Interventions for sustainable consumption

Experiments in Behavioral Environmental Economics

Inauguraldissertation

zur Erlangung des Grades eines Doktors
der Wirtschaftswissenschaften

durch

die Rechts- und Staatswissenschaftliche Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Anna Schulze Tilling

aus Köln

2024

Dekan:	Prof. Dr. Jürgen von Hagen
Erstreferent:	Prof. Dr. Lorenz Götte
Zweitreferent:	Prof. Dr. Sebastian Kube
Tag der mündlichen Prüfung:	22. May 2024

Acknowledgements

I would like to thank my supervisors Lorenz Götte, Sebastian Kube, and Thomas Dohmen for supporting me throughout my PhD and making it both instructive and enjoyable. Every step of the journey brought with it new tasks and challenges, and it was thanks to you that I experienced these as valuable opportunities to grow. I was blessed with support throughout the whole faculty. I would especially like to thank Holger Gerhardt for TeXnical assistance, valuable feedback on my projects and advice and practical help on my experiments, and Julia Mink, Botond Köszegi, Amelie Schiprowski, Hanna Schwank, Florian Zimmermann, and Peter Andre for their advice and support. I learned so much from my co-authors Mark Andor, Lukas Tomberg, Mike Price, Charlotte Klatt, Elena Shvartsmann, Si Chen and Yana Radeva and it has been a pleasure to work with you! Hanna Fuhrmann-Riebel, thank you for being part of my PhD journey, I am glad that we maintained contact for so long and hope that our regular chats will continue.

I would like to thank Ilona Krupp and Simone Jost for helping me through any administrative struggles, as well as the administration at CRC TR 224, EconTribute, and the BGSE.

My PhD colleagues have been and are my trusted companions, and I am most lucky to be part of such a fantastic group that continuously supports each other. My cohort has been my BGSE family from the start. I am not sure how I would have gotten through the first years without Jakob and Lenard, and I am glad that Andrea, Martin, Dominik, Ege, Miguel, Robin, Christina and Alex shared the adventures of the PhD and the job market with me. My office mates often made my day, thank you to Johannes, Oleksii, Yana, and Chen, and to Alina and Theresa for accompanying me in various sporty activities! I am also especially thankful for the friends I made during my exchange at LSE, especially to Sanchayan, Manuel, Glen and Julien. Luca, Paul, Timo, Chui Yee, Ximeng, and Radost: Thank you for your friendship and advice, I was very lucky to benefit from your experiences and views a couple of steps ahead from where I was. I am especially thankful to Camille for always having an open ear for my research and struggles. The biggest thank you goes to Simon, without whom my PhD experience would surely have been very different. Listing all the ways in which you have supported and continue to support me could alone fill a book.

I am most thankful to my family, who has encouraged and supported me throughout my entire education. The six-year-old you sent to school 25 years ago might have liked it just a bit too much and ended up staying… and staying… and staying.

Anna Schulze Tilling
Bonn, April 2024

Contents

List of Figures

List of Tables

Introduction

The fight against climate change is one of the greatest social challenges of our time. Almost twenty years ago, Nicholas Stern described greenhouse gas emissions as "the biggest market failure the world has seen" (Stern, 2008). Their impact requires a rethinking of production and consumption systems, and many economists have contributed their expertise to evaluating ideal policies (e.g. European Association of Environmental and Resource Economists, 2019), quantifying the (expected) costs of climate change, and identifying areas of urgent action.

A relatively recent development has been the incorporation of insights from Behavioral Economics into these efforts. With it has come the development of behavioral interventions designed to promote sustainable behavior. Such interventions would be ineffective on fully rational and self-interested (Croson and Treich, 2014) individuals with full information and willpower (Mullainathan and Thaler, 2000). However, individuals who do not fulfill these criteria—who perhaps do not have perfect information and willpower or do not behave in a fully rational and self-interested manner—could respond to, and even benefit from, these initiatives. Allcott, Mullainathan, and Taubinsky (2014) refer to this benefit as an "internality dividend": Behavioral tools can help an individual choose the action that is not only better for the environment, but also for the individual themself, as it helps them avoid monetary, psychological, or health costs. In some cases, employing such behavioral interventions to tackle externalities can even be better for overall welfare than employing taxes (List et al., 2023).[1] Insights from Behavioral Economics have more recently also been helpful in understanding variations in public support for more traditional measures to curb carbon emissions (e.g. Andre et al., 2021; Woerner et al., 2023), and in understanding how environmental policies may affect consumers even after they are removed (e.g. Byrne et al., 2024) or beyond their initial policy target (e.g. Goetz, Mayr, and Schubert, 2022).

This dissertation contributes to this evolving strand of literature. The three chapters stand independently, but all aim to employ insights from Behavioral Economics to help provide solutions for environmental problems. The first two chapters evaluate three popular behavioral interventions to decrease climate externalities. The final chapter examines the post-intervention effects of a command-and-control measure to decrease environmental externalities, and draws on the behavioral economics toolkit to provide explanations for consumer behavior.

1. The focus of this dissertation lies on debiasing interventions. However, List et al. (2023) find that these conclusions often also hold when interventions are deceptive rather than debiasing.

One behavioral intervention that has increasingly received attention from academia[2], regulatory agencies[3], and private companies[4] is carbon labeling. Chapter 1 examines carbon labels as a policy tool for reducing the carbon footprint of consumers' meal choices. To date, little is known about the effectiveness of carbon labels relative to other policy instruments and the channels through which they affect behavior. Through a series of experiments, including two framed field experiments ($N = 289$ and $N = 444$, respectively) and a natural field experiment (involving over 120,000 purchase decisions by over 10,000 customers) conducted across three student canteens, this chapter provides causal evidence that carbon labels affect consumption behavior. It evaluates the effectiveness of labels relative to a carbon tax, both through direct elicitation in the framed field experiment and through price variation in the natural field experiment. In both settings, the overall effectiveness of the labels is similar to that of a carbon tax of €120 per tonne. Moreover, complementary evidence from both settings conveys that the labels, on average, create psychological benefits for consumers. In the second framed field experiment, the behavioral channels that drive label effectiveness are identified using variations in the treatment conditions. The findings convey that carbon labels affect consumers primarily by drawing attention to carbon emissions, rather than by correcting consumers' perceptions of carbon footprints. Using a structural model and data from the second framed field experiment, it is estimated that, on average, carbon labels increase consumer welfare.

Behavioral interventions can also be effective tools to foster resource conservation. Chapter 2 examines the behavioral and welfare effects of two such interventions. Specifically, this chapter evaluates behavioral interventions that help consumers reduce the length of their showers, as showering is an energy-intensive and correspondingly emission-intensive activity: A typical shower in the sample analyzed in this chapter uses 1.6 times the daily electricity consumption of a typical household for lighting, or 2.5 times the daily consumption of a refrigerator.[5] The first intervention examined in this chapter is social comparison reporting, which primarily provides consumers with information to motivate behavior change. Specifically, the weekly reports convey how much resources the consumer on average uses per shower and how this amount compares to the consumption of other households. The second intervention is real-time feedback, which primarily provides consumers with information that facilitates behavior change. Specifically, the consumer's shower head glows in different colors when showering, signaling how much resources they had already used during the shower. In a field experiment with about 1,000 participants, the social comparison reports reduce warm water consumption per shower by 9.4%, real-time feedback by 28.8%, and the combination of both interventions by 35.0%. Participants' willingness to pay for real-time feedback and the combination is higher than for social comparison reports. The findings convey that all interventions increase welfare.

While Chapter 1 and Chapter 2 analyze behavioral interventions, Chapter 3 shows that insights from Behavioral Economics can also be useful for a better understanding of how

2. See Reisch et al. (2021) for an overview.

3. For example, the Obama administration issued an executive order on behavioral science, and the European Commission includes carbon labels in its Farm to Fork Strategy (Obama, 2015; European Commission, 2023).

4. For example, Oatly, an oat milk producer, Just Salad, a restaurant chain, Panera Bread, and Allbirds, a shoe brand (Wolfram, 2021) all engage in carbon labeling.

5. A typical individual in the sample uses 33 liters of hot water per shower, which requires an average of 1.6 kWh to heat. In comparison, the average household in the European Union uses 1.0 kWh per day for lighting (Faberi et al., 2015), and a modern refrigerator uses 0.63 kWh per day (Michel, Attali, and Bush, 2016).

non-behavioral interventions impact consumers. It examines the post-intervention effects of a command-and-control intervention to reduce meat consumption. Global meat consumption is a significant contributor to climate change: Poore and Nemecek (2018) estimates that meat and dairy provide only 18% of calories consumed, while producing 60% of agricultural greenhouse gas emissions. Yet interventions to reduce meat consumption are often implemented for short periods of time, and it is unclear how they might have lasting effects. This chapter combines student canteen consumption (over 270, 000 purchases made by over 4, 500 guests) and survey data ($N > 800$) to examine how a one-month intervention to reduce meat consumption affects consumer behavior during the intervention and in a two-month post-intervention period. During the intervention period, meat meals were eliminated from the menu of the treatment canteen, while the two control canteens were unaffected. A difference-in-difference estimation conveys that that guests usually frequenting the treatment canteen did not significantly reduce their visits to the canteen during or after the intervention. In the two months following the intervention, they were still 4% less likely to choose the meat option when visiting the canteen, relative to baseline. The intervention itself is best understood as a command-and-control instrument, and the observation that it decreases meat consumption during the intervention period can be well explained within a rational agent framework. However, the mechanisms behind the observed post-intervention effects are less clear. This chapter thus also draws on evidence on possible behavioral drivers. Results suggest that a large part of the post-intervention effect can be explained by guests learning about the quality of the canteen's vegetarian meals and by habit formation. There is little to no evidence that the intervention changed the perception of social norms.

All three chapters are similar in the methods they employ. To provide evidence based on real-world behavior and applications, all three chapters make use of field experiments that observe subjects' behavior with real-world goods and in real-world contexts.[6] These are supplemented with more controlled elements and surveys to provide evidence on the mechanisms driving effects. To this end, Chapter 1 combines framed and natural field experiments, while Chapter 2 combines a framed field experiment with surveys and incentivized elicitations, and Chapter 3 combines rich data from a natural field experiment with surveys conducted in the field.

References

Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky. 2014. "Energy policy with externalities and internalities." *Journal of Public Economics* 112: 72–88. [1]

Andre, Peter, Teodora Boneva, Felix Chopra, and Armin Falk. 2021. "Fighting climate change: The role of norms, preferences, and moral values." Working Paper, Working Paper Series 14518. IZA Discussion Paper. https://docs.iza.org/dp14518.pdf. [1]

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. [1]

Croson, Rachel, and Nicolas Treich. 2014. "Behavioral environmental economics: promises and challenges." *Environmental and Resource Economics* 58: 335–51. [1]

European Association of Environmental and Resource Economists. 2019. *Economists' Statement on Carbon Pricing.* https://www.eaere.org/statement/. Accessed: 2024-03-25. [1]

European Commission. 2023. *Farm to Fork strategy.* https://food.ec.europa.eu/horizontal-topics/farm-fork-strategy_en. Accessed: 2023-01-30. [2]

6. I follow the field experiment definition and classifications by (Harrison and List, 2004).

Faberi, Stefano, Loriana Paolucci, Bruno Lapillonne, and Karine Pollier. 2015. *Trends and policies for energy savings and emissions in transport.* Report, ODYSSEE-MURE project. https://www.odyssee-mure.eu/publications/archives/energy-efficiency-trends-policies-transport.pdf, accessed on October 12, 2023. [2]

Goetz, Alexander, Harald Mayr, and Renate Schubert. 2022. "Beware of side effects? Spillover evidence from a hot water intervention." Working Paper. https://congress-files.s3.amazonaws.com/2022-07/BewareOfSideEffects.pdf. [1]

Harrison, Glenn W, and John A List. 2004. "Field experiments." *Journal of Economic literature* 42 (4): 1009–55. [3]

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. [1]

Michel, Anette, Sophie Attali, and Eric Bush. 2016. *Energy efficiency of white goods in Europe: monitoring the market with sales data.* Final report. Study realised on behalf of ADEME by SOWATT and Bush Energie. https://storage.topten.eu/source/files/Market-Monitoring-2016-EN-Topten.eu.pdf, accessed on October 12, 2023. [2]

Mullainathan, Sendhil, and Richard H Thaler. 2000. "Behavioral Economics." Working Paper, Working Paper Series 7948. National Bureau of Economic Research. https://doi.org/10.3386/w7948. [1]

Obama, Barack. 2015. *Executive order – Using Behavioral Science insights to better serve the American people.* Media Release, White House, Washington, DC, September 15, 2015. [2]

Poore, Joseph, and Thomas Nemecek. 2018. "Reducing food's environmental impacts through producers and consumers." *Science* 360 (6392): 987–92. [3]

Reisch, Lucia A., Cass R. Sunstein, Mark A. Andor, Friederike C. Doebbe, Johanna Meier, and Neal R. Haddaway. 2021. "Mitigating climate change via food consumption and food waste: A systematic map of behavioral interventions." *Journal of Cleaner Production* 279: 123717. [2]

Stern, Nicholas. 2008. "The economics of climate change." *American Economic Review* 98 (2): 1–37. [1]

Woerner, Andrej, Taisuke Imai, Davide Pace, and Klaus Schmidt. 2023. "How to Increase Public Support for Carbon Pricing." Working Paper, Working Paper Series 489. Ludwig-Maximilians Universität München und Humboldt-Universität zu Berlin, Collaborative Research Center Transregio 190 - Rationality and Competition. https://www.econstor.eu/bitstream/10419/282180/1/489.pdf. [1]

Wolfram, Jessica. 2021. "Companies bet carbon labels can help the planet. Will consumers catch on?" Accessed January 30, 2023. https://www.washingtonpost.com/climate-solutions/2021/06/17/carbon-footprint-emissions-label. [2]

Chapter 1

Changing consumption behavior with carbon labels: Causal evidence on behavioral channels and effectiveness*

1.1 Introduction

The recent IPCC report makes it clear that strong political action is necessary to limit global warming to 1.5°C (IPCC, 2023). However, public support for traditional policy tools, such as carbon taxes and command-and-control measures, varies strongly between economic sectors. For instance, support for such measures is especially low in the food sector (Dechezleprêtre et al., 2022), which is responsible for 26%–34% of global greenhouse gas emissions.[1] Clark et al. (2020) predict that even if fossil fuels were banned immediately, emissions from the global food system alone would make it impossible to limit warming to 1.5°C. Shifting towards diets with lower carbon footprints[2] would greatly reduce these emissions.[3] Yet, the global effort to mitigate greenhouse gas emissions in the food sector has been weak,[4] likely due to the political challenges faced by traditional policy measures.

* 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, Nava Ashraf, Sanchayan Banerjee, Simon Block, Stefano DellaVigna, Eugen Dimant, Raphael Epperson, Ximeng Fang, Holger Gerhardt, Luca Henkel, Michael Kosfeld, Chris Krekel, Matthew Levy, Paul Lohmann, Julia Mink, Simone Quercia, Alex Rees-Jones, Paul Schäfer, Ganga Schreedhar, Marie Claire Villeval, Jakob Wegmann, and audiences at the London School of Economics, the University of Bonn, IDOS, Maastricht University, the Briq Workshop on Climate Change, the CRC TR 224 Retreat, the Third Winter Workshop on the Behavioral Economics of Food Consumption, the 15th RGS Doctoral Conference in Economics, the 2023 Mannheim Conference on Energy and Environment, the IDOS, the 2023 AERE Summer Conference, the 2023 SABE-IAREP Conference, the 2023 M-BEPS Conference, the 2023 EAERE Conference, the 2023 EEA Conference, the 2023 VfS Conference for their comments and suggestions.

1. See e.g. Poore and Nemecek (2018) and Crippa et al. (2021). The largest contribution to this amount comes from agriculture and land use/land-use change activities, while supply chain activities make up a smaller proportion.

2. I use the term "carbon footprint" to refer to all greenhouse gas emissions. In my calculations, I transfer gases other than CO_2 to CO_2 equivalents.

3. See e.g. Poore and Nemecek (2018), Kim et al. (2020), Grummon et al. (2023), and Scarborough et al. (2023). For example,Scarborough et al. (2023) study a UK sample and estimate the dietary impact of vegans as 25.1% of those of high meat-eaters. (Grummon et al., 2023) study a US sample and find that simple changes such as substituting chicken for beef can already reduce the dietary carbon footprint by more than 25% .

4. See e.g. OECD (2019). For example, the agricultural sector is excluded from the EU-ETS trading scheme and the USA does not have a carbon tax on the agricultural sector.

A more politically feasible way to influence consumption behavior in such a context could be to remove the behavioral and informational frictions that prevent consumers from making carbon-friendly consumption decisions. If emission-heavy choices are caused by consumers lacking knowledge about the carbon footprint of different options[5] or by consumers paying insufficient attention to these factors at the moment of choice, behavioral interventions could correct these frictions. One intervention that has received attention from academia[6], regulatory agencies[7], and private companies[8] is carbon labeling.

This paper quantifies the effectiveness of carbon labels in reducing carbon emissions relative to a carbon tax. It then examines the behavioral channels through which carbon labels impact behavior. Understanding these channels is relevant for assessing under which circumstances carbon labels are effective, as well as for understanding which frictions are impeding consumers from taking the more carbon-friendly choice in the absence of any intervention. Moreover, I make use of these insights to estimate the effect of carbon labels on consumer welfare.

Results are based on two framed field experiments ($N = 289$ and $N = 444$, respectively) and one natural field experiment (more than 120,000 purchase decisions by $N > 10{,}000$ customers).[9] To allow for comparability across experiments, I conduct all three experiments in a student canteen context. While the student canteen context in itself offers potential for reducing emissions on a large scale[10], the findings are also relevant for related settings, such as corporate canteens, grocery shopping, or other settings in which the carbon footprint caused by different items can be calculated and labeled, for example, shopping for toiletries or clothes.

Experiment 1, the first of the two framed field experiments, and Experiment 2, the natural field experiment, jointly establish that carbon labels affect consumption behavior and allow me to estimate the magnitude of the effect. In particular, I estimate the magnitude of a change in prices (a carbon tax) that would produce similar changes in purchase quantities as are brought about by the carbon labels. In the framed field experiment, I directly elicit an estimate based on how participants' willingness to pay for meals changes when shown carbon labels. In the natural field experiment, I rely on the combination of a carbon labeling intervention and variations in pricing. I estimate the effectiveness in both experimental settings to provide a precise and externally valid estimate: While the framed field experiment trumps the natural field experiment in terms of precision and clean causal identification, the natural field experiment provides evidence that the framed field estimates can be reconciled with student canteen purchasing behavior observed over longer time periods and across a large number of customers. Experiment 3 ($N = 444$) then returns to a framed field experimental set-up to assess the significance of both removing information frictions and addressing attention frictions in shaping consumers' responses to carbon labels.

5. Camilleri et al. (2019) and Imai et al. (2022)

6. See Reisch et al. (2021) for an overview, as well as Ho and Page (2023), Lohmann et al. (2022), Bilén (2022), and Imai et al. (2022)

7. For example, the Obama administration issued an executive order on Behavioral Science and the European Commission includes carbon labels in its Farm to Fork Strategy (Obama, 2015; European Commission, 2023).

8. For example, Oatly, an oat milk producer, Just Salad, a restaurant chain, Panera Bread and Allbirds, a shoe brand (Wolfram, 2021) all engage in carbon labeling.

9. My classification as framed or natural field experiment follows the Harrison and List (2004) taxonomy.

10. In Germany, 2.9 million individuals classified as students in 2021 (Federal Statistical Office (Germany), 2023), of which around 54% eat in the student canteen at least once a week (Federal Ministry of Education and Research (Germany), 2023).

Experiment 1 ($N = 289$) is a framed field experiment examining how willingness to pay for typical student canteen meals changes when participants are shown labels. The carbon labels I test include both an ordinal (traffic light system) and a quantitative ranking (greenhouse gas emissions in kg). This has been identified as an effective combination in previous studies (Potter et al., 2021; Taufique et al., 2022). The willingness to pay values participants indicate in the experiment influence the meal they receive at payout. I examine how participants' willingness to pay for specific meals changes when they are shown carbon labels, and compare these with the respective carbon footprints of those meals. I find that, on average, there is a decrease of €0.12 for each kilogram of CO_2 emissions caused by a meal. A decrease in average willingness to pay for a meal should have the same effect on the total quantity purchased as an equivalent increase in meal price, and I thus conclude that carbon labels produce a similar decrease in carbon emissions as a carbon tax of €120 per tonne.[11]

Experiment 2 is a natural field experiment ($N > 120{,}000$ choices from over 10,000 guests), showing that the consumption reactions I observe to carbon labels in Experiment 1 are reconcilable with behavior observed outside of a one-shot consumption setting. One of Bonn's university canteens is equipped with carbon labels for seven weeks, while the two other canteens serve as control restaurants, allowing for a difference-in-differences estimation of label effectiveness. The carbon labels I test are similar to those used in Experiment 1. I find that the labels decrease consumption of the higher carbon option by 2 percentage points or 5 percent of baseline consumption. The student canteen was still open for three weeks after the intervention period before closing for summer break. The effect of the labels persists in these three weeks. I compare the effect with an estimate of the effect of a carbon tax, which I estimate based on consumption choices and variations in canteen prices. I find that the effect of the carbon labeling intervention is comparable to that of a price change resulting from a carbon tax of €80 to €120 per tonne in the same setting. This is similar to the effect magnitude I observe in Experiment 1.

This quantification enables me to compare carbon labels with other policy tools and allows us to better understand the magnitude of the effect. €120 per tonne is about three times the current German carbon tax on petrol. Further, the current carbon price in the EU-ETS trading scheme is around €70 per tonne. At the same time, €120 per tonne is still lower than many estimates of the social cost of carbon (e.g. €160 per tonne in Rennert et al., 2022). This suggests that the labels are neither inefficiently "over-correcting" behavior, nor in itself sufficient as a sole policy to internalize the social cost of carbon.

To determine which behavioral channels drive consumers' reactions to carbon labels, I set up a simple theoretical model of consumption behavior in the presence of carbon emissions. My model features two main characteristics: First, consumers may prefer meals with lower carbon emissions, but may not be attentive to emissions at the moment of choice. Second, consumers may lack knowledge of the carbon emissions caused by different meals, that is, misperceive the carbon footprint. Behavioral interventions such as carbon labels make the consumer both attentive and informed. [12]

11. This is €120 per metric tonne. One metric tonne ≈ 1.1 short Tons/US Tons.

12. These modeling choices are based on the existing literature: I focus on attention focus as an important factor impacting decision making, as conceptualized by Bordalo, Gennaioli, and Shleifer (2022) and identified as relevant in the tax salience and resource consumption context (Chetty, 2009; DellaVigna, 2009; Byrne et al., 2024), as well as in suggestive empirical evidence from the food consumption context (Ho and Page, 2023). I focus on misperceptions

Experiment 3 ($N = 444$) is a framed field experiment seeking to quantify the relevance of each of these two possible channels. Are carbon labels mainly changing behavior by correcting misperceptions about carbon footprints or are they primarily changing behavior by increasing consumers' attention? I elicit participants' meal valuations in different treatment conditions, prior beliefs of the carbon footprints of different meal options, and participants' willingness to pay for seeing or avoiding carbon labels in the student canteen context. The meal valuations participants indicate influence the meal they receive after completing the experiment, and the elicitation of prior beliefs of carbon footprints as well as willingness to pay to see or avoid the labels is also incentivized. I observe meal valuations in different treatment conditions: first, in the absence of any behavioral intervention, second, with a behavioral intervention increasing attention (asking consumers to guess emissions), third, with a behavioral intervention increasing attention and correcting misperceptions (carbon labels) and, finally, when carbon emissions are removed (carbon offsetting).

