The Effectiveness of Carbon Labels [*]

Anna Schulze-Tilling

University of Glasgow

First version: May 18, 2023

This version: December 10, 2025 Click here for the most recent version

Behavioral climate interventions are often promoted as politically feasible substitutes for carbon taxes, yet we lack a way to value them in tax-equivalent terms. I develop and field-test a framework that benchmarks such interventions against a carbon tax. In a German student canteen, a framed field (N = 289) and a natural field experiment (125,000 purchases) show that carbon labels replicate the effect of a €120/t CO_2 tax. A third experiment (N = 444) and a structural model show that the effect is driven mainly by increased salience rather than misperception correction, and that the labels raise consumer welfare.

JEL Classification: C93, D83, H23, Q54, Q58

Keywords: carbon labels, climate policy, behavioral interventions, salience, consumer welfare

[*]**Contact:** Anna Schulze-Tilling; University of Glasgow, anna.schulzetilling@glasgow.ac.uk.

Acknowledgements: I am indebted to Lorenz Götte, Sebastian Kube, and Thomas Dohmen for their continuous support, and am grateful to the Studierendenwerk Bonn for their great cooperation. I am thankful to Hunt Allcott, Peter Andre, Sanchayan Banerjee, Simon Block, Stefano DellaVigna, Eugen Dimant, Katrin Gödker, Raphael Epperson, Ximeng Fang, Nicola Gennaioli, Holger Gerhardt, Luca Henkel, Michael Kosfeld, Chris Krekel, Matthew Levy, Paul Lohmann, Julia Mink, Davide Pace, Eugenio Proto, Simone Quercia, Alex Rees-Jones, Matthias Rodemeier, Paul Schäfer, Ganga Schreedhar, Marie Claire Villeval, Jakob Wegmann, and audiences at the LSE, Bonn, Maastricht, the UIB, Geneva, PSE, Glasgow,the Briq Workshop on Climate Change, BEEF, MCEE, AERE, SABE-IAREP, M-BEPS, EAERE, EEA, VfS, ECBE for their comments.

Funding: Support by the German Research Foundation (DFG) through CRC TR 224 (Project B07) and Germany's Excellence Strategy – EXC 2126/1-390838866, and from the European Research Council - HEUROPE - 2022-ADG GA 101097578 - MemDec is gratefully acknowledged.

Ethics approval: The experiments obtained ethics approval from the German Association for Experimental Economic Research (#BnyKWhHQ and #RSzBqAcE)

Pre-registration: The experiments were pre-registered at the AEA RTC registry and aspredicted.org (#AEARCTR-0007858,#AEARCTR-0008435 and #95108).

1 Introduction

Many economists agree that implementing carbon taxes is the most effective policy to combat climate change (European Association of Environmental and Resource Economists, 2019).[1] However, their implementation is not always feasible. For instance, even modest levies on food, a sector contributing an estimated 26% to 34% of global greenhouse gas emissions (Poore and Nemecek, 2018; Crippa et al., 2021), face considerable political resistance.[2] Behavioral climate interventions such as carbon labels have emerged as a seemingly promising and easier-to-implement alternative. Carbon labels are included in the EU's Farm-to-Fork Strategy and are increasingly voluntarily adopted by producers. Yet we lack an evidence-based benchmark of the labels' effectiveness, comparing them to the carbon tax they might proxy for. To gauge their true policy value, we need to understand: What per-tonne tax can the labels substitute for, and what is the net welfare impact of this substitution on consumers?

This paper provides the first estimate of this benchmark, using data from a student canteen in Bonn, Germany. I estimate how high a carbon tax the labels can proxy for by applying two complementary methods: a framed field experiment and a large-scale natural field experiment. Each method has distinct strengths and limitations that together reinforce the robustness of the benchmark estimate. In the framed field experiment ($N = 289$), I elicit participants' willingness to pay (WTP) for typical student canteen meals, with and without carbon labels. This allows me to directly observe how labels shift the demand curve. The WTP elicitation is incentivized by providing participants with their meal choice upon completing the experiment. In the natural field experiment, I use guest-level canteen purchase data to analyze a 7-week carbon labeling intervention in one of Bonn's canteens, using difference-in-differences to compare changes in purchasing behavior to those in Bonn's other canteens. I then compare estimated effect sizes to demand fluctuations induced by variations in the canteens' pricing.

While the framed field experiments provides more precision and control, the natural field experiment better captures long-term, real-world behavior. Crucially, both methods point to the same

1. I use the term carbon tax to refer to price increases passed on to consumers that reflect the carbon emissions associated with a product, whether imposed through a direct tax (e.g., Germany's fuel tax) or through market mechanisms like the EU Emissions Trading System (EU ETS). Such carbon costs are typically embedded in the final price and are not explicitly communicated to consumers at the point of purchase as a distinct "carbon tax".

2. Dechezleprêtre et al. (2022) show opposition to a regulation of the food sector in global survey data. Additionally, Douenne and Fabre (2020) document considerable opposition to meat taxes in France, and Fesenfeld (2023) outlines the political challenges of implementing such taxes in Germany. Dietary shift towards lower-emission foods would substantially reduce food emissions (Poore and Nemecek, 2018; Kim et al., 2020; Grummon et al., 2023; Scarborough et al., 2023).

benchmark: carbon labels induce demand shifts equivalent to a carbon tax of €120 per tonne. This estimate exceeds average 2024 EU ETS prices by over 150% and is roughly triple the German carbon tax for fossil fuels for transport and heating in 2024.[3] I estimate that carbon labels reduce emissions by about 4%, consistent with findings from related field studies.[4]

Next, I compare the impact of carbon labels and a revenue-neutral carbon tax on consumer welfare. For this purpose, I develop a simple discrete choice model describing the impact of carbon labels and the benchmarked tax on consumer behavior and welfare, estimate model parameters with my experimental data and then simulate the behavioral and welfare effects of both policies.

In the model, a diner chooses the meal that maximizes their perceived utility. I model this perceived utility as a function of taste and price, the diner's prior estimate of the meal's CO_2 emissions, the psychological cost (or "guilt") associated with each kilogram of emissions, and the salience of those emissions at the moment of choice. Carbon labels affect perceived utility in two ways: they correct misperceptions about a meal's carbon footprint and they increase the salience of emissions at the moment of choice.[5] The benchmarked carbon tax affects perceived utility and thereby consumption decisions via a change in prices.

The carbon labels affect consumer welfare in two ways. First, by changing consumption decisions and thus directly influencing realized utility. Second, by generating psychological costs or benefits independent of any behavioral change. This second term captures any effects on consumer welfare stemming from the labels having effects on consumers other than those driving behavioral changes. For example, a consumer might feel annoyed by the presence of labels or may suffer from experiencing increased social pressure. Conversely, a consumer might experience satisfaction from increased information or a stronger sense of "warm glow" even if the labels do not alter their decisions (e.g. because they were already previously making environmentally friendly choices).[6] The benchmarked

3. The average EU-ETS price in 2024 was €65, €84 in 2023, and €80 in 2022 (German Environment Agency (UBA), 2024; International Carbon Action Partnership (ICAP), 2025). The German carbon tax was €45 in 2024, and €30 in 2023 and 2022 (Wettengel, 2024). The agricultural sector is not included in either of these schemes.

4. Comparable studies include Lohmann et al. (2022), which reports a 4.3% reduction in British student canteens, and Brunner et al. (2018), which finds a 3.6% reduction in a Swedish student canteen.

5. My focus on these two channels is motivated by my reduced-form results suggesting that misperception correction and salience alone explain a large share of the observed behavioral response. These are therefore the focus of my model of consumer behavior. However, secondary channels may still affect consumer welfare, even if they do not alter consumer behavior. I thus adopt a more agnostic approach when modeling welfare effects.

6. My model focuses on the direct effects of labels on the individual consumer and excludes second-order welfare effects, such as feeling good because others might change their behavior. Any warm glow from others' emission reductions would apply equally to the introduction of carbon labels and introduction of the benchmarked tax, and is therefore less relevant for the label vs. tax comparison. As the experimental data used to estimate the model concerns participants'

carbon tax affects consumer welfare by impacting realized utility and via a lump-sum redistribution of tax proceedings.

To quantitatively estimate my model, I conduct another framed field experiment ($N = 444$). This study mirrors the first benchmarking experiment but additionally introduces treatments for carbon offsetting and for a pure increase in salience, and elicits participants' own emissions estimates. Reduced-form results suggest that only a small share of the labels' effect is attributable to them correcting misperceptions; the majority is driven by the increased salience of carbon emissions. Results also suggest that the bulk of the labels' effect on behavior is attributable to these two channels, supporting my modeling choices. Structural estimation of the model yields that approximately 80% of the labels' impact stems from an increased salience. Further, participants are on average willing to pay €0.21 to see carbon labels when making a meal choice. Incorporating these data into the welfare analysis, my simulations yield that, in this context, labels are welfare-improving, delivering higher net welfare gains than the benchmarked carbon tax.

My quantitative estimates contribute to assessing the policy relevance of carbon labels in student canteens—a particularly promising setting due to the ease of implementation and the large student population.[7] Beyond this immediate application, the findings yield three broader contributions.

First, this paper shows the value of benchmarking behavioral climate interventions against carbon taxation. The emission reductions I estimate are similar to those reported in earlier studies,[8] but my interpretation differs. Previous work often labeled such reductions as "small" seemingly because they were single-digit (see e.g. Brunner et al., 2018). By contrast, my benchmarked carbon tax equivalent allows effect sizes to be interpreted against established policy instruments, offering a different perspective on the labels' effectiveness than prior research. The divergence in resulting assessments suggests that one should be wary of judging a policy's effectiveness solely against ideal emission targets; it should also be evaluated relative to feasible alternatives. The benchmarking approach,

private decisions, made in isolation from other consumers, such second-order effects should not influence participants' choices.

7. In Germany alone, 2.9 million individuals were classified as students in 2021 (Federal Statistical Office (Germany), 2023), with approximately 54% dining in student canteens at least once per week (Federal Ministry of Education and Research (Germany), 2023). It is relatively cost-efficient to implement carbon labels in student canteens since a single facility often caters to many guests, which means that calculating carbon labels on a single menu impacts many guests. For example, student canteens in Bonn centralize their menu planning and cater to over 2,000 guests daily. Further, most student canteens in Germany operate as non-profit institutions aimed to serve students. This may make the introduction of carbon labels feel like a more natural step than in the case of commercial restaurants, especially if desired by a majority of the student population.

8. See, e.g., Lohmann et al. (2022) and Brunner et al. (2018), who find reductions of 4.3% and 3.6%, respectively.

particularly the framed field experiment, can readily be applied to other consumption settings and interventions.

Second, this study provides the first experimental evidence that carbon labels primarily operate by increasing the salience of carbon emissions at the point of decision-making, rather than by correcting misperceptions. This challenges a prevailing interpretation in the literature (e.g., Shewmake et al., 2015; Camilleri et al., 2019; Imai et al., 2022) and helps reconcile previously mixed results. For instance, Imai et al. (2022) find no behavioral effect when testing carbon labels in a design that isolates the information-correction channel. My findings show that this null result does not imply that carbon labels are ineffective, but rather that their main effect arises through salience, a mechanism absent from their design. While the relative importance of salience likely varies across settings and populations, my evidence demonstrates that it is a key channel that should be explicitly incorporated into theoretical and empirical models of carbon label effects.

Third, I provide novel evidence that carbon labels can generate psychological benefits for consumers, even when they do not change behavior. This complements existing discussions, which often emphasize potential psychological costs. By incorporating these effects into a structural welfare analysis, I show that net psychological effects may be positive, an aspect that deserves greater consideration in both academic and policy debates.

Correspondingly, this paper contributes to three key strands of literature. First, my findings inform our understanding of the effectiveness of carbon labels on food consumption products. Previous field studies (most closely related are Brunner et al., 2018; Bilén, 2022; Lohmann et al., 2022; Ho and Page, 2023) report emission reductions of 1–5%. Reasons estimates differ across studies can be differences in setting (e.g., how substitutable high- and low-carbon options are) and population (e.g., baseline carbon intensity of diets or concern for emissions). Findings also differ by study design, with online studies (e.g. Imai et al., 2022; Banerjee et al., 2023; Lohmann, Gsottbauer, et al., 2024) sometimes finding smaller, sometimes larger effect sizes depending on incentivization and study design.[9] This paper adds to this literature by providing evidence from a German canteen, the first causal estimates of the labels' post-intervention effects, and the first benchmark and welfare comparison to a carbon tax.

More broadly, my method for establishing a tax benchmark offers a framework for comparing behavioral interventions to price instruments in a consistent, internally valid way. Rather than relying on elasticity estimates elicited within a different context and population, my benchmarking approach compares a single population's responses to both carbon labels and price changes, in a single setting. This removes potential noise due to price elasticities and consumer behavior varying

9. See Lohmann, Pizzo, et al. (2024) for a recent meta-analysis of demand-side interventions for food consumption.

across populations and contexts (also see Lusk and Tonsor, 2016). Further, the benchmarking procedure in the framed field experiment does not rely on any price variation, avoiding concerns about the endogeneity of prices. This approach can be extended to benchmark other interventions aimed at reducing carbon emissions, including those in energy, water, and household consumption (relating to e.g., Allcott, 2011; Brent, Cook, and Olsen, 2015; Tiefenbeck et al., 2018), as well as food purchasing behavior (relating to e.g., Jalil, Tasoff, and Bustamante, 2023; Lohmann, Gsottbauer, et al., 2024).

Second, my findings on the relative role of salience and misperception correction – providing first experimental evidence on the role of salience in driving carbon label effectiveness – contribute to a growing literature on salience and attention in consumption decisions (e.g., Chetty, Looney, and Kroft, 2009; Busse et al., 2013; Taubinsky and Rees-Jones, 2018), particularly in environmental contexts (Allcott and Taubinsky, 2015; Tiefenbeck et al., 2018; Rodemeier and Löschel, 2022). I quantify the relative contributions of salience and misperception correction, offering new insight into the cognitive mechanisms behind behavioral policy tools.

Finally, I contribute to the literature on the welfare implications of nudges and information policies. One strand of this literature applies generalized frameworks to evaluate a range of interventions across settings (e.g., List et al., 2023; Hahn et al., 2024; Allcott et al., 2025), while another develops more detailed models for specific interventions (e.g., Allcott and Kessler, 2019; Andor et al., 2023). My paper fits within the latter strand, offering a detailed welfare evaluation of carbon labels relative to taxes. However, my findings may also inform the former strand, for example by providing empirical guidance on modeling assumptions concerning behavioral channels and consumer welfare. In particular, the dominant role of salience and the presence of psychological effects independent of behavior change may generalize to other labeling interventions, such as those related to health or animal welfare.

The remainder of this paper proceeds as follows. Section 2 describes Experiment 1, which benchmarks the effectiveness of carbon labels using a framed field experiment. Section 3 presents Experiment 2, which benchmarks the effectiveness of carbon labels using a natural field experiment. Section 4 examines the behavioral channels driving the observed effects, drawing on data from Experiment 3. Section 5 shows reduced-form evidence from all three experiments concerning consumers' preferences towards the presence of the label. Section 6 introduces a simple theoretical model describing the labels' effect on consumer behavior and welfare, which I structurally estimate using data from Experiment 3. Finally, section 7 concludes.

2 Benchmarking method 1: Framed field experiment

This section presents the first method I use for benchmarking the effectiveness of carbon labels relative to a carbon tax. For this method, I use a framed field experiment (Experiment 1) as it provides precision and control. Subsection 2.1 describes the experimental design and data, and subsection 2.2 shows the empirical strategy and results.

2.1 Experimental design

Overview. I designed the framed field experiment such that it, by design, benchmarks the effectiveness of carbon labels against a carbon tax. The key design feature enabling this benchmark is that I observe participants' willingness to pay (WTP) for the same meal both with and without carbon labels, keeping all else equal. Below, I summarize key design choices and provide details in the following paragraphs.

(1) Participants' lunch choices are moved to an online survey, completed just before lunchtime on the experiment day. Shortly after, they go to campus to receive their experiment payment and the meal corresponding to their choices.

(2) In the survey, participants indicate their WTP for different meals, totaling to 15 meal purchase decisions. One decision is implemented at payout.

(3) Participants first state their WTP for four meals without seeing carbon labels, to capture their WTP at baseline. Then, if they are in the Label condition, they state their WTP for the same four meals while seeing labels. If they are in the Control condition, they simply repeat the WTP elicitation without seeing labels, to capture any effect of eliciting WTP twice. The BDM mechanism incentivizes truthful responses in the WTP elicitations.

(4) WTP is elicited relative to an alternative lunch: In each of the 15 decisions, participants state their WTP for a given meal relative to a cheese sandwich, reflecting the real-world fact that humans must eat: Deciding not to eat one item implies needing to eat an alternative item. By directly eliciting WTP relative to an alternative lunch, I observe this alternative item, giving me more precision and control when interpreting results. Labels are always either shown or not shown for both options.

(5) Carbon labels display greenhouse gas emissions (in kg), a traffic-light-style ordinal rating, and the equivalent driving distance (in km) to produce the same emissions (see Figure 4). I co-designed the labels with Bonn's student canteens to reflect a format they would realistically

implement. A similar design was in fact adopted when the canteens later introduced permanent carbon labels (see Appendix E.9). Combining quantitative and ordinal information has been shown to be effective in prior studies (Taufique et al., 2022; Potter et al., 2021).

(6) WTP to see or avoid carbon labels is also elicited: Before the final three purchase decisions (three new meals), participants choose whether they want to see carbon labels and indicate their WTP to enforce their choice. This elicitation is incentivized with a BDM mechanism.[10] These results are discussed in Section 5.

Experiment timeline. The experiment timeline is shown in Figure 1. First, participants receive an explanation of the WTP elicitation and answer comprehension questions.[11] Next, they indicate their baseline WTP for four canteen meals, incentivized by a BDM mechanism.[12] To create buffer time before the second WTP elicitation, participants answer incentivized guessing questions on unrelated topics (e.g., the length of a popular running route in Bonn). In the second WTP elicitation, the framing of the decisions depends on the randomly assigned treatment:

- Control: Decisions are identical to the baseline elicitation.
- Label: Participants see carbon labels as they make their decisions.

For additional insights, WTP is elicited a third time,[13] with altered treatment conditions:

- Participants previously in the Label condition are in the Offset condition: Participants are informed that meal emissions (meal or sandwich) will be offset. This serves as a robustness check for Experiment 3 and is detailed in online Appendix D.4, with reduced-form results in Table A.8.
- Half of the participants previously in the Control condition receive the Label condition, and half of the participants previously in the Control condition repeat the Control condition. Afterward, this group estimates emission values.[14]

Each round includes four meal purchasing decisions (12 total). Additionally, three final decisions involve previously unseen meals. Before making these final decisions, participants indicate whether they want to see carbon labels for these decisions and state their WTP to enforce their choice, in-

10. See online Appendix D.3 for details.
11. Participants must answer correctly to proceed. Those requiring more than five attempts are excluded, as pre-registered.
12. See online Appendix D.3 for details.
13. Analyses control for third-round data. Main results replicate using only the first two rounds, see Table A.23.
14. Used in Figure 11. As these questions occur after the third WTP elicitation, they do not affect results in this section.

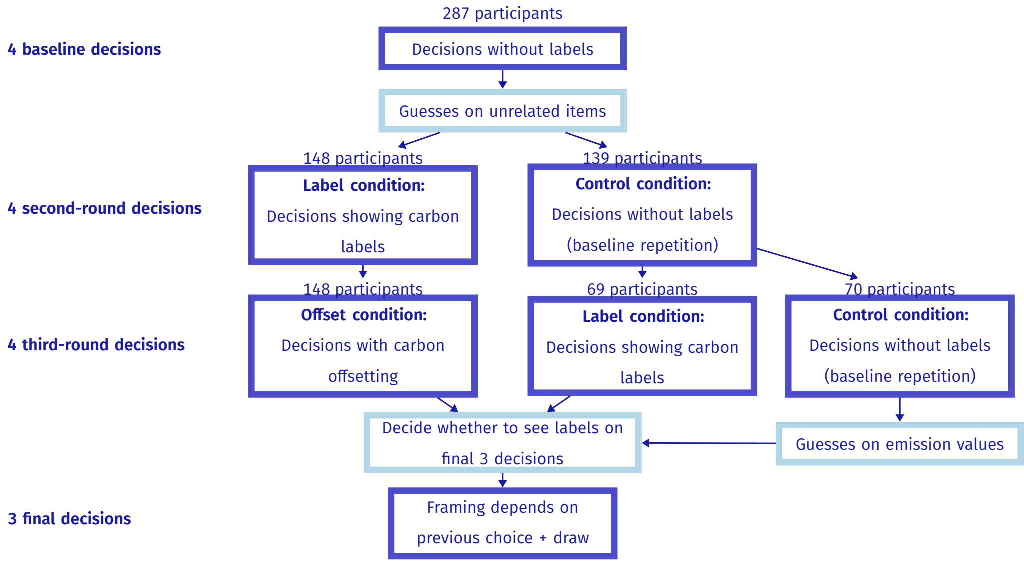

Figure 1. Experiment 1 schedule and participant numbers after data cleaning

centivized via a BDM mechanism. Finally, participants answer questions on environmental attitudes and psychology,[15] and guess the calories of each meal.[16]

Details on the meal purchasing decisions. In each of the 15 decisions, participants first choose between a specific meal and a cheese sandwich. An example decision is shown in Figure 2. The left option changes across decisions to cycle through four meals, while the right option (cheese sandwich) remains constant.[17] Participants who indicate they are vegetarian see only vegetarian meals.[18]

After selecting a preferred option, a second window appears where participants indicate how much of their experiment payment they would forego to secure their preference (Figure 3). The question wording in this second step differs depending on whether a participant selected the meal or sandwich as their preferred option. If they preferred the meal, they are asked to indicate their WTP to receive it instead of the sandwich. If they preferred the sandwich, they are asked to indicate

15. I collect these for the heterogeneity analyses in online Appendix C.1.

16. I collect these for a test whether the labels affect participants' perception of the nutritional content of meals. I do not find evidence of this being the case (Table A.7).

17. The cheese sandwich option remains constant to ease comparison across WTP elicitations. If a participant rather likes or dislikes cheese sandwiches, this will affect all their decisions in the same manner, still enabling a clean comparison between WTP values. To control for left-right effects, positioning is reversed in half of the sessions. Meal order is randomized.

18. Meals are detailed in online Appendix D.2. Participants with stricter dietary requirements (vegan, gluten-intolerant, lactose-intolerant, or halal) are excluded.

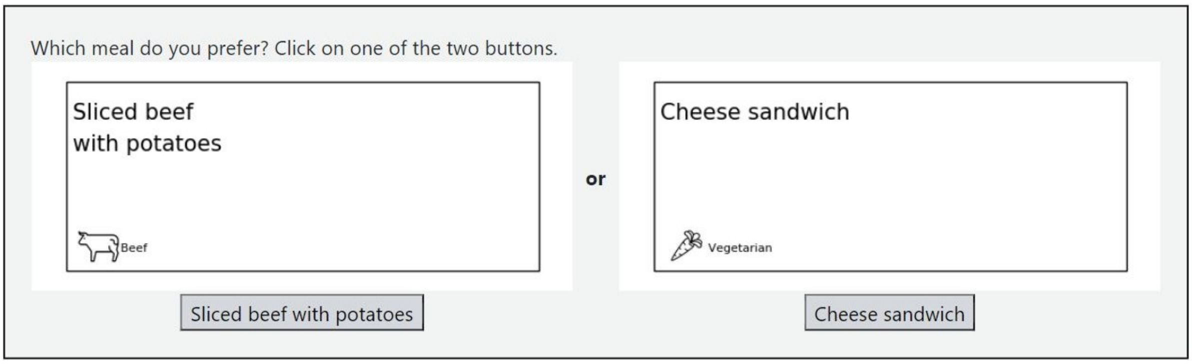

Figure 2. Meal purchase decision example step 1

Notes: Step 1 of the purchasing decision. Depending on participants' choice, Step 2 (Figure 3) asks for their WTP to receive or avoid the warm meal.

their WTP to receive it instead of the meal. Responses are made on a slider in five-cent intervals, between €0.00 and €3.00.[19]

This procedure effectively captures participants' WTP for the meal relative to the cheese sandwich, allowing for negative values: If a participant states they prefer the sandwich and have a positive WTP to secure their choice, I multiply this amount by -1 and interpret it as negative WTP for the meal. Participants are incentivized to report their true WTP using a BDM mechanism, as detailed in Appendix D.3.

In the four baseline decisions, participants see only the meal name and main ingredient, without carbon labels (Figure 2). This mirrors how meals are typically displayed on the student canteen website. The four second-round and four third-round decisions resemble the baseline, except for framing differences across treatments. In the Label condition, emission values are added to both meal options (Figure 4).[20] In the Control condition, framing remains unchanged. In the Offset condition, participants are informed that the emissions of the consumed meal option will be offset through a donation to the nonprofit carbon offsetting service Atmosfair. The Offset condition is not further discussed here, but details are in online Appendix D.4, with results in Table A.8.

19. €3.00 is the maximum price students pay for any meal in the canteen. WTP values at the interval boundaries (±€3.00) occurred in less than 4% of observations. Figure A.1 in the online Appendix shows the distribution of baseline WTP values.

20. Meal emissions were calculated using the application Eaternity Institute (2020). The student canteen provided meal recipes for these calculations.

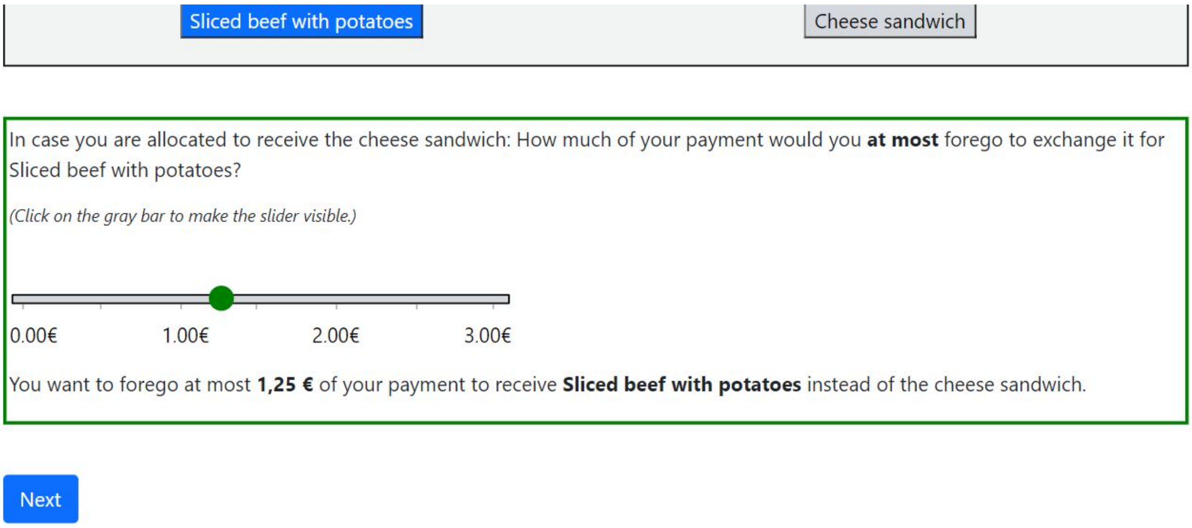

Figure 3. Meal purchase decision example step 2

Notes: Step 2 of the purchasing decision, as shown to a participant preferring the meal in Step 1. For a participant preferring the cheese sandwich, the wording is changed to "In case you are allocated to receive the Sliced beef with potatoes: How much of your payment would you at most forego to exchange it for the cheese sandwich? – You want to forego at most €1,25 of your payment to receive the cheese sandwich instead of the Sliced Beef with potatoes." Note that the €1.25 is an example of what a participant might indicate. The handle on the slider and the text below the slider only appear once participants click on it.

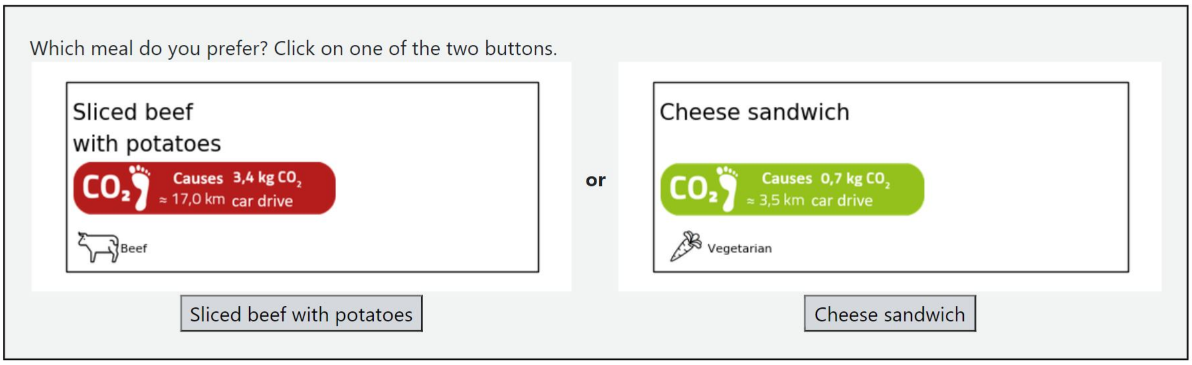

Figure 4. Meal purchase decision example: Decisions with labels

Participants and data. 304 participants from the BonnEconLab, the University of Bonn's behavioral experimental lab, took part in one of eight experimental sessions between October 26 and November 5, 2021. I pre-registered the experiment design and main outcomes shown in this section (Schulze Tilling, 2021b).[21]

Participants are informed that the experiment is conducted online but that they must visit campus directly afterward to collect their cash payment and lunch. They receive no further details on the experiment's purpose. The experiment is conducted using oTree software (Chen, Schonger, and Wickens (2016)).

21. See Table A.8 in the online Appendix for all pre-registered main results.

Meals are provided by the student canteen. While all experiment meals are regularly offered by the canteen, they are not available on experiment days, meaning the canteen prepares them exclusively for participants. Participants receive a warm, ready-to-eat meal based on their online experiment choices and can consume it on-site.

Of the 304 participants, I exclude the 3% fastest respondents and those failing the comprehension check after five attempts, as pre-registered.[22] One incomplete response is also dropped, leaving 287 participants who were computer-randomized into treatments (randomization check in Appendix A.1). Participants are on average 24 years old; 67% are female, 80% are students, and 25% are vegetarians. The sample is roughly representative of regular student canteen guests, as discussed in online Appendix A.2, and results hold when restricting the sample to only students or only non-vegetarians (online Appendix A.8). Baseline WTP distributions are shown in online Appendix A.3: 22% of WTP values are 0 (indicating indifference between the meal and sandwich), 17% are negative (preference for the sandwich), and the rest are positive, with some bunching around €1. Less than 4% of WTP values are at the boundaries of the -€3 to €3 interval.

2.2 Estimation strategy and results

I conduct two main analyses using the experimental data. First, I test whether carbon labels significantly affect participants' willingness to pay (WTP), thereby establishing their effectiveness in this setting. Second, I benchmark the labels' effect against an equivalent carbon tax to assess their policy relevance.

In both analyses, the outcome variable is the within-subject *change* in WTP for each meal – that is, the difference between participants' baseline WTP and their subsequent WTP in the Label or Control treatment. This differencing approach controls for individual- and meal-specific factors (e.g., taste, hunger, mood) that should remain constant across both elicitation rounds.

I then compare how changes in WTP differ across treatments. In the Control group, changes reflect only the effect of repeated elicitation. In the Label group, they additionally capture the effect of seeing carbon labels. As a robustness check, I estimate an alternative specification using WTP as the outcome and including individual–meal fixed effects. This yields similar results (Appendix A.10).

Effect of carbon labels on WTP. I first test whether carbon labels significantly affect WTP for meals, by comparing the change in WTP in the Label group to that in the Control group.

22. See Schulze Tilling (2021a). Dohmen and Jagelka (2024) find that fast respondents are more likely to give random answers.

The effect of the labels can be expected to differ depending on the relative carbon emissions of the meal compared to the alternative lunch, the cheese sandwich. If the meal has higher emissions than the sandwich, the labels should reduce WTP as participants shift toward the lower-emission alternative. Conversely, if the meal has lower emissions, the labels should increase WTP as participants shift away from the sandwich and toward the more climate-friendly meal.

I thus distinguish between high- and low-emission meals in the regression:

$$\Delta WTP_{ijm} = \beta_1 High_m + \beta_2 Low_m + \delta_1(Label_{ij} \times High_m) + \delta_2(Label_{ij} \times Low_m) + ThirdRound_j + \varepsilon_{ijm} \quad (1)$$

Here, ΔWTP_{ijm} is the change in WTP of participant i in round j for meal m, relative to their baseline WTP for that same meal. All WTP values are expressed relative to the cheese sandwich, which serves as the constant alternative lunch option.

The indicators $High_m$ and Low_m denote whether a given meal has higher or lower emissions than the sandwich.[23] These terms capture any effect of repeated WTP elicitation. The interaction terms $(Label_{ij} \times High_m)$ and $(Label_{ij} \times Low_m)$ identify the causal effect of seeing the carbon labels for high- and low-emission meals, respectively. Finally, $ThirdRound_j$ controls for whether the decision occurred in round 3,[24] ensuring comparability across rounds.

Table 1, Specification (1), presents the results. In the Label condition, WTP for lower-emission meals increases by €0.14, while WTP for higher-emission meals decreases by €0.31. In contrast, changes in the Control group are small, not statistically significant, and move in the opposite direction, suggesting that repeated WTP elicitation alone does not significantly affect WTP. Figure 5 visualizes these average WTP changes by treatment group and emission category.

Benchmarking labels against a CO_2 tax. I now benchmark the effect of carbon labels against an equivalent carbon tax. The core analysis is a regression that estimates how much WTP shifts in response to labels, depending on a meal's relative emissions. To build intuition, I first illustrate how labels shift demand similarly to a carbon tax. I estimate empirical demand curves using participants' WTP. For every possible price on the -3 to 3 interval, I compute participants' demand as the share of participants with a WTP equal or higher to the price, and then estimate the best-fitting line (gray line based on baseline WTP; red line based on WTP with labels). Figure 6 (bottom-left) shows that for high-emission meals, carbon labels shift the demand curve downward. For low-emission meals,

23. For non-vegetarians, three of four meals had higher emissions; for vegetarians, two of four. See Appendix D for details.

24. Since some participants experienced the Label condition in round two and others in round three, this control ensures comparability. Dropping round 3 decisions yields similar results; see Table A.23 in the Online Appendix.

demand shifts upward with the labels (top-left), as labels make them more attractive relative to the sandwich alternative.

How high of a carbon tax would generate a comparable shift as produced by the labels? To provide an estimate, I next analyze by how much WTP decreases in reaction to the labels, depending on the difference in emission intensity between meal m and the sandwich:

$$\Delta WTP_{ijm} = \beta_1 Emi_m + \delta_1 (Label_{ij} \times Emi_m) + ThirdRound_j + \varepsilon_{ijm} \tag{2}$$

where Emi_m is the difference in emissions, in kg, between meal m and the sandwich. The interaction term captures how the labels' effect scales with this difference in emissions. Table 1, Specification (2), estimates that the labels reduce WTP by €0.12 per kg (€120 per tonne) of CO_2. This result establishes the €120 per tonne tax benchmark I use throughout the paper.

To visualize the equivalence, I additionally estimate the effect of a €120/t CO_2 tax: for each meal, I subtract the resulting tax from participants' baseline WTP and re-estimate demand curves (Figure 6, right panel).[25] The resulting demand shift closely mirrors that induced by the labels.[26]

The key experimental feature enabling this straight-forward benchmarking analysis is my incentive-compatible elicitation of WTP. Capturing participants' WTP, rather than a single binary purchasing decision at a given set of prices, allows for a clean comparison between a behavioral intervention (carbon labels) and a price instrument (carbon tax), within the same population and context. The design is portable to other interventions and domains. Moreover, combining within-subject WTP changes with between-subject treatment variation boosts statistical power, an advantage when working with incentive-compatible designs and the often resulting moderate sample sizes.

Online Appendix A.11 provides further intuition behind the benchmark. Appendix A.5 simulates participants' canteen choices using experiment data and estimates that labels reduce average emissions by 4.8%.

25. The effect of a tax can be visualized as a downwards shift in the demand curve. This is most intuitive to understand thinking of a point-of-sale tax: Those consumers willing to purchase a product for a given product price now need to be willing to pay the tax as well. This decreases the proportion of consumers willing to purchase, and thus shifts the demand curve downwards. However, the tax does not need to be deducted at sale for this logic to hold. Any carbon price that affects the final consumer-facing price, whether upstream or at checkout, would produce the same shift.

26. Both the demand-curve estimations and the regression are approximations for the purpose of benchmarking the labels' effect size. Later sections provide a more detailed model of how labels affect consumers. Confidence intervals in Figure 6 reflect estimation uncertainty.

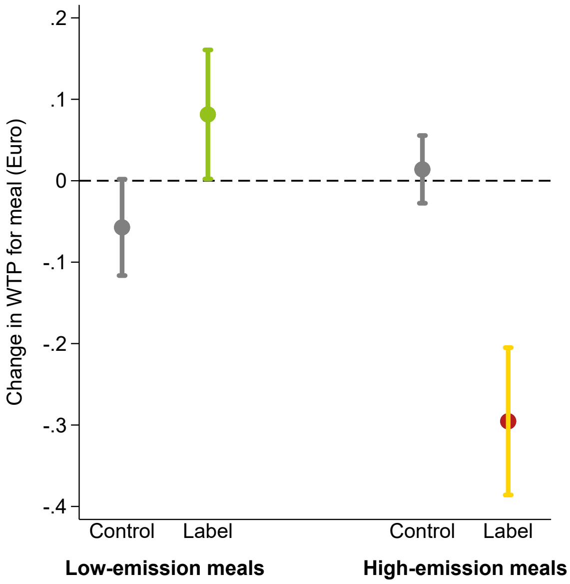

Figure 5. Within-subject change in WTP for meals by treatment condition.

Notes: Bars indicate 95% confidence intervals. Figure visualizes spec. (1) in the Table to the right.

Table 1. Within-subject change in WTP for meals

	Change in WTP compared to baseline	
	(1)	(2)
High emission meal × Shown label	-0.31^{***}	
	(0.05)	
Low emission meal × Shown label	0.14^{***}	
	(0.04)	
High emission meal	0.01	
	(0.02)	
Low emission meal	-0.06^{*}	
	(0.03)	
Emissions(kg) × Shown label		-0.12^{***}
		(0.03)
Emissions(kg)		0.02
		(0.01)
Shown label		-0.08^{**}
		(0.03)
Participants control	139	139
Participants treated	217	217
Observations	1,704	1,704

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors are clustered at the individual level. Both columns additionally include a control for third-round decisions, and Col. (2) includes a constant term.

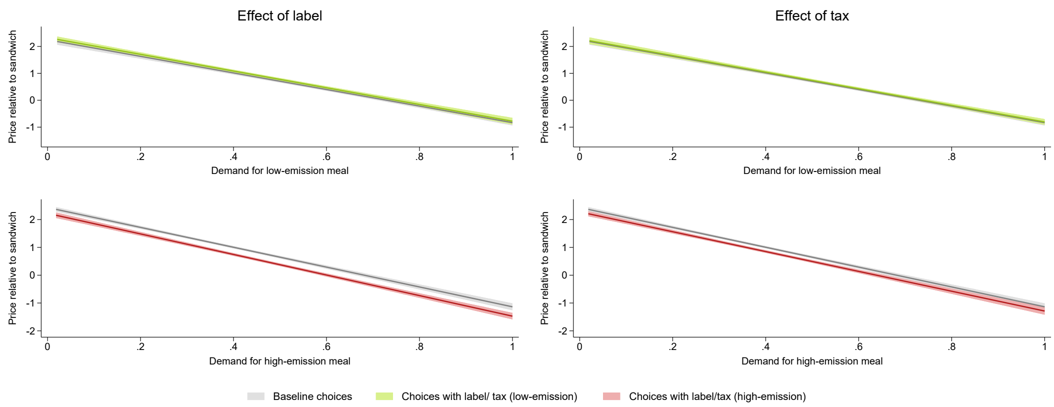

Figure 6. Demand curve shifts with labels vs. a carbon tax

Notes: Demand curves for low-emission meals (top, N=265) and high-emission meals (bottom, N=603) estimated using data from participants in the LABEL condition. Gray lines show baseline WTP, while green and red lines show WTP with carbon labels (left) and net WTP after deducting a carbon tax of €120 per ton (right). Shaded areas represent 95% confidence intervals using robust standard errors.

3 Benchmarking method 2: Natural field experiment

This section presents a second, independent approach to benchmarking the effectiveness of carbon labels against a carbon tax. This approach serves as a validity check on the previous benchmark and adds credibility by drawing on longer-term behavior observed in a field context. In subsection 3.1, I use a field experiment in student canteens to analyze how the introduction of carbon labels affects purchases of high-emission meals. In subsection 3.2, I use field data from the same setting to examine how variations in the canteens' pricing policy affect purchases of high-emission meals. By comparing the magnitudes of these two effects, I derive a second benchmark of the label's effectiveness relative to a carbon tax.

3.1 The effect of labels

Experimental design. To evaluate the effectiveness of carbon labels in a natural field setting, I partnered with Bonn's three student canteens. Key features of the experimental design are summarized below; further details are provided in online Appendix E. The experiment and primary outcomes were pre-registered.[27]

(1) The experiment follows a difference-in-differences design with three phases: (i) pre-intervention (4 weeks, no labels), (ii) intervention (7 weeks, labels added in the treatment canteen but not in the two control canteens), and (iii) post-intervention (3 weeks, no labels).

(2) Carbon labels resemble those in Experiment 1, combining quantitative and ordinal information. In the treatment canteen, labels are shown on the online menu, digital billboards, and paper signage at meal counters (examples in Figure7).[28]

(3) Labels are applied only to the main meals, of which usually two are offered daily, one of these vegetarian and one meat-based. Labels are not applied to sides or desserts, for reasons of clarity and feasibility (details in online Appendix E). The main meals account for roughly 70% of total lunchtime emissions. During the sample period, the meat option always had higher or roughly identical emissions compared to the vegetarian option.

(4) The main analysis is based on canteen sales data. Because 69% of purchases are made with personalized payment cards, I can track individuals' purchasing behavior over time.

27. See AsPredicted#95108. Pre-registered outcomes include meat/vegetarian consumption, green-labeled meal uptake, greenhouse gas emissions, and canteen visits during and after the intervention. Full analyses are reported in online Appendix B.2.

28. Emissions are based on canteen recipes and the Eaternity Institute (2020) database, consistent with Experiment1.

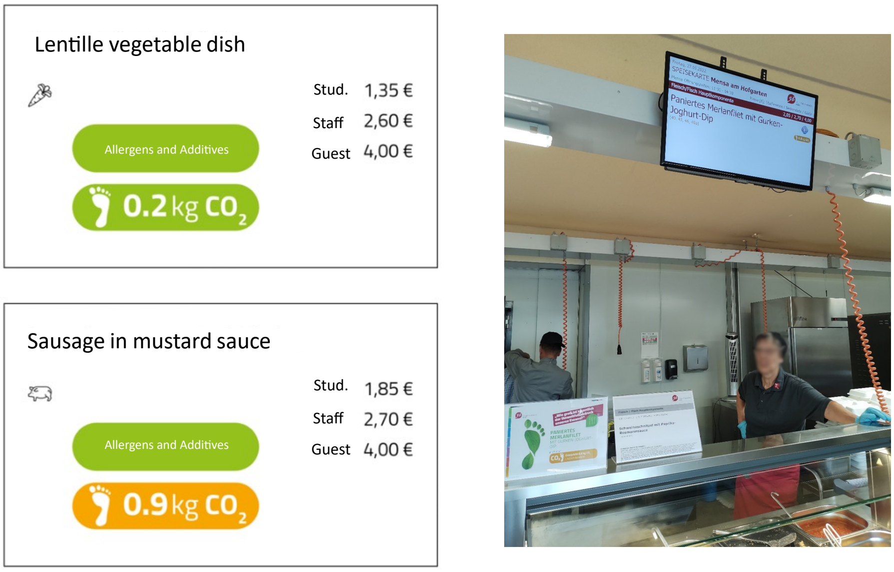

Figure 7. Labels in the canteen

Notes: Labels online (left, menu translated from German) and in the student canteen (right)

(5) Pre- and post-intervention surveys (pre: N>1,700, post: N>900) additionally capture demographic data and opinions. Survey responses are linkable to the purchasing data. See online Appendix E.8 for details.

Estimation strategy. To estimate the causal effect of carbon labels on meal choice in the student canteen, I use a difference-in-differences (DiD) approach:

$$Y_{it} = \alpha + \delta_1 (Treat_{it} \times LabelPeriod_t) + \delta_2 (Treat_{it} \times PostPeriod_t) + \gamma Treat_{it} + \tau + X_{it} + \epsilon_{it} \quad (3)$$

where Y_{it} is the outcome variable varying by specification: the carbon emissions of the meal chosen by individual i on day t; whether i chose a green-, yellow-, or red-labeled option; and whether i chose the higher-emission (i.e. the meat-based) option offered on a given day. Notably, the emissions and label color of the chosen meal are only somewhat under the individual's control, since the emissions of the meals on offer and the offer of green, yellow or red labeled meals differs across days. Choice of the higher-emission (i.e. meat-based) option is, in contrast, fully under the individual's control since there is a meat-based and a lower-emission vegetarian option offered every day. The main text thus focuses on choice of the meat-based option as the main outcome, with Y_{it} taking on a value of 1 if the meat-based option, and 0 if the vegetarian option is chosen. Results using the other outcome variables are consistent with these results, and shown in online Appendix B.2.

$LabelPeriod_t$ is an indicator for the seven-week intervention period, $PostPeriod_t$ captures the post-intervention period, and $Treat_{it}$ is an indicator for purchases made in the treatment canteen. Accordingly, the interaction term ($Treat_{it} \times LabelPeriod_t$) identifies the DiD estimate of the treatment effect during the labeling period, while ($Treat_{it} \times PostPeriod_t$) captures post-intervention effects. τ describes time controls, included in most specifications as week and day-of-the-week controls, and X_{it} describes further control variables differing by specification. Most specifications additionally include binary day- and canteen- specific controls for the two main meals on offer, and a control for whether the purchase is made in the smaller or larger control canteen. Guest-level intent-to-treat specifications additionally allow for the inclusion of guest fixed effects.

Data and results. I use transaction data from April 4 to July 8, 2022, after which the treatment canteen closed for an irregular two-week break, followed by temporary closures of the control canteens. The three canteens usually centralize their meal planning and offer the same two main meals on a given day. I exclude purchases by Ukrainian refugees who received free meals in the canteens starting in week 9. In the main analysis, I also exclude week 5, which coincided with a Healthy Campus" event in all canteens,[29] and the last week of the observation period to avoid effects caused by the irregular closure. Results are robust to alternative exclusion criteria (see Appendix E.5). The final sample contains 124,830 transactions from over 2,000 guests. For each transaction, I observe the meal, price, location, time, and whether the guest is a student (80%) or employee (17%).

Approximately 69% of transactions are linked to personalized payment cards, enabling individual-level tracking. For the intent-to-treat (ITT) analyses, I restrict the sample to these transactions, and additionally to guests who (i) visited any canteen at least five times during the pre-intervention period and (ii) made at least 80% of their visits to the same canteen. This classification is based solely on pre-intervention behavior and is robust to alternative thresholds (Appendix E.6).

Table 2 reports regression results. To avoid results being influenced by variations in the canteens' offering, I focus on choice of the higher-emission meal on offer (i.e. the meat meal) as my main outcome. Results using greenhouse gas emissions and label color of the chosen meal as outcome variables are consistent with these results, and shown in Col. 6 of Table 2 and online Appendix B.2.[30]

29. The Healthy Campus" week featured an additional, rarely chosen ($\approx$ 10% of sales) premium-priced meal and light promotional activities (e.g., nutrition quizzes). Since the event week affected all canteens, it should not confound the treatment effect unless its impact differed across locations. To be conservative, I omit this week.

30. The reasoning behind this choice is that there is a meat and a vegetarian meal on offer every day, and the choice of meat or veg. meal is thus always available to guests. In contrast, there is not a green-labeled, yellow-labeled, or red-

Column (1) estimates the most basic DiD model from equation 3, while Col. (2) additionally includes week and day-of-week fixed effects to flexibly control for time trends (e.g., seasonal or semester-specific effects), and includes 39 binary controls for the main meat and vegetarian meal on offer on a given day in a given canteen. Since meal planning is centralized across the three canteens, these binary controls take on the same values across the three canteens on a given day for $\approx$ 90% of observations.[31] On occasion, a canteen offers a second vegetarian and/ or a second meat main meal.[32] Column (2) thus also controls for whether an additional vegetarian or meat option is offered on the day and in the canteen in which the purchase is made.

Both Col. (1) and (2) estimate that the labels decreased guests' likelihood of purchasing the higher-emission meat meal by 2 percentage points, a 5% decline relative to baseline meat consumption in the treatment group. This decrease in meat consumption seems to persist even in the two weeks after labels were removed. Table B.3 shows that results are similar when controls are gradually added.

Since the three canteens are located over 1.7 km apart, only a minority of guests regularly visit more than one, and I do not find any evidence of treatment-induced sorting (see online Appendix E.3).[33] To nevertheless ensure that results are not driven by different-minded guests sorting into Control and Treatment canteens, I restrict the analysis to the ITT sample in Columns (3) to (5), and assign variable values based on guest-level ITT. Estimation results are consistent with the previous specifications. To additionally rule out that results are driven by different-minded guests changing the frequency with which they visit the canteens, I include guest fixed effects in Columns (4) and (5). Col. (5) additionally includes date fixed effects, which absorb common time trends in a more fine-grained manner than the weekly controls. Again, results are consistent with previous results.

Figure 8 shows similar pre-trends across groups, and relatively stable treatment effects over time. Further, I estimate negative coefficients for the two weeks following the intervention. These post-intervention effects do not persist into the following semester (see Figure B.2 in the Appendix). A short-term persistence of effects might reflect a temporarily increased salience of carbon emissions

labeled meal on offer every day, and the emissions of the meals offer strongly vary across days and weeks. Thus, due to mechanical reasons, students may not be able to choose a green-labeled or very low emission meal even if they wanted to. I show greenhouse gas emissions as alternative outcomes in online Appendix B.2.

31. For the remaining observations, there are small deviations in menu planning, e.g. the smaller control canteen deviates in one of the mains it offers. The binary meal controls take these deviations into account.

32. Specifically, 18% of observations pertain to canteen-day instances where a second vegetarian main was offered, and 7% to instances where a second meat main was offered.

33. Results are similar when dropping instances of canteen guests frequenting a different canteen than usual. This decreases the number of observations by 1%. Results are shown in Table B.8 in the online Appendix.

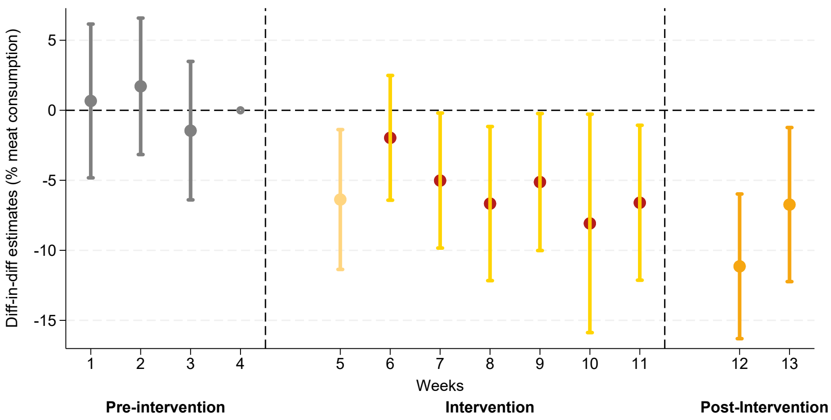

Figure 8. Event study: Difference-in-difference estimates

Notes: Difference-in-difference estimates of the likelihood of consuming the meat option (percentage points), using week 4 of the pre-intervention phase as the baseline. Weeks 1–4 represent the pre-intervention phase, weeks 5–11 the intervention phase, and weeks 12–14 the post-intervention phase. The regression specification follows specification (3) in Table 2 but estimates weekly effects. Controls are assigned according to ITT classification. Table B.1 and B.2 show estimated coefficients. Weeks 5 is excluded from the main estimation in Table 2, as effects cannot be clearly attributed to carbon labels (see main text and Appendix E for details). Standard errors are clustered at the guest level, and bars represent 95% confidence intervals.

in guests' consumption decisions (aligning with the findings of Byrne et al., 2024) or a change in consumption patterns due to learning about vegetarian meals (aligning with Klatt and Schulze-Tilling, 2024).

In the pre-intervention period, the average emissions of meat meals in the treatment canteen were 1.9 kg, compared to 0.4 kg for vegetarian meals. A back-of-the-envelope calculation (1.5×0.02) suggests that, absent any changes in menu composition, the intervention would have reduced emissions by 30 grams per meal, or 3% of baseline emissions (1 kg). Col. (6) in Table 2 analyzes the change in greenhouse gas emissions directly, using a Poisson Pseudo Maximum Likelihood specification, and estimates a 3% decrease in emissions caused by the labels. Online Appendix B.2 provides alternative specifications and evidence that consumption of green-labeled meals increased, while consumption of yellow-labeled meat meals and red-labeled meals decreased due to the intervention.

3.2 The effect of a carbon tax

Setting. To benchmark the effect of carbon labels against a potential carbon tax in the natural field setting, I examine how canteen guests respond to price variations and compare these responses to the labeling effects estimated in the previous subsection.

Table 2. Field estimates of the effect of carbon labels on meat consumption

	Meat choice (pp.)					% Δ GHGE
	Full	Full	ITT	ITT	ITT	Full
Treatment × Label period	-0.02***	-0.02***	-0.03***	-0.02**	-0.02**	-2.77*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(1.42)
Treatment × Post period	-0.05***	-0.04***	-0.04***	-0.04***	-0.04**	-5.14***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(1.36)
Treatment	-0.10***	-0.07***				-15.29***
	(0.01)	(0.01)				(1.39)
Label period	0.01**					1.48**
	(0.00)					(0.73)
Post period	-0.01					2.83***
	(0.00)					(0.77)
Constant	0.51***	0.64***	0.64***	0.65***	0.46***	-46.73***
	(0.00)	(0.03)	(0.04)	(0.04)	(0.01)	(0.46)
Week fixed effects	No	Yes	Yes	Yes	Yes	No
Control for offer	No	Yes	Yes	Yes	Yes	Yes
Guest fixed effects	No	No	No	No	Yes	No
Guests control			1,383	1,383	1,383	
Guests treated			474	474	474	
Observations	124,830	124,830	36,380	36,380	36,380	124,830

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: All specifications are estimated as linear probability models (LPMs), except Column (6), which uses a Poisson Pseudo-Maximum-Likelihood (PPML) estimator. The outcome is meat likelihood of meat consumption (Col. 1-5) and GHGE of the meal consumed (Col. 6). Specifications (1), (2), and (6) assign treatment and control status at the canteen level, whereas Specifications (3)–(5) assign treatment and control status at the individual (guest) level (ITT). Correspondingly, (1), (2), and (6) make use of the full final data set, while spec. (3)–(5) use the ITT data set. Data set construction is described in the main text. Standard errors in Columns (1), (2), and (6) are heteroskedasticity-robust, while those in Columns (3)–(5) are clustered at the guest level. Columns (2)–(5) include week fixed effects, day-of-the-week fixed effects, binary controls for the specific main vegetarian and meat meal option on offer, and controls for whether a second vegetarian or meat option is available. Columns (4)–(5) further include guest fixed effects, and Column (5) includes date fixed effects. Column (6) analyzes the effect of the treatment on the average greenhouse gas emissions (GHGE) of the meals consumed, using a Poisson Pseudo-Maximum-Likelihood (PPML) specification that additionally controls for the emissions of the meals on offer on each day. Column (6) reports coefficients and standard errors transformed to represent percentage changes in the outcome, using the delta-method transformation $100(\exp(\beta) - 1)$. Appendix B.2 presents details and robustness checks.

If the student canteens were to implement a carbon tax, the most straightforward way to do so would be to increase the price difference between meat and vegetarian options without raising the overall price level.[34] Depending on the design of the tax, one could either apply a uniform price increase to all meat meals or vary the magnitude of the price adjustment by meat type.

To approximate the effect of a uniform or differentiated carbon tax, I exploit naturally occurring variation in the daily price difference between the main meat and the main vegetarian option.[35] Price differences arise because individual dishes have different base prices.[36] When a given meat dish is paired with different vegetarian dishes, the resulting price difference between the two options changes. A price adjustment in October 2022 further increases this variation. The adjustment was not a uniform markup: prices increased by about €0.50 on average for meat meals and €0.35 for vegetarian meals, with the exact change differing by dish. As the canteen is a non-profit organization, prices are set to cover ingredient costs and a share of operating costs (Studierendenwerk Bonn, 2025b). Thus, price differences primarily reflect cost variation across dishes rather than strategic pricing or demand considerations. Additional robustness checks and alternative specifications are reported in Online Appendix B.4.

Estimation strategy. I first estimate the effect of a general meat price increase by regressing the probability that a guest purchases the meat rather than the vegetarian option on the price difference between the two options:

$$M_{it} = \alpha + \beta_1 \Delta Price_{it} + X_{it} + \tau + \epsilon_{it}, \tag{4}$$

where M_{it} equals 1 if purchase i on day t is of the meat option and 0 otherwise. $\Delta Price_{it}$ is the price difference between the meat and vegetarian option in the canteen and on the day purchase it is made, and is the main variable of interest. Because this difference may correlate with unobserved attributes influencing meal popularity (e.g., meals with higher-quality ingredients are more costly and might also be more attractive), I include a rich set of controls X_{it} to account for menu characteristics. Specifically, I include 55 binary indicators for the specific meat meal offered in the canteen and on the day on which purchase it is made, and three indicators capturing the type of vegetarian

34. Meals containing meat offered in the canteens emit on average 1.5 kg CO_2e, compared to 0.4 kg for vegetarian meals (over the period April–July 2022). While one could in principle imagine pairings of a low-emission meat meal and a high-emission vegetarian meal, such combinations do not occur within the canteens' actual range of offerings. Across all menus, meat dishes consistently have higher or comparable emissions to the vegetarian alternative. Moreover, since the canteens are non-profit institutions mandated to provide affordable meals to students (see Section E for details on the institutional setup), raising the general price level would be inconsistent with their mission. The most practical way for them to introduce a carbon tax would therefore be to apply a linear tax based on emissions while simultaneously reducing the price of vegetarian meals, thereby generating a relative price increase for meat meals.

35. Each day, one main meat and one main vegetarian option are offered. See Online Appendix E for further details.

36. Meat prices range from €1.85 to €2.50, vegetarian options from €1.35 to €2.40.

dish (fried/breaded, oven-baked, and curry/stir-fry).[37] I also control for the labeling intervention in the treatment canteen as well as week and day-of-the-week effects.

In a second specification, I interact $\Delta Price_{it}$ with the type of meat offered (chicken, pork, beef, or fish) to estimate type-specific price responses. This allows me to approximate the effect of a differentiated carbon tax that varies by meat type in proportion to emissions.

Data and results. For this analysis, I use student canteen consumption data from April 2022 to March 2023.[38] I restrict the sample to purchases made by students, as employees and external guests face higher prices. Additionally, I drop 5% of purchases that do not pertain to the standard two meal options offered on a given day (e.g. because they are meals leftover from the previous day) and 0.1% of purchases that pertain to canteen-day combinations for which offer and price of the two main meals on offer are unclear. This yields one main meat and one main vegetarian meal price per canteen and day, resulting in a single relevant price difference driving meat versus vegetarian purchases.

Table 3, Col. (1) directly follows equation 4 to provide the basis for estimating the effect of an overall tax on meat, i.e. a crude carbon tax. I estimate that a €0.10 increase in the meat meal price relative to the vegetarian meal is associated with a 0.9 percentage point reduction in guests' likelihood of purchasing the meat rather than the vegetarian option. Col. (2) estimates meat-type specific price responses. Canteen guests show the strongest reaction for pork meals, a moderate reaction for beef and chicken, and no reaction for fish meals.[39] Results are similar when restricting the sample to the period before the price change, and I do not find significant evidence of price reactions differing by canteen or by general price level (Table B.10). Effects are less pronounced when not controlling for time trends or meal attractiveness (Table B.9).

To provide context to these estimates and compare them with previous literature, I calculate price elasticities. I estimate an own-price elasticity of -0.4 for meat meals in general, and an own-price elasticity of -0.3 for chicken, -0.9 for pork, -0.4 for beef and 0 for fish meals in particular.

37. Among the 64 vegetarian meals offered during the observation period, 19 are classified as fried/breaded, 9 as oven-baked, and 36 as curry/stir-fry.

38. This larger data set combines two data sets provided by the student canteen: one covering April to July 2022 (the main data set for this paper) and another covering August 2022 to July 2023 (see Klatt and Schulze-Tilling, 2024). Data from April to July 2023 is excluded due to another intervention during this period (analyzed in Klatt and Schulze-Tilling, 2024). Unfortunately, individual guests cannot be linked across the two data sets, and emissions data is only available for meals served from April to July 2022.

39. Fish is almost exclusively served on Fridays, possibly making it a routine (and thus demand-inelastic) choice for some students. Effects may also be influenced by pescetarians or fish's healthy image, reducing price sensitivity.

Overall elasticity levels are similar to Roosen, Staudigel, and Rahbauer (2022) estimates for low-income German households using household-reported consumption. While Roosen, Staudigel, and Rahbauer (2022) do not observe such a stark difference between elasticity for pork and for other meat, Liu and Ansink (2024) observe a similarly start contrast using Dutch supermarket scanner data, which is more similar in nature to my canteen purchasing data.[40]

In the framed field experiment in section 2, the effect of carbon labels is similar to that of a carbon tax of €120 per tonne. To understand whether my effect estimates in the natural field experiment (section 3.1) yield a similar equivalence, I use my regression estimates to approximate the effect of a €120 per tonne carbon tax in the student canteen. First, assuming the carbon tax is implemented as a crude general meat tax, the average emissions difference between meat and vegetarian meals (1.1 kg) translates to a price increase of €0.12 (€0.11 × 1.2kg). Using the Col. (1) estimate, this implies a 1.1 percentage point decrease in meat consumption (€0.12 × −0.09). Second, assuming a more sophisticated tax differentiating by meat type, the average difference in emissions for each meat meal type translates into a price increase of €0.08 for chicken, €0.13 for pork, €0.38 for beef, and €0.04 for fish.[41] Using the Col. (2) estimates to approximate the demand reaction for each meat type, and then weighing each effect by the proportion of days on which the relevant meat type is offered[42], I estimate an overall decrease in meat consumption of 1.6 percentage points in reaction to the tax.

My estimates thus imply that a carbon tax of €120 per tonne would translate to a decrease in meat consumption of approx. 2 percentage points, and is thus similar to my field estimate of carbon label effectiveness (also ≈2 pp., Table 2). I thus also conclude in the natural field setting that the effect of carbon labels is similar to that of a carbon tax of €120 per tonne, corroborating my result from section 2.

40. See online Appendix B.4 for a more detailed discussion.

41. Based on an average emission difference of 0.7 kg between a chicken and a veg. meal, 1.1 for pork, 3.2 for beef, and 0.3 for fish. Also see Table B.12 in online Appendix B.4.

42. Chicken is offered on 31% of days, pork on 37%, beef on 13%, and fish on 19%. Also see Table B.12 in online Appendix B.4.

Table 3. Comparison of effects: labels vs. "carbon tax"

	Likelihood of consuming meat	
	(1)	(2)
	Grouping all meat	By meat type
Price difference (in €)	-0.09***	
	(0.01)	
Price difference (in €) x Chicken		-0.08***
		(0.02)
Price difference (in €) x Pork		-0.20***
		(0.02)
Price difference (in €) x Beef		-0.09**
		(0.04)
Price difference (in €) x Fish		0.02
		(0.02)
Treatment restaurant x Label period	-0.03***	-0.03***
	(0.01)	(0.01)
Treatment restaurant x Post period	-0.06***	-0.06***
	(0.01)	(0.01)
Treatment restaurant	-0.03***	-0.03***
	(0.00)	(0.00)
2nd control restaurant	-0.00	0.00
	(0.00)	(0.00)
Weekly time controls	Yes	Yes
Control for exact meat meal	Yes	Yes
Control for veg. meal type	Yes	Yes
Observations	362,686	362,686

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Dependent variable: 0/1 indicator for consumption of the meat option. Linear probability regression using student canteen data from April 2022–March 2023. The variable "Price difference" describes the price difference between the meat and vegetarian main meal. Both columns include binary controls for week, day-of-the-week, and the specific meat meal offered, as well as two binary variables for the vegetarian meal type. See Table B.9 in the online Appendix for an estimation without controls. Standard errors are robust.

4 Experiment 3: Behavioral channels

Having established that carbon labels can be similarly effective as a substantive carbon tax, I next look into what drives this effectiveness: Why do consumers react to carbon labels? Experiment 3 explores two behavioral channels: (1) labels inform consumers about emissions, correcting misperceptions; and (2) labels increase attention toward emissions at the moment of choice. Subsection 4.1 describes the experimental design, subsection 4.2 the experimental data, and subsection 4.3 the estimation strategy and results. The reduced-form results shown in this section support my modeling choices in section 6, and the experimental data described here is used to structurally estimate the model.

4.1 Experimental design

Overview. Experiment 3 investigates two mechanisms potentially driving the effectiveness of carbon labels: (1) labels correct misperceptions by informing consumers about the emissions caused by different items, and (2) labels direct attention toward emissions during the decision-making process.[43] To analyze these channels, I conduct a framed field experiment similar to Experiment 1 with two key differences:

(1) To assess the role of correcting misperceptions, I additionally track participants' initial estimates of meals' carbon footprints. If label effectiveness was driven by a correction of misperceptions, effects should be differently pronounced depending on participants' initial misperceptions. I thus analyze how participants' reaction to the carbon labels correlates with their initial misperceptions.

(2) To evaluate the role of attention, I include an experimental condition that increases attention toward carbon emissions without providing any emissions information. If label effectiveness was driven by an increase in the salience of carbon emissions in the decision process, we should observe similar effects for an intervention that is similar to the labeling intervention in increasing salience, but does not provide information. I analyze effects for such an intervention and compare these to carbon labeling effects.

43. See section 1 for literature supporting the role of either of these two mechanisms. Note that consumption decisions in Experiments 1 and 3 are made in private, making social pressure as a main mechanism unlikely. In Experiment 2, consumption decisions are made publicly, but the labels' effects are similar to those in Experiments 1 and 3, indicating that the addition of social pressure did not substantially increase the labels' effect on behavior.

Experiment timeline. The experiment timeline is visualized in Figure 9 and follows a structure similar to Experiment 1, with one key distinction. After the baseline WTP elicitation as in Experiment 1, participants in Experiment 3 estimate the carbon footprints of various meals, including the meals on which they make purchasing decisions and six additional meals (see Figure 11 for a list). Participants receive emissions information for a single reference meal (Red Thai Curry with pork and rice, emitting 1.7 kg of CO_2) to guide their guesses. An example guessing screen is displayed in Figure 10. These guessing tasks are incentivized and timed.[44]

The experiment then varies by treatment group, assigned via computer randomization. All participants again indicate their WTP for the four meals, but decision framing differs by treatment condition:

- Attention: WTP elicitation is as in the baseline elicitation. The only difference is that participants, after completing the guessing questions, should be more attentive to the carbon emissions caused by the different meals. WTP elicitation mirrors the baseline; however, prior carbon footprint guessing may enhance attention toward emissions during choices.
- Attention+Label: Participants view carbon labels while indicating their WTP. They should be both more attentive and informed of emissions (see Figure 4).
- Attention+Offset: Participants are informed that meal emissions will be offset, rendering choices carbon-neutral.[45]

To increase statistical power and gather additional insights, WTP for the same meals is elicited a third time,[46] with altered treatment conditions:

- Participants previously in the Attention+Label condition are in the Attention+Offset and vice versa.
- Participants in the Attention condition remain in the Attention condition.

Subsequent procedures, the design of meal purchase decisions and the incentivization of the WTP elicitations are as in Experiment 1.

44. Participants answer each of the ten guessing questions on separate screens, presented in random order. The emissions of the reference meal are consistently displayed. See online Appendix D.5 for screenshots of the guessing instructions. The emission guessing questions take the place of the unrelated guessing questions in Experiment 1, i.e. Experiment 3 participants do not additionally respond to unrelated guessing questions as Experiment 1 participants do. Timing and incentivization of the guessing questions works as in Experiment 1.

45. Labels indicate carbon neutrality through CO_2 offsetting. The Offset condition is detailed in online Appendix D.4 and results are in Table A.8. This condition supports the structural estimation in Section 6 as discussed in Section 6.3.

46. Analyses control for third-round elicitations. Main results are consistent when using data from only the first two rounds.

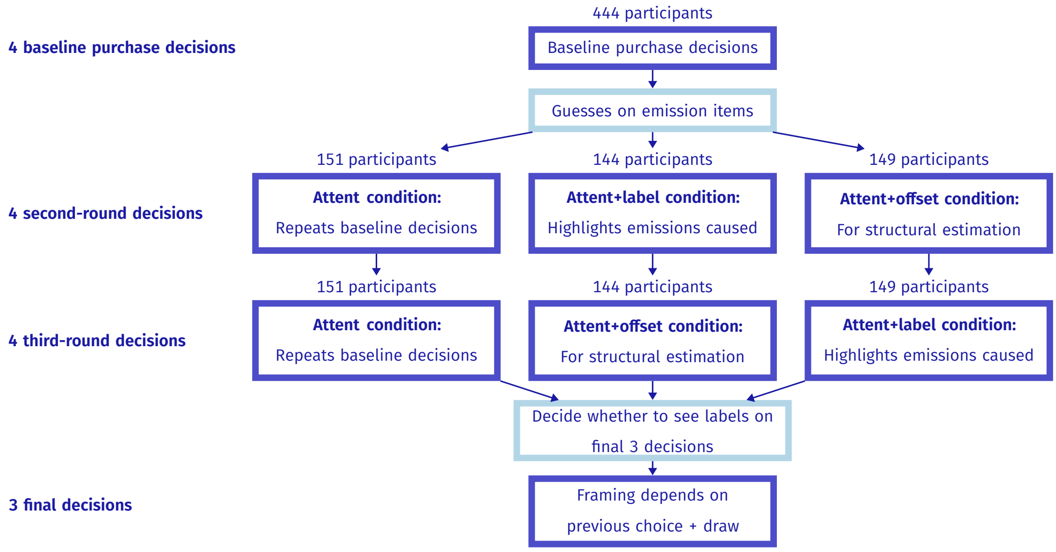

Figure 9. Experiment 3 schedule and treatment groups

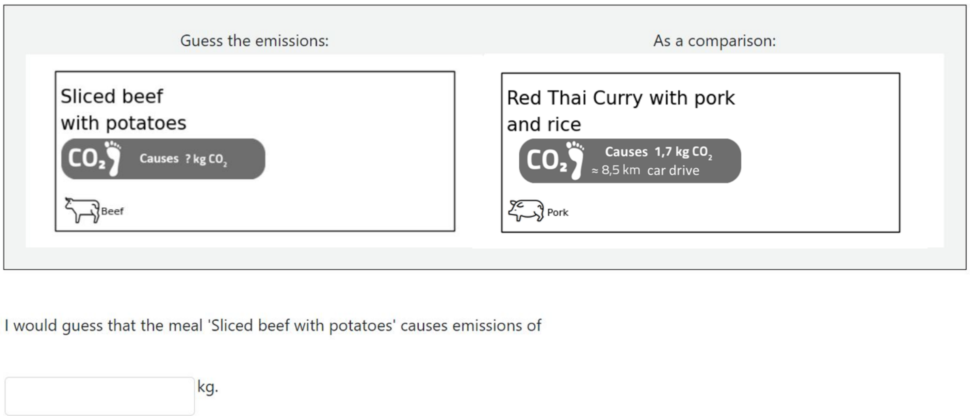

Figure 10. Example guessing questions

Participants and Set-Up. 476 participants from the BonnEconLab pool at the University of Bonn take part in one of 12 experimental sessions between June 22 and July 8, 2021. The experiment design, sample restrictions, and key analyses, i.e. Figure 13 and Table 5, are pre-registered.[47] Participant recruitment and setup are as in Experiment 1.

47. See Schulze Tilling (2021b). All pre-registered main analyses are in Tables A.9 and A.10; Figure 11 was pre-registered as an additional analysis. Figure 12 and Table 4.3 were not pre-registered.

4.2 Data

I exclude the 3% fastest participants and those failing the comprehension check after five attempts, as pre-registered. The remaining 444 participants are computer-randomized into treatments.[48] The sample is on average 26 years old, 55% female, 69% students, and 24% vegetarians. It is roughly representative of regular student canteen guests, as discussed in online Appendix A.2. Results are similar when restricting the sample to only (non-)students or (non-)vegetarians (online Appendix A.8).

4.3 Estimation Strategy and Results

I analyze two main channels through which carbon labels may influence behavior:

(1) Correcting misperceptions about carbon footprints.

(2) Directing attention to emissions at the moment of choice.

Descriptive statistics, estimation strategy and reduced-form results for each of these analyses are shown below.

Correcting Misperceptions as a Treatment Channel. This section provides reduced-form evidence on whether treatment effects align with a correction of misperceptions as the main mechanism driving effects. The analysis draws on participants' carbon footprint guesses for different meals (step 2 in the experiment timeline visualized in Figure 9, example screen in Figure 10). As descriptive evidence, Figure 11 shows that participants tend to underestimate emissions for high-emission meals (blue dots) and overestimate emissions for low-emission meals (black dots).[49]

The misperceptions we see in Figure 11 suggest that there is scope for labels to affect behavior by correcting misperceptions. One empirical pattern that would convincingly demonstrate that such a correction of misperceptions indeed drives the labels' effects would be that we see the labels' treatment effects correlating with participants' previous misperception of emissions. I thus use data from the Attention+Label condition to examine whether participants who underestimated emissions react differently to those labels than those who overestimated emissions:[50]

$$\Delta WTP_{ijm} = \beta_1(Label_{ij} \times Under_{im}) + \beta_2(Label_{ij} \times Over_{im}) + ThirdRound_j + \varepsilon_{ijm} \tag{5}$$

48. See online Appendix A.1 for a randomization check.

49. Further statistics on under- and overestimation appear in online Appendix A.12, including accuracy in meal ranking by carbon footprint.

50. For 55 participant-meal combinations, participants were correct in their estimate. Also in these instances, participants on average adjusted WTP downward, see Table A.31 in the online Appendix.

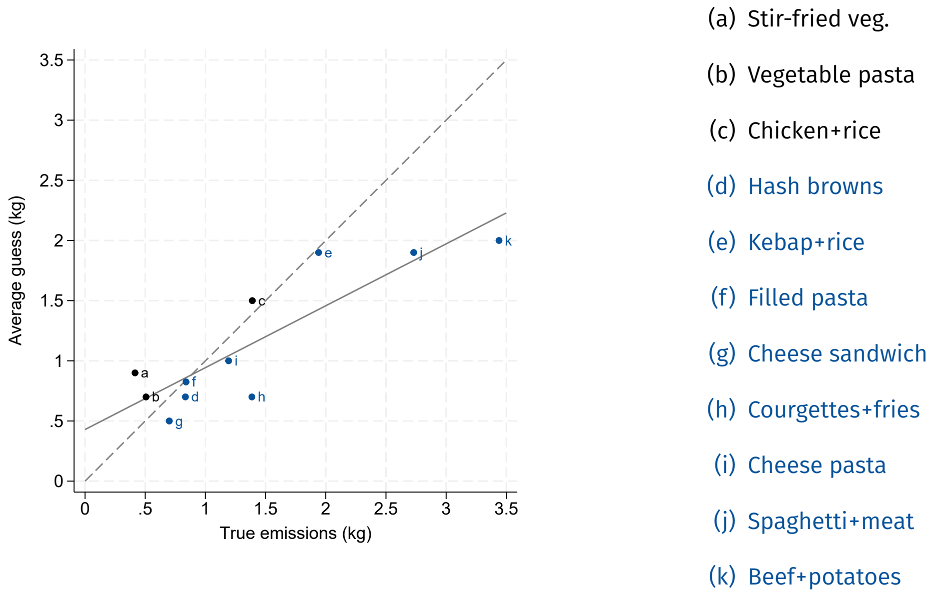

Figure 11. Median guess of the emissions caused by a given meal. Guesses are plotted against calculated emissions. Guesses closer to the dashed line are closer to the correct emission values. Meals corresponding to black scatter points are on average overestimated in their emissions, while meals corresponding to blue scatter points are on average underestimated. Values are based on guesses made by the participants of Experiment 3, and 71 participants in the "Control, then Control" group in Experiment 1. The meal "Spaghetti with meat" was only guessed by the 71 participants of Experiment 1 guessing emissions. Graph based on 5,210 observations from 515 participants.

where ΔWTP_{ijm} is the change in willingness to pay (WTP) for meal m by individual i in round j, relative to their baseline WTP for the same meal, as in the estimation for Experiment 1 (Eq. 2).[51] Thus, the dependent variable directly captures subject- and meal-specific treatment effects.

$Under_{im}$ is an indicator for whether the respective individual underestimated the difference in emissions between meal m and the cheese sandwich. This is determined by comparing the individual's estimated difference in emissions with the true difference.[52] β_1 captures the effect of the labels for participant-meal combinations where the participant previously underestimated the emissions of the meal.

Correspondingly, $Over_{im}$ is an indicator for whether emissions are overestimated, and β_2 captures the effect of the labels for participant-meal combinations in which the participant overestimated emissions.[53]

51. See Section 2.2 and online Appendix A.10 for details on the dependent variable.

52. This refers to the signed difference, not the absolute difference. For example, if a meal generates 0.2 kg more emissions than the cheese sandwich, but a participant estimates it to produce 0.3 kg less, this constitutes an underestimation. Results are similar when conditioning only on participants' estimates of meal emissions and not on the difference to sandwich emissions (Figure A.5 in the online Appendix).

53. In 55 instances (5% of observations) guesses are correct. Also in these instances, participants on average adjust WTP downward, see Table A.31 in the online Appendix.

$ThirdRound_j$ controls for whether round j is the third decision round.[54]

This specification provides reduced-form evidence on whether correcting misperceptions is the main channel driving treatment effects. If it were, β_1 should be negative as participants learn that emissions are higher than they thought when they see the labels, and should decrease their WTP for the meal accordingly. β_2 in turn should be positive, since participants learn that emissions are lower than they thought, and should increase their WTP for the meal accordingly.

Table 4.3, Spec. (1), presents OLS estimates of equation 5. Participants who underestimated meal *m*'s emissions relative to the cheese sandwich decrease their WTP by €0.23 when shown carbon labels. This coefficient is in line with the correction of misperceptions driving treatment effects. However, participants who overestimated meal *m*'s emissions also decrease their WTP for meal *m*, albeit to a significantly lesser extent than in the case of underestimation, by an average of €0.10. This contradicts that a correction of misperceptions is the main channel driving treatment effects.

Spec. (2) estimates $\Delta WTP_{ijm} = \gamma_1(Label_{ij} \times Under(kg)_{im}) + ThirdRound_j + \varepsilon_{ijm}$ where $Under(kg)$ is the participants' underestimation in kg. I estimate that an additional kg of underestimation is correlated with a €0.07 decrease in WTP for the meal. However, the large constant term of €-0.16 is striking. This constant term captures factors driving WTP change that are not related to participants' misperceptions.

Overall, results provide evidence that a correction of misperceptions plays *some* role in driving treatment effects, since we observe that treatment effects are significantly larger if a participant underestimated a meal's emissions. However, the observed pattern suggests that another mechanism—beyond misperception correction—is driving treatment effects.

54. This accounts for mixing data from second- and third-round decisions, as some participants experienced the ATTENT+OFFSET condition in round two and others in round three. Excluding third-round observations yields similar results (Table A.24 in the online Appendix).

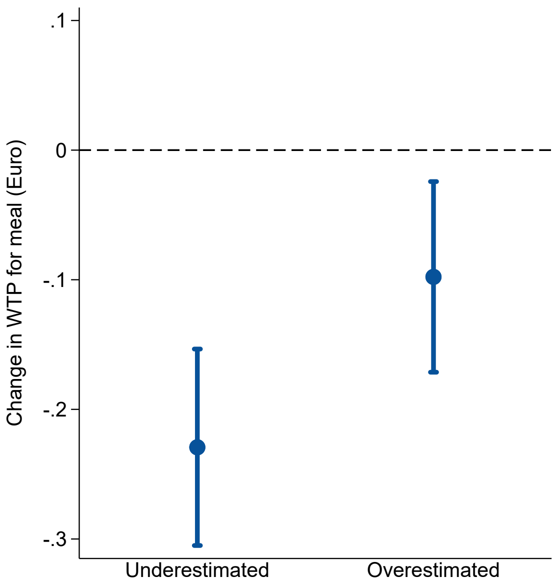

Figure 12. Within-subject change in WTP for meals when shown carbon labels, depending on previous estimation. *Notes:* Bars indicate 95% confidence intervals. Figure visualizes spec. (1) in the Table to the right.

Table 4. Within-subject change in WTP for meals when shown carbon labels, depending on underestimation.

	Change in WTP compared to baseline	
	(1)	(2)
Underestimated emissions × Shown label	-0.23^{***}	
	(0.04)	
Overestimated emissions × Shown label	-0.10^{***}	
	(0.04)	
Underestimation (in kg) × Shown label		-0.07^{***}
		(0.02)
Shown label		-0.16^{***}
		(0.03)
Participants	293	284
Obs. underestimate	555	515
Obs. overestimate	562	493
Observations	1,117	1,063

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Analysis uses data from the ATTENTION+LABEL condition. For Spec. (1), I drop 55 participant-meal combinations where emissions were correctly estimated. Also in these instances, WTP is adjusted downward (Table A.31, Col. 1). The difference between the under- and overestimation coefficient is significant at the 1% level (Table A.31, Col. 2). For each meal in spec. (2), the 10% most extreme guesses of the difference in emissions to the cheese sandwich (in terms of deviation from the true emission difference) are dropped. Both columns additionally include a control for third-round decisions. Standard errors are clustered at the individual level.

Increasing Salience as a Treatment Channel. This subsection examines whether treatment effects can be explained by an increased salience of carbon emissions. To estimate the magnitude of an attention effect, I analyze data from the ATTENTION and ATTENTION+LABEL conditions. Specifically, I estimate:

$$\begin{aligned}\Delta WTP_{ijm} &= \beta_1(Attent_{ij} \times High_m) + \beta_2(Attent_{ij} \times Low_m) \\ &\quad + \delta_1(Attent_{ij} \times Label_{ij} \times High_m) + \delta_2(Attent_{ij} \times Label \times Low_m) + ThirdRound_j + \varepsilon_{ijm} \qquad (6)\end{aligned}$$

where ΔWTP_{ijm} is defined as above. $Attent_{ij} \times High_m$ indicates that participant i in round j was made attentive to the emissions of meal m that has higher emissions than the cheese sandwich, while $Attent_{ij} \times Low_m$ indicates that i was made attentive to the emissions of meal m that has lower emissions than the cheese sandwich. $Label_{ij}$ is an indicator for whether individual i also saw carbon labels in round j. The interaction terms capture the additional effect of labels beyond increasing attentiveness.

Table 5 presents the results, and Figure 13 illustrates average WTP changes in the ATTENTION and ATTENTION+LABEL treatments. Simply directing attention to carbon emissions reduces WTP for high-emission meals by €0.10 on average. Providing labels in addition further decreases WTP by €0.10 for high-emission meals. This effect in the ATTENTION condition is primarily driven by decisions where participants already had a relatively accurate perception of the meal's emissions, as visualized in Figures A.6 and A.7 in the online Appendix. These findings suggest that increased attention alone accounts for a significant share of the treatment effect. Section 6 quantitatively assesses the relative contributions of attention and misperception correction to the label's overall effect.

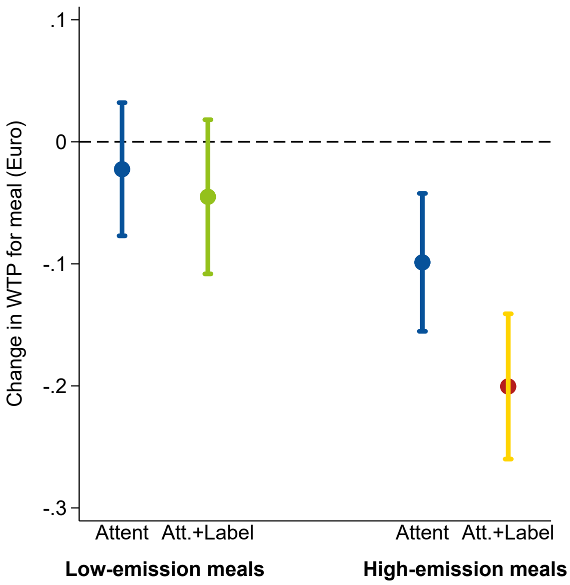

Figure 13. Within-subject change in WTP for meals in the Attention vs. Attention+Label condition.
Notes: Standard errors are clustered at the individual level. Bars indicate 95% confidence intervals. Includes control for third round.

	Change in WTP compared to baseline
	(1)
High emission meal × Attent × Shown label	-0.10^{***}
	(0.04)
Low emission meal × Attent × Shown label	-0.02
	(0.04)
High emission meal × Attent	-0.10^{***}
	(0.03)
Low emission meal × Attent	-0.02
	(0.03)
Participants attent	151
Participants label	293
Observations	2,380

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5. Within-subject change in WTP for meals in the Attention vs. Attention+Label condition

Notes: Analysis uses data from the ATTENTION and ATTENTION+LABEL condition. Both columns additionally include a control for third-round decisions. Standard errors are clustered at the individual level.

5 Consumer Preferences for Carbon Labels

This section presents experimental evidence on consumer preferences for carbon labels in consumption decisions. Section 5.1 examines findings from Experiments 1 and 3, while Section 5.2 discusses results from Experiment 2. The reduced-form results shown in this section motivate how I model the labels' effect on consumer welfare in section 6, and Experiment 3 participants' willingness to pay for the presence or absence of carbon labels, as described in subsection 5.1 is used to structurally estimate the model.

5.1 Evidence from the Framed Field Experiments

In Experiments 1 and 3, participants state their willingness to pay (WTP) for the presence or absence of carbon labels during their final consumption decisions. These elicitations are incentivized using a BDM mechanism.[55] About 50% of participants report a WTP of zero, indicating no strong preference. Fewer than 5% prefer labels to be absent (negative WTP). The remaining participants are willing to pay for labels, with 21% willing to pay €0.50 or more. WTP values are similar across treatment groups, though slightly higher among those who had not previously encountered labels during the experiment.[56] The average WTP for seeing the labels is 20.3 cents. While Thunström (2019) finds that calorie labels impose greater psychological costs on individuals with low self-control, I find no evidence of such a correlation, as shown in Table A.33 in the online Appendix. However, I find a weak positive correlation between preferences for the labels and perceived social norms for reducing carbon emissions in food choices, as well as between preferences for the labels and self-reported willingness to use the information.

5.2 Evidence from the Natural Field Experiment

Following Experiment 2, student canteen guests participated in a follow-up survey assessing their preferences for permanent carbon label installation. As detailed in online Appendix E.8, the survey was framed as an opportunity for guests to provide feedback on various aspects of the canteen, with carbon labels being one of multiple topics. Carbon labels were not mentioned in the survey advertisement. Survey respondents were aware that their responses could impact student canteen policies, incentivizing them to report truthfully.

Among the 284 respondents who visited the treatment canteen at least once during the study period noticed the labels, 75% supported making the labels permanent, 16% were unsure, and only 9% opposed the measure. By comparison, a revenue-neutral carbon tax of unspecified magnitude (operationalized as aligning prices with the label colors) received 62% support, with 14% unsure and 24% opposed. While both measures have majority approval, opposition to the tax is more than twice as high as opposition to the labels. This difference suggests that carbon labels face markedly less resistance, making their implementation more politically feasible.

55. See online Appendix D.3 for details.

56. See Figure A.8 for WTP distribution and Table A.32 for differences across treatments.

6 Structural Model

To formalize how the two behavioral mechanisms shown in section 4 drive consumers' responses to carbon labels, and provide a quantitative estimate of the relative importance of each of the two channels, I introduce a simple discrete choice model of meal selection, which I structurally estimate using data from Experiment 3. I then extend this model to estimate the effect of carbon labels versus a carbon tax on consumer welfare.

In the model, a consumer chooses from a set of meals and selects the meal that maximizes her perceived utility. In general, the perceived utility of a meal may depend on a multitude of meal attributes. The main attribute of interest in this model is the consumers' expectation of the carbon emissions caused by each meal. Ceteris paribus, the consumer has a higher valuation for a meal that causes fewer carbon emissions. How much the consumer cares about emissions depends on two parameters: the salience of carbon emissions at the moment of choice and the guilt the consumer perceives per kg of carbon emitted.[57]

6.1 Basic model

There is a finite set of meals $\mathscr{M}$ and a single consumer. The consumer chooses a meal $m \in \mathscr{M}$ which maximizes her *perceived utility*

$$u(m) = v_m - p_m - \theta \gamma e_m. \tag{7}$$

Here, v_m is the *consumption utility* of meal m that is independent of emissions[58], p_m is the *price* of meal m, and e_m is the consumer's *estimate of emissions* caused by meal m at the moment of choice.[59]

The *salience* of carbon emissions $\theta \in [0, 1]$[60] and the consumer's *environmental guilt per perceived kg of emissions* γ jointly determine how much weight the consumer puts on carbon emissions when deciding.

57. Instead of speaking of guilt, one can also re-formulate the model for the consumer to experience warm glow for every kg of emissions less caused by the chosen option relative to the option highest in emissions. Results would only differ in the interpretation of the parameter γ in the structural estimation.

58. For the purposes of this paper, it is sufficient to consider v_m as being exogenously given for each meal. However, one can also think of v_m being derived from a vector of other observable attributes x_m and an unobservable taste shock ε_m, so that $v_m = \beta^T x_m + \varepsilon_m$.

59. Similar to Imai et al. (2022) I assume in this formulation that consumers' perceived utility is additively separable in v_m and perceived environmental guilt.

60. I hereby use a similar formulation as used in the literature on attentiveness to taxes and resource consumption (Chetty, 2009; DellaVigna, 2009; Byrne et al., 2024). In the framework of Bordalo, Gennaioli, and Shleifer (2022), a straight-forward reason why emissions might not be fully salient to consumers is a lack of prominence, as a meal's emissions are usually not featured at the moment of choice.

The consumer's prior estimate of emissions caused by meal m is denoted by e_m^{prior}, which may differ from the true emissions, denoted by e_m^{true}. If the consumer is *informed*, her updated estimate of emissions is

$$e_m^{\text{info}} \;=\; (1-\kappa)e_m^{\text{true}} + \kappa e_m^{\text{prior}}. \tag{8}$$

Hence, the parameter $\kappa \in [0,1]$ is a measure of the stickiness of the consumers' prior estimate of emissions, e.g. due to a lack of trust in the carbon footprint information provided.[61] If the consumer is *attentive* to emissions, this sets $\theta = 1$.[62] Introducing *carbon labels* makes the consumer both informed and attentive. Introducing a *carbon tax* affects perceived utility via changes in price p_m.

6.2 Extension to consumer welfare

Introducing *carbon labels* makes the consumer both informed and attentive. Her perceived utility then becomes more similar to her *true utility* for meal m,

$$u^{True}(m) \;=\; v_m - p_m - \gamma e_m^{\text{true}} \tag{9}$$

Accordingly, carbon labels increase the likelihood of the consumer choosing the meal m that maximizes her true utility.[63] If the consumer can make a choice $P \in 0,1$ on the presence of carbon labels in her decisions, the *utility change she experiences from the presence of labels* is

$$u(P \;=\; 1) - u(P \;=\; 0) \;=\; u^{True}(m^L) - u^{True}(m^{prior}) + F \tag{10}$$

Here, $u^{True}(m^L)$ is the true utility the consumer would realize from the meal she chooses in the presence of the labels, while $u^{True}(m^{prior})$ is the true utility she would realize from the meal she chooses in the absence of labels. F denotes a *fixed psychological cost or benefit* the consumer experiences as a result of seeing the labels, independent of any behavioral change provoked by the carbon labels.

Carbon taxes affect consumer welfare via changes in meal choices, changes in the true utility of the chosen meal (via changes in p_m), as well as via a lump-sum redistribution which I assume would occur, for reasons of practicality, via a uniform discount on all meals.

61. The above formulation leans on the evidence-informed framework proposed by Epstein, Noor, and Sandroni (2008) to model non-Bayesian updating. Bouchaud et al. (2019) use the same updating rule to study under-reaction in financial markets.

62. This is just a normalization, for any other value $x > 0$ under attention, one could redefine $\theta = \theta/x$ and $\gamma = \gamma x$.

63. The consumers' true valuation of the emissions caused by the meal is not influenced by a lack of salience or misperceptions of the carbon impact. By modeling utility in this manner, I assume that consumers will at some point in their lives find out about the true emissions caused by their consumption decisions, and will experience ex-post regret accordingly (e.g. such as consumers might have experienced ex-post regret about previous decisions to take a plane as the general public became more aware of environmental impact, coining the term "flight shame").

6.3 Identification of parameters

Experiments 1 and 3 represent a special case with binary choice set $\mathscr{M} = m, o$, where m is the meal option and o is the outside option (cheese sandwich). The WTP to exchange meals is given by:

$$u(m) - u(o) \; = \; v_m - v_o - \theta\gamma(e_m - e_o),$$

The treatment conditions yield four equations with four unknowns:[64]

In the absence of any treatment, I assume $\theta \in [0, 1]$.

Baseline (No Treatment):

$$WTP^B \; = \; v_m - v_o - \theta\gamma(e_m^{\text{prior}} - e_o^{\text{prior}}) \tag{11}$$

The treatment condition Attention directs participants' attention towards carbon emissions without providing information. Assuming this sets $\theta = 1$,

Attention Treatment:

$$WTP^A \; = \; v_m - v_o - \gamma(e_m^{\text{prior}} - e_o^{\text{prior}}) \tag{12}$$

Presenting carbon labels directs participants' attention towards carbon emissions, but also provides information on true carbon emissions. I assume this sets $\theta = 1$ and the participant updates as described in equation 8.[65]

Label Treatment (Informed and Attentive):

$$WTP^{A+L} \; = \; v_m - v_o - \gamma\big((1-\kappa)(e_m^{\text{true}} - e_o^{\text{true}}) + \kappa(e_m^{\text{prior}} - e_o^{\text{prior}})\big) \tag{13}$$

In the Attention+Offset treatment, participants are informed that the carbon emissions caused by their choice, regardless of whether they choose the cheese sandwich or warm meal, will be offset. Assuming this sets $\theta = 1$, $e_o = 0$ and $e_m = 0$:[66]

64. $v_m - v_o$ is treated as a single parameter; e_m^{prior}, e_o^{prior} are elicited, and e_m^{true}, e_o^{true} are known.

65. In Experiment 3, participants seeing carbon labels experience the Label treatment on top of the Attention treatment. I assume the Attention+Label, Label and Attention treatment all set salience $\theta = 1$, without any additional attention-directing effect occurring from a combination of attention and labeling treatments. This assumption is in line with a comparison of effect sizes across experiments 1 and 3, where I see similar to larger treatment effects in the Label treatment in Experiment 1 than in the Attention+Label treatment in Experiment 3 (Table A.8 in the online Appendix).

66. Appendix D.4 provides reduced-form evidence indicating that the Attention+Offset treatment substantially reduces environmental guilt, although it may not eliminate it entirely. Appendix A.15.6 repeats the structural estimation assuming that only 70% of guilt is removed. Parameter magnitudes increase mechanically, yet predicted choices and emission reductions remain unchanged, and the relative pattern of consumer-welfare effects across interventions is very similar to the baseline estimates.

Attention+Offset Treatment (No emissions considered):

$$WTP^{A+O} = v_m - v_o \tag{14}$$

Finally, I assume that participants' WTP for the presence of carbon labels reflects their expectation of the true utility they obtain from the choice they make with carbon labels, m^L, versus the choice they make when merely attentive, m^A.[67]

WTP for the presence of labels:

$$WTP^P = \mathbb{E}[u(P = 1) - u(P = 0)] = \mathbb{E}[u^{True}(m^L) - u^{True}(m^A)] + F. \tag{15}$$

Rewriting these equations for structural estimation (Appendix A.15), I estimate parameters via GMM using Experiment 3 data, assuming γ, κ, and θ are homogeneous across participants. To mitigate the influence of outliers, I exclude, for each meal, the 10% of observations corresponding to the most extreme 10% of $e_m^{\text{prior}} - e_o^{\text{prior}}$ values.

6.4 Results

Estimated parameters are similar across the basic and extended model. I estimate θ at 16%, implying that participants behave as if carbon footprints are only 16% of their true size. This estimate is not statistically different from zero, suggesting emissions might not factor into decisions at all absent intervention. κ is estimated at 0.21 and insignificant, indicating that participants indeed take into account the emissions information provided on the labels. γ is estimated at €0.12 per kg CO_2 and significant at the 1% level. F is estimated at 20.5 ¢, while WTP to see labels, averaged across observations in the structural estimation sample, is 21.2 ¢. The proximity of the two values conveys that only a small part of participants' WTP to see the labels can be attributed to the expected improvement in consumption decisions. The remaining amount seems to be a fixed benefit, independent of the labels' impact on consumption decisions, and could be explained by curiosity, or expected additional warm glow from the labels confirming that one is already making an environmentally-friendly choice. Estimates are similar across different specifications (See Table A.35 in the online Appendix).

I simulate how consumers would make choices under different types of canteen interventions, using their stated meal preferences, prior emissions estimates, the estimated behavioral parameters

67. Allcott and Kessler (2019) and Butera et al. (2022) take a similar interpretation of their experiment participants' WTP to experience interventions. I assume that introducing participants to the concept of carbon labels and asking them for their WTP to see them will make them attentive to carbon emissions, so the relevant figure a participant will consider is choices with labels compared to choices when merely attentive.

θ,γ, and κ, and assuming a typical student canteen offer and pricing structure.[68] The interventions I consider are:

(1) Knowledge intervention: Providing emissions information without increasing attention. Note that this is a fictional intervention for the purpose of decomposing the treatment effects of the labels. In practice, it is not possible to provide information without increasing attention.

(2) Attention intervention: Increasing attention without providing emissions information.

(3) Label intervention: Combining both knowledge and attention effects.

(4) Carbon tax: Pricing emissions at €120 per ton with lump-sum redistribution.

(5) Meat ban: Eliminating meat options from the menu.

Table 6 presents the simulation results, with Col. (1) to (3) showing the choices expected under each of the interventions. Col. (4) estimates that the attention intervention reduces greenhouse gas emissions by an average of 27 grams per meal. In contrast, the pure knowledge intervention does not reduce emissions.[69] The label intervention, which increases both attention and knowledge, produces a larger decrease in emissions than the attention intervention alone (34 grams), indicating that improved knowledge can reduce emissions when it is coupled with higher salience. The benchmarked carbon tax produces a similar decrease in emissions (31 grams).

The final four columns of Table 6 estimate how consumer welfare changes under each of the interventions, focusing on changes in consumer welfare brought about by differences in the meals consumed under each of the policies. Carbon taxes additionally affect consumer welfare via a change in prices and receipt of a uniform discount on all meals (functioning as a lump-sum transfer of tax proceedings), taken into account here. The additional choice-independent benefits accruing for carbon labels (F in Eq. 10) are not taken into account here.

I find that carbon labels improve consumer welfare by an average of 0.18 ¢ per meal, or 10 ¢ per affected meal. The carbon tax also increases welfare to a roughly similar extent, while the meat ban decreases welfare. If we additionally take into account the increase in consumer welfare accruing from carbon labels independent of their effect on meal choice (estimated at 20.5¢), the carbon labels produce a much larger increase in welfare than the carbon tax (20.68¢vs. 0.13¢).

68. Details are shown in online Appendix A.15.4.

69. In some cases, updating beliefs moves participants toward the higher-emission option. This occurs when a participant overestimates the emissions of the low-emission meal and underestimates those of the high-emission meal, so that learning the true values increases the attractiveness of the higher-emission meal. Whether such an instance occurs depends on how meals are paired in the simulation, and on correlations between baseline WTP and baseline knowledge for these meals.

Table 6. Simulation of policies based on model estimates

	# of choices			Δ GHGE	Δ consumer welfare			
Intervention	sandwich	veg.	meat	Mean	Mean	SD	Min	Max
None	73.1%	18.1%	8.8%					
Attention	74.4%	18.1%	7.4%	−0.0267	0.0010	0.0160	−0.0849	0.2456
Knowledge	73.1%	17.3%	9.6%	0.0027	−0.0002	0.0068	−0.1910	0.0211
Labels	74.1%	18.6%	7.3%	−0.0337	0.0018	0.0164	−0.0023	0.2456
Carbon tax	72.4%	19.9%	7.7%	−0.0310	0.0013	0.0676	−0.3125	0.2648
Meat ban	78.3%	21.7%		−0.1472	−0.0350	0.1728	−1.3934	0.2456

Notes: Notes: Estimated change in consumption choices, greenhouse gas emissions and consumer welfare which would be caused by different types of interventions. Change in greenhouse gas emissions is in kg per meal and change in consumer welfare is in €per meal. Simulations are based on estimated model parameters, experiment data, and canteen offer and price structure.

As a robustness check, I relax the assumption that the Attention and Label treatments purely increase salience and correct misperceptions without imposing psychological costs. Instead, I assume they double all psychological costs consumers incur due to causing carbon emissions. Under this assumption, the labels create an estimated psychological cost of 5¢ per choice (see online Appendix A.15.5). In this scenario, welfare losses occur even for choices unaffected by the labels, as psychological costs rise regardless of behavior. However, the estimated 20.5¢ choice-independent psychological benefit from labels still outweighs these costs, leaving the net welfare effect positive, though smaller than in the baseline case. For carbon taxes, there is a slight welfare loss under this assumption, driven by a lower estimated environmental guilt parameter γ.[70] With lower perceived guilt per kg of emissions, shifting to lower-emission foods has a less positive effect on consumer welfare than in the baseline model.

7 Discussion

This paper demonstrates how benchmarking behavioral climate interventions against a carbon tax evaluates their effectiveness, by providing a first benchmark estimate of the effectiveness of carbon labels relative to a carbon tax. Across three complementary experiments in student canteens in Bonn, I find that the labels induce demand shifts equivalent to a tax of about €120 per tonne, reduce

70. This is a mechanical result of the assumption that θ is set to 2 in the Attention and Label treatments, see model specification above.

emissions by roughly 4%, and generate net welfare gains for consumers. Most of the behavioral effect is explained by salience rather than misperception correction.

Carbon taxation remains the most effective instrument for reducing emissions, and its implementation is essential for meeting long-run climate targets (European Association of Environmental and Resource Economists, 2019). Yet political resistance makes food taxation difficult to implement, even at modest levels.[71] In this context, carbon labels represent a politically feasible second-best policy. My €120 per tonne benchmark estimate conveys that labels are likely not sufficient as a stand-alone policy to address the climate crisis,[72] but still represent a meaningful reduction within the scope of politically viable interventions and in the context of the well-documented difficulty of shifting food consumption patterns (see, e.g., Guthrie, Mancino, and Lin, 2015).

Does the €120 benchmark extend to other contexts? Although existing studies do not report effects in tax-equivalent terms, their outcomes are consistent with mine. Emission reductions in other student canteens are of similar magnitude (Brunner et al., 2018; Lohmann et al., 2022), and my elasticity estimates align with prior estimates for low-income consumers (Roosen, Staudigel, and Rahbauer, 2022). These parallels suggest that responses to labels and prices in my sample are not idiosyncratic to the Bonn canteen but informative of student canteens more broadly.

Student canteens are a promising setting for carbon label implementation. They serve large populations – 2.9 million students in Germany alone (Federal Statistical Office (Germany), 2023) – and a single menu influences thousands of daily choices (about 2,000 in Bonn). Their non-profit nature in many countries (e.g., Germany, France, Italy, Norway) also makes adoption consistent with institutional missions when supported by students. In Bonn, canteens permanently installed labels in early 2025 by merging emission values from the Eaternity Institute (2020) database with their internal systems, ensuring automatic calculation for all meals (Studierendenwerk Bonn, 2025a, and author's conversations with the canteens).

Evidence on the effectiveness of carbon labels outside canteen contexts is limited. A notable exception is Ho and Page (2023), who test similar labels in the context of Hello Fresh meal kits in the United States and find smaller emission reductions (0.6%). A plausible explanation to the differences we observe between studies lies in the decision settings.[73] The magnitude of an inter-

71. See Douenne and Fabre (2020), Dechezleprêtre et al. (2022), and Fesenfeld (2023) for evidence on the unpopularity of agricultural taxation policies. The agricultural sector is excluded from the EU ETS and German carbon taxation.

72. Estimates for the social cost of carbon range from 50 USD (€49) per tonne and lower (some scenarios in Barrage and Nordhaus, 2024), €160 per tonne (e.g. Rennert et al., 2022), to substantially higher estimates (Bilal and Känzig, 2024; Moore et al., 2024). A carbon tax of €120, or a policy yielding similar effects, thus seems unlikely to suffice as a stand-alone policy.

73. An alternative explanation concerns differences in sample composition. For example, one might expect the larger effects in my setting to result from a higher level of environmental consciousness among students. However, I find no

vention's effect depends not only on how strongly an intervention shifts attitudes, but also on how easily the choice environment allows those attitudes to translate into action. When available options are close substitutes, even modest attitude changes can lead to observable behavioral shifts. Conversely, when options are less substitutable, equally strong attitude changes may not alter behavior. Substitutability may depend on several factors: the number and type of available options, whether the decision-maker is also the consumer or chooses on behalf of others, and the temporal distance between decision and consumption.[74]

Benchmarking treatment effects relative to a carbon tax abstracts from these differences in decision settings and isolates the relative strength of a policy, enabling more reliable cross-study comparisons. In settings where options are close substitutes, both labels and small taxes will yield large reductions; in settings with low substitutability, both will yield smaller reductions. More benchmarked evidence across contexts would allow treatment effect differences to be more clearly attributed to demographic factors or intervention-specific characteristics rather than to decision environments, leaving scope for future research.

My results indicate that roughly 80% of the effect of carbon labels is driven by increased salience alone. I find correcting misperceptions to only be effective in reducing emissions when combined with an increase in salience. These findings clarify the mechanism through which carbon labels operate and help reconcile earlier mixed evidence. Imai et al. (2022), for instance, report null effects when isolating the information-correction channel in a U.S. online experiment, initially interpreting this as evidence that carbon labels are ineffective. My results show instead that labels can meaningfully influence behavior through heightened salience even when they provide no new information, an interpretation later adopted by Imai et al. (2022), who cite my findings as supporting evidence. While other mechanisms may contribute to the labels' effectiveness, they appear not to account for the bulk of the observed effect.[75]

evidence that environmental attitudes are a major source of treatment effect heterogeneity. One likely reason is that environmentally conscious individuals already consume less meat *ex ante*, mechanically limiting their potential for further reductions. Appendix C.1 provides suggestive evidence for this mechanism and for heterogeneity across my experiments.

74. As an illustrative example, consider a consumer in two different contexts. In the first, she is in a canteen choosing between similarly priced and similarly appealing options for her own lunch. In the second, she is in a supermarket shopping for her family's dinner, facing a much larger price difference between vegetarian and meat options. While carbon labels may affect her underlying preferences in both settings to a similar degree, those preferences are more easily translated into behavior in the first setting, where the available options are closer substitutes.

75. A natural alternative explanation is social norms. However, in my framed field experiment, choices were private and unobservable, whereas in the natural field experiment they were observable, yet the estimated effects are of similar magnitude. This makes it unlikely that social pressure is the primary driver (see also Andreoni and Petrie, 2004).

Beyond their behavioral impact, carbon labels generate net consumer welfare gains in my setting, with estimated gains exceeding those of the equivalent tax. High consumer support across the three experiments reinforces this result.[76] Taken together, the results show that behavioral interventions can induce sizable changes in consumption, and that benchmarking them against a tax provides a unified and transparent metric for comparing the behavioral and welfare effects of different climate policy tools.

References

Allcott, Hunt. 2011. "Social norms and energy conservation." *Journal of Public Economics* 95 (9-10): 1082–95. [6]

Allcott, Hunt, Daniel Cohen, William Morrison, and Dmitry Taubinsky. 2025. "When do "nudges" increase welfare?" *American Economic Review* 115 (11): 3448–93. https://doi.org/10.1257/aer.20231304. [6]

Allcott, Hunt, and Judd B. Kessler. 2019. "The welfare effects of nudges: A case study of energy use social comparisons." *American Economic Journal: Applied Economics* 11 (1): 236–76. [6, 38]

Allcott, Hunt, and Dmitry Taubinsky. 2015. "Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market." *American Economic Review* 105 (8): 2501–38. [6]

Andor, Mark A, Lorenz Goette, Michael K Price, Anna Schulze-Tilling, and Lukas Tomberg. 2023. "Differences in how and why Social Comparison and Real-Time Feedback impact resource use: Evidence from a field experiment." Working Paper, Working Paper Series 31845. National Bureau of Economic Research. https://doi.org/10.3386/w31845. [6]

Andreoni, James, and Ragan Petrie. 2004. "Public goods experiments without confidentiality: a glimpse into fundraising." *Journal of Public Economics* 88 (7): 1605–23. https://doi.org/https://doi.org/10.1016/S0047-2727(03)00040-9. [42]

Banerjee, Sanchayan, Matteo M Galizzi, Peter John, and Susana Mourato. 2023. "Sustainable dietary choices improved by reflection before a nudge in an online experiment." *Nature Sustainability* 6 (12): 1632–42. [5]

Barrage, Lint, and William Nordhaus. 2024. "Policies, projections, and the social cost of carbon: Results from the DICE-2023 model." *Proceedings of the National Academy of Sciences* 121 (13): e2312030121. [41]

Bilal, Adrien, and Diego R Känzig. 2024. "The Macroeconomic Impact of Climate Change: Global vs. Local Temperature." Working Paper, Working Paper Series 32450. National Bureau of Economic Research. https://doi.org/10.3386/w32450. [41]

Bilén, David. 2022. "Do carbon labels cause consumers to reduce their emissions? Evidence from a large-scale natural experiment." Working paper. Mimeo, Gothenburg University. [5]

76. While support in my sample may exceed that in the general population, recent EU surveys still report around 80% approval for carbon labels (European Investment Bank (EIB), 2023; Yara International and IPSOS, 2023; see also John, Martin, and Mikołajczak, 2023).

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2022. “Salience.” *Annual Review of Economics* 14 (1): 521–44. [35]

Bouchaud, Jean-Philippe, Philipp Krueger, Augustin Landier, and David Thesmar. 2019. “Sticky expectations and the profitability anomaly.” *Journal of Finance* 74 (2): 639–74. [36]

Brent, Daniel A, Joseph H Cook, and Skylar Olsen. 2015. “Social comparisons, household water use, and participation in utility conservation programs: Evidence from three randomized trials.” *Journal of the Association of Environmental and Resource Economists* 2 (4): 597–627. [6]

Brunner, Florentine, Verena Kurz, David Bryngelsson, and Fredrik Hedenus. 2018. “Carbon label at a university restaurant–label implementation and evaluation.” *Ecological Economics* 146: 658–67. [3–5, 41]

Busse, Meghan R, Nicola Lacetera, Devin G Pope, Jorge Silva-Risso, and Justin R Sydnor. 2013. “Estimating the effect of salience in wholesale and retail car markets.” *American Economic Review* 103 (3): 575–79. [6]

Butera, Luigi, Robert Metcalfe, William Morrison, and Dmitry Taubinsky. 2022. “Measuring the welfare effects of shame and pride.” *American Economic Review* 112 (1): 122–68. [38]

Byrne, David P, Lorenz Goette, Leslie A Martin, Lucy Delahey, Alana Jones, Amy Miles, Samuel Schob, Thorsten Staake, Verena Tiefenbeck, et al. 2024. “How nudges create habits: Theory and evidence from a field experiment.” Working Paper. https://ssrn.com/abstract=3974371. [20, 35]

Camilleri, Adrian R., Richard P. Larrick, Shajuti Hossain, and Dalia Patino-Echeverri. 2019. “Consumers underestimate the emissions associated with food but are aided by labels.” *Nature Climate Change* 9 (1): 53–58. [5]

Chen, Daniel L., Martin Schonger, and Chris Wickens. 2016. “oTree—An open-source platform for laboratory, online, and field experiments.” *Journal of Behavioral and Experimental Finance* 9 (C): 88–97. https://doi.org/10.1016/j.jbef.2015.12.00. [11]

Chetty, Raj. 2009. “Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods.” *Annual Review of Economics* 1 (1): 451–88. [35]

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. “Salience and taxation: Theory and evidence.” *American economic review* 99 (4): 1145–77. [6]

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. [2]

Dechezleprêtre, Antoine, Adrien Fabre, Tobias Kruse, Bluebery Planterose, Ana Sanchez Chico, and Stefanie Stantcheva. 2022. “Fighting climate change: International attitudes toward climate policies.” Working Paper, Working Paper Series 30265. National Bureau of Economic Research. https://doi.org/10.3386/w30265. [2, 41]

DellaVigna, Stefano. 2009. “Psychology and economics: Evidence from the field.” *Journal of Economic Literature* 47 (2): 315–72. [35]

Dohmen, Thomas, and Tomáš Jagelka. 2024. “Accounting for Individual-Specific Reliability of Self-Assessed Measures of Economic Preferences and Personality Traits.” *Journal of Political Economy Microeconomics* 2 (3): 399–462. [12]

Douenne, Thomas, and Adrien Fabre. 2020. “French attitudes on climate change, carbon taxation and other climate policies.” *Ecological Economics* 169: 106496. [2, 41]

Eaternity Institute. 2020. *Eaternity data base.* https://eaternity.org/. Accessed: 2020-08-26. [10, 16, 41]

Epstein, Larry G., Jawwad Noor, and Alvaro Sandroni. 2008. "Non-Bayesian updating: a theoretical framework." *Theoretical Economics* 3 (2): 193–229. [36]

European Association of Environmental and Resource Economists. 2019. *Economists' Statement on Carbon Pricing.* Https://doi.org/10.1257/rct.8435. [2, 41]

European Investment Bank (EIB). 2023. *The EIB Climate Survey: Government action, personal choices and the green transition.* https://www.eib.org/attachments/lucalli/20230098_climate_survey_europe_en.pdf. Accessed: 2025-08-25. [43]

Federal Ministry of Education and Research (Germany). 2023. *The student survey in Germany: 22nd Social Survey - Executive Summary.* https://www.studierendenwerke.de/fileadmin/user_upload/22._Soz_Exec_Summary_EN_barrierefrei.pdf. Accessed: 2023-10-04. [4]

Federal Statistical Office (Germany). 2023. *Hochschulen - Hochschulstatistik.* https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/_inhalt.html. Accessed: 2023-10-04. [4, 41]

Fesenfeld, Lukas. 2023. "The political economy of taxing meat." *Nature food* 4 (3): 209–10. [2, 41]

German Environment Agency (UBA). 2024. *New record revenue in emissions trading: More than 18 billion euros for climate protection.* https://www.umweltbundesamt.de/en/press/pressinformation/new-record-revenue-in-emissions-trading-more-than. Accessed: 2025-08-25. [3]

Grummon, Anna H, Cristina JY Lee, Thomas N Robinson, Eric B Rimm, and Donald Rose. 2023. "Simple dietary substitutions can reduce carbon footprints and improve dietary quality across diverse segments of the US population." *Nature Food* 4 (11): 966–77. [2]

Guthrie, Joanne, Lisa Mancino, and Chung-Tung Jordan Lin. 2015. "Nudging consumers toward better food choices: Policy approaches to changing food consumption behaviors." *Psychology & Marketing* 32 (5): 501–11. [41]

Hahn, Robert W, Nathaniel Hendren, Robert D Metcalfe, and Ben Sprung-Keyser. 2024. "A welfare analysis of policies impacting climate change." Working paper. National Bureau of Economic Research. [6]

Ho, Lisa, and Lucy Page. 2023. "Got beef with beef? Evidence from a large-scale carbon labeling experiment." Working Paper. MIT. https://economics.mit.edu/sites/default/files/inline-files/HF_paper_draft.pdf. [5, 41]

Imai, Taisuke, Davide D Pace, Peter Schwardmann, and Joël J van der Weele. 2022. "Correcting consumer misperceptions about CO 2 emissions." Working Paper, Working Paper Series 10138. CESifo Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4307482. [5, 35, 42]

International Carbon Action Partnership (ICAP). 2025. *EU Emissions Trading System (EU ETS).* https://icapcarbonaction.com/en/ets/eu-emissions-trading-system-eu-ets. Accessed: 2025-08-25. [3]

Jalil, Andrew J, Joshua Tasoff, and Arturo Vargas Bustamante. 2023. "Low-cost climate-change informational intervention reduces meat consumption among students for 3 years." *Nature Food* 4 (3): 218–22. [6]

John, Peter, Aaron Martin, and Gosia Mikołajczak. 2023. "Support for behavioral nudges versus alternative policy instruments and their perceived fairness and efficacy." *Regulation & Governance* 17 (2): 363–71. [43]

Kim, Brent F., Raychel E. Santo, Allysan P. Scatterday, Jillian P. Fry, Colleen M. Synk, Shannon R. Cebron, Mesfin M. Mekonnen, Arjen Y. Hoekstra, Saskia De Pee, Martin W. Bloem, et al. 2020. "Country-specific dietary shifts to mitigate climate and water crises." *Global Environmental Change* 62: 101926. [2]

Klatt, Charlotte, and Anna Schulze-Tilling. 2024. “Tastes Better than Expected: Post-Intervention Effects of a Vegetarian Month in the Student Canteen.” Working Paper, Working Paper Series 315. ECONtribute. https://www.econtribute.de/RePEc/ajk/ajkdps/ECONtribute_315_2024.pdf. [20, 23]

List, John A, Matthias Rodemeier, Sutanuka Roy, and Gregory K Sun. 2023. “Judging nudging: Understanding the welfare effects of nudges versus taxes.” Working Paper, Working Paper Series 31152. National Bureau of Economic Research. https://doi.org/10.3386/w31152. [6]

Liu, Zhaoxin, and Erik Ansink. 2024. “Price elasticities of meat, fish and plant-based meat substitutes: Evidence from store-level Dutch supermarket scanner data.” Working paper. Tinbergen Institute Discussion Paper. [24]

Lohmann, Paul M, Elisabeth Gsottbauer, Anya Doherty, and Andreas Kontoleon. 2022. “Do carbon footprint labels promote climatarian diets? Evidence from a large-scale field experiment.” *Journal of Environmental Economics and Management* 114: 102693. https://doi.org/https://doi.org/10.1016/j.jeem.2022.102693. [3–5, 41]

Lohmann, Paul M, Elisabeth Gsottbauer, James Farrington, Steve Human, and Lucia A Reisch. 2024. “Choice architecture promotes sustainable choices in online food-delivery apps.” *PNAS nexus* 3 (10): pgae422. [5, 6]

Lohmann, Paul M, Alice Pizzo, Jan Michael Bauer, Tarun M Khanna, Max Callaghan, Jan Minx, and Lucia A Reisch. 2024. “Demand-side interventions for sustainable food systems: a meta-analysis of food-policy interventions targeting food consumption and waste behaviours.” *Available at SSRN 4811931.* [5]

Lusk, Jayson L, and Glynn T Tonsor. 2016. “How meat demand elasticities vary with price, income, and product category.” *Applied Economic Perspectives and Policy* 38 (4): 673–711. [6]

Moore, Frances C, Moritz A Drupp, James Rising, Simon Dietz, Ivan Rudik, and Gernot Wagner. 2024. “Synthesis of evidence yields high social cost of carbon due to structural model variation and uncertainties.” Working paper. CESifo. [41]

Poore, Joseph, and Thomas Nemecek. 2018. “Reducing food’s environmental impacts through producers and consumers.” *Science* 360 (6392): 987–92. [2]

Potter, Christina, Anastasios Bastounis, Jamie Hartmann-Boyce, Cristina Stewart, Kerstin Frie, Kate Tudor, Filippo Bianchi, Emma Cartwright, Brian Cook, Mike Rayner, et al. 2021. “The effects of environmental sustainability labels on selection, purchase, and consumption of food and drink products: a systematic review.” *Environment and Behavior* 53 (8): 891–925. [8]

Rennert, Kevin, Frank Errickson, Brian C. Prest, Lisa Rennels, Richard G. Newell, William Pizer, Cora Kingdon, Jordan Wingenroth, Roger Cooke, Bryan Parthum, et al. 2022. “Comprehensive evidence implies a higher social cost of CO2.” *Nature* 610 (7933): 687–92. [41]

Rodemeier, Matthias, and Andreas Löschel. 2022. “Information nudges, subsidies, and crowding out of attention: Field evidence from energy efficiency investments.” Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4213071. [6]

Roosen, Jutta, Matthias Staudigel, and Sebastian Rahbauer. 2022. “Demand elasticities for fresh meat and welfare effects of meat taxes in Germany.” *Food Policy* 106: 102194. [24, 41]

Scarborough, Peter, Michael Clark, Linda Cobiac, Keren Papier, Anika Knuppel, John Lynch, Richard Harrington, Tim Key, and Marco Springmann. 2023. “Vegans, vegetarians, fish-eaters and meat-eaters in the UK show discrepant environmental impacts.” *Nature Food* 4 (7): 565–74. [2]

Schulze Tilling, Anna. 2021a. *Additional data collection to: Quantifying the role of greenhouse gas emissions in consumption choice.* Https://doi.org/10.1257/rct.8435. [12]

Schulze Tilling, Anna. 2021b. *Quantifying the role of greenhouse gas emissions in consumption choice.* Https://doi.org/10.1257/rct.7858. [11, 28]

Shewmake, Sharon, Abigail Okrent, Lanka Thabrew, and Michael Vandenbergh. 2015. "Predicting consumer demand responses to carbon labels." *Ecological Economics* 119: 168–80. [5]

Studierendenwerk Bonn. 2025a. *CO_2-Ampel in den Mensen.* https://www.studierendenwerk-bonn.de/essen-trinken/co2-ampel-in-den-mensen. Accessed on 2025-08-27. [41]

Studierendenwerk Bonn. 2025b. *Essen & Trinken — „Unsere Preise“ (Studierendenwerk Bonn).* https://www.studierendenwerk-bonn.de/essen-trinken#c181. Please click on the "Unsere Preise" tab (site is in German). [22]

Taubinsky, Dmitry, and Alex Rees-Jones. 2018. "Attention variation and welfare: theory and evidence from a tax salience experiment." *Review of Economic Studies* 85 (4): 2462–96. [6]

Taufique, Khan M.R., Kristian S. Nielsen, Thomas Dietz, Rachael Shwom, Paul C. Stern, and Michael P. Vandenbergh. 2022. "d2022randomized." *Nature Climate Change* 12 (2): 132–40. [8]

Thunström, Linda. 2019. "Welfare effects of nudges: The emotional tax of calorie menu labeling." *Judgment and Decision Making* 14 (1): 11. [34]

Tiefenbeck, Verena, Lorenz Goette, Kathrin Degen, Vojkan Tasic, Elgar Fleisch, Rafael Lalive, and Thorsten Staake. 2018. "Overcoming salience bias: How real-time feedback fosters resource conservation." *Management Science* 64 (3): 1458–76. [6]

Wettengel, Julian. 2024. *Germany's carbon pricing system for transport and buildings.* https://www.cleanenergywire.org/factsheets/germanys-planned-carbon-pricing-system-transport-and-buildings. Accessed: 2025-08-25. [3]

Yara International and IPSOS. 2023. *Strong European consumer demand for sustainable food.* https://www.yara.com/corporate-releases/strong-european-consumer-demand-for-sustainable-food/. Accessed: 2025-08-25. [43]