\beta_{\text{Status quo}}$	-4.02*** (0.23)	-4.56*** (0.38)	-3.77*** (0.28)	-4.23*** (0.34)
Interactions				
MWTP _{Carbon neutral} X MWTP _{Organic}	-0.21 (0.19)	-0.32 (0.32)	0.11 (0.30)	-0.53 (0.30)
MWTP _{Carbon neutral} X MWTP _{Ethical trade}	-0.22 (0.17)	-0.31 (0.26)	0.01 (0.28)	-0.48 (0.24)
MWTP _{Organic} X MWTP _{Ethical trade}	-0.34 (0.18)	-0.29 (0.39)	-0.54 (0.29)	-0.26 (0.25)
MWTP _{Carbon neutral} X MWTP _{Organic} X MWTP _{Ethical trade}	-1.29*** (0.29)	-1.18*** (0.38)	-1.58*** (0.45)	-0.78 (0.37)
Standard deviations of random parameters				
$\sigma_{\text{Carbon neutral}}$	0.75*** (0.28)	0.50* (0.27)	0.74*** (0.14)	0.60*** (0.08)
σ_{Organic}	1.63*** (0.10)	1.54*** (0.43)	1.61*** (0.12)	1.64*** (0.14)
$\sigma_{\text{Ethical trade}}$	1.28*** (0.11)	1.25*** (0.13)	1.37*** (0.16)	1.31*** (0.27)
σ_{Price}	0.80*** (0.08)	0.85*** (0.09)	0.72*** (0.07)	0.92*** (0.10)
$\sigma_{\text{Status quo}}$	2.36*** (0.13)	2.73*** (0.29)	2.27*** (0.21)	2.50*** (0.25)
Covariances				
Cov (Carbon neutral, Organic)	0.69*** (0.14)	0.52*** (0.12)	0.81*** (0.14)	0.67*** (0.09)
Cov (Carbon neutral, Ethical trade)	0.91*** (0.21)	0.95*** (0.20)	0.88*** (0.13)	0.76*** (0.08)
Cov (Organic, Ethical trade)	0.41*** (0.10)	0.63*** (0.12)	0.42*** (0.11)	0.34*** (0.08)
Log Likelihood	-8302.66	-2721.48	-2813.28	-2746.64
AIC	16639.32	5476.96	5660.55	5527.28
BIC	16762.83	5581.94	5765.46	5631.92
Pseudo R ²	0.28	0.30	0.28	0.28
Number of observations	10568	3552	3536	3480
Number of participants	1321	444	442	435

Table A.37: Mixed logit model in WTP space, with sustainability label interactions, correlated errors

This table presents the choice model results from the MXL model estimated in WTP space.

Robust standard errors are reported in parentheses.

Holm-Bonferroni adjusted p-values (for 6 comparisons) and significance levels:

*** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

WTP estimates: MNL model, WTP space, Org=0, ET=0 (Table A.36)			
	Full sample	Control	Treatment 1 Treatment 2
WTP _{CN} (Org=0,ET=0)	0.91*** (0.10)	1.02*** (0.20)	0.86*** (0.17) 0.83*** (0.17)
Poe test: Control vs Treatment 1		0.29 (p = 0.570)	
Poe test: Control vs Treatment 2		0.25 (p = 0.494)	
Poe test: Treatment 1 vs Treatment 2			0.45 (p = 0.907)
WTP estimates: MNL model, WTP space, Org=0, ET=1 (Table A.36)			
	Full sample	Control	Treatment 1 Treatment 2
WTP _{CN} (Org=0,ET=1)	0.40*** (0.08)	0.55*** (0.12)	0.44*** (0.14) 0.30 (0.15)
Poe test: Control vs Treatment 1		0.29 (p = 0.581)	
Poe test: Control vs Treatment 2		0.09 (p = 0.188)	
Poe test: Treatment 1 vs Treatment 2			0.24 (p = 0.476)
WTP estimates: MNL model, WTP space, Org=1, ET=0 (Table A.36)			
	Full sample	Control	Treatment 1 Treatment 2
WTP _{CN} (Org=1,ET=0)	0.26** (0.10)	0.29 (0.19)	0.40* (0.19) 0.11 (0.18)
Poe test: Control vs Treatment 1		0.66 (p = 0.685)	
Poe test: Control vs Treatment 2		0.25 (p = 0.493)	
Poe test: Treatment 1 vs Treatment 2			0.14 (p = 0.274)
WTP estimates: MNL model, WTP space, Org=1, ET=1 (Table A.36)			
	Full sample	Control	Treatment 1 Treatment 2
WTP _{CN} (Org=1,ET=1)	-0.25* (0.14)	-0.18 (0.26)	-0.02 (0.28) -0.43 (0.26)
Poe test: Control vs Treatment 1		0.66 (p = 0.672)	
Poe test: Control vs Treatment 2		0.25 (p = 0.495)	
Poe test: Treatment 1 vs Treatment 2			0.14 (p = 0.282)

Table A.38: Poe test results for competing label scenarios

WTP estimates are reported with standard errors in parentheses.

Significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

Poe test statistics are provided with p-values in parentheses, which are Holm-Bonferroni adjusted for three comparisons.

CN stands for carbon neutral, Org stands for Organic and ET for Ethical Trade. A value of 0 indicates that the label is present on the choice card, while 1 indicates that it is absent.

WTP estimates: MNL model, WTP space, Org=0, ET=0 (Table A.37)			
	Full sample	Control	Treatment 1 Treatment 2
WTP _{CN} (Org=1,ET=0)	0.74*** (0.20)	0.69*** (0.22)	1.00*** (0.25) 0.45** (0.20)
Poe test: Control vs Treatment 1		0.82 (p = 0.355)	
Poe test: Control vs Treatment 2		0.21 (p = 0.422)	
Poe test: Treatment 1 vs Treatment 2			0.04 (p = 0.088)
WTP estimates: MNL model, WTP space, Org=0, ET=1 (Table A.37)			
	Full sample	Control	Treatment 1 Treatment 2
WTP _{CN} (Org=0,ET=1)	0.73*** (0.19)	0.71*** (0.23)	0.90*** (0.21) 0.52*** (0.18)
Poe test: Control vs Treatment 1		0.73 (p = 0.540)	
Poe test: Control vs Treatment 2		0.25 (p = 0.505)	
Poe test: Treatment 1 vs Treatment 2			0.08 (p = 0.167)
WTP estimates: MNL model, WTP space, Org=1, ET=0 (Table A.37)			
	Full sample	Control	Treatment 1 Treatment 2
WTP _{CN} (Org=1,ET=0)	0.74*** (0.20)	0.69*** (0.22)	1.00*** (0.25) 0.45** (0.20)
Poe: Control vs Treatment 1		0.82 (p = 0.355)	
Poe: Control vs Treatment 2		0.21 (p = 0.422)	
Poe: Treatment 1 vs Treatment 2			0.04 (p = 0.088)
WTP estimates: MNL model, WTP space, Org=1, ET=1 (Table A.37)			
	Full sample	Control	Treatment 1 Treatment 2
WTP _{CN} (Org=1,ET=1)	-0.77*** (0.16)	-0.80*** (0.25)	-0.56** (0.27) -0.80*** (0.21)
Poe test: Control vs Treatment 1		0.74 (p = 0.524)	
Poe test: Control vs Treatment 2		0.49 (p = 0.984)	
Poe test: Treatment 1 vs Treatment 2			0.24 (p = 0.475)

Table A.39: Poe test results for competing label scenarios

WTP estimates are reported with standard errors in parentheses.

Significance levels: *** $p \leq 0.01$, ** $0.01 < p \leq 0.05$, * $0.05 < p \leq 0.1$

Poe test statistics are provided with p-values in parentheses, which are Holm-Bonferroni adjusted for three comparisons.

CN stands for carbon neutral, Org stands for Organic and ET for Ethical Trade. A value of 0 indicates that the label is present on the choice card, while 1 indicates that it is absent.

D Confirmatory and exploratory analyses

This section outlines the deviations from the pre-analysis plan and clarifies the distinction between confirmatory and exploratory analyses.

The study is pre-registered on the AEA RCT Registry (ID: 12520-2.0), with the main objective of estimating the effect of transparency regarding CO₂ offsetting and reduction on consumers' WTP for a carbon-neutral label. The following deviations from the pre-registration occurred:

First, although the pre-registration stated that 1,200 participants would be recruited and payment was made for 1,200 responses, the survey company provided data for a total of 1,339 participants.

Second, while the original plan indicated the use of a preference-space model, the final main analysis uses an MXL model estimated in WTP space. This deviation was motivated by the methodological advantages of WTP-space models, such as direct estimation of MWTP values and better interpretability. All preference-space model results are reported in the robustness checks.

Third, although the variable 'warm glow' variable was originally planned to be included in the estimation exploring underlying mechanisms, it was excluded due to its high correlation with guilt and social approval variables. Its role is examined separately in the robustness checks.

Fourth, to detect differences among experimental groups, both the Poe et al. (2005) and Swait and Louviere (1993) tests were pre-registered. However, only the Poe test was used in the analysis, as the main model was estimated using an MXL specification that accounts for individual heterogeneity. The Swait and Louviere (1993) test, by contrast, requires an MNL model.

Fifth, since the primary goal of the paper is not to estimate WTP for a carbon-neutral label per se, but rather to assess the effect of transparency, the translation of WTP into CO₂ reduction or offset equivalents, and its comparison with the social cost of carbon, is not discussed. This decision responds to criticisms that the study focuses on a single product, which may not be representative of all products and services, and relies on a stated preference method that may be

subject to hypothetical bias.

Fourth, although the pre-registration did not mention multiple hypothesis correction, all p-values for WTP estimates and hypothesis tests are corrected using the Holm-Bonferroni procedure. This correction has become a standard practice in experimental economics, and strengthens the reliability of analyses. Results without Holm-Bonferroni correction are reported in the robustness checks.

Fifth, I explore a set of additional factors that underlie my findings, such as familiarity with carbon-neutral labels and survey consequentiality, as robustness tests. I distinguish the pre-registered set from the additional ones explicitly in the main body.

Finally, while the exploration of the role of different sustainability labels was pre-registered, the interaction effects between sustainability labels are considered exploratory analyses.

All deviations and additional analyses are transparently reported, and robustness tests confirm that the main findings remain consistent across alternative specifications and samples.

E 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 (CO₂ 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
			
CO ₂ neutral label	☒ No label	☒ CO ₂ neutral	
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		
Organic label	☐ No label		
Ethical trade label	☐ No label		
	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
			
	X No label	X No label	
	Organic	X No label	
	X 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
			
	☐ No label	☐ No label	
	☐ No label	☒ Organic	
	☒ 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		
Organic label	☒ No label		
Ethical trade label	☒ No label		
	Price of tea product A £0.9	Price of tea product B £4.9	Price £0
I choose...			

CHOICE CARD 6/8

	Tea product A	Tea product B	None of the two
			
CO ₂ neutral label		☐ No label	
Organic label		☐ No label	
Ethical trade label		☐ No label	
	Price of tea product A £6.9	Price of tea product B £0.9	Price £0
I choose...			

CHOICE CARD 7/8

	Tea product A	Tea product B	None of the two
			
CO ₂ neutral label	☒ No label	☒ No label	
Organic label	☒ Organic	☒ No label	
Ethical trade label	☒ Ethical trade	☒ No label	
	Price of tea product A £2.9	Price of tea product B £0.9	Price £0
I choose...			

CHOICE CARD 8/8

	Tea product A	Tea product B	None of the two
			
			
			
			
	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.

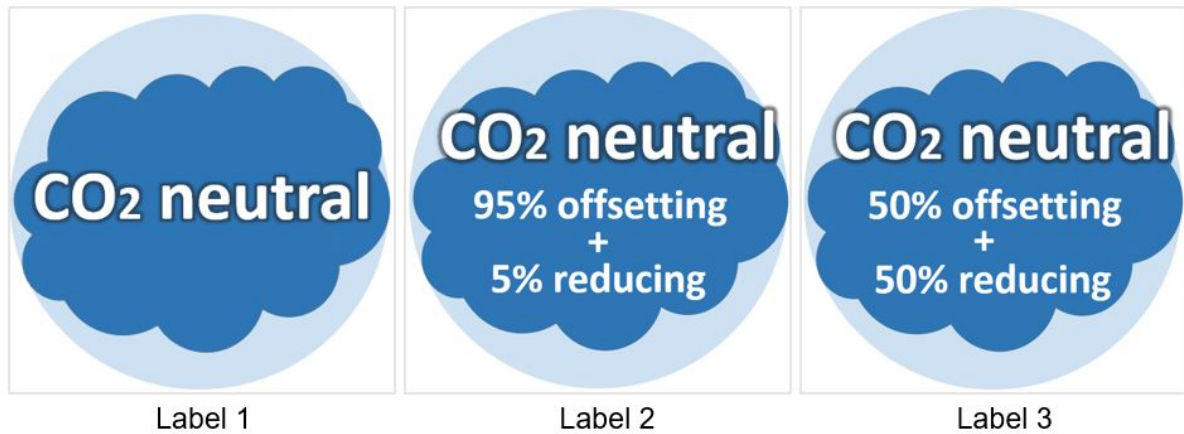
16. Please indicate to which degree you agree or disagree with the following statements.

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 lifecycle 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.							



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.ozdemiruluk@eawag.ch).

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

F 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 $\text{sd}(y)$ 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, t_1 is the t-value corresponding to the significance level (0.05) of the test, t_2 is the t-value corresponding to the power of the design (0.80), and MDEs is the minimum detectable effect size. I assume $\text{sd}(y)$ 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., 2025) 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.

Generative AI usage declaration

During the preparation of this thesis, the author used ChatGPT, Claude, Gemini, and Grammarly to assist with language editing and coding. All outputs were subsequently reviewed and revised by the author, who assumes full responsibility for the content of the thesis.

Begum Ozdemir-Oluk

CONTACT

July 2026

Affiliation	Swiss Institute for International Economics and Applied Economic Research (SIAW-HSG), University of St. Gallen, Switzerland
Email	begum.ozdemiroluk@unisg.ch
Website	begumozdemiroluk.github.io
Pronouns	she/her

EDUCATION

2020–2026	Ph.D. in Economics & Finance , <i>magna cum laude</i> , University of St. Gallen (HSG) , Switzerland. Advisors: <i>Sabrina Eisenbarth, Stefano Carattini, and Ivana Logar.</i>
2018–2020	M.S. in Economics; Minor in Mathematics , University of North Texas , Denton, Texas, USA.
2016–2018	M.A. in Economics , Bilkent University , Ankara, Turkiye. Advisor: <i>Bilin Neyapti.</i>
2011–2016	B.Sc. in Management (Finance specialization) , Bilkent University , Ankara, Turkiye.

RESEARCH POSITIONS

2020–2025	Swiss Federal Institute of Aquatic Science & Technology, Department of Environmental Social Sciences, Dubendorf, Switzerland.
-----------	---

WORKING PAPERS

Transparency of carbon-neutral labels: evidence from a choice experiment – **Revise and Resubmit** at *Environmental and Resource Economics (ERE)*. Job market paper.

Demand for carbon-neutral products, with Stefano Carattini, Fabian Dvorak, and Ivana Logar – **Revise and Resubmit** at *Journal of Environmental Economics and Management (JEEM)*.

PUBLICATIONS (PRE-DISSERTATION)

Neyapti, B., & Özdemir Oluk, B. (2021). *Fiscal Transfers in Turkey: Do Politics Matter?* *Economic Systems*, 45(3), 100909.

PEER-REVIEW SERVICE

2025–present Ad hoc reviewer, *Journal of Behavioral and Experimental Economics*.

SEMINARS

- 2026 Seminar, Institute of Responsible Innovation, Sustainability and Energy (RISE), University of St. Gallen, Switzerland.
- 2025 Colloquium, Department of Environmental Social Sciences (ESS), Eawag, Dubendorf, Switzerland.
- 2024 Seminar, School of Economics and Political Science (SEPS), University of St. Gallen, Switzerland.

CONFERENCES AND WORKSHOPS

- 2025 30th Annual Conference of the European Association of Environmental and Resource Economists (EAERE), Bergen, Norway.
- Annual Congress of the Swiss Society of Economics and Statistics (SSES), ETH Zurich, Switzerland.
- ETH Domain Conference: Social Sciences on Environment, Technology and Sustainability, Dubendorf, Switzerland.
- HSG Academic Poster Session, HSG Responsibility & Sustainability Services, University of St. Gallen, Switzerland.
- JEEM/REE Workshop on Nonmarket Valuation, Bergen, Norway.

- 2024 29th Annual Conference of the European Association of Environmental and Resource Economists (EAERE), Leuven, Belgium.
 Choice Experiment Clinics, VU Amsterdam, Netherlands.
 European Association of Environmental and Resource Economists (EAERE) Summer School: Data Science in Environmental Economics Research, Graz, Austria.
 Eawag Symposium, Dubendorf, Switzerland.
 Monte Verità Conference on Sustainable Resource Use and Economic Dynamics (SURED), Ascona, Switzerland.
 Workshop on Empirical Methods in Energy Economics (EMEE), ETH Zurich, Switzerland.
 Workshop on Experimental Economics for the Environment (EE4E), Bochum, Germany.
- 2023 Graduate Program in Economics and Finance PhD Day, University of St. Gallen, Switzerland.
- 2022 Eawag Symposium, Dubendorf, Switzerland.
- 2021 Eawag Symposium, Dubendorf, Switzerland.

TEACHING EXPERIENCE

- 2019–2020 Teaching Assistant, [University of North Texas](#), Denton, Texas, USA.
 Advanced Microeconomic Theory (ECON 5340); Intermediate Microeconomics (ECON 3550).
- 2016–2018 Teaching Assistant, [Bilkent University](#), Ankara, Türkiye.
 Macroeconomic Theory II (ECON 206); Monetary Economics (ECON 322); History of Economic Thought (ECON 363); Econometrics (ECON 301).

INDUSTRY EXPERIENCE

- 2014–2015 Audit Intern, Ernst & Young, Istanbul, Türkiye; Accounting Intern, Ford Otosan, Eskisehir, Türkiye; Accounting Intern, Salim Arabaci Certified Public Accountant, Eskisehir, Türkiye.

SELECTED COURSEWORK

- 2021 [University of Zurich](#): Topics in Experimental and Behavioral Economics.
- 2021 [Statistical Horizons](#): Meta-Analysis.
- 2020–2021 [University of St. Gallen](#): Topics in Advanced Microeconometrics; Causal Machine Learning; Econometric Methods for Social Spillovers and Networks; Regression Analysis for Spatial Data.

HONORS & AWARDS

- 2019–2020 [University of North Texas](#)
Texas Public Education Grant; Department of Economics Chair’s Award.
Toulouse Graduate School Tuition Benefit Program.
- 2011–2018 [Bilkent University](#)
Scholarship for M.A. in Economics (full tuition and monthly stipend).
Scholarship for B.Sc. in Management (2015–2016: full tuition; 2011–2015: 50% tuition).

TECHNICAL & LANGUAGE SKILLS

- Software [R](#); [Stata](#); [Ngene](#); [L^AT_EX](#).
- Languages [Turkish](#) (native); [English](#) (advanced); [German](#) (elementary).