Reduced-form results suggest that the labels primarily impact consumers by directing their attention toward carbon emissions, while improving consumers' knowledge about carbon impact plays a secondary role. Participants on average underestimate the emissions caused by high-emission meals and overestimate the emissions caused by low-emission meals. Correcting these misperceptions significantly impacts consumption choices: Consumers react to carbon labels with a stronger demand reduction if emissions were previously underestimated. However, a large part of the carbon labels' treatment effect is independent of previous under- or overestimation. The treatment effects observed for the intervention that merely increases attention without correcting misperceptions suggest that a large part of the remaining effect can be explained by an allocation of attention.

Using data from Experiment 3, I structurally estimate my model. Based on the estimated model parameters, I simulate how solely removing attentional biases or solely correcting consumers' misperceptions would impact carbon emissions and consumer welfare in the student canteen context. The former is more than seven times as effective as the latter, both in increasing consumer welfare and in decreasing carbon emissions. The combination of the two interventions (carbon labels) is most effective, and also more effective than the sum of the two single interventions, indicating important complementarities.

Data from all three experiments suggests that carbon labels are creating an overall psychological benefit to consumers. Experiments 1 and 3 elicit participants' willingness to pay to see or avoid carbon labels in a direct and incentive-compatible manner. The vast majority of participants report a zero (50%) or positive (45%) willingness to pay to see carbon labels. This evidence speaks against carbon labels imposing a net psychological cost on consumers. In the structural estimation of my model, I show that carbon labels on average create a psychological benefit for consumers independent of their impact on consumers' decisions. I estimate that carbon labels increase consumer welfare by on average 0.18¢ per consumption decision,[13] similar to the impact of a revenue-neutral carbon tax. Additionally, they create a fixed psychological benefit of 21¢ independent of their impact on consumption decisions. These results are further

of carbon impact as an important factor impeding optimal decision-making based on suggestions in recent papers on carbon labeling (Shewmake et al., 2015; Camilleri et al., 2019; Imai et al., 2022).

13. This figure averages over all consumption decisions, including the large majority of decisions unchanged by the labels. Consumer welfare increases on average by 10¢ per consumption decision changed by the labels.

supported by a post-intervention survey conducted after Experiment 2, the natural field experiment in the student canteen ($N = 234$). 73% of the guests exposed to the labels report that they would like the labels to be installed permanently (18% do not know, 9% against).

My contributions to the literature are three-fold: First, I contribute to the literature on the role of attentional biases in consumption decisions. The effectiveness of interventions that potentially direct consumers' attention has been shown in environmentally relevant contexts such as resource consumption (e.g. Allcott and Taubinsky, 2015; Tiefenbeck et al., 2018) and the purchase of environmentally durable goods (e.g. Rodemeier and Löschel, 2022). I add to this literature by experimentally estimating and quantifying the extent to which consumers exhibit attentional biases towards environmental externalities and, correspondingly, the relevance of a direction of attention in explaining consumers' responses to an intervention. This paper is thus similar in spirit to Taubinsky and Rees-Jones (2018) who use variations in experimental conditions to quantify a lack of consumer attention to sales taxes. It differs from previous literature on carbon labels that has so far mainly conceptualized carbon labels as a tool for correcting consumer misperceptions (Shewmake et al., 2015; Camilleri et al., 2019; Imai et al., 2022) or provided suggestive evidence on attentional biases playing a role in driving treatment effects (Lohmann et al., 2022; Ho and Page, 2023). Importantly, the variations in my treatment conditions allow me to structurally estimate a model describing the role of attentional biases and misperceptions as possible impediments to optimal decision-making, and to quantify the relevance of carbon labels in addressing each of these frictions. I find a high statistical and economic relevance of the attention channel in driving treatment effects, providing quantitative evidence on the role of attentional biases influencing consumption choice in the presence of greenhouse gas emissions. My results can also be understood more broadly in the context of recent research providing evidence in an abstract setting on the role information interventions can take in directing consumers' attention (Conlon, 2024).

Second, I contribute to the relatively recent literature on the consumer welfare impact of behavioral interventions. One strand of this literature derives consumer welfare from structural models or sufficient statistics (Chetty, 2009; DellaVigna, List, and Malmendier, 2012; DellaVigna et al., 2016; Rodemeier, 2021; Allcott et al., 2022; Goldin and Reck, 2022; Rodemeier and Löschel, 2022; Barahona, Otero, and Otero, 2023; List et al., 2023), while a second strand experimentally elicits consumers' willingness to pay to receive a behavioral intervention. Such an elicitation takes into account possible psychological costs and benefits arising to the consumer as a result of a change in consumption behavior induced by the intervention, as well as possible psychological costs and benefits arising independent of an impact on behavior (Allcott and Kessler, 2019; Thunström, 2019; Butera et al., 2022; Andor et al., 2023).[14] I take an approach in line with this second strand, providing first experimental evidence of consumers' preferences for the presence of carbon labels. I elicit consumers' willingness to pay for the presence of carbon labels directly in framed field experiments 1 and 3 and, for robustness, conduct an opinion survey at the end of the natural field experiment 2. Based on my theoretical framework and experiment 3 data, I provide an estimate of the effect carbon labels have on consumer welfare and compare it to alternative interventions.

14. Within my study context, one could think of such costs and benefits arising independent of a change in behavior as, for example, a change in feelings towards an unaffected choice or a change in the decision-making experience.

Finally, I contribute to the literature on the effectiveness of carbon labels on food consumption. Lohmann et al. (2022) use a difference-in-difference experiment design and estimate that labels decrease the probability of selecting a high-carbon meal by approximately 2.7 percentage points in Cambridge student canteens. Brunner et al. (2018) study a similar context but only observe changes over time in a single restaurant. They find a decrease in sales of red-labeled meat dishes by 2.4 percentage points. Bilén (2022) studies the introduction of carbon labels in the grocery shopping context and estimate a 2.5 percentage point reduction in carbon emissions caused by the carbon labels, employing a difference-in-difference estimation. Ho and Page (2023) find a 0.6% decrease in emissions of Hello Fresh consumers' meal package choices when carbon labels are installed. Correlational evidence (Spaargaren et al., 2013; Vlaeminck, Jiang, and Vranken, 2014; Visschers and Siegrist, 2015) and evidence from hypothetical decisions (Osman and Thornton, 2019; Banerjee et al., 2023) further corroborate the finding that carbon labels can reduce carbon emissions. [15] Other studies examine consumer behavior in the lab, asking consumers to make a decision for consumption happening at some point in the future. Camilleri et al. (2019) and Panzone et al. (2021) find carbon labels effective, while Imai et al. (2022) do not find an effect. In a broader sense, this paper also adds to environmental interventions in the restaurant context in general (e.g. Jalil, Tasoff, and Bustamante, 2020).

Previous studies examining the effectiveness of carbon labels estimate effect sizes in terms of percentage changes in consumption behavior, which are arguably difficult to compare across consumption contexts and policy instruments. In Experiment 1, I provide a first experimental estimate of the effectiveness of carbon labels relative to a price change resulting from a carbon tax, yielding a useful context for interpreting effect sizes. Within-subject designs as used here and in other structural behavioral studies (Taubinsky and Rees-Jones, 2018) can easily be adapted to other experiment populations, consumption environments, or other behavioral interventions, making intervention effects comparable across various domains. The experimental design is further validated by the large-scale natural field experiment (Experiment 2) producing effect estimates in line with the results of Experiment 1. Further, Experiment 2 provides the—to my knowledge first—estimate of the post-intervention effects of a carbon labeling intervention.

The rest of this paper is structured as follows. Section 1.2 describes how Experiment 1 quantifies the effectiveness of carbon labels using direct elicitation in a framed field experiment. Section 1.3 describes the design and results of Experiment 2, which is the natural field experiment corroborating my Experiment 1 estimate. Section 1.4 outlines a simple theoretical model describing possible behavioral biases influencing consumption behavior in the presence of environmental externalities, and the channels through which I expect a behavioral intervention such as carbon labels to impact behavior. Section 1.5 describes Experiment 3, the framed field experiment I conducted to examine the relevance of each of these channels, and discusses reduced-form evidence. Section 1.6 structurally estimates the theoretical model using data from Experiment 3. Section 1.7 discusses the impact of behavioral interventions on consumer welfare, drawing on data from all experiments. Finally, Section 1.8 concludes.

15. Also see Rondoni and Grasso (2021) for a review.

1.2 Experiment 1: Quantifying the effectiveness of labels in a framed field experiment

Experiment 1 quantifies the effectiveness of carbon labels in a framed field experiment. Subsection 1.2.1 describes the experimental design, 1.2.2 describes the empirical strategy, and subsection 1.2.3 describes data and results.

1.2.1 Experimental design

Overview

To cleanly measure the impact of carbon labels and elicit how their effectiveness quantifies relative to a carbon tax, willingness to pay of a given individual for a given meal should best be observed, at the same time, once in the absence of carbon labels and once in the presence of carbon labels. Experiment 1 is designed accordingly. I summarize the most important design choices below and add details in the following subsections.

(1) For this experiment, I move participants' lunch consumption decision to an online survey, which they fill out just before lunchtime on the experiment day. Participants make their way to the university campus shortly after completing the survey and receive the experiment payment and lunch option corresponding to the choices they made in the survey.

(2) In the survey, experiment participants indicate their willingness to pay for different meals multiple times, totaling to 15 meal purchase decisions. One of these is implemented at payout.

(3) I allocate participants to the Label or the Control condition: Participants in the Label condition first indicate willingness to pay for four meals in the absence of carbon labels and shortly after indicate willingness to pay for the same four meals in the presence of carbon labels. Participants in the Control condition make the same decisions but do not see any carbon labels in the second elicitation.

(4) Willingness to pay for meals is elicited relative to an alternative lunch: In each of the 15 meal purchase decisions, participants first decide whether they prefer a given meal or a cheese sandwich.[16] They then indicate how much they are willing to pay to receive the given meal rather than the cheese sandwich, and vice versa if they prefer receiving the cheese sandwich. Willingness to pay for a given meal is thus always measured relative to the cheese sandwich (reflecting the real-world fact that the alternative to not eating something is eating something else). The dependent variable of interest in the analysis is the **change** in relative willingness to pay between the first and second elicitation.

(5) Carbon labels show a quantitative and ordinal ranking (see Figure 1.4 for an example). The carbon labels I test include greenhouse gas emissions in kg, as calculated based on the quantity of each meal ingredient and its average greenhouse gas emissions. It also includes an ordinal ranking using a traffic light system, ranking the meal relative to other meals typical of Bonn's student canteens. Combining an ordinal and a quantitative ranking has

16. All the meals are typical student canteen meals and a cheese sandwich is also a typical lunch choice in Germany. Meals are further described in section 1.D.

been identified as an effective combination in previous literature (see Taufique et al., 2022 and Potter et al., 2021). Further, I designed the labels in cooperation with Bonn's student canteens to ensure that I am testing labels that they would be willing to implement and thus to ensure comparability to Experiment 2. The labels also indicate the distance a car would need to be driven (in kilometers) to produce an equivalent level of CO_2 emissions.

(6) Willingness to pay to see or avoid carbon labels is also elicited: Before the final three meal purchase decisions (three new meals), participants indicate whether they would like to see carbon labels on these final decisions, and indicate their willingness to pay to enforce their choice. This elicitation is incentivized.[17] I discuss these results in Section 1.7.

Timeline

The survey timeline is visualized in Figure 1.1. First, the elicitation of willingness to pay is explained to participants and they are shown how their payout and the meal they receive will depend on the choices they make throughout the experiment. They then answer four comprehension questions, which they must answer correctly before proceeding. Any participant taking more than five attempts in doing so is excluded from the analysis, as pre-registered. Second, participants indicate their baseline willingness to pay for four meals (four questions). Third, participants answer several incentivized and timed[18] guessing questions on unrelated issues (e.g. on the length of a popular running route in Bonn).

The experiment then proceeds differently depending on the treatment group participants are assigned to by computer randomization. All participants are again asked to indicate their willingness to pay for the four meals, but the framing of the decision and some characteristics of the decision depend on the treatment condition:

- In the Control condition, decisions are exactly as in the first, baseline elicitation.
- In the Label condition, participants see carbon labels.

To increase power and elicit further information, participants' willingness to pay for the same four meals is elicited a third time[19], with partly changed treatment conditions:

- Participants previously in the Label condition are in the third round assigned to the Offset condition: Participants are informed that the emissions caused by their lunch choice (be it the meal or the sandwich) will be offset.[20]
- Half of the participants previously in the Control condition are in the third round assigned to the Label condition, and half of the participants previously in the Control condition repeat the Control condition. Afterward, before proceeding with the experiment, this group guesses emission values.[21].

17. See section 1.D.2 for details.

18. For each question for which participants answer a number within 30% of the true value, €0.10 is added to participants' pay-out. Further, each question is restricted to 60 seconds of answering time to ensure that participants can not search for answers online.

19. In the analyses, I control for whether observations stem from a third-round elicitation. All the main results replicate including only data from the first two rounds, see Table 1.B.20.

20. The details and results of the Offset condition are shown in section 1.D.3 and in Table 1.B.6. As pre-registered, the Offset condition serves as a robustness check of the results of the Attention+Offset condition in Experiment 3, which is used as input for the structural estimation described in Section 1.6.

21. This data is used for the analysis shown in Figure 1.14. As these guessing questions occur after the first, second, and third willingness to pay elicitation, they do not affect the results displayed in this section.

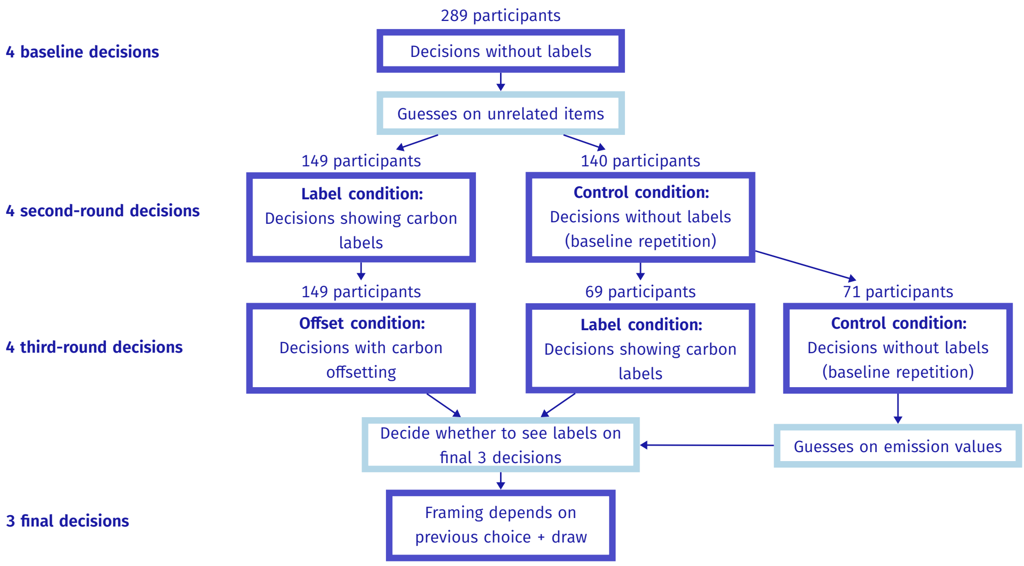


Figure 1.1. Experiment schedule and treatment groups

Note: Participants repeat the same four meal purchase decisions three times, with the decision framing differing across rounds. Treatments are described in more detail in the "Experiment timeline" paragraph above. The results of the OFFSET condition are not further discussed in this section, but details are described in section 1.D.3, and results are shown in Table 1.B.6. As pre-registerd, the OFFSET condition serves as a robustness check of the results of the ATTENTION+OFFSET condition in Experiment 3, which is used as input for the structural estimation described in Section 1.6.

The three rounds include four meal purchasing decisions each, constituting a total of 12 decisions. Additionally, three final purchase decisions revolve around three not previously seen meals. Before seeing these final decisions, participants are asked whether they would like to see carbon labels for these decisions and indicate how much they are willing to pay such that their preferred display option is implemented. This elicitation is incentivized as detailed below.

In the final steps, participants answer questions concerning their environmental attitude and psychology, and participants' guesses of the calories contained in each meal are elicited for further robustness checks.[22]

Details on the meal purchasing decisions

Participants make a total of 15 meal-purchasing decisions in the course of the experiment (4 baseline, 4 first-round, 4 second-round, and 3 final decisions). The 12 first decisions revolve around the same 4 meals, and the final 3 decisions around 3 other, not previously seen meals. Participants who indicate that they are vegetarian are shown only vegetarian meals.[23] In each decision, participants first choose whether they prefer consuming a certain meal or a cheese sandwich. An example of a baseline decision is shown in Figure 1.2. The left option in the example changes across decisions to indicate one of the four meals, while the option on the right, the cheese sandwich, stays constant for all decisions.[24]

22. See section 1.B.11 and 1.D.4 for results and experiment screenshots.

23. Meals are detailed in Section 1.D.1. Participants with stricter dietary requirements (vegan, gluten-intolerant, lactose-intolerant, or halal) are not permitted to participate in the experiment.

24. To ensure that results are not driven by a left-right effect, the left-right positioning of the two options is reversed in half of the experiment sessions. The order in which meals are shown is randomized.

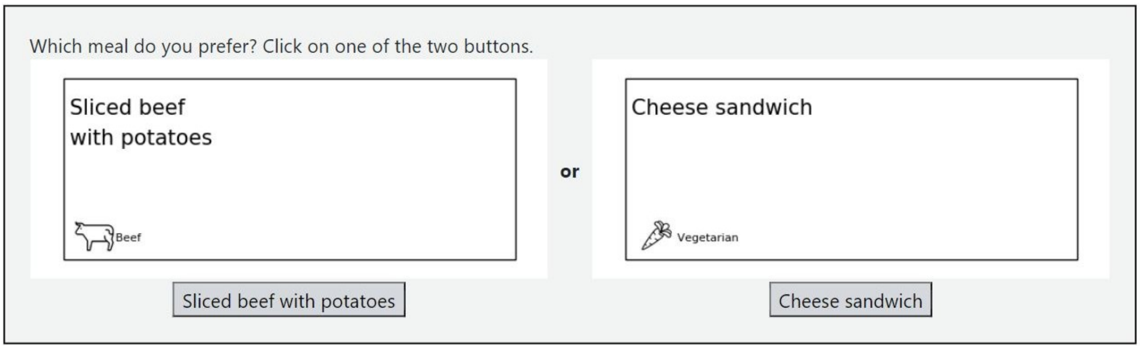


Figure 1.2. Meal purchase decision example step 1

Note: Step 1 of the purchasing decision. Depending on the participants' decision in Step 1 of the decision, Step 2 (Figure 1.3 asks participants for their willingness to pay to receive or avoid the warm meal.

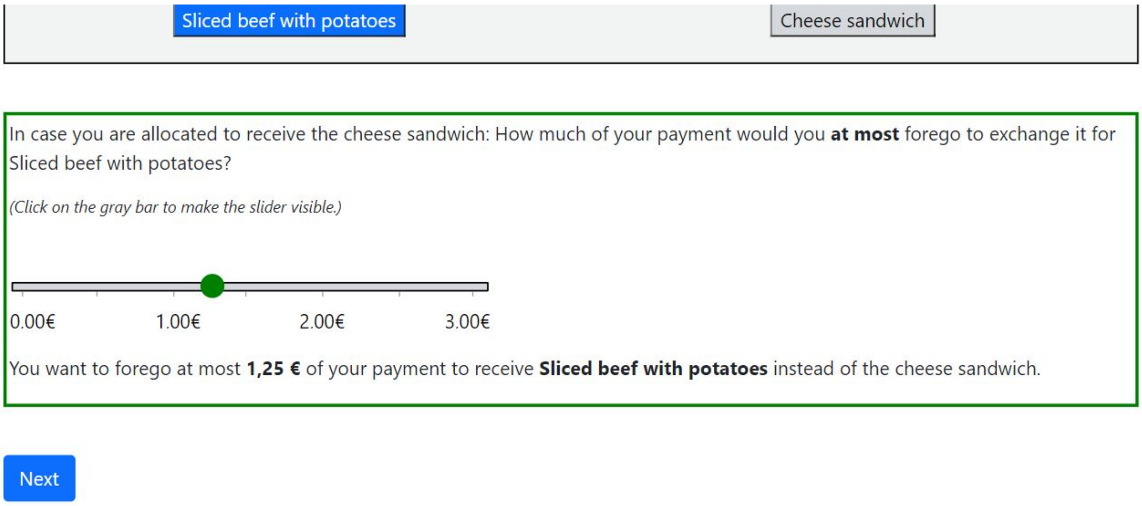


Figure 1.3. Meal purchase decision example step 2

Note: Step 2 of the purchasing decision. If participants indicate in Step 1 that they prefer the warm meal, Step 2 is as shown above. If participants indicate in Step 1 that they prefer the cheese sandwich, Step 2 asks participants how much they are at most willing to forego to receive the cheese sandwich instead of the warm meal.

Once participants indicate their preference for one of the two options, a second window appears and they indicate how much of their experiment payment they would at most be willing to forego to ensure their preference (see example in Figure 1.3 in which the participant indicated a preference for Sliced beef in the first step). If participants prefer the specific meal, they indicate how much they are willing to forego to ensure they receive this meal instead of the cheese sandwich. If participants prefer the cheese sandwich, they indicate how much they are willing to forego to ensure they receive the cheese sandwich instead of the specific meal. Any amount between €0.00 Euro and €3.00 can be indicated on a slider in five-cent intervals. [25]

This meal-purchasing procedure captures participants' willingness to pay for the specific meal, relative to the cheese sandwich. If participants indicate in the first step that they prefer the

25. I chose €3.00 as the maximum amount since this is the maximum price a student would pay to purchase any of the meals in the student canteen. A willingness to pay of €3.00 or −€3.00 was indicated in less than 3% of all observations. Figure 1.B.1 shows the distribution of baseline willingness to pay values indicated.

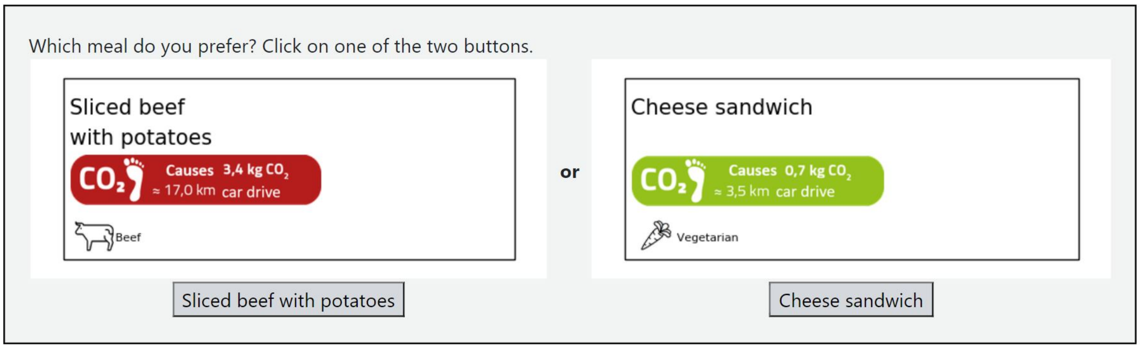


Figure 1.4. Meal purchase decision example: Decisions with labels

Note: Carbon labels include both an ordinal (traffic light system) and a quantitative ranking (greenhouse gas emissions in kg). This has been identified as an effective combination in previous literature (Potter et al., 2021; Taufique et al., 2022).

specific meal, the amount they indicate in the second step can be interpreted as willingness to pay to receive the meal. If participants indicate in the first step that they prefer the cheese sandwich, the amount they indicate in the second step can be interpreted as willingness to pay to avoid the meal, i.e. negative willingness to pay for the meal.

Decision framing differs across treatment conditions. In the four baseline decisions, participants do not see any carbon labels but are merely shown the meal name and the meal's main ingredient (see Figure 1.2 for an example)[26]. The four second-round and four third-round decisions are very similar to the baseline decisions, with the exception that the framing of the decision changes for some of the participants. For participants in the Label condition, emission values are added to the meal options. An example is shown in Figure 1.4.[27] For participants in the Control condition, there is no change in framing relative to the baseline decisions. For participants in the Offset condition, participants are told that the emissions caused by the meal will be offset with a donation to Atmosfair. An example is shown in Figure 1.5.[28]

Participants and set-up

289 experiment participants are recruited from the participant pool of the BonnEconLab, the behavioral experimental lab of the University of Bonn, to participate in one of eight experimental sessions taking place between the 26th of October and the 5th of November 2021. I preregistered the experiment design and the main outcomes shown in this section (Schulze Tilling, 2021b). Participants are informed in the experiment invitation that vegetarian participants

26. I chose this display to reflect exactly how a meal would be displayed on the student canteen website, see Figure 1.9 for an example of implementation in the field.

27. I calculated the emissions caused by each meal with the application Eaternity Institute (2020). The student canteen in Bonn catered the meals and provided me with recipes for the emissions calculation.

28. The results of the Offset condition are not further discussed in this section, but details are described in section 1.D.3, and results are shown in Table 1.B.6. As pre-registerd, the Offset condition serves as a robustness check of the results of the Attention+Offset condition in Experiment 3, which is used as input for the structural estimation described in Section 1.6.

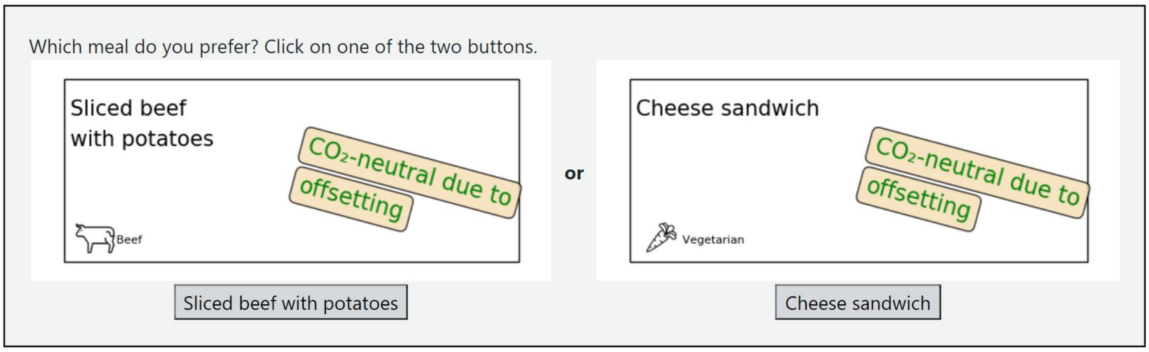


Figure 1.5. Meal purchase decision example: Decisions with carbon offsetting

Note: If the meal is payout-relevant, I offset emissions of the selected meal with a corresponding donation to Atmosfair.

Figure 1.6. Gazebo set-up on University campus

Note: Set-up to provide participants with their payment in cash and a lunch corresponding to one of their choices. While completing the experiment, participants do not know which meal is payout-relevant.

are permitted, but not participants with stricter dietary requirements (vegan, gluten-intolerant, lactose-intolerant, or halal). Participants are informed that the experiment will be conducted online, but that they are required to make their way to campus afterward to collect their payment in cash and a lunch. They are not given any further information on the purpose of the experiment. The experiment is conducted using oTree software (Chen, Schonger, and Wickens (2016)).

Meals are catered by the student canteen. All experiment meals are in regular intervals offered by the student canteen, but they are not offered in the canteen on the particular experiment day, i.e. the student canteen prepared the meals only for our experiment participants. When participants pick up their meal, it is warm, ready to eat, and can be consumed on the spot, as shown in Figure 1.6.

Incentivization

At the beginning of the experiment, participants are informed that one of the meal purchase decision will be implemented, but it is not known to participants before the end of the experiment for which decision this will be the case. For the relevant decision, a random price draw and the willingness to pay a participant indicated jointly determine whether the participant receives the meal or the cheese sandwich and which amount is deducted from his payment. Participants' decision to see or avoid carbon labels on their final three decisions is incentivized with a similar

BDM mechanism, with a random price draw and participants' indicated willingness to pay to see or avoid the labels jointly determining whether carbon labels are shown on the final three decisions. The details of the incentivization mechanisms are explained in Appendix 1.D.2.

1.2.2 Estimation strategy

In my estimation, I want to identify the effect of carbon labels on meal choices while controlling for any other factors that might influence meal choices (e.g. individual's tastes, hunger level or mood). I thus examine the **change** in willingness to pay of a certain individual for a certain meal as the outcome variable: Instead of directly examining an individual's willingness to pay for a meal in the Label or Control condition, I subtract the individual's baseline willingness to pay for the same meal from this amount, and then examine the remaining difference. This is the change in willingness to pay occurring due to being shown carbon labels (theLabel condition) or due to merely being asked for willingness to pay a repeated time (the Control condition).

In this manner, I control for individual-specific meal tastes at baseline. This includes any inclinations toward the warm meals or the cheese sandwich, and any other factors that might influence meal choice apart from the carbon labels. All of these factors should stay constant between the baseline elicitation and subsequent elicitations. One can also interpret the outcome variable as denoting individual- and meal-specific within-subject treatment effects, which I compare between treatment groups. An alternative approach would be to use willingness to pay as the dependent variable and include a fixed effect for every individual-specific meal choice. This approach yields similar results, as shown in Section 1.B.7.

My basic specification is:

$$Diff_{ijm} = \beta_1 High_m + \beta_2 Low_m + \delta_1 (Label_{ij} \times High_m) + \delta_2 (Label_{ij} \times Low_m) + ThirdRound_j + \varepsilon_{ijm} \tag{1.1}$$

where $Diff_{ijm}$ describes the difference between willingness to pay of individual i in round j for meal m and individual i's baseline willingness to pay for meal m.

$High_m$ is an indicator of whether the meal causes higher emissions than the sandwich.[29] Low_m is an indicator of whether the meal causes lower emissions than the sandwich. Together, these variables capture any effect that the mere act of asking participants for their willingness to pay multiple times might have. I differentiate between meals with emissions lower than the sandwich and meals with emissions higher than the sandwich because I expect participants to respond to carbon labels differently depending on how the emissions of the two options compare: For meals with emissions lower than the cheese sandwich, the participant can reduce his expected emissions if he adjusts his willingness to pay for these meals upward. For meals with emissions higher than the sandwich, the participant can reduce his expected emissions if he adjusts his willingness to pay for these meals downward.

$(Label_{ij} \times High_m)$ interacts the high-emission indicator with an indicator for whether individual i saw carbon labels in round j. This describes the average causal effect of carbon labels on willingness to pay for a meal that is high in carbon emissions. $(Label_{ij} \times Low_m)$ describes the

29. For non-vegetarians, these were three of the four meals. For vegetarians, these were two of the four meals. See section 1.D for details.

average causal effect of carbon labels on willingness to pay for a meal that is low in carbon emissions. $ThirdRound_j$ is an indicator of whether it was the third round of decisions.[30]

1.2.3 Data and results

I use data from Experiment 1 to estimate equation 1.1, but exclude the 3% fastest participants as well as participants not passing the comprehension check after five attempts, as pre-registered[31]. The remaining 289 experiment participants are computer-randomized into treatments. Section 1.B.1 shows a randomization check. 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 in terms of these characteristics, as discussed in Section 1.B.2, and results hold when restricting the sample to only students or only non-vegetarians, as shown in Section 1.B.5. Section 1.B.3 shows the baseline distribution of relative willingness to pay for meals.

Table 1.1 Spec. (1) shows the results of the OLS estimation of equation 1.1, clustering standard errors at the individual level.[32] For meals with lower emissions than the cheese sandwich, willingness to pay increases by €0.14 on average due to the labels. For meals with higher emissions than the cheese sandwich, willingness to pay decreases by €0.31 due to the labels. Changes in willingness to pay for participants in the Control condition are not significant, and, coefficient-wise, move in opposite directions. Thus, the mere act of asking participants for their willingness to pay multiple times does not seem to significantly impact their willingness to pay. Figure 1.7 illustrates effects by showing average changes in willingness to pay for the Control and Label groups, for low-emission and high-emission meals.

Specification (2) in Table 1.1 does not group the four meals into low-emission and high-emission meals but instead regresses the change in willingness to pay on the difference in emissions between the warm meal and cheese sandwich. This specification estimates that on average, willingness to pay decreases by €0.12 for every additional kg of emissions that the warm meal causes on top of the cheese sandwich. This result can be interpreted as—assuming that a downward shift in the demand curve results in the same effect on quantity purchased as an upward shift in the supply curve—the carbon labels producing a similar impact in this setting as would result from a carbon tax of €0.12 per kg or €120 per tonne[33]. This is four-fold the current German CO_2 tax on petrol (€30 per tonne).

Section 1.B.9 shows the distribution of participants' reactions to carbon labels by baseline willingness to pay for meals. Section 1.B.10 shows suggestive evidence of heterogeneity in treatment effects. Section 1.B.11 tests whether the carbon labels influenced participants' perception of meals' calories (as a proxy for other meal characteristics). I do not find any evidence for this being the case.

30. An alternative approach to controlling for possible third-round effects is excluding third-round decisions entirely. This yields similar results (Table 1.B.20).

31. See Schulze Tilling (2021a). Dohmen and Jagelka (2023) find that fast respondents are more likely to not pay attention and give random answers.

32. The results of the Offset condition are not further discussed in this section, but details are described in section 1.D.3 and results are shown in Table 1.B.6. As pre-registered, the Offset condition serves as a robustness check of the results of the Attention+Offset condition in Experiment 3, which is used as input for the structural estimation described in Section 1.6.

33. See section 1.B.8 for an illustration of the intuition behind this comparison.

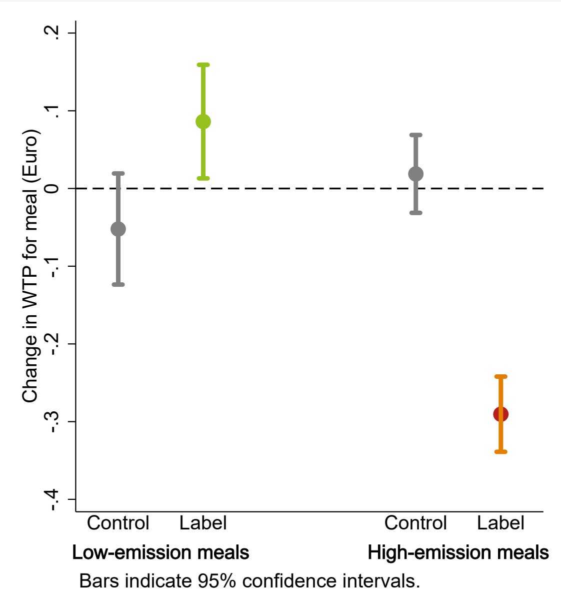


Figure 1.7. Within-subject change in willingness to pay for meals

	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)
Control for third round	0.01	0.02
	(0.03)	(0.03)
Constant		−0.02
		(0.02)
Participants control	140	140
Participants treated	218	218
Observations	1,716	1,716

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

Table 1.1. Within-subject change in willingness to pay for meals

1.3 Experiment 2: Quantifying the effectiveness of labels in a natural field experiment

The results in Section 1.2 show that carbon labels change consumer behavior and reduce emissions in a one-shot consumption setting. Experiment 2 tests the external validity of this result: It investigates in a natural field experiment in the student canteens in Bonn whether effects are similar if carbon labels are installed over longer time periods. Subsection 1.3.1 describes the experiment design, subsection 1.3.2 describes the estimation strategy, and subsection 1.3.3 describes data and results.

1.3.1 Experimental design and setting

Overview. To identify the causal effect of carbon labels in the field, I make use of the fact that there are multiple student canteens in Bonn that centralize their meal planning, i.e. on a given day roughly the same meals are offered in all canteens. I summarize the most important details below and describe the student canteen setting in Bonn more in detail in Section 1.E. I preregistered the experiment design and main outcomes.[34]

(1) I use a difference-in-difference design, as illustrated in Figure 1.8: Purchasing behavior in all three student canteens is first observed in the absence of labels (pre-intervention phase, 4 weeks), then labels are installed in the treatment student canteens (intervention phase, 7 weeks). After the removal of the labels, I observe consumption behavior until the end of the semester (post-intervention phase, 3 weeks).

34. AsPredicted#95108

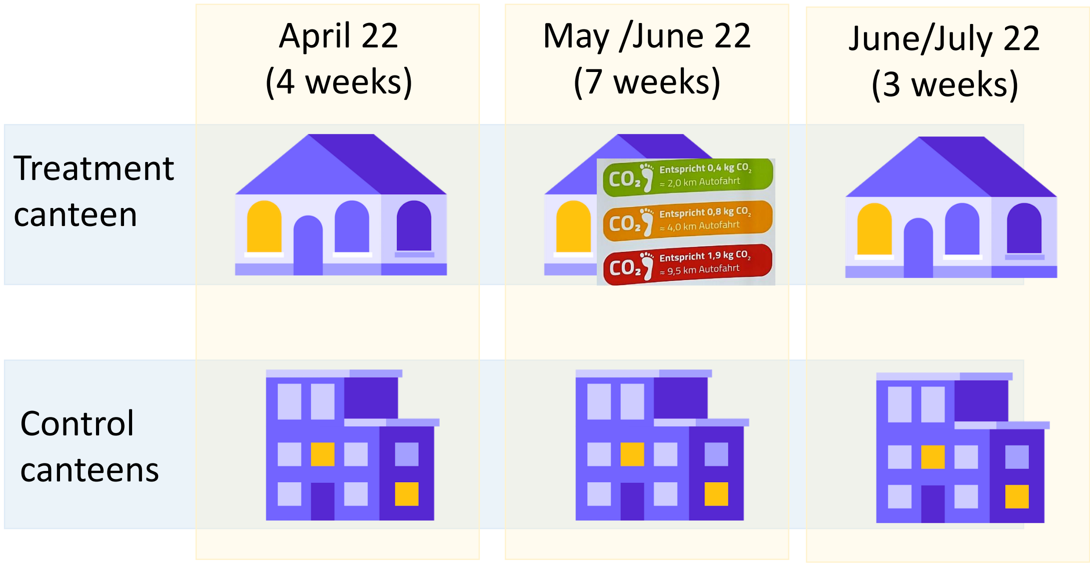


Figure 1.8. Timeline Experiment 2

Note: The start of the pre-intervention phase and the end of the post-intervention phase correspond with the beginning and end of the summer semester in Bonn. After the post-intervention phase, the student canteens closed for summer break. During the treatment phase, carbon labels were shown in the treated canteen, on the online menu, digital billboards in the canteen, and signs on meal counters.

(2) Carbon labels show a quantitative and ordinal ranking, and are similar to the carbon labels used in Experiment 1. In the treatment canteen, they are added to the online menu, to the digital billboards in the student canteen, and to the paper signs on top of the meal counters. Examples are shown in Figure 1.9. Emissions are calculated based on student canteen recipes and Eaternity Institute (2020) emission values.

(3) Carbon labels are installed for the two main meal components sold by the treatment canteen, but not for sides and desserts, for ease of implementation and interpretability (see 1.E for details). A typical student canteen meal consists of one meal component and one or two sides, with the main meal component on average causing 70% of the emissions caused by a typical meal. The two main meal components on offer always consist of one vegetarian and one meat-based component, which is higher in carbon emissions than the vegetarian option.

(4) I accompany the natural field experiment with a pre-intervention (N>1,700) and post-intervention survey (N> 900) in the field. These capture students' demographic characteristics (connectable to canteen purchasing data) and opinions on the carbon labels. These surveys are described in more detail in Section 1.E.6.

1.3.2 Estimation strategy

To estimate the causal effect of carbon labels in the student canteen, I use a difference-in-difference estimation using choice of emission-heavy main meal component as the outcome vari-

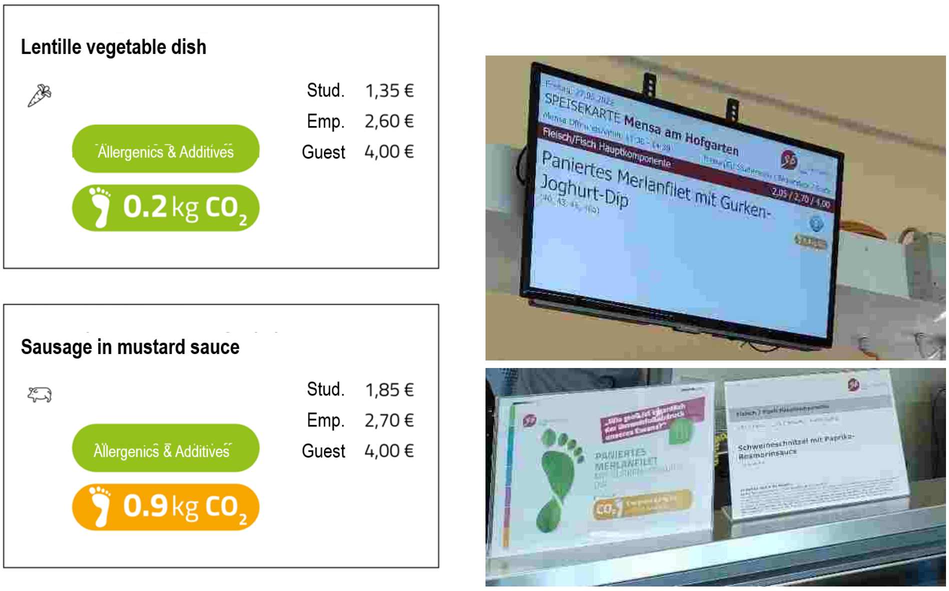


Figure 1.9. Labels in the canteen

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

able. Before showing the estimation strategy in detail, I would like to address possible threats to identification and the design choices I take accordingly:

- The proportion of meat main components sold in the treatment and control canteen differ pre-intervention (41% and 51%). I thus use a difference-in-difference design to control for baseline differences in meat consumption.
- Identification in the difference-in-difference framework hinges on the assumption of parallel trends between the canteens in the absence of treatment. While this assumption cannot be tested directly, I provide evidence suggesting that this is a reasonable assumption. Figure 1.10 shows that student canteen sales develop similarly throughout the entire 14-week period, the event plot in Figure 1.11 shows reasonable pre-trends and Figure 1.C.2 shows that trends are parallel in the semester following the intervention.
- Identification also relies on students not increasingly switching between canteens due to the carbon labels. This seems unlikely due to the canteens being located over 1.7 km (1.1 miles) apart. Section 1.E discusses possible switching in detail, using pre-intervention individual-level purchase data to identify a guest's "home" canteen and then tracking "non-home" visits throughout the period. There is no clear time trend in switching attributable to the labels, and the proportion of meat purchases made by switchers does not increase throughout the period, which also makes an intervention-motivated switching from treatment to control canteen seem unlikely. Col.(5) in Table 1.2 shows that results are robust to employing an intent-to-treat analysis.
- One could imagine treatment effects from the treatment canteen spilling over to the control canteen. This would lead to an understimation of effect sizes. However, section 1.C.5 shows — based on survey evidence — that such spillover effects are limited.
- The student canteens offer one vegetarian and one meat main meal component every day, with the vegetarian main meal component always causing lower emissions than the meat

main meal component. The meal offer changes daily, and emissions caused thus largely fluctuate across days. The main analysis thus focuses on changes in the proportion of meat main meal components purchased.[35]

My most basic difference-in-difference specification is:

$$Meat_{it} = \alpha + \beta_1 LabelPeriod_t + \beta_2 PostPeriod_t + \gamma Treat_{it} + \\ + \delta_1 (Treat_{it} \times LabelPeriod_t) + \delta_2 (Treat_{it} \times PostPeriod_t) + \epsilon_{it} \quad (1.2)$$

The variable $Meat_{it}$ is a binary outcome describing whether the main meal component purchased by individual i on day t is meat-based, i.e. $Meat_{it}$ equals 1 if the meat-based main meal component is purchased, and 0 if the vegetarian main meal component is purchased. $LabelPeriod_t$ is an indicator of whether this purchase occurred during the intervention period (May/June 22), and $PostPeriod_t$ is an indicator of whether this purchase occurred in the three weeks following the intervention period, before the canteens went into summer break (June/July 22). $Treat_{it}$ is an indicator of whether the purchase occurred in the treatment canteen.

$(Treat_{it} \times LabelPeriod_t)$ is the variable of interest identifying the difference-in-difference estimate of any change in purchasing behavior occurring during the labeling period in the treated canteen relative to the control canteens. $(Treat_{it} \times PostPeriod_t)$ identifies possible post-intervention effects.

1.3.3 Data and results

I include purchase data from April 1st (beginning of the semester) to July 8th (end of the semester) in my analysis. For each purchase, I observe the meal purchased, the price paid, and the location, day, and time of the purchase. I observe whether the purchase is made by a student (81% of purchases) or by an employee (17% of purchases). Further, around 2/3 of sales are made with a personalized payment card, allowing me to track individuals across time.

I drop data from seven days on which the treatment and the larger control canteen did not offer the same main meal components. I also drop all consumption of Ukrainian refugees, who received free meals in the student canteens from week 9 of the sample period. For my main analysis, I additionally drop data from the first week of the label period (week 5), since a "Healthy Campus" week occurred simultaneously and it is not clear whether the carbon labels or this event are driving the increased vegetarian consumption identified for week 5 in the event plot shown in Figure 1.11. The main results are robust to these exclusions, as discussed more in detail in Section 1.E. The final sample includes 120,093 observations, made by over 6,000 guests.

35. Average carbon footprints at baseline differ between treatment and control canteens, and average emissions of the gastronomic offer also differ between the baseline and the intervention period. As further explained in Section 1.C.3, an analysis of the full sample using carbon footprints as the outcome variable could mistakenly attribute changes in emissions due to changes in the gastronomic offer to the carbon label. Section 1.C.3 shows an analysis with emissions as the main outcome variable on a sample restricted for this analysis. I find that the labels reduce emissions by 25 grams per meal, or 3% of baseline emissions.

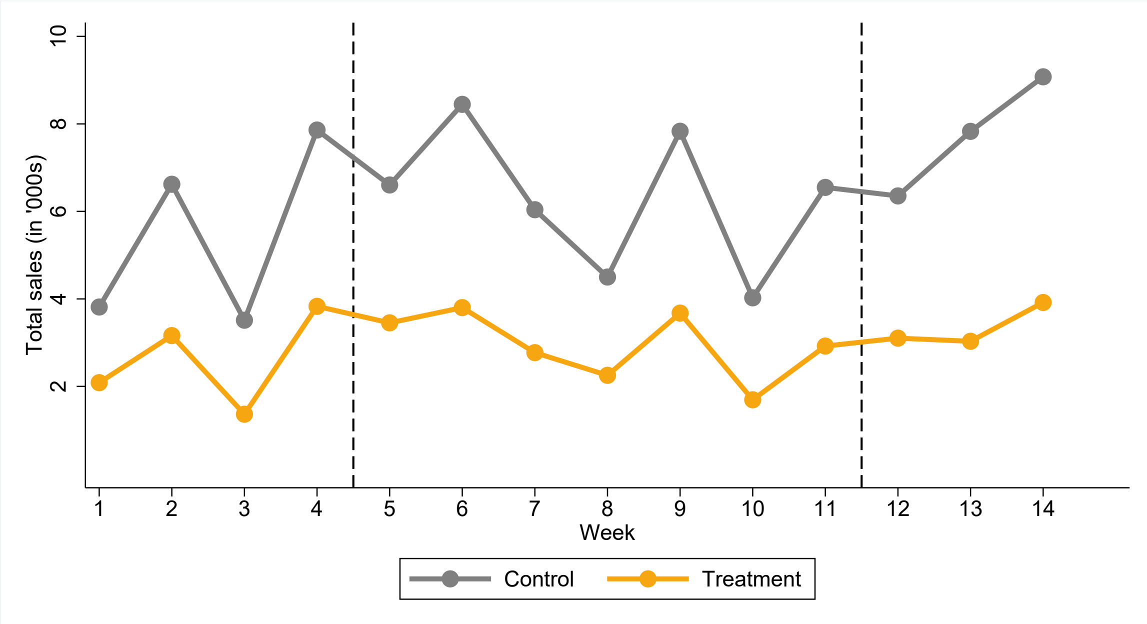


Figure 1.10. Weekly student canteen sales of main meal components

Note: Sales after data cleaning described in Section 1.3, but including week 5. $N = 130{,}132$. Weeks 1–4 are the pre-intervention period (April 2022), weeks 5–11 are the intervention period (May to Mid-June 2022), and weeks 12–14 are the post-intervention period (last week of June and two weeks of July 22). The drop in sales in week 10 is likely due to the one-week Pentecost holidays, during which no classes took place.

Col. (1) in Table 1.2 estimates specification 1.2 in a linear probability model. It estimates that the carbon labels decrease the probability that a purchased main meal component is meat-based by 2 percentage points or 5% of the baseline likelihood. Post-intervention effects are estimated at 7 percentage points or 17% of the baseline likelihood. Figure 1.11 shows an estimation of weekly treatment effects, using week 4 as the baseline period. Effect sizes during the intervention period seem to increase over time. One explanation for this might be that perhaps canteen guests do not notice the carbon labels immediately, but only on their second or third visit to the student canteen. The large effect estimated for the post-intervention period is similar to that estimated for the final weeks of the intervention period.

Col. (2) of Table 1.2 drops the $LabelPeriod_t$ and $PostPeriod_t$ time controls and instead includes date-fixed effects. This allows for a more fine-grained control for time trends (e.g. semester times, seasonal trends) and changes in the gastronomic offer, since the offer changes daily but control and treatment canteens coordinate on meal offers). Estimated effect sizes are similar.

Col. (3)–Col. (5) examines whether treatment effects are caused by a change in canteen guests' behavior, as opposed to selection effects. I restrict the sample to canteen guests paying with their individual payment card, visiting the student canteen regularly pre-intervention (at least five times within four weeks) and at least once during the intervention phase, and pre-dominantly visiting the same canteen pre-intervention (at least 80% of pre-intervention visits to the same canteen). Col. (3) applies the same regression specification as in Col. (2) to the restricted sample for comparison purposes. Col. (4) includes individual fixed effects in the regression, and Col. (5) shows an intent-to-treat analysis: Here, I fix a value of the "Treatment restaurant" indicator for each individual, depending on consumption behavior in the four-week pre-intervention period. For individuals mainly going to the treatment restaurant in the pre-intervention period, "Treatment restaurant" is set to 1, while it is set to 0 for individuals mainly going to the control restaurants during the pre-intervention period.

To assess whether the strong post-intervention effects last, Table 1.C.1 includes data from the semester following the intervention (Oct. 22–Jan. 23) in the difference-in-difference estimation. There is no evidence of this being the case, and the time trends in Figure 1.C.2 suggest—if at all—an upwards-sloping pattern. It, therefore, seems unlikely that treatment effects may have in fact persisted among the canteen guests who visited the student canteen in May 2022, and that my null effects are entirely attributable to incoming new and never-treated students.[36]

Post-intervention effects thus seem rather short-lived, in line with the attention-habit model described in Byrne et al. (2024): The pattern could be explained by the intervention drawing consumers' attention toward the issue of carbon emissions, and consumers making a short-lived habit out of paying attention to the issue. A similar pattern is observed for an attention-directing intervention in the resource conservation context described in Byrne et al. (2024).

Section 1.C discusses additional results. Drawing on a larger data set of consumption data from April 22 to March 23, I identify a rough estimate of how demand for meat meals would react to a carbon tax in the student canteen. Using this estimate, I approximate that a carbon tax of €80 per tonne to €120 per tonne would result in a similar demand reaction as is produced by the carbon labels. This is reconcilable with the assessment that carbon labels are as effective as a carbon tax of €120 per tonne, as elicited in Experiment 1. Results are shown in Section 1.C.2. Further, I estimate the decrease in average greenhouse gas emissions caused by the labels at 25g per meal on average. This is around 3% of the average emissions of a meal consumed at baseline (Section 1.C.2). Examining heterogeneity in treatment effects, I find similar treatment effects when restricting the sample to only employees, only off-peak hours, only payments made by individual payment cards, or only frequent canteen guests. Combining the purchase data with demographic data I elicited in the field surveys described in Section 1.E.6, I find suggestive evidence of treatment effects being larger for female guests, for younger guests, and for guests indicating that environmental aspects play an important role in their consumption choices (Section 1.C.4). Section 1.C.5 draws on survey data to provide suggestive evidence on how the carbon labels influenced canteen guests in the field (e.g. visibility of the labels, effect of the labels on other attitudes).

36. Unfortunately, I can track individuals' payment card IDs from April to July 2022 or from August 2022 to January 2023, but not across the entire time frame.

Table 1.2. Field estimates of the effect of carbon labels on meat consumption

	Full sample		Restricted sample			
	Base	Date FE	Date FE	+Guest FE	ITT	Probit
Treatment restaurant × Label period	-0.02***	-0.02***	-0.04***	-0.02**	-0.04***	-0.02***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
Treatment restaurant × Post period	-0.07***	-0.07***	-0.11***	-0.04***	-0.12***	-0.07***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)
Treatment restaurant	-0.10***	-0.10***	-0.02	-0.08***	-0.01	-0.10***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)
Label period	0.01					0.01
	(0.00)					(0.00)
Post period	0.01*					0.01*
	(0.00)					(0.00)
Constant	0.51***	0.48***	0.42***	0.45***	0.42***	
	(0.00)	(0.01)	(0.02)	(0.02)	(0.02)	
Date fixed effects	No	Yes	Yes	Yes	Yes	No
Guest fixed effects	No	No	No	Yes	No	No
Guests control	6,924	6,924	875	875	875	6,924
Guests treated	2,810	2,810	339	339	339	2,810
Observations	120,082	120,082	25,372	25,372	25,372	120,082

Standard errors in parentheses

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

Notes: Note: Dependent variable: 0/1 indicator for consumption of the meat option. Col.(1) corresponds to Equation 1.2. The Constant term describes the proportion of meat meals sold in the Control canteens pre-intervention. Specifications (2)–(5) include date fixed effects to control for the daily changing offer of main meal components, which is common across canteens. The "Post period" and "Label period" indicators are thus dropped due to collinearity. Specifications (3)–(5) restrict the sample to canteen guests paying with their individual payment card, visiting the student canteen regularly pre-intervention (at least five times within four weeks) and at least once during the intervention phase, and pre-dominantly visiting the same canteen pre-intervention (at least 80% of pre-intervention visits to the same canteen). Specification (4) includes individual fixed effects, and specification (5) estimates ITT effects. Specification (6) reports the marginal effects of a probit regression of spec. (1). The standard errors of Col.(1)-(2) are robust. The standard errors of Col.(3)-(5) are clustered at the individual level.

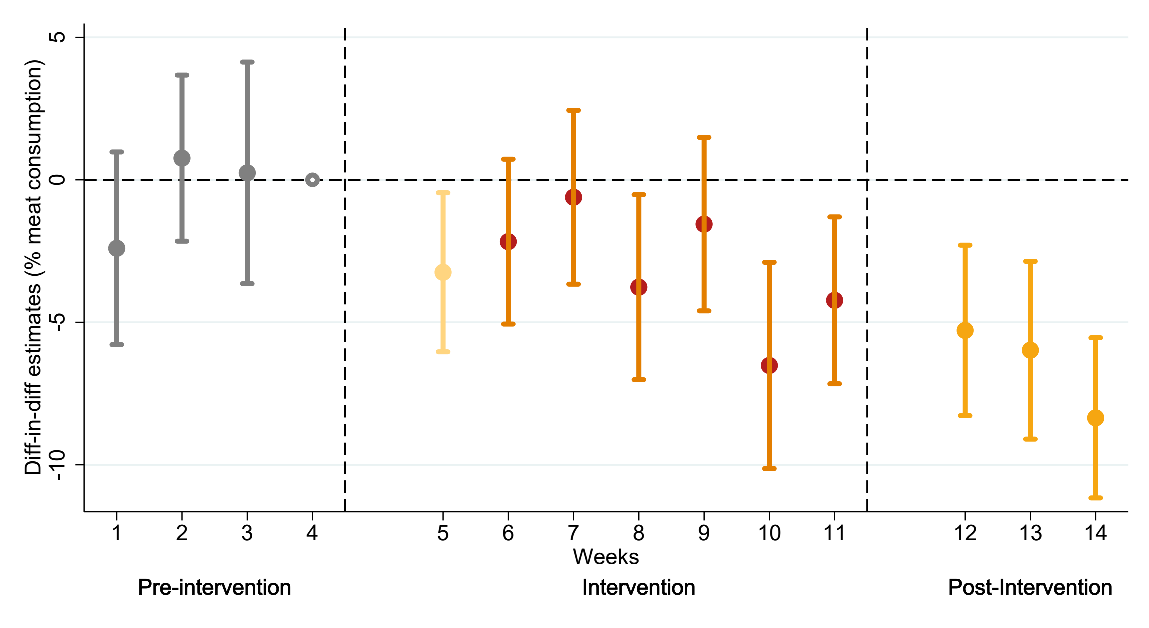


Figure 1.11. Event study: Difference in difference estimates

Note: Difference in difference estimates of the likelihood of consuming the meat option (in percentage points), using week 4 of the pre-intervention phase as a baseline. Weeks 1–4 constitute the pre-intervention phase, while weeks 5–11 constitute the intervention phase, and weeks 12–14 the post-intervention phase. The regression specification follows specification (1) in Table 1.2, but estimates weekly effects and controls for weekly time trends and student canteen offer, as detailed in regression table 1.C.1. Figure 1.C.2 shows an event plot without controls for student canteen offer. Week (5) is excluded from the main estimation in Table 1.2, because effects cannot be clearly attributed to the carbon labels, as described more in detail in Appendix 1.E. Bars indicate 95% confidence intervals.

1.4 Structural Model

To provide insights into the behavioral mechanisms driving consumers' responses to carbon labels, I introduce a simple discrete choice model of meal selection in this section. In section 1.6 I will report model parameters using data from Experiment 3 which will be described in Section 1.5.

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.[37]

1.4.1 Model

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

$$u(m) \ = \ v_m - p_m - \theta\gamma e_m. \tag{1.3}$$

Here, v_m is the *consumption utility* of meal m that is independent of emissions[38], p_m is the *price* of meal m, and e_m is the consumers *estimate of emissions* caused by meal m at the moment of choice.[39]

The *salience* of carbon emissions $\theta \in [0, 1]$[40] 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.

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

$$e_m^{\text{info}} \ = \ (1-\kappa)e_m^{\text{true}} + \kappa e_m^{\text{prior}}. \tag{1.4}$$

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.[41] If

37. 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.

38. 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 $v_m = \beta^T x_m + \varepsilon_m$.

39. 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.

40. 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)).

41. 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.

the consumer is *attentive* to emissions, this sets $\theta = 1$.[42] Introducing *carbon labels* makes the consumer both informed and attentive.

1.4.2 Identification of parameters

The setting of experiments 1 and 3 corresponds to a special case of the model with a binary choice set $\mathcal{M} = \{m, o\}$ with m being the meal option and o being the outside option of a cheese sandwich. The willingness to pay to exchange meals corresponds to

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

where the values of θ, e_m and e_o depend on the treatment condition. The parameters θ, γ, and κ can be estimated from Experiment 3 data. I directly elicit e_m^{prior} and e_o^{prior}, as participants guess carbon footprints at the start of the experiment. Further, the treatment conditions yield four equations with four unknowns[43] as follows. First, in the absence of any treatment (elicitation at baseline), participants' willingness to pay is

$$WTP^B \ = \ v_m - v_o - \theta\gamma(e_m^{\text{prior}} - e_o^{\text{prior}}) \tag{1.5}$$

where I assume $\theta \in [0, 1]$. The treatment condition, Attention directs participants' attention towards carbon emissions without providing information. Assuming this sets $\theta = 1$,

$$WTP^A \ = \ v_m - v_o - \gamma(e_m^{\text{prior}} - e_o^{\text{prior}}) \tag{1.6}$$

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 1.4. In Experiment 3, participants seeing carbon labels experience the Attention treatment on top of the Label treatment. This direction of attention has no effect on top of the direction of attention induced by the carbon labels,[44] and willingness to pay indicated in the Attention+Label condition can thus be described as

$$WTP^{A+L} \ = v_m - v_o - \gamma\Big(\kappa e_m^{\text{true}} + (1-\kappa)e_m^{\text{prior}}\Big) \tag{1.7}$$

where I assume $\kappa \in [0, 1]$. The treatment condition Attention+Offset removes the carbon emissions caused by both meal options. Assuming this sets $\theta = 1$, and $e_m = 0$:

$$WTP^{A+O} \ = \ v_m - v_o \tag{1.8}$$

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

43. I treat $v_m - v_o$ as a single parameter in the estimation, i.e. I only identify the difference and not the individual values of v_m and v_o. e_m^{prior},e_o^{prior} are directly elicited, and e_m^{true} and e_o^{true} are known.

44. Specifically, I assume an Attention+Label, Label and Attention treatment would all set salience $\theta = 1$, without any additional attention-directing effect occurring from a combination of treatments. This assumption is in line with a comparison of effect sizes across experiments 1 and 3, where I see similar treatment effects across the Label treatment in Experiment 1 and the Attention+Label treatment in Experiment 3. These are shown side-by-side in Tables 1.B.6 and 1.B.7.

1.5 Experiment 3: Behavioral channels

Experiment 3 provides framed field experiment evidence on the respective relevance of each of the two behavioral channels proposed in the theoretical model in Section 1.4. Subsection 1.5.1 describes the experimental design. Subsection 1.5.2 describes data and reduced-form results. Experiment 3 data is also used to estimate the parameters of the structural model, as detailed in 1.4.2. Results of the structural estimation are discussed in Section 1.6.

1.5.1 Experimental design

Overview

The theoretical framework in Section 1.4 proposes that carbon labels impact consumers by making consumers 1) informed, and 2) attentive. To investigate the relevance of each of the two channels, I conduct a framed field experiment similar to Experiment 1 apart from two key differences:

(1) To identify the extent to which an information effect drives consumers' reactions to carbon labels, I track participants' initial estimates of meals' carbon footprints. In the reduced-form analysis, I compare initial misperceptions with participants' reactions to carbon labels.
(2) To identify the extent to which an attention effect drives consumers' reactions to carbon labels, I include a separate experimental condition increasing attention towards carbon emissions without providing any information on carbon footprints. In the reduced-form analysis, I estimate treatment effects for this condition.

Experiment timeline

The experiment timeline is visualized in Figure 1.12. It proceeds very similarly to Experiment 1. Recall that in Experiment 1, experiment participants answer guessing questions on unrelated items after completing the four baseline purchase decisions (e.g. on the length of a popular running route in Bonn). In contrast, Experiment 3 participants do not answer these questions, but instead, guess the carbon footprints of different meals. These questions concern the four meals around which the meal purchasing decisions revolve, as well as six further meals (see Figure 1.14 for a list). Participants answer each of the ten guessing decisions on separate screens, shown to participants in a random order. On each screen, they are shown the emissions of the same reference example meal (Red Thai Curry with pork and rice, causes 1.7 kg of CO_2). This reference meal is not included in any willingness to pay elicitations. An example of a guessing screen is shown in Figure 1.13 and section 1.D.4 shows screenshots of the guessing instructions. The guessing questions are incentivized and timed as in Experiment 1.

The experiment then proceeds differently depending on the treatment group participants are assigned to by computer randomization. All participants are again asked to indicate their willingness to pay for the four meals, but the framing of the decision and some characteristics of the decision depend on the treatment condition:

- In the Attention condition, the willingness to pay elicitation is exactly as in the first, baseline elicitation. However, since participants completed the carbon footprint guessing task between the two elicitations, they have now spent time thinking about the issue of greenhouse gas emissions, and are thus arguably more attentive.

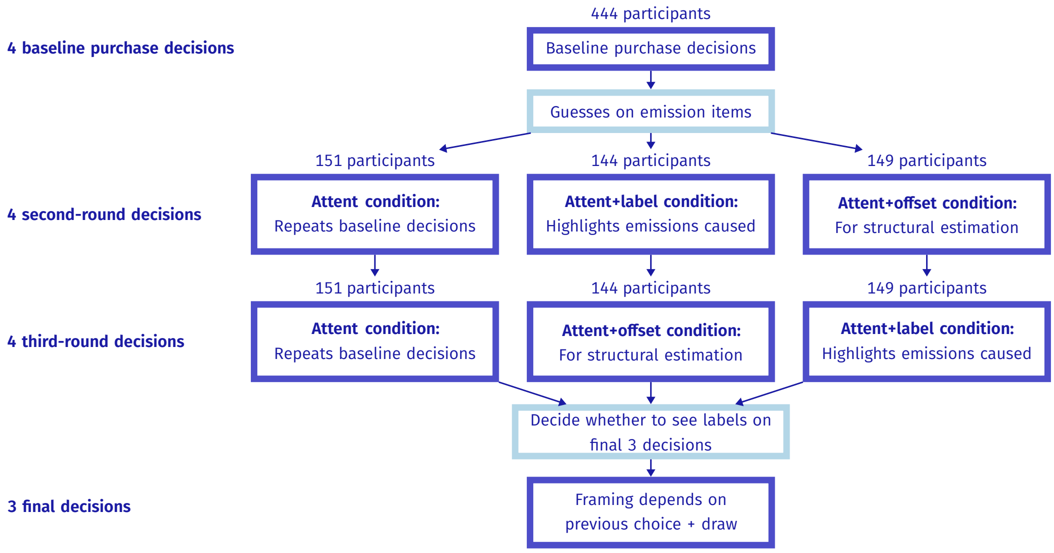


Figure 1.12. Experiment schedule and treatment groups

Note: Participants repeat the same four meal purchase decisions three times, with the decision framing differing across rounds. Treatments are described in more detail in the "Experiment timeline" paragraph above. The results of the ATTENTION+OFFSET condition are not further discussed in this section, details are described in section 1.D.3, and results are shown in Table 1.B.7. It is used as input for the structural estimation described in Section 1.6.

- In the ATTENTION+LABEL condition participants are now shown carbon labels when indicating their willingness to pay. An example is shown in Figure 1.4. They are thus attentive and informed.
- In the ATTENTION+OFFSET condition, participants are informed that the emissions caused by their lunch choice (be it the meal or the sandwich) will be offset.[45]

To increase power and elicit further information, participants' willingness to pay for the same four meals is elicited a third time[46], with partly changed treatment conditions:

- Participants previously in the ATTENTION+LABEL condition are now assigned to the ATTENTION+OFFSET condition and vice versa.
- Participants previously in the ATTENTION condition remain in the ATTENTION condition.

The experiment then proceeds as in Experiment 1. The design of the meal purchase decisions and their incentivization, as well as the incentivization of the elicitation of willingness to pay for seeing carbon labels, is as in Experiment 1.

Participants and set-up

444 experiment participants are recruited from the participant pool of the BonnEconLab, the behavioral experimental lab of the University of Bonn, to participate in one of 12 experimental

45. The results of the OFFSET condition are not further discussed in this section, details are described in section 1.D.3 and results are shown in Table 1.B.7. The OFFSET condition serves as input for the structural estimation described in Section 1.6, as detailed in Section 1.4.2.

46. In the analyses, I control for whether observations stem from a third-round elicitation. All the main results replicate including only data from the first two rounds.

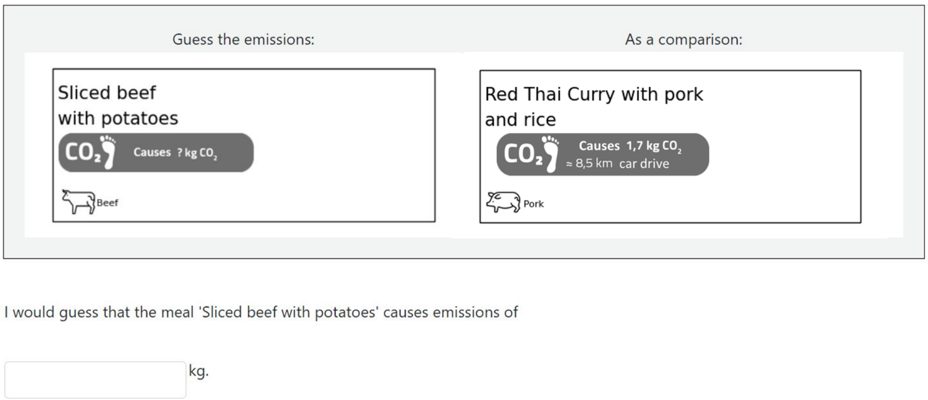


Figure 1.13. Example guessing questions

Note: After completing the baseline purchase decisions and before the second round of decisions, all participants answer incentivized guessing questions in which they estimate the carbon footprint of ten different meals. The carbon footprint of the meal Red Thai Curry with pork and rice is always shown as a reference meal. Participants do not learn the carbon footprint of any other meal at this stage of the experiment.

sessions taking place between the 22nd of June and the 8th of July 2021. I pre-registered experiment design, sample restrictions, the analysis shown in Figure 1.16 and Table 1.4, and, roughly, the structural estimation.[47] Participant invitation and experiment set-up are as in Experiment 1.

1.5.2 Data and reduced-form results on channels

I exclude the 3% fastest participants and participants not passing the comprehension check after five attempts, as pre-registered [48]. The remaining 444 participants are computer-randomized into treatments. Section 1.B.1 shows a randomization check. Participants are on average 26 years old, 55% are female, 70% are students and 24% are vegetarians. The sample is roughly representative of regular student canteen guests in terms of these characteristics, as discussed in Section 1.B.2, and results hold when restricting the sample to only students or only non-vegetarians, as shown in Section 1.B.5.

The effect of correcting misperceptions

This subsection provides reduced-form evidence on whether treatment effects are reconcilable with a correction of misperceptions about carbon impact being the main channel driving treatment effects. This analysis draws on Experiment 3 participants' guesses of the carbon footprints of different meals. Figure 1.14 displays how average guesses deviated for each of the meals. On average, participants rather underestimate emissions (green-colored dots) for high-emission meals and overestimate emissions for low-emission meals (red-colored dots). Section 1.B.12 shows further descriptive statistics on under- and overestimation of emissions, such as a comparison of the number of under- and overestimations by meal and participants, as well as the accuracy of the ranking of meals by carbon footprint which can be inferred from participants' guesses.

47. Schulze Tilling (2021b)
48. Schulze Tilling (2021b)

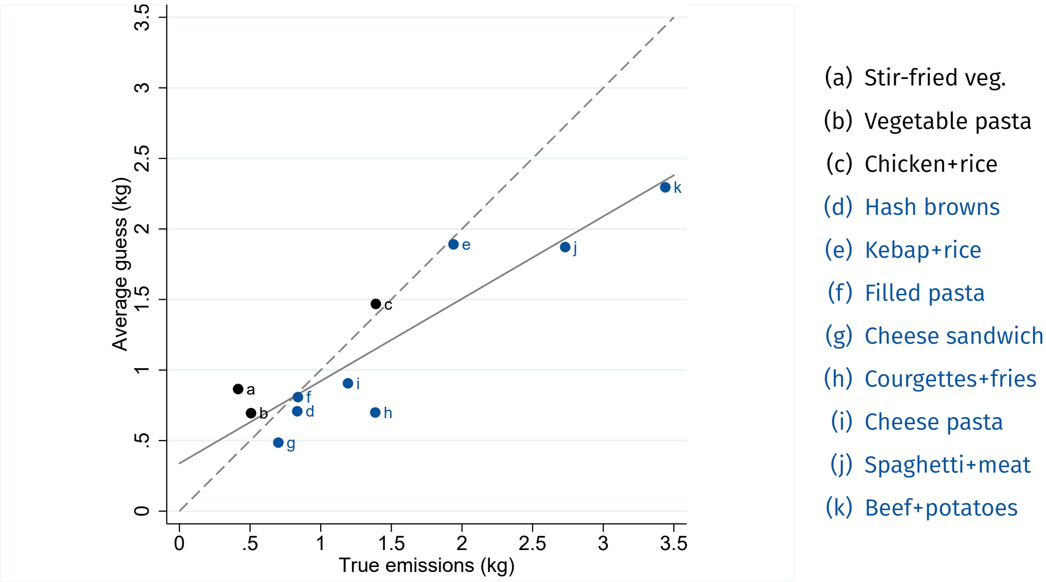


Figure 1.14. Average guess of the emissions caused by a given meal

Note: Guesses are plotted against calculated emissions. Guesses closer to the dashed line are closer to calculated emission values. Meals corresponding to orange scatter points are on average overestimated in their emissions, while meals corresponding to green scatter plots are on average underestimated. The dashed fitted line is described by $y = 0.39 + 0.57x$, with both the intercept and the coefficient significant at $p < 0.01$. Values are based on guesses made by the participants of Experiment 3. Further, the 71 participants in the "Control, then Control" group in Experiment 1 also estimated greenhouse gas emissions towards the end of the experiment (see Section 1.2 for details). This data is included in this graph, but not in any other analyses shown in this section. The meal "Spaghetti with meat" was only guessed by the 71 participants of Experiment 1 guessing emissions. For each meal, the 10% most extreme guesses (in terms of deviation from the true emission value) are dropped. This leaves a total of 4,261 observations made by 490 participants.

In the next step of the analysis, I combine individual and meal-specific treatment effects with participants' emission estimates for the respective meals. I estimate

$$Diff_{ijm} = \alpha + \delta_1 Under_{im} + ThirdRound_j + \varepsilon_{ijm} \qquad (1.9)$$

where $Diff_{ijm}$ describes the difference between willingness to pay of individual i in round j for meal m and individual i's baseline willingness to pay for meal m, as in Experiment 1.[49] I estimate this specification including only data from the ATTENTION+LABEL condition. Thus, my dependent variable directly captures subject- and meal-specific treatment effects for carbon labels. $Under_{im}$ is an indicator of whether the individual underestimated the difference in emissions between meal m and the cheese sandwich.[50] I calculate this indicator by comparing the difference between the individual's guess for the emissions of meal m and her guess for the cheese sandwich with the true difference in emissions. $ThirdRound_j$ is an indicator of whether it was the third round of decisions. I use this specification to examine in reduced-form the role of a correction of misperceptions in driving treatment effects. If a correction of misperceptions was the main channel driving effects, one would expect treatment effects to be proportional to a subjects' underestimation of emissions.

49. Please see Section 1.2.2 and 1.B.7 for details on this specification.

50. This refers to the signed, not the absolute difference. For example, if a meal causes 0.2 kg of emissions more than the cheese sandwich, and the participants estimate that the meal causes 0.3 kg of emissions less than the cheese sandwich, this is an underestimation of the difference in emissions.

Table 1.3, Spec. (1) shows the results of the OLS estimation of equation 1.9. If an individual underestimated the emissions of meal *m* relative to the cheese sandwich, presenting her with carbon labels on average leads to her decreasing her willingness to pay by an additional €0.13. This suggests that part of the effect of the labels can be explained through a correction in misperceptions on carbon impact: The labels inform participants that the meal has a higher relative carbon footprint than they previously expected, and they react accordingly. Spec. (2) in Table 1.3 does not group observations by previous under- or overestimation but instead regresses the change in willingness to pay on the degree of underestimation (in kg). This specification suggests that seeing labels on average decreases willingness to pay by €0.16, with an additional decrease of €0.07 for each kg by which emissions were underestimated.

The large negative constant term in both specifications is striking. In spec. (1), a decrease in willingness to pay of €0.10 is independent of a previous underestimation of emissions. Spec. (2) estimates a decrease in willingness to pay independent of previous underestimation of €0.16. Figure 1.15 shows average effects split by previous under- or over-estimation of emissions and visualizes that participants on average also significantly adjust their willingness to pay downward for meals for which they previously *over*estimated emissions. In these cases, the labels inform participants that the meal has a lower relative carbon footprint than they previously expected. If a correction of misperceptions were the sole effect induced by the label, one would expect participants to adjust their willingness to pay upwards in such a situation, and not downwards. The pattern we see in Figure 1.15 is thus evidence against this being the case and in favor of a second mechanism driving treatment effects.

I replicate the analysis in Figure 1.15 including (a) only individuals who did an above-average job at guessing the relative emissions of at least three of the four meals correctly (Figure 1.B.5 in the Appendix) and (b) only individuals who did an above-average job in guessing emission magnitudes (Figure 1.B.7 in the Appendix). Patterns look similar to Figure 1.15.

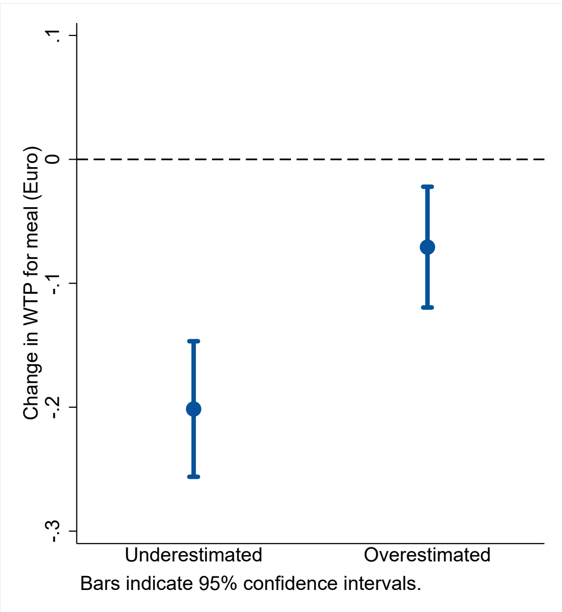


Figure 1.15. Within-subject change in willingness to pay for meals when shown carbon labels, depending on previous estimation

	Change in WTP compared to baseline	
	(1)	(2)
Underestimated emissions	−0.13***	
	(0.04)	
Underestimation (in kg)		−0.07***
		(0.02)
Control for third round	0.05	0.07
	(0.05)	(0.05)
Constant	−0.10***	−0.16***
	(0.04)	(0.03)
Participants	293	262
Obs. underestimate	555	515
Obs. overestimate	562	494
Observations	1,117	1,009

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

Table 1.3. Within-subject change in willingness to pay for meals when shown carbon labels, depending on participants' estimation of emissions

The effect of directing attention

This subsection provides reduced-form evidence on whether treatment effects are reconcilable with a direction of attention towards carbon emissions driving treatment effects. For the purpose of this analysis, I examine data from the Attention and Attention+Label conditions to estimate the magnitude of a possible attention effect. I estimate

$$Diff_{ijm} = \alpha + \beta_1 High_m + \beta_2 Low_m + \delta_1(Label_{ij} \times High_m) + \delta_2(Label \times Low_m) + ThirdRound_j + \varepsilon_{ijm} \tag{1.10}$$

where $Diff_{ijm}$ is defined as above, and $High_m$ and Low_m are indicators for meal m's footprint relative to the cheese sandwich, while $Label_{ij}$ is an indicator for whether individual i sees carbon labels in round j, additionally to being made attentive.

Results are shown in Table 1.4, Figure 1.16 illustrates average changes in willingness to pay for the Attention and the Attention+Label treatment. Simply directing attention towards carbon emissions decreases willingness to pay for high-emission meals by €0.08, on average. Providing labels on top of increasing attention leads to an additional decrease of €0.10 for high-emission meals. The decrease in willingness to pay for high-emission meals in the Attention condition is driven by decisions for which participants had a relatively good idea of the emissions caused by the meal in question. This is visualized in Figures 1.B.10 and 1.B.11 in the Appendix. These results highlight that an increase in attention alone can explain a large proportion of the treatment effect. Section 1.6 provides an estimation of the quantitative relevance of each of the two channels in driving the label's treatment effect.

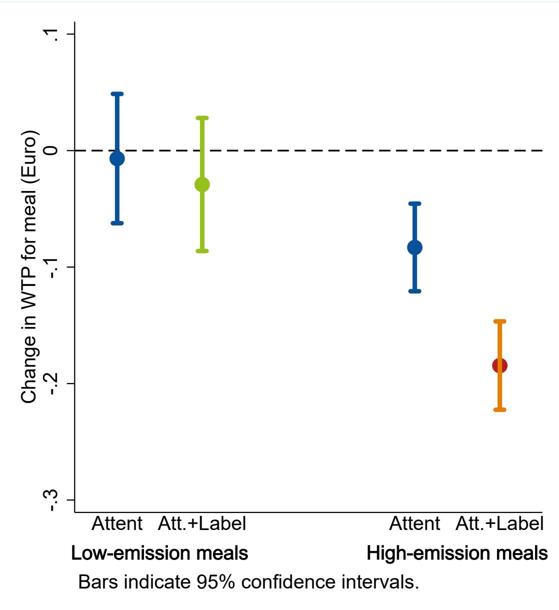


Figure 1.16. Within-subject change in willingness to pay for meals in the Attention vs. Attention+Label condition

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

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

Table 1.4. Within-subject change in willingness to pay for meals in the Attention vs. Attention+Label condition

1.6 Structural estimation

The structural estimation complements the reduced-form results from Section 1.5.2. Section 1.6.1 estimates the parameters of the theoretical model described in Section 1.4 using data from Experiment 3. Section 1.6.2 then simulates the effects of different types of interventions in the student canteen context and compares effects on carbon footprints and consumer welfare. This allows for a direct comparison of the importance of misperception-correcting and attention-directing effects of carbon labels identified in Section 1.5.2.

1.6.1 Results

I rewrite the four equations in Section 1.4.2 for the structural estimation, as shown in Section 1.A.3, and estimate parameters with GMM. I assume that the parameters γ, κ, and θ are homogeneous across participants.

Results are shown in Table 1.5, Col.(1). θ, the average attentiveness to greenhouse gas emissions in the absence of carbon labels, is estimated at 16%. This estimate implies that on average, individuals in my study react to the carbon footprint they perceive as if it was only 16% its size. The estimate is not significantly different from zero, implying that the true level of attentiveness might also be zero. This would imply that individuals do not react at all to the perceived carbon footprint in the absence of any intervention.

κ, the stickiness of the average consumers' prior estimate of a meal's carbon footprint, is estimated at 0.21 and insignificant. This suggests that individuals on average place a relatively large weight $(1-\kappa)$ on the carbon footprint information shown on the carbon labels when revising their carbon footprint estimate upon seeing the labels.

γ describes how the emissions of one kg of greenhouse gas emissions affect an individual's utility. This is estimated as a decrease in monetized utility of Euro 0.12 per kg of emissions caused by the meal chosen, i.e. individuals on average experience guilt equivalent to a monetary cost of 0.12 per kg of perceived emissions.

Columns (2)–(6) show that estimates are similar in alternative specifications of the model. In column (2), I re-estimate the model imposing that $\kappa = 0$, i.e. that individuals completely trust the emissions information. In column (3), I re-estimate the model imposing that $\theta = 0$, i.e. that individuals are completely inattentive to carbon emissions in the absence of an intervention. In column (4), I impose $\theta = \kappa = 0$. In column (5), I impose $\theta = 1$, assuming that consumers are fully attentive to carbon emissions, even in the absence of labels.

To provide an estimate of the effect carbon labels have on consumer welfare, I expand the theoretical model to make predictions on the labels' effect on consumer welfare, as detailed in Section 1.A.1. Essentially, I assume that consumer welfare is a function of the true—and not the perceived—emissions caused by the meal consumed. Thus, the carbon labels by construction increase consumer welfare by helping consumers make the choice that maximizes consumer welfare. I also assume that carbon labels have a psychological effect on consumers independent of their effect on consumption decisions. The sign of this fixed effect F is a priori undetermined and may reflect psychological costs or benefits accruing to consumers as a result of seeing the carbon labels. In the estimation shown in column (6) I add a fifth equation describing these effects to my GMM estimation and include participants' willingness to see or avoid labels on their final three consumption decisions in the estimation. Through the lens of the model, I interpret these values as an estimate of the labels' effect on consumer welfare, taking a similar interpretation as e.g. Allcott and Kessler (2019) and Butera et al. (2022). This allows me to estimate F. I estimate this figure at the monetary equivalent of €0.21 and significantly different from

Table 1.5. Structural estimates of model parameters

	(1)	(2)	(3)	(4)	(5)	(6)
Theta	0.16	0.03				0.18
	(0.18)	(0.17)				(0.17)
Gamma	-0.12^{***}	-0.10^{***}	-0.10^{***}	-0.10^{***}	-0.12^{***}	-0.12^{***}
	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
Kappa	0.21		0.12		0.12	0.23
	(0.20)		(0.19)		(0.21)	(0.20)
F						0.21^{***}
						(0.01)
Observations	3,216	3,216	3,216	3,216	3,216	3,216

Standard errors in parentheses

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

Notes: Note: Analysis is based on data from Experiment 3. For each meal, the observations corresponding to the 10% most extreme guesses (in terms of deviation from the true emission value) are dropped. Regression does not include a constant, since the estimation follows the model outlined in Section 2. Column (1) shows the main estimation, based on equations 1.A.9, 1.A.8, 1.A.10. Columns (2)–Column (7) each modify the model in Column (1) as follows: Column (2) imposes $\kappa = 0$. Column (3) imposes $\theta = 0$. Column (4) imposes $\theta = \kappa = 0$. Column (5) imposes $\theta = 1$. Column (6) includes equation 1.A.7 in the estimation.

zero, suggesting that consumers on average experience a psychological benefit from seeing the carbon labels independent of their effect on consumption decisions.

1.6.2 Comparison of interventions based on estimated parameters

In the model described in Section 1.4, introducing carbon labels affects consumers by making them both informed and attentive. Using estimated parameters, I can compare the importance of each of these two effects in driving consumers' responses to carbon labels. I simulate how experiment participants would react to different interventions in the student canteen context: 1) a KNOWLEDGE intervention making them informed, but not attentive, 2) an ATTENTION intervention making them attentive, but not informed, and 3) a LABEL intervention making them both attentive and informed. This simulation is based on participants' tastes for different student canteen meals as elicited in Experiment 3, participants' prior estimates of emissions as elicited in Experiment 3, my estimates of θ, γ, and κ which I assume are homogenous across participants, the model specification shown in Section 1.4, and some assumptions on what constitutes a typical student canteen offer and pricing structure. These assumptions and the simulation are discussed in more detail in Section 1.A.4. Table 1.6 shows simulation results.

For all three interventions, the interventions do not impact the vast majority of consumption decisions, with 98% to 99% of consumption decisions not affected by the interventions. This intuitively makes sense—Interventions will typically only affect decisions that were at the margin, to begin with. This is in line with my findings from the natural field experiment (Experiment 2) in which the labeling intervention also affects only 2% of consumption decisions, and correspondingly leaves 98% of consumption decisions unaffected. Participants' valuation for the student canteen meals in Experiment 3 is, in over 70% of cases, lower than the student canteen price. This is also in line with observations from the field experiment that an average student canteen guest does not visit the student canteen every day. An average student canteen guest visits the student canteen 20 times during the 14-week sample period, i.e. on 29% of possible occasions. On the remaining 71% of occasions, he will also opt towards an alternative lunch (e.g. taking a sandwich with him).

The ATTENTION, KNOWLEDGE, and LABEL intervention all decrease the consumption of the meat option. In the ATTENTION and KNOWLEDGE intervention, consumption of the cheese sandwich increases. In the LABEL intervention, consumption of the cheese sandwich and the vegetarian option increases. The ATTENTION intervention decreases the carbon footprint of an average meal by 27 grams, while the KNOWLEDGE intervention decreases carbon by 4 grams, and the LABEL intervention decreases carbon by 34 grams. The average effect of the ATTENTION intervention is thus around 7-fold that of the KNOWLEDGE intervention. Further, there are some synergies between the ATTENTION and KNOWLEDGE intervention, leading to the LABEL intervention producing a greater decrease in emissions than the sum of its parts.

In the extension of my model to consumer welfare specified in Section 1.A.1, consumer welfare resulting from a meal choice is a function of the true—and not the perceived—emissions resulting from the meal choice. Carbon labels thus, by moving perceived emissions closer to true emissions, increase the likelihood of a consumer choosing the option maximizing his welfare. The final four columns of Table 1.6 estimate how consumer welfare changes accordingly under each of the interventions. Importantly, these estimates account for the fact that a change in meal choice also leads to a change in consumption utility. For example, if a consumer switches from a meat to a vegetarian meal as a result of the label, but enjoys the taste of the meat meal more, the calculations account for this. They are thus considerably lower than a mere multiplication of the average reduction in greenhouse gas emissions with the average guilt perceived per kg of emissions.

I estimate that carbon labels improve consumer welfare by the monetary equivalent of 0.18¢ per choice (averaging over choices affected and not affected by the labels), or 10¢ on average for every choice affected by the labels. Synergies between the ATTENTION and KNOWLEDGE intervention are more sizable here, with the effects of the other two interventions merely summing to 0.1¢. Section 1.A.5 examines the distribution of welfare effects. Both the ATTENTION and the KNOWLEDGE intervention in some cases result in considerable decreases in consumer welfare. This can be the case if a consumer with large misperceptions of carbon impact is made attentive, or if a consumer who generally overestimates emissions and is very inattentive towards emissions is made knowledgeable of emissions. Welfare changes are thus in both cases more dispersed than for the LABEL intervention.

To provide comparability with other possible policy interventions, I also estimate the impact of a carbon tax of €120 per ton.[51] I assume that the proceeds from this tax are uniformly redistributed among individuals.[52] I estimate that such a measure would lead to a higher decrease in average carbon footprint than the carbon labels—42 g per meal vs. 34 g per meal—and a similar increase in consumer welfare. I also examine a meat ban, which would lead to much higher emissions reductions (147 g per meal) but also to a higher loss in consumer welfare. Importantly, these estimations assume that restaurant guests have no choice but to eat at the

51. I use a value of €120 per tonne for comparability with framed field experiment 1 results.

52. Whether and how the tax proceeds are redistributed among consumers does not affect the amount of greenhouse gas emissions avoided, but affects consumer welfare estimations.

Table 1.6. Estimated effect of different policies in the student canteen

	# of choices			Δ GHGE	Δ consumer welfare			
Intervention	sandwich	veg.	meat	Average	Average	SD	Min	Max
None	73.1%	18.1%	8.8%					
Attention	74.4%	18.1%	7.4%	-.0267	.0010	.0160	-.0849	.2456
Knowledge	73.8%	18.1%	8.1%	-.0036	.0001	.0043	-.0657	.0583
Labels	74.1%	18.6%	7.3%	-.0338	.0018	.0164	-.0022	.2456
Carbon tax	74.4%	18.7%	7%	-.0423	.0017	.0653	-.3130	.2626
Meat ban	78.3%	21.7%		-.1473	-.0350	.1728	-1.3935	.2456
Beef ban	78.3%	20.4%	1.4%	-.0800	-.0128	.1047	-1.3827	.2456

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

student canteen. The emission savings are thus rather an upper bound estimation and welfare effects a lower bound estimation.

1.7 Consumer preferences for the presence of carbon labels

This section discusses experimental evidence of consumers' preferences for the presence of carbon labels in their consumption decisions. Section 1.7.1 discusses evidence from experiments 1 and 3, and Section 1.7.2 discusses evidence from Experiment 2. Section 1.7.3 discusses possible determinants of consumers' willingness to see or avoid carbon labels.

1.7.1 Evidence from the framed field experiments

In both framed field experiments, participants indicate their willingness to pay for carbon labels being present or absent during their final set of consumption decisions. These elicitations are incentivized as described in Section 1.2. The frequency distribution of willingness to pay values is visualized in Figure 1.17. About 50% of participants have a willingness to pay of 0, meaning they have no strong preference for the presence or absence of labels. Less than 5% have a negative willingness to pay, meaning they prefer the labels being absent. The remaining participants are willing to pay for the presence of labels, with 21% of the sample willing to pay €0.50 and above. Values barely differ between treatment groups, although willingness to pay seems to be slightly higher among those who have not yet seen labels in the course of the experiment, as shown in Table 1.B.32.

Table 1.7 shows a correlation analysis between willingness to pay for the presence of carbon labels and individual characteristics. Willingness to pay for seeing labels is strongly positively correlated with participants' approval of carbon labels being shown in the student canteen and participants' interest in using this information. It is also weakly positively correlated with participants' perceived strength of social norms for avoiding carbon emissions in food consumption, as measured by adapting the procedure developed by Krupka and Weber (2013). Further, participants' self-control in eating behavior (as elicited using the questionnaire developed by Haws, Davis, and Dholakia (2016)) is not correlated with willingness to pay to see emission values. This implies that the concerns Thunström (2019) identified for calorie labels—They might take a particularly high emotional toll on individuals with low self-control—do not seem to play a meaningful role in the context of carbon labels.

Remarkably, participants' reactions to carbon labels are strongly positively correlated with their willingness to pay for the presence of carbon labels (Table 1.B.33). Participants who react strongly to the labels also have a stronger preference for seeing them in their decisions.

1.7.2 Evidence from the natural field experiment

After the natural field experiment is completed, student canteen guests are asked in a follow-up survey whether they would like the labels to be installed permanently. The details of this survey and the measures I took to limit non-response bias are described in Section 1.D. 73% of the 234 participants are in favor of installing the labels permanently, 18% are not sure, and 9% against the measure. A revenue-neutral carbon tax of an unspecified amount[53] , in contrast, is only favored by 60% of students, while 14% do not know and 26% are against. Carbon labels thus seem to enjoy greater support than carbon taxes, making an implementation more feasible.

1.7.3 Discussion of possible preference drivers

Subsections 1.7.1 and 1.7.2 show evidence of a positive reception towards carbon labels. In the framed field experiment setting, 95% of participants would either like to see the carbon labels or are indifferent. In the natural field experiment, 91% of survey participants expressed a preference for or no clear opinion towards carbon labels. Importantly, these indications can not be described as "cheap talk" in either of the two settings. In the framed field setting, participants' responses directly influence the presentation of carbon labels on their final choices and affect their compensation in the experiment. In the natural field setting, survey participants expect the survey results to be communicated to the student canteen and to impact the future presence of the carbon labels.

These results raise the question of why consumers have a preference for the presence of carbon labels and how we should interpret these results. The evidence suggests that consumers benefit from the presence of carbon labels. Instead of incurring a net psychological cost from the labels, they rather seem to derive a net psychological benefit. Why might this be the case? First, consumers might find the information itself intriguing, offering insights into the environmental impact of different food choices. Second, consumers might notice that they are more prone to take the environmentally friendly option in the presence of carbon labels and choose to see carbon labels as a type of commitment device. The carbon labels then remind them of self-set goals to decrease emission-heavy consumption. Third, consumers might appreciate that the labels help them make the environmentally-friendly choice, providing them to experience a feeling of "warm glow" or avoid a feeling of "cold prickle". Fourth, for those already inclined towards eco-friendly choices the presence of the labels might amplify the experience of warm glow. All of these four dynamics fit well into the model's framework. The first two factors relate to costs or benefits created by the labels independent of their impact on consumption behavior. The third factor relates to increased utility from label-influenced choices, while the fourth factor

53. Specifically, I asked survey participants if they would be in favor of canteen prices being in line with the carbon labels (green-labeled meals being least expensive, red-labeled meals being most expensive).

relates to increased attentiveness towards emissions, θ, increasing the experienced intensity of warm glow for carbon-friendly (or cold prickle for emission-heavy) choices.

A fifth factor might be that consumers see benefits in other consumers seeing the carbon labels. I do not consider this factor in the structural model, since participants' decisions in the framed field setting are anonymous and their decision on the presence or absence of carbon labels only influences whether they see the labels, and does not impact anyone else. The setting is thus not apt to estimate such a parameter. In the natural field setting, however, a consumer might be in favor of carbon labels since he believes that this will lead to other consumers changing their behavior. Consumers might believe that the carbon labels make the social norm more salient to themselves and to other consumers. In many cases, consumers might derive utility from behaving according to the norm. Thus, all of these interpretations speak towards the carbon labels creating a benefit for consumers.

One might argue that behaving according to the norm can also, in some cases, decrease utility for consumers. Consumers might feel observed in their choices and thus—out of social pressure—make the socially desirable choice, although it is not in line with their true interests. In the framed field experiment, this is a bit difficult to imagine: Participants do not know the experimenter, make their choices anonymously, and suffer real consequences if their choice is not in line with their true interests. This interpretation would also imply that they are willing to pay money to see the carbon labels on their final choices to "look good" in front of an experimenter they do not know and who will only look at their choices in an anonymized manner. In the natural field experiment, the interpretation seems more plausible: The carbon labels make the socially desirable choice (as designated by the student canteen) visible to all canteen guests and guests may fear being judged by other canteen guests if they choose differently. They may be making consumption choices that do not maximize their true utility, and suffer a utility loss through the labels. In the survey I conduct in the field, they however have the chance to change their fortune: By indicating a preference against the carbon labels being continued, they can alleviate the social pressure and return to the consumption choices maximizing their true utility. Thus, if these forces were driving consumers' responses, I would expect a sizable proportion of canteen guests to voice their disapproval of the labels. It is unlikely that canteen guests are willing to incur great costs (since they expect the survey results to influence canteens' choices), only to "look good" in front of an experimenter they do not know and who will only look at their choices in an anonymized manner. There is a large discrepancy between this prediction and the survey responses I observe. Only 9% of survey respondents communicate that they would not like the carbon labels to be installed permanently. I thus do not consider it likely that carbon labels mainly affect the average consumer by moving him towards a socially desirable choice that is not in line with his true interests. However, the above interpretation might of course align with the experience of a select few consumers.

In general, my results point towards the carbon labels on average creating a net psychological benefit to consumers. However, it is important to acknowledge that, although a large majority of individuals seem to incur psychological benefits with carbon labels, there is a small proportion of individuals who prefer to avoid carbon labels. If a policymaker is strongly concerned about these individuals, it might be worthwhile to explore technological solutions that allow consumers to decide whether or not to see carbon labels.

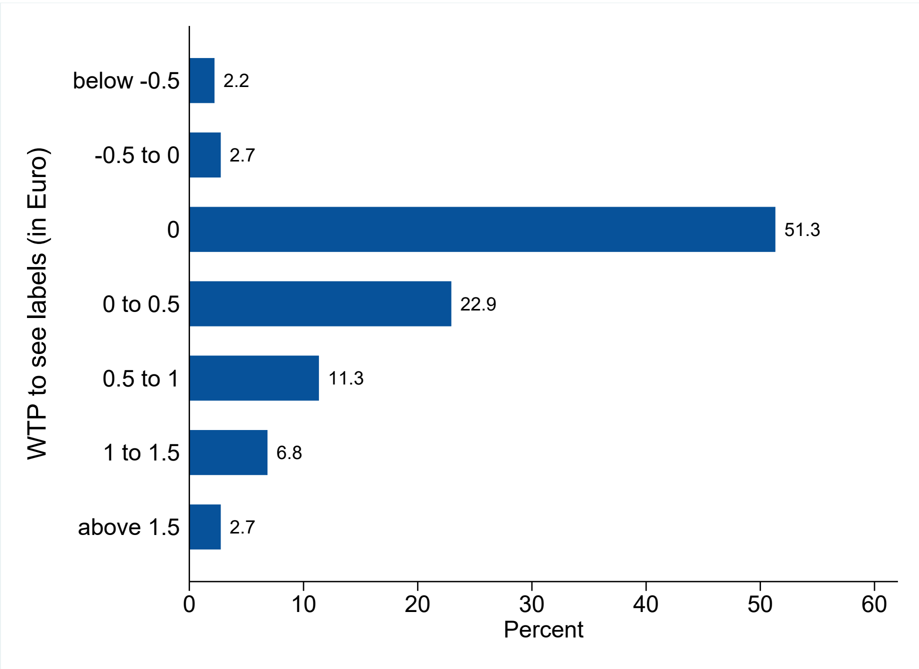


Figure 1.17. Distribution of willingness to pay indicated to see carbon labels on the final three consumption decisions

Note: In Euro. Based on Experiments 1 and Experiment 3. Includes data from all 733 participants.

Table 1.7. Correlation between willingness to pay for seeing carbon labels and individual characteristics

	WTP for the presence of carbon labels				
	(1)	(2)	(3)	(4)	(5)
Perceived strength of social norms	0.01*				
	(0.01)				
In favor of labels in student restaurant		0.03***			
		(0.01)			
Self-reported willingness to use info			0.03***		
			(0.01)		
Self-reported confidence in own knowledge				−0.01	
				(0.01)	
Eating self-control					0.00
					(0.01)
Constant	0.15***	−0.03	0.03	0.18***	0.21***
	(0.03)	(0.06)	(0.04)	(0.02)	(0.02)
Observations	732	732	732	732	732

Standard errors in parentheses

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

Notes: Note: Dependent variable: Willingness to pay for seeing labels (in Euro) for the final three consumption decisions. "In favor of labels in student canteen" is measuring using approval of the statement "I would appreciate if the student canteen would introduce such a measure". "Self-reported willingness to use info" is measured using approval of the statement "I would include this information in my decision.". "Self-reported confidence in own knowledge" is measured with two questions: (1) approval of the statement "I already know without labels which emissions are caused by different meals.", and (2) "I think this information will partially surprise me." The perceived strength of social norms is measured using the procedure developed by Krupka and Weber (2013). Eating self-control is measured using the questions developed by Haws, Davis, and Dholakia (2016).

1.8 Discussion

This paper provides evidence from the student canteen setting that carbon labels causally impact consumption behavior, estimating the effectiveness of carbon labels in reducing emissions as similar to that of a carbon tax of €120 per tonne. The labels primarily impact consumers by directing their attention toward carbon emissions, and their presence on average increases consumer welfare.

These results speak towards attention frictions playing an important role in impeding consumers from behaving in a carbon-friendly manner. While a lack of attention has been shown to play an important role in impeding sustainable behavior in the energy and resource consumption context (Allcott and Taubinsky, 2015; Taubinsky and Rees-Jones, 2018; Tiefenbeck et al., 2018), this is a new result in the food consumption context. The food consumption context differs from the resource consumption context in two ways. First, reducing energy and resource consumption usually also creates financial benefits for consumers, while reducing emissions in food consumption does not. Second, resource consumption is a continuous choice while food consumption is a discrete choice decision. Results showcase that increasing attention can also be effective in a discrete choice context, and open the door to examining related discrete choice consumption contexts.

While results corroborate previous findings on the emission-saving potential of attention-directing behavioral interventions in general, they also support the potential of carbon labels in particular. Findings are likely also relevant for related food contexts, such as corporate canteens or grocery shopping. Identifying carbon labels as a promising policy tool is especially relevant in the food sector. Carbon taxes for this sector are still widely uncommon (e.g. the agricultural sector is excluded from the EU-ETS trading scheme) and Dechezleprêtre et al. (2022) identify agriculture-targeted policies as among the least popular policies to reduce carbon emissions. In such a setting, alternative policy tools are especially called for.

Further, there are other discrete choice contexts in which the carbon footprint caused by different items could be calculated and labeled, e.g. shopping for toiletries or clothing. Future research could test the effectiveness and consumer welfare impact of carbon labels in these other consumption contexts, and also among other target populations. One way of doing so would be an adaptation of the design of Experiment 1—we would then be able to compare effects across domains and populations.

Further research would also be beneficial to assess whether carbon labels affect consumers in other domains apart from the target behavior. Suggestive evidence from a field survey I conducted to accompany the natural field experiment (Experiment 2) provides no evidence of the carbon labels affecting consumers' attitudes towards political measures to decrease carbon emissions (see section 1.E.6 and Table 1.C.6). However, spillovers may appear if labels are installed over longer time periods, or spillovers might affect other domains. Since the carbon labels mainly affect behavior by directing attention, attentional spillovers as described by Nafziger (2020) are also thinkable.

Appendix 1.A Additional material on theoretical model and structural estimation

1.A.1 Extension of the model to consumer welfare impact

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

$$u^T(m) \;=\; v_m - p_m - \gamma e_m^{\text{true}} \tag{1.A.1}$$

Accordingly, carbon labels increase the likelihood of the consumer choosing the meal m that maximizes her true utility.[54] 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

$$u(P) \;=\; u^{True}(m^L) - u^{True}(m^{prior}) + F \tag{1.A.2}$$

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.

1.A.2 Quantification of welfare impact in the experiment setting

In the experiment setting, the mere act of showing example carbon labels to participants and asking consumers for their willingness to pay to see carbon labels in a decision will make participants attentive to emissions (provided they have not already been made attentive of emissions earlier in the experiment). Thus, the difference in utility consumers' experience in the presence of carbon labels, $u(P=1)$ relative to utility in the absence of labels, $u(P=0)$, is

$$u(P \;=\; 1) - u(P \;=\; 0) \;=\; u^{True}(m^{*L}) - u^{True}(m^{*A}) + F \tag{1.A.3}$$

and the true utility the consumer reaps from meal m in the experiment context is

$$u^{True}(m) \;=\; v_m - o_m - \gamma(e_m^{\text{true}} - e_o^{\text{true}}) - p_m - p_o \tag{1.A.4}$$

In the experiment setting, there are only two possible cases in which $u^T(m^{*L}) - u^T(m^{*A}) \neq 0$:

(1) The WTP which the participant indicates when seeing labels, WTP^{A+L} is higher than the price $p_m - p_o$ to receive meal m rather than the outside option o, but $WTP^A < p_m - p_o$

(2) The WTP which the participant indicates merely attentive, WTP^A is higher than the price $p_m - p_o$ to receive meal m rather than the outside option o, but $WTP^{A+L} < p_m - p_o$

54. 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").

In the experiment context, equation 1.A.3 thus transforms to:

$$\begin{aligned} u(P=1)-u(P=0) &= \mathbb{1}\Big(WTP^{A+L} \geq p_m - p_o\Big) \\ &\Big(v_m - o_m - \gamma(e_m^{\text{true}} - e_o^{\text{true}}) - E[p_m - p_o | WTP^{A+L} \geq p_m - p_o]\Big) \\ -\mathbb{1}\Big(WTP^{A} \geq p_m - p_o\Big)&\Big(v_m - o_m - \gamma(e_m^{\text{true}} - e_o^{\text{true}}) - E[p_m - p_o | WTP^{A} \geq p_m - p_o]\Big) + F \end{aligned} \tag{1.A.5}$$

When the participant indicates her willingness to pay for the presence of labels, she weights each event with the probability of it occurring:

$$\begin{aligned} WTP^P = &\ Prob\Big(WTP^{A+L} \geq p_m - p_o\Big)\Big(v_m - o_m - \gamma(e_m^{\text{true}} - e_o^{\text{true}}) - E[p_m - p_o | \hat{V}_m^L \geq p_m - p_o]\Big) \\ &- Prob\Big(WTP^{A} \geq p_m - p_o\Big)\Big(v_m - o_m - \gamma(e_m^{\text{true}} - e_o^{\text{true}}) - E[p_m - p_o | \hat{V}_m^A \geq p_m - p_o]\Big) \\ &+ F \end{aligned} \tag{1.A.6}$$

In the experiment, relative meal prices $p_m - p_o$ are drawn from a uniform distribution, with each value between -3 and 3 being equally likely, in five-step intervals. Thus, $Prob(p \leq x) = (x+3)/6$. Similarly, $E[p|p \leq x] = (x-3)/2$. Inserting this above:

$$\begin{aligned} WTP^P = &\ \Big((WTP^{A+L}+3)/6\Big)\Big(v_m - o_m - \gamma(e_m^{\text{true}} - e_o^{\text{true}}) - (WTP^{A+L}-3)/2\Big) \\ &- \Big((WTP^{A}+3)/6\Big)\Big(v_m - o_m - \gamma(e_m^{\text{true}} - e_o^{\text{true}}) - (WTP^{A}-3)/2\Big) + F \end{aligned} \tag{1.A.7}$$

When I ask experiment participants for their willingness to pay for the presence of labels on their three final meals, I do not tell them in advance which meals these will be, and only tell them that these will be three new meals which they have not seen in the experiment previously. Thus, participants are not able to compute the first two terms of the above equation for the new meal. Participants in the ATTENT+LABEL condition have, however, seen carbon labels on the meals shown to them in the second or third round of their main choices. I would assume that participants indicate their willingness to pay for the presence of labels somewhat along the lines of "Based on the value I previously derived from the carbon labels, my willingness to pay to see carbon labels on a choice is XYZ." The willingness to pay for the presence of labels thus enters my estimation as a type of ex-post willingness to pay to see carbon labels on the four main meal decisions, for participants in the ATTENT+LABEL condition.

Participants in the ATTENTION condition have not seen emission labels before indicating their willingness to pay for the presence of labels, and would thus have to form a less informed expectation over the first two terms in 1.A.7. I thus do not include them in the main estimation of F (Col.(6) in Table 1.5 in the main text and Table 1.A.1). Col. 7 in Table 1.A.1 includes these observations and finds estimates similar to the previous specification. Table 1.B.32 shows that the average willingness to pay indicated for the presence of carbon labels does not differ across treatments.

1.A.3 Equations for structural estimation

$$WTP^{A+L} - WTP^{B} = \gamma(e_{im}^{prior} - e_{io}^{prior})(\kappa - \theta) + \gamma(e_{im} - e_{io})(1-\kappa) \quad (1.A.8)$$

$$WTP^{A} - WTP^{B} = \gamma(e_{im}^{prior} - e_{io}^{prior})(1-\theta) \quad (1.A.9)$$

$$WTP^{A+O} - WTP^{A+L} = -\gamma(e_{im}^{true} - e_{io}^{true})(1-\kappa) - \gamma(e_{im}^{prior} - e_{io}^{prior})\kappa \quad (1.A.10)$$

$$\begin{aligned} WTP^{P} = & \Big((WTP^{A+L}+3)/6\Big)\Big(v_m - o_m - \gamma(e_m^{\text{true}} - e_o^{\text{true}}) - (WTP^{A+L}-3)/2\Big) \\ & -\Big((WTP^{A}+3)/6\Big)\Big(v_m - o_m - \gamma(e_m^{\text{true}} - e_o^{\text{true}}) - (WTP^{A}-3)/2\Big) + \hat{F} \end{aligned} \quad (1.A.11)$$

Table 1.A.1. Structural estimates of model parameters including data on willingness to pay for the presence of carbon labels

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Theta	0.16	0.03				0.18	0.12
	(0.18)	(0.17)				(0.17)	(0.20)
Gamma	-0.12^{***}	-0.10^{***}	-0.10^{***}	-0.10^{***}	-0.12^{***}	-0.12^{***}	-0.11^{***}
	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Kappa	0.21		0.12		0.12	0.23	0.17
	(0.20)		(0.19)		(0.21)	(0.20)	(0.22)
F						0.21^{***}	0.20^{***}
						(0.01)	(0.01)
Observations	3,216	3,216	3,216	3,216	3,216	3,216	3,216

Standard errors in parentheses

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

Notes: Analysis is based on data from Experiment 3. For each meal, the observations corresponding to the 10% most extreme guesses (in terms of deviation from the true emission value) are dropped. Regression does not include a constant, since the estimation follows the model outlined in Section 2. Column (1) shows the main estimation, based on equations 1.A.9, 1.A.8, 1.A.10. Columns (2)–Column (7) each modify the model in Column (1) as follows: Column (2) imposes $\kappa = 0$. Column (3) imposes $\theta = 0$. Column (4) imposes $\theta = \kappa = 0$. Column (5) imposes $\theta = 1$. Column (6) includes equation 1.A.7 in the estimation. Column (7) includes values for willingness to pay for the presence of labels indicated by participants in the Attention treatment.

1.A.4 Details on the comparison of interventions

I use Experiment 3 data to deduce how experiment participants would make typical student canteen choices in the absence of any intervention, as well as under different interventions. Based on the willingness to pay which participants indicated for each of the four meals at baseline, I can deduce how experiment participants would make their consumption choice in a typical canteen setting, i.e. with a meal offer and pricing structure typical at the university of Bonn. In the next step, I additionally make use of the results from Col.(6) of the structural estimation shown in Table 1.5, participants' emission guesses, and true emissions to estimate the consumption choices that participants would make if they were experiencing an intervention.

I assume the following meal offer and pricing structure for the simulations. Specifically, I simulate how participants would choose on the following four exemplary days:

- Day 1: Canteen offers Filled courgettes with potato croquettes or Chicken Schnitzel with rice at a price of €3.05 each, as well as a cheese sandwich at a price of €1.50
- Day 2: Canteen offers Filled courgettes with potato croquettes or Beef ragout with potatoes at a price of €3.05 each, as well as a cheese sandwich at a price of €1.50
- Day 3: Canteen offers Italian vegetable ragout with pasta (€2.75) or Chicken Schnitzel with rice (€3.05), as well as a cheese sandwich at a price of €1.50
- Day 4: Canteen offers Italian vegetable ragout with pasta (€2.75) or Beef ragout with potatoes (€3.05), as well as a cheese sandwich at a price of €1.50

I chose the meals because these are the four meals I use in the baseline purchase decisions in Experiment 3 and I know participants' taste preferences for these meals accordingly.[55] The student canteen in Bonn always offers one meat meal and one vegetarian meal, so I designed the four days to cover all possible combinations of the four meals. The four meals are regularly offered in the student canteen, and I use the student canteen's prices for these meals in the simulations. Further, the student canteen always offers cheese sandwiches and prices these at €1.50, so this is included on all days as a third option.

To calculate choices in the absence of any intervention—line 1 in Table 1.6—I calculate for each participant and day the difference in the participant's baseline willingness to pay for the option and the canteen price. Since willingness to pay for each meal is in the experiment elicited relative to a cheese sandwich, I add €1.50 to all relative willingness to pay values. €1.50 is the price of a cheese sandwich in the canteen and also the average value of what experiment participants indicated in a hypothetical question as the amount they were willing to pay to receive a cheese sandwich. Accordingly, I set the willingness to pay for a cheese sandwich of all participants to €1.50. I assume the participant would decide to take the option with the largest difference between the two, allowing her to realize the highest consumer surplus.

To calculate choices with an intervention solely increasing attention—line 2 in Table 1.6—I use participants' baseline willingness to pay and prior emission estimates as well as the estimated model parameters to calculate an ATTENTION willingness to pay for each participant and meal, according to equations 1.5 and 1.6. I then simulate meal choices as in the previous calculation.

To calculate choices with an intervention solely increasing knowledge—line 3 in Table 1.6—I use participants' baseline willingness to pay, prior emission estimates, and estimated model parameters to calculate a KNOWLEDGE willingness to pay for each participant and meal. This is based on 1.5 and equation 1.A.12 below. A KNOWLEDGE treatment is assumed to lead to the consumer updating her emissions estimate according to 1.4 without directing attention.

$$WTP^K = v_m - v_o - \theta\gamma(e_m^{\text{info}} - e_o^{\text{info}}) \tag{1.A.12}$$

I then simulate meal choices as in the previous calculation.

To calculate choices with a carbon label—line 4 in Table 1.6—I use participants' baseline willingness to pay, prior emission estimates, and estimated model parameters to calculate a LABEL willingness to pay for each participant and meal based on equations 1.5 and 1.6. I then simulate meal choices as in the previous calculation.

55. This is the case for non-vegetarian participants. For vegetarians, the two meat meals are exchanged in the experiment for two vegetarian meals. They are dropped from the simulation.

To calculate choices with a carbon tax—line 5 in Table 1.6—I repeat the analysis in line 1, but use adjusted canteen prices, increasing prices by €0.12 for every kg of emissions caused by an option (i.e. using a carbon tax of €120 per tonne). To calculate choices with a meat ban—line 6 in Table 1.6—I similarly remove the meat option from each of the daily choices. For the beef ban—line 7 in Table 1.6—I only removed the beef option on days 2 and 4.

1.A.5 Additional simulation results: Distribution of welfare changes

Figure 1.A.1 shows that the change in utility achieved through making consumers solely attentive is more dispersed than with the combined intervention. In some instances, utility change is even slightly negative. This is mainly attributable to meals for which participants on average overestimated emissions. In this case, increasing attention without providing information can make consumers avoid meals that are in fact low in emissions. Solely increasing knowledge can also decrease consumption utility. These negative effects are also attributable to meals for which participants overestimated emissions, but explicable with a different channel: In the absence of any behavioral intervention, the overestimation can partly compensate for participants' lack of attention towards carbon emissions, and move participants more toward the optimal choice. When the misperception is removed, participants move further away from the optimal choice. There is also one case in which the carbon labels decreased consumption utility: This can be caused by special cases due to slow updating described by the κ parameter.

Providing both interventions prevents either of the two interventions from backfiring: Solely raising attention might lead to a decrease in consumer utility if meals are on average overestimated by consumers. Further, one can think of parameter combinations with very low consumer attention but high overestimation in which solely correcting misperceptions would also lead to a decrease in consumer utility. Providing both interventions simultaneously can prevent both of these situations.

Figures 1.A.2 and 1.A.3 additionally compare the distribution of welfare effects resulting from carbon labels to that of a ban on beef, meat, and a carbon tax.

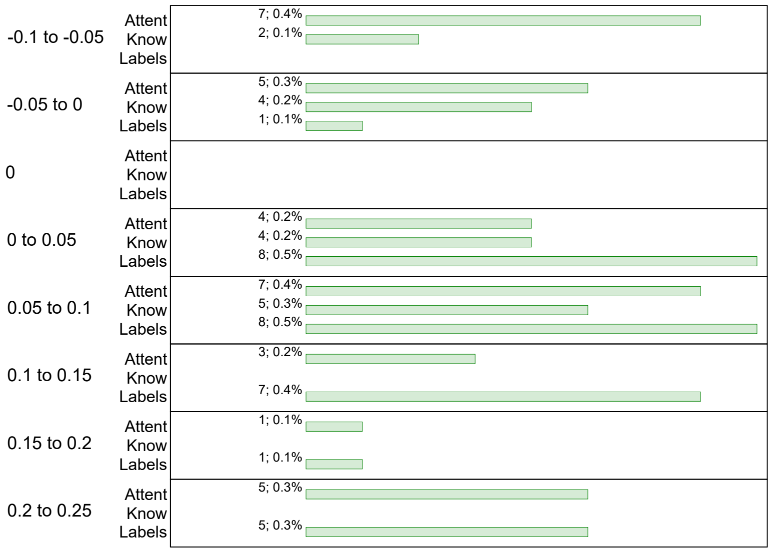


Figure 1.A.1. Estimated change in consumer welfare per meal

Note: Estimated welfare change which would be caused by solely raising attention, solely correcting misperceptions or the combination of both (labels), in Euro. The Figure shows utility changes for instances in which the interventions lead to behavioral change (otherwise change in utility is 0). For the intervention raising attention, this is 2% of instances, for the intervention increasing knowledge it is 0.9% of instances, and for the carbon labels, it is 1.9% of instances. Read numbers e.g. as: With the Attention intervention, there are 7 instances in which a participant would on one of the four simulated canteen days experience a welfare loss equal to a monetary equivalent of €−0.1 and €−0.05 due to the Attention intervention. That is, he would choose a meal due to the Attention intervention that decreases his true utility by this amount. These are 0.4% of all consumption cases. There are 2 instances (0.1% of all consumption cases) in which a similar welfare loss would be incurred due to a Knowledge intervention, etc.

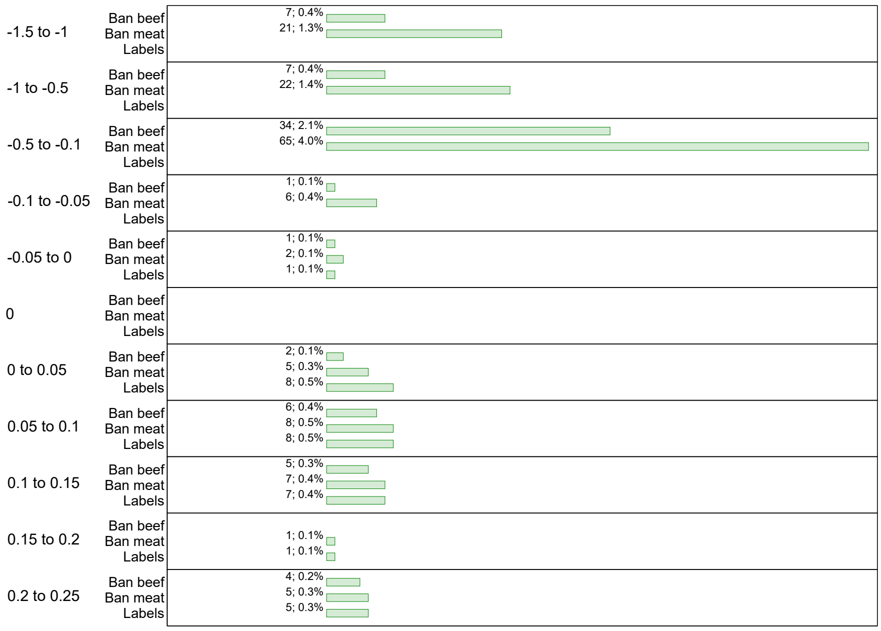


Figure 1.A.2. Estimated change in consumer welfare which would be caused by a ban

Note: Estimated welfare change which would be caused by a ban of beef or a ban of meat compared to carbon labels, in Euro. Estimation based on experiment data and student canteen prices and offer structure. The Figure shows utility changes for instances in which the interventions lead to behavioral change (otherwise change in utility is 0). For the beef ban, this is 4.2% of instances, for the meat ban it is 8.8% of instances, and for the carbon labels, it is 1.9% of instances. Example of how to read this table: With the Beef ban intervention, there are 7 instances in which a participant would on one of the four simulated canteen days experience a welfare loss equal to a monetary equivalent between €–1.5 and €–1 due to the Beef ban intervention. That is, he would choose a meal due to the Beef ban intervention that decreases his true utility by this amount. These are 0.4% of all consumption cases. There are 21 instances (1.3% of all consumption cases) in which a similar welfare loss would be incurred due to a Meat ban intervention, etc.

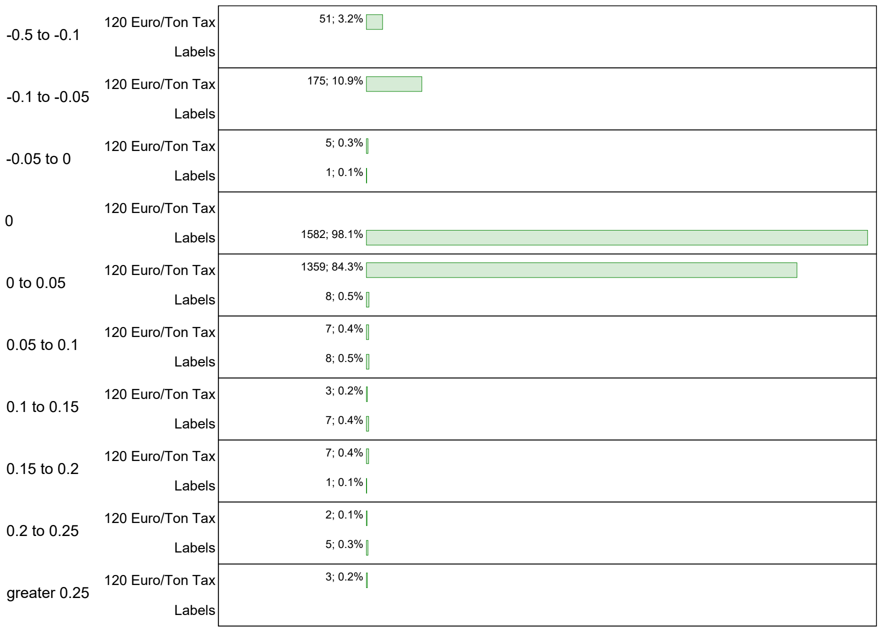


Figure 1.A.3. Estimated change in consumer welfare which would be caused by a carbon tax

Note: Estimated welfare change which would be caused by a carbon tax of €120 per tonne compared to carbon labels, in Euro. Proceedings from the tax are re-distributed equally to all consumers. Estimation based on experiment data and student canteen prices and offer structure. Read numbers e.g. as: With the CARBON TAX intervention, there are 51 instances in which a participant would on one of the four simulated canteen days experience a welfare loss equal to a monetary equivalent between €–0.5 and €–0.1 due to the CARBON TAX intervention. That is, he would choose a meal due to the CARBON TAX intervention that decreases his true utility by this amount, relative to choice and prices in the absence of a tax. These are 3.2% of all consumption cases.

Appendix 1.B Experiments 1 and 3: Additional tables and figures

1.B.1 Randomization checks

Table 1.B.1 shows a randomization check for participants of Experiment 1. Participants are computer assigned into one of the following three groups: 1) Label condition in the second round and Offset condition in the third round, 2) Control condition in the second round and Label condition in the third round, 3) Control condition in the second round and Control condition in the third round. Table 1.B.1 tests whether there are significant differences between these three groups in age, gender, student status, employment, vegetarianism, and hunger at the time of the experiment. There is a higher proportion of non-vegetarians in the group "Control, then Control" (significant at the 5% level), but the groups do not significantly vary otherwise.

To test whether the higher proportion of non-vegetarians impacts results, I perform the main analysis separately for vegetarian and non-vegetarian participants. These analyses should not be influenced by the higher proportion of non-vegetarians in the control group. Results are shown in Table 1.B.8 and Table 1.B.9. Results only including non-vegetarians are similar in coefficient size to the main results. I thus do not believe that the higher proportion of non-vegetarians in the "Control, then Control" group poses a reason for concern.

Table 1.B.1. Randomization Experiment 1

	Average value					
	(1) Age	(2) Male	(3) Student	(4) Working	(5) Non-vegetarian	(6) Hungry
Control, then Control	-0.53	-0.00	0.08	0.05	-0.14**	0.06
	(1.08)	(0.07)	(0.06)	(0.07)	(0.06)	(0.37)
Control, then Label	-0.75	-0.01	0.00	0.10	-0.08	-0.03
	(1.08)	(0.07)	(0.06)	(0.07)	(0.06)	(0.38)
Constant	24.56***	0.33***	0.78***	0.58***	0.81***	5.15***
	(0.62)	(0.04)	(0.03)	(0.04)	(0.04)	(0.21)
Control, then Control	61	71	71	71	71	71
Control, then Label	62	69	69	69	69	69
Label, then Offset	127	149	149	149	149	149
Observations	250	289	289	289	289	289

Standard errors in parentheses

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

Notes: The analysis checks whether there are significant differences in any of the six variables between treatment groups. The group "Label, then Offset" is the baseline category. I do not have full observations for the variable "age", since some participants reported unrealistic numbers Summary statistics for each variable are shown in Table 1.B.3.

Table 1.B.2. Randomization Experiment 3

	Average value					
	(1) Age	(2) Male	(3) Student	(4) Working	(5) Non-vegetarian	(6) Hungry
Attention+Offset, then Attention+Labels	0.04	-0.01	-0.00	0.00	0.03	0.27
	(0.88)	(0.06)	(0.05)	(0.05)	(0.05)	(0.29)
Attention+Labels, then Attention+Offset	-0.53	0.02	0.01	-0.04	0.04	0.10
	(0.89)	(0.06)	(0.05)	(0.05)	(0.05)	(0.30)
Constant	25.93***	0.45***	0.69***	0.75***	0.74***	4.73***
	(0.63)	(0.04)	(0.04)	(0.04)	(0.03)	(0.21)
Attention, then Attention	124	151	151	151	151	151
Attention+Label, then Attention+Offset	126	144	144	144	144	144
Attention+Offset, then Attention+Label	131	149	149	149	149	149
Observations	381	444	444	444	444	444

Standard errors in parentheses

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

Notes: The analysis checks whether there are significant differences in any of the six variables between treatment groups. The group "Attention, then Attention" is the baseline category. I do not have full observations for the variable "age", since some participants reported unrealistic numbers Summary statistics for each variable are shown in Table 1.B.4.

1.B.2 Representativeness of the sample

Tables 1.B.3 and 1.B.4 report descriptive statistics for experiments 1 and 3. Table 1.B.5 reports descriptive statistics elicited in a survey among student canteen guests, as described in Section 1.E.6. In terms of age, participants of experiments 1 and 3 are slightly older than the student canteen guests (average age of 24 and 26 vs. an average age of 23 in the survey). The proportion of males is slightly lower in Experiment 1 (33%) and slightly higher in Experiment 3 (45%) than in the survey (40%), while the proportion of non-vegetarians is similar across all three data sources (70%–76%). In the student canteen purchase data analyzed in Experiment 2, 66% of guests paying with an individual payment card make at least one non-vegetarian purchase during the sample period.

The proportion of students is higher in the survey (93%)than in experiments 1 and 3 (80% and 69%). However, it is likely that my survey over-proportionally surveyed student canteen guests who are students. In the student canteen purchase data analyzed in Experiment 2, 17% of guests paying with an individualized payment card are employees, 81% are students and 2% are non-student and non-employee.[56]

Overall, these statistics suggest that the participants of experiments 1 and 3 are fairly representative of student canteen guests. The largest difference between the experiment sample and survey and student canteen data is the proportion of non-students present in both. Tables 1.B.10, 1.B.18 and 1.B.14 thus repeat the main analyses from experiments 1 and 3 including only students. Results are similar to those reported in sections 1.2 and 1.5, so results are not driven by a higher proportion of non-students.

56. This is the only demographic characteristic reported in the student canteen purchase data. I thus rely on the survey data for the other characteristics.

Table 1.B.3. Socio-economic summary statistics for Experiment 1

Variable	Explanation	Mean	Std. Dev.
Age	Age of participant	24.16	7.05
Male	Dummy: 1 if participant is a man	0.33	–
Student	Dummy: 1 if participant is a student	0.80	–
Working	Dummy: 1 if participant is working in some form	0.62	–
Non-vegetarian	Dummy: 1 if participant eats meat	0.75	–
Hungry	Hunger on scale of 1 to 10 beginning experiment	4.16	2.58
N	289		

Notes: Table shows average socio-economic summary statistics for participants of Experiment 1.

Table 1.B.4. Socio-economic summary statistics for Experiment 3

Variable	Explanation	Mean	Std. Dev.
Age	Age of participant	25.77	7.02
Male	Dummy: 1 if participant is a man	0.45	–
Student	Dummy: 1 if participant is a student	0.69	–
Working	Dummy: 1 if participant is working in some form	0.74	–
Non-vegetarian	Dummy: 1 if participant eats meat	0.76	–
Hungry	Hunger on scale of 1 to 10 beginning experiment	4.85	2.54
N	444		

Notes: Table shows average socio-economic summary statistics for participants of Experiment 3.

Table 1.B.5. Socio-economic summary statistics for student canteen guests

Variable	Explanation	Mean	Std. Dev.
Age	Age of participant	22.90	–
Male	Dummy: 1 if participant is a man	0.41	–
Student	Dummy: 1 if participant is a student	0.94	–
Non-vegetarian	Dummy: 1 if participant eats meat	0.68	–
N	1,451		

Notes: Statistics are based on a survey I conducted among student canteen guests in April. I include only survey respondents who visited a student canteen at least once in the 14-week study period and paid with their individual payment cards. See 1.E.6 for details on the survey design. To preserve anonymity (since I also asked these survey participants about their study field), I elicited age in intervals. To reach an estimation of the mean age, I set the age equal to the midpoint of each interval. For 13% of respondents, I have the information that they are below 20. For the calculation, I estimate their age at 18. For 53% of respondents, I have the information that they are between 20 and 23 (which I set to 21.5 for the estimation), 23% of respondents are between 24 and 27 (set to 25.5), 6% of respondents are between 28 and 31 (set to 30), and 4% of respondents are 32 or older (set to 35). I did not directly elicit vegetarianism, but I elicited how much of a role animal rights play in participants' consumption decisions. I code participants reporting the highest degree of importance as vegetarians.

1.B.3 Descriptive statistics on baseline willingness to pay for meals

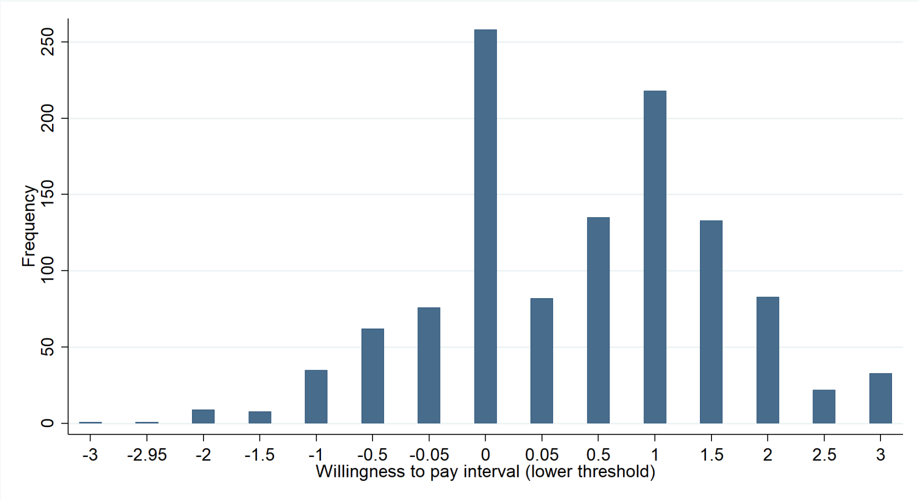


Figure 1.B.1. Willingness to pay indicated for meals in the baseline purchase decisions in Experiment 1

Note: $N = 1,156$ (289 participants making 4 baseline decisions each).

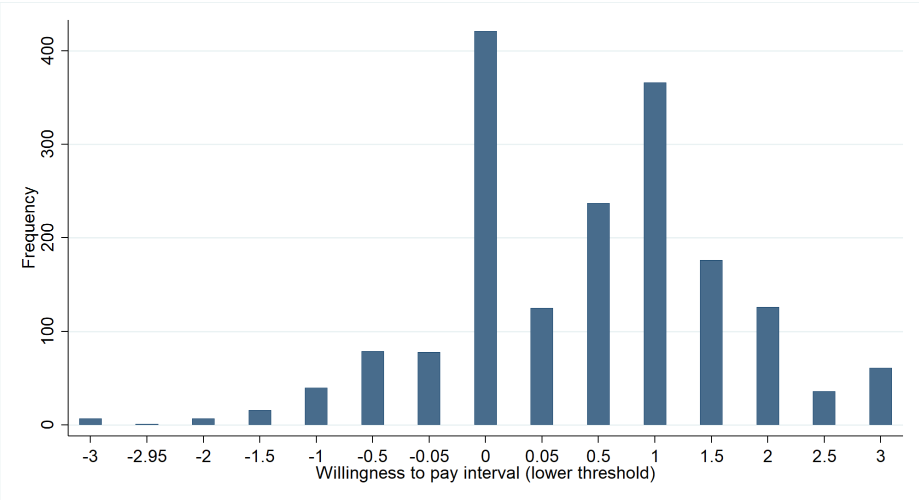


Figure 1.B.2. Willingness to pay indicated for meals in the baseline purchase decisions in Experiment 3

Note: $N = 1,776$ (444 participants making 4 baseline decisions each).

1.B.4 Comparison of effects Exp. 1 and Exp. 3

Table 1.B.6. Comparison of effects in Experiment 1

	Change in WTP					
	(1) Con.	(2) Con.	(3) La.	(4) La.	(5) Of.	(6) Of.
High emission meal	0.01		-0.15***		0.13***	
	(0.01)		(0.03)		(0.03)	
Low emission meal	-0.05*		0.08*		-0.04	
	(0.03)		(0.04)		(0.03)	
Emissions(kg)		0.01		-0.10***		0.09***
		(0.01)		(0.02)		(0.02)
Control for third round	-0.01	-0.01	0.03	0.02		
	(0.02)	(0.02)	(0.04)	(0.04)		
Constant		-0.01		-0.05**		0.04***
		(0.01)		(0.02)		(0.02)
Participants	140	140	218	218	149	149
Observations	1,452	1,452	1,485	1,485	1,009	1,009

Standard errors in parentheses

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

Notes: Dependent variable: within-subject change in willingness to pay for a meal, compared to baseline (Col. 1-4) or WTP with carbon labels (Col. 5-6). Regression specifications follow 1.1. Col.(1) and (2) include only participants in the Control condition, Col.(3) and (4) only participants in the Label condition, and Col.(5) and (6) only participants in the Offset. Effects are split into effects for meals with low emissions (defined as meals with emissions lower than that of the alternative option, the cheese sandwich) and meals with high emissions (meals with emissions higher than the sandwich). To analyze the Offset condition, I use as the dependent variable participants' WTP for meals with offsetting minus WTP with carbon labels. This is to isolate the effect of the offsetting, keeping the attention (Informing participants about carbon offsetting draws attention to emissions) and information effect (With offsetting, participants are informed that emissions are zero) of carbon offsetting constant. Spec. (5) and (6) do not control for the third round of decisions, since Experiment 1 has an Offset condition only in the third round and not in the second round of decisions. Standard errors are clustered at the individual level. Bars indicate 95% confidence intervals.

Table 1.B.7. Comparison of effects in Experiment 3

	Change in WTP					
	(1)	(2)	(3)	(4)	(5)	(6)
	At.+La.	At.+La.	At.	At.	At.+Of.	At.+Of.
High emission meal	-0.10***		-0.04***		0.08***	
	(0.02)		(0.01)		(0.02)	
Low emission meal	-0.04		-0.01		0.02	
	(0.03)		(0.03)		(0.03)	
Emissions(kg)		-0.08***		-0.06***		0.07***
		(0.01)		(0.02)		(0.01)
Control for third round	0.03	0.03	0.01	0.01	-0.01	-0.01
	(0.03)	(0.03)	(0.01)	(0.01)	(0.02)	(0.02)
Constant		-0.04**		-0.00		0.03**
		(0.02)		(0.01)		(0.02)
Participants	293	293	151	151	293	293
Observations	2,051	2,051	2,114	2,114	2,051	2,051

Standard errors in parentheses

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

Notes: Dependent variable: within-subject change in willingness to pay for a meal, compared to baseline (Col. 1-4) or WTP with carbon labels (Col. 5-6). Regression specifications follow 1.1. Col.(1) and (2) include only participants in the Attention+Label condition, Col.(3) and (4) only participants in the Attention condition, and Col.(5) and (6) only participants in the Attention+Offset. Effects are split into effects for meals with low emissions (defined as meals with emissions lower than that of the alternative option, the cheese sandwich) and meals with high emissions (meals with emissions higher than the sandwich). To analyze the Attention+Offset condition, I use as the dependent variable participants' WTP for meals with offsetting minus WTP with carbon labels. This is to isolate the effect of the offsetting, keeping the attention (Informing participants about carbon offsetting draws attention to emissions) and information effect (With offsetting, participants are informed that emissions are zero) of carbon offsetting constant. Standard errors are clustered at the individual level. Bars indicate 95% confidence intervals.

1.B.5 Results split by (non-) vegetarians and (non-) students

Experiment 1

	Change in WTP compared to baseline	
	(1)	(2)
High emission meal x Shown label	−0.26***	
	(0.05)	
Low emission meal x Shown label	0.17***	
	(0.06)	
High emission meal	−0.00	
	(0.02)	
Low emission meal	−0.10**	
	(0.05)	
Emissions(kg) x Shown label		−0.12***
		(0.03)
Emissions(kg)		0.03**
		(0.01)
Shown label		−0.04
		(0.04)
Control for third round	0.01	0.01
	(0.04)	(0.04)
Constant		−0.05*
		(0.03)
Participants control	97	97
Participants treated	170	170
Observations	1,256	1,256

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

Table 1.B.8. Replication of Table 1.1 including only non-vegetarians

	Change in WTP compared to baseline	
	(1)	(2)
High emission meal x Shown label	−0.53***	
	(0.11)	
Low emission meal x Shown label	0.11	
	(0.07)	
High emission meal	0.06	
	(0.05)	
Low emission meal	−0.02	
	(0.04)	
Emissions(kg) x Shown label		−0.75***
		(0.18)
Emissions(kg)		0.08
		(0.08)
Shown label		−0.08
		(0.05)
Control for third round	0.04	0.04
	(0.04)	(0.04)
Constant		0.00
		(0.02)
Participants control	43	43
Participants treated	48	48
Observations	460	460

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

Table 1.B.9. Replication of Table 1.1 including only vegetarians

	Change in WTP compared to baseline	
	(1)	(2)
High emission meal x Shown label	-0.29***	
	(0.05)	
Low emission meal x Shown label	0.15***	
	(0.05)	
High emission meal	-0.01	
	(0.02)	
Low emission meal	-0.08**	
	(0.03)	
Emissions(kg) x Shown label		-0.13***
		(0.03)
Emissions(kg)		0.01
		(0.01)
Shown label		-0.05
		(0.04)
Control for third round	0.01	0.01
	(0.03)	(0.03)
Constant		-0.04**
		(0.02)
Participants control	115	115
Participants treated	170	170
Observations	1,384	1,384

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

Table 1.B.10. Replication of Table 1.1 including only students

	Change in WTP compared to baseline	
	(1)	(2)
High emission meal x Shown label	-0.41***	
	(0.09)	
Low emission meal x Shown label	0.03	
	(0.07)	
High emission meal	0.12**	
	(0.06)	
Low emission meal	0.08	
	(0.07)	
Emissions(kg) x Shown label		-0.08
		(0.08)
Emissions(kg)		0.02
		(0.03)
Shown label		-0.22***
		(0.07)
Control for third round	0.05	0.05
	(0.09)	(0.09)
Constant		0.10*
		(0.06)
Participants control	25	25
Participants treated	48	48
Observations	332	332

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

Table 1.B.11. Replication of Table 1.1 including only non-students

Experiment 3

	Change in WTP compared to baseline	
	(1)	(2)
Underestimated emissions	−0.11**	
	(0.04)	
Underestimation (in kg)		−0.06**
		(0.03)
Control for third round	0.05	0.05
	(0.05)	(0.05)
Constant	−0.12***	−0.16***
	(0.04)	(0.04)
Participants	227	206
Obs. underestimate	451	420
Obs. overestimate	418	364
Observations	869	784

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

Table 1.B.12. Replication of Table 1.3 including only non-vegetarians

	Change in WTP compared to baseline	
	(1)	(2)
Underestimated emissions	−0.21***	
	(0.07)	
Underestimation (in kg)		−0.14**
		(0.06)
Control for third round	0.05	0.13
	(0.10)	(0.09)
Constant	−0.02	−0.18**
	(0.09)	(0.07)
Participants	66	60
Obs. underestimate	104	96
Obs. overestimate	144	130
Observations	248	226

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

Table 1.B.13. Replication of Table 1.3 including only vegetarians

	Change in WTP compared to baseline	
	(1)	(2)
Underestimated emissions	-0.18***	
	(0.04)	
Underestimation (in kg)		-0.10***
		(0.03)
Control for third round	0.10*	0.11**
	(0.05)	(0.06)
Constant	-0.12**	-0.21***
	(0.05)	(0.04)
Participants	203	184
Obs. underestimate	383	360
Obs. overestimate	391	344
Observations	774	704

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

Table 1.B.14. Replication of Table 1.3 including only students

	Change in WTP compared to baseline	
	(1)	(2)
Underestimated emissions	-0.00	
	(0.05)	
Underestimation (in kg)		-0.02
		(0.04)
Control for third round	-0.06	-0.06
	(0.08)	(0.09)
Constant	-0.05	-0.05
	(0.06)	(0.05)
Participants	90	81
Obs. underestimate	172	158
Obs. overestimate	171	153
Observations	343	311

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

Table 1.B.15. Replication of Table 1.3 including only non-students

	Change in WTP compared to baseline
	(1)
High emission meal x Shown label	−0.10**
	(0.04)
Low emission meal x Shown label	−0.06
	(0.05)
High emission meal	−0.11***
	(0.03)
Low emission meal	−0.01
	(0.04)
Control for third round	0.04
	(0.03)
Participants attent	112
Participants label	227
Observations	1,804

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

Table 1.B.16. Replication of Table 1.4 including only non-vegetarians

	Change in WTP compared to baseline
	(1)
High emission meal x Shown label	−0.12
	(0.08)
Low emission meal x Shown label	0.03
	(0.06)
High emission meal	−0.05
	(0.04)
Low emission meal	−0.04
	(0.04)
Control for third round	0.02
	(0.04)
Participants attent	39
Participants label	66
Observations	576

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

Table 1.B.17. Replication of Table 1.4 including only vegetarians

	Change in WTP compared to baseline
	(1)
High emission meal x Shown label	-0.17***
	(0.04)
Low emission meal x Shown label	-0.02
	(0.05)
High emission meal	-0.08***
	(0.03)
Low emission meal	-0.03
	(0.03)
Control for third round	0.05*
	(0.03)
Participants attent	104
Participants label	203
Observations	1,644

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

Table 1.B.18. Replication of Table 1.4 including only students.

	Change in WTP compared to baseline
	(1)
High emission meal x Shown label	0.04
	(0.08)
Low emission meal x Shown label	-0.03
	(0.08)
High emission meal	-0.14**
	(0.06)
Low emission meal	-0.00
	(0.06)
Control for third round	-0.01
	(0.04)
Participants attent	47
Participants label	90
Observations	736

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

Table 1.B.19. Replication of Table 1.4 including only non-students

1.B.6 Replication excluding round 3 observations

Table 1.B.20. Replication of Table 1.1 excluding round 3 observations

	Change in WTP compared to baseline	
	(1)	(2)
High emission meal x Shown label	-0.34***	
	(0.06)	
Low emission meal x Shown label	0.15**	
	(0.06)	
High emission meal	0.02	
	(0.02)	
Low emission meal	-0.05*	
	(0.03)	
Emissions(kg) x Shown label		-0.15***
		(0.04)
Emissions(kg)		0.03**
		(0.01)
Shown label		-0.07*
		(0.04)
Control for third round		
Constant		-0.02
		(0.02)
Participants control	140	140
Participants treated	149	149
Observations	1,156	1,156

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

Table 1.B.21. Replication of Table 1.3 excluding round 3 observations

	Change in WTP compared to baseline	
	(1)	(2)
Underestimated emissions	-0.12**	
	(0.05)	
Underestimation (in kg)		-0.06*
		(0.03)
Constant	-0.10**	-0.17***
	(0.04)	(0.03)
Participants	144	133
Obs. underestimate	269	248
Obs. overestimate	281	248
Observations	550	496

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

Table 1.B.22. Replication of Table 1.4 excluding round 3 observations

	Change in WTP compared to baseline
	(1)
High emission meal x Shown label	-0.11**
	(0.05)
Low emission meal x Shown label	-0.06
	(0.05)
High emission meal	-0.09***
	(0.03)
Low emission meal	-0.01
	(0.03)
Participants attent	151
Participants label	144
Observations	1,180

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

1.B.7 Exp. 1: Alternative econometric specifications

Alternatively to the estimation approach described in Section 1.2.2, one could instead estimate the following specification:

$$\begin{aligned} WTP_{ijm} = \alpha_{im} + \beta_1(High_m \times Post_j) + \beta_2(Low_m \times Post_j) + \delta_1(High_m \times Post_j \times Label_{ij}) \\ + \delta_2(Low_m \times Post_j * Label_{ij}) + ThirdRound_j + \varepsilon_{ijm} \quad (1.B.1) \end{aligned}$$

This specification is more similar to a classic diff-in-diff approach. Instead of directly using the difference between indicated willingness to pay for a meal and baseline willingness to pay as the dependent variable (as in 1.1), I use raw willingness to pay of individual i in round j for meal m as the dependent variable. Accordingly, I also include observations from the baseline elicitation round in the regression.

α_{im} are individual and meal-specific fixed effects. These are 1156 fixed effects in total: 289 participants × 4 meals. These fixed effects control for individual-specific baseline tastes. Note that it would not make much sense to include merely a single fixed effect for each individual. A single fixed effect would capture the average willingness to pay of each individual across the four meals. However, I expect the effect of the carbon labels to differ across meals. Willingness to pay for low-emission meals should increase as a result of the label, while willingness to pay for high-emission meals should decrease. It is thus insufficient to control for individuals' willingness to pay averaged across meals. To illustrate with an example, imagine I only had two meals, one low-emission and one high-emission meal. An individual has a willingness to pay of €1.00 for the low-emission meal and a willingness to pay of €3.00 for the high-emission meal. When the individual sees the carbon labels, he adjusts his willingness to pay for the low-emission meal upward to €2.00 euros, and his willingness to pay for the high-emission meal downward to €2.00 euros. Treatment effects are thus sizable. However, his average willingness to pay for the two meals did not change, and a regression including a single individual fixed effect term would falsely not identify a treatment effect.

$(High_m \times Post_j)$ is an indicator variable for whether the meal causes higher emissions than the sandwich, and interacted with the elicitation round $j > 1$, i.e. it being the second or third round of elicitations and not the baseline round. $(Low_m \times Post_j)$ is the equivalent indicator for low-emission meals. Note that all meals classified are classified either as Low_m or $High_m$. The two variables thus together capture the $Post_j$ effect, and a separate $Post_j$ indicator would be dropped due to collinearity. I also do not include separate controls for Low_m and $High_m$ since meal characteristics are captured by the α_{im} fixed effects.

$(High_m \times Post_j \times Label_{ij})$ interacts the high-emission and $Post_j$ indicator with an indicator for whether individual i saw carbon labels in round j. This describes the average causal effect of carbon labels on willingness to pay for a meal that is high in carbon emissions. $(Low_m \times Post_j \times Label_{ij})$ describes the average causal effect of carbon labels on willingness to pay for a meal that is low in carbon emissions. $ThirdRound_j$ is an indicator of whether it was the third round of decisions. Standard errors are clustered at the individual level.

Spec. (1) in Table 1.B.23 shows regression results. They are very similar to those reported in the main text. Spec. (2) replicates Spec. (2) of Table 1.1 with a fixed effect approach and also finds similar results as reported in the main text.

Table 1.B.23. Replication of Experiment 1 results with fixed effects approach

	WTP	
	(1)	(2)
High x Post x Label	−0.30***	
	(0.04)	
Low x Post x Label	0.09**	
	(0.04)	
High x Post	0.01	
	(0.02)	
Low x Post	−0.03	
	(0.04)	
Emissions(kg) x Post x Label		−0.12***
		(0.03)
Emissions(kg) x Post		0.01
		(0.01)
Post x Label		−0.08***
		(0.03)
Post		−0.02
		(0.02)
Control for third round	0.01	0.01
	(0.03)	(0.03)
Constant	0.65***	0.65***
	(0.01)	(0.01)
Participants control	140	140
Participants treated	218	218
Observations	2,872	2,872

Standard errors in parentheses

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

Notes: Table replicates the estimation in Table 1.1 using willingness to pay for meals directly as the outcome variable, instead of taking the difference. Spec. (1) corresponds to Equation 1.B.1 and includes individual× meal fixed effects. It does not include a "Post" or a "Post× Label" variable, because "Low emissions meal" and "High emissions meal" are mutually exclusive. In spec. (2), emissions (kg) are defined as the emissions caused by the meal relative to the cheese sandwich. This is positive for "high-emission" and negative for "low-emission" meals. Standard errors are clustered at the individual level.

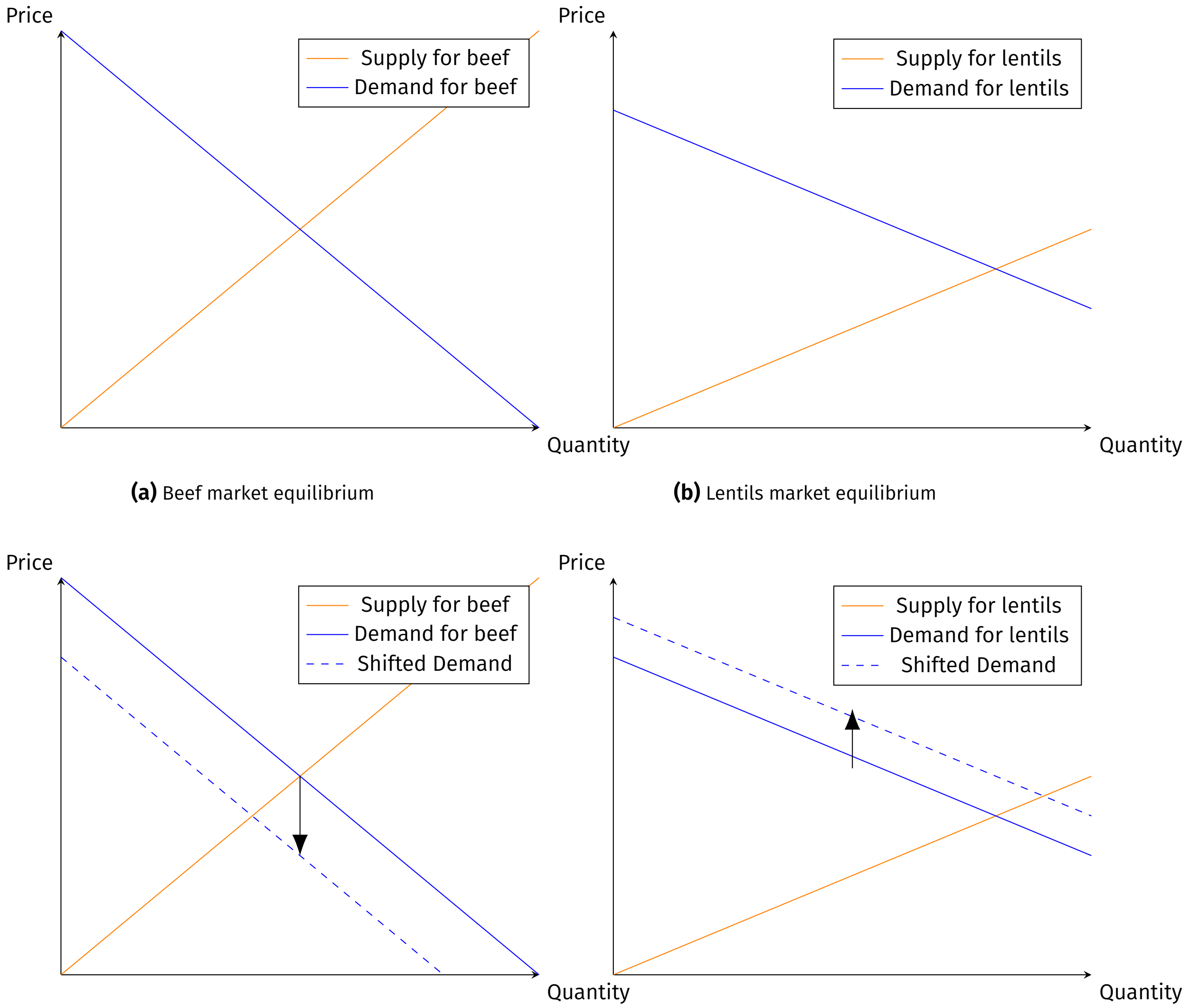
1.B.8 Exp. 1: Intuition behind expressing effect sizes in terms of a carbon tax

One of the main results shown in section 1.2.3 is that carbon labels in Experiment 1 produce a similar impact as would result from a carbon tax of €0.12 per kg or €120 per tonne. The underlying assumption for this comparison is that a shift in the demand curve due to the installation of carbon labels affects total quantity similarly as a would a shift in the demand curve due to the installation of a carbon tax.

To illustrate this point, I first show in Figure 1.B.3 how carbon labels and a carbon tax would affect price and quantity purchased in two specific product markets: beef and lentils. Images (a) and (b) show a stylized illustration of how the current market equilibrium in the beef market and the lentils market might look like. In each market, the equilibrium price and quantity is determined by the intersection of the supply and demand curves. Image (c) shows how the beef market would be affected by a downward shift in the demand curve. This shift in the demand curve could either result from consumers being willing to pay less for beef due to carbon labels, or consumers being willing to pay less because a carbon tax will be added to their purchase. The downward shift in the demand curve leads to the demand curve and supply curve now intersecting at a lower price and a lower quantity. Image (d) shows how the lentils market would be affected by an upward shift in the demand curve. This shift could again either result from consumers being willing to pay more for lentils as they recognize their good environmental performance on the carbon labels, or consumers being willing to pay more because there will be no carbon tax added to their purchase. The upward shift in the demand curve leads to the demand curve and supply curve now intersecting at a higher price and a higher quantity.

More generally, one could think of demand for emission-heavy goods in a more abstract sense, with there being some demand curve describing consumer demand for different items as a function of how much emissions result from their production. A carbon tax would shift this demand curve downward, just as would carbon labels. My analysis in section 1.2.3 quantifies the shift occurring through the labels in terms of which height of a carbon tax would be required to shift this demand curve downward by the same extent. Note that my estimate of €0.12 per kg averages over all participants, i.e. it already incorporates that some consumers might be reacting to the labels more strongly than other consumers.

Importantly, my €120 per tonne equivalence result describes participant behavior in Experiment 1, i.e. it is specific to a certain population group and consumption context. To reach a carbon tax equivalence estimate for e.g. the entire German or European market, data from other population groups and consumption contexts is needed.



(a) Beef market equilibrium

(b) Lentils market equilibrium

(c) Shifted demand in beef market

(d) Shifted demand in lentils market

Figure 1.B.3. Comparison of supply and demand in beef and lentils markets

1.B.9 Exp. 1: Reaction to carbon labels by baseline WTP

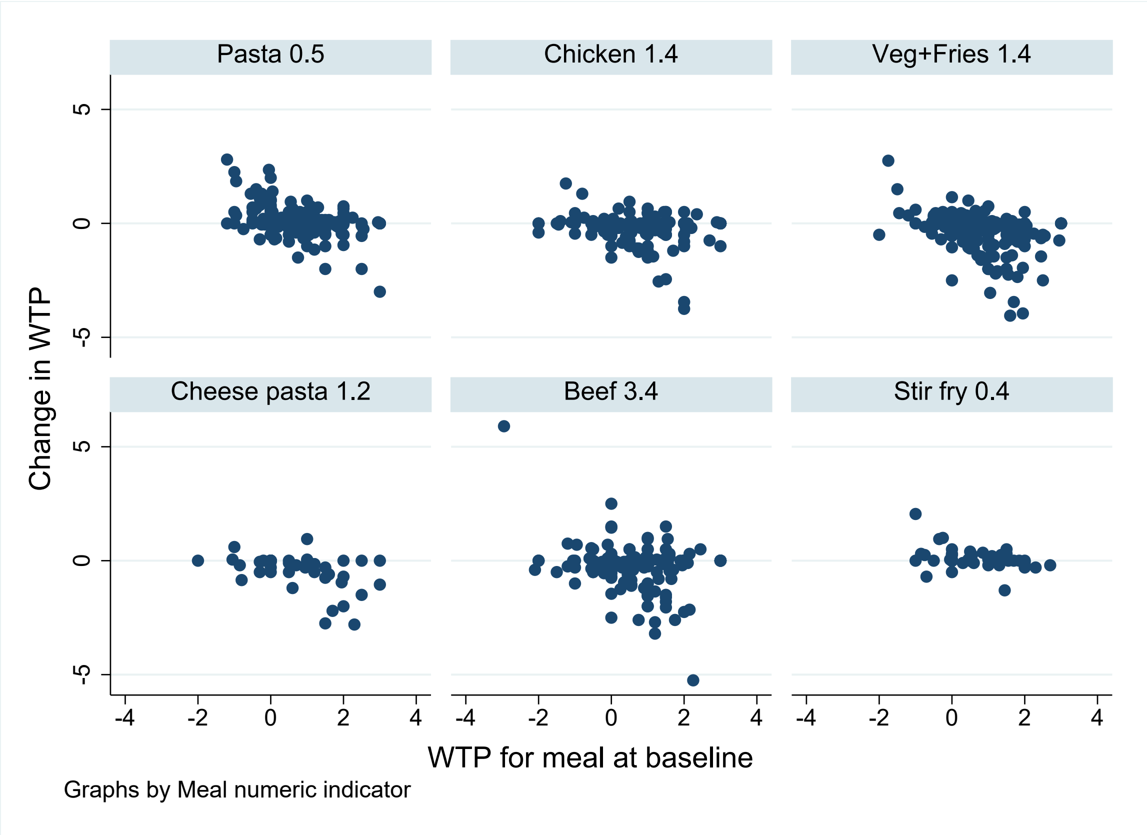


Figure 1.B.4. Scatterplot of participants' change in willingness to pay for meals when shown carbon labels

Note: Shown by baseline willingness to pay. $N = 218$ for meals shown to vegetarians and non-vegetarians: Vegetable pasta (emissions 0.5 kg), Courgettes with fries (emissions 1.4 kg). $N = 170$ for meals shown only to non-vegetarians: Chicken with rice (emissions 1.4 kg) and beef with potatoes (3.4 kg). $N = 48$ for meals only shown to vegetarians: Cheese pasta "Spätzle" (1.2 kg) and Stir-fried vegetables (0.4 kg). Plots show that individuals across a variety of baseline WTP categories are reacting to the carbon labels—It is not a couple of individuals alone driving effects.

1.B.10 Exp. 1: Heterogeneity in treatment effects

Table 1.B.24. Heterogeneity analysis using same items as heterogeneity analysis in the field (Table 1.C.4)

	Change in WTP compared to baseline			
	(1)	(2)	(3)	(4)
	All	Female	Below 24	Env. important
High EM x Shown label	−0.31***	−0.36***	−0.32***	−0.39***
	(0.05)	(0.06)	(0.07)	(0.07)
Low EM x Shown label	0.14***	0.10*	0.12**	0.17***
	(0.04)	(0.05)	(0.06)	(0.06)
High emission meal	0.01	0.00	−0.02	−0.00
	(0.02)	(0.02)	(0.03)	(0.03)
Low emission meal	−0.06*	−0.04	−0.05	−0.08**
	(0.03)	(0.03)	(0.04)	(0.03)
Control for third round	0.01	0.03	0.03	0.04
	(0.03)	(0.04)	(0.05)	(0.04)
Participants control	140	95	80	90
Participants treated	218	147	118	123
Observations	1,716	1,160	952	1,040

Standard errors in parentheses

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

Notes: Dependent variable: within-subject change in willingness to pay for a meal, compared to baseline. Specifications correspond to Equation 1.1 and do not include a constant, because “Low emissions meal” and “High emissions meal” are mutually exclusive. Col.(1) includes all data, and Col.(2) includes only females. Col.(3) includes only under 24-year olds. Col.(5) includes only survey participants who report an above-average importance of environmental aspects in their food consumption decisions. Standard errors are clustered at the individual level. Table 1.C.5 reports evidence from Experiment 2 for the same heterogeneity factors.

Table 1.B.25. Heterogeneity in Experiment 1

	Change in WTP compared to baseline					
	(1) All	(2) Social circle	(3) Hungry	(4) Strong l.o.c.	(5) Low income	(6) Price sens.
High EM x Shown label	−0.31***	−0.36***	−0.40***	−0.40***	−0.33***	−0.27***
	(0.05)	(0.05)	(0.07)	(0.07)	(0.07)	(0.05)
Low EM x Shown label	0.14***	0.14**	0.08	0.18***	0.12*	0.19***
	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)	(0.07)
High emission meal	0.01	0.04	0.05	0.04	0.05*	0.05*
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Low emission meal	−0.06*	−0.04	−0.02	−0.05	−0.01	−0.12**
	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)	(0.06)
Control for third round	0.01	0.03	0.02	0.03	0.00	0.01
	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)
Participants control	140	79	66	84	54	60
Participants treated	218	123	104	118	82	109
Observations	1,716	968	816	988	656	800

Standard errors in parentheses

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

Notes: Dependent variable: within-subject change in willingness to pay for a meal, compared to baseline. Specifications correspond to Equation 1.1 and do not include a constant, because "Low emissions meal" and "High emissions meal" are mutually exclusive. Regressions include only individuals who report above-average values for the respective items. "Social circle" is measured using approval of the statement "My friends and acquaintances would approve if I try to reduce my carbon footprint.". "Hungry" is measured on a 10-point scale using the question "How hungry are you feeling now, in this moment?". "L.o.c. (Locus of control)" is measured with approval of the statement "I believe I can contribute to the solution of the climate crisis by reducing my carbon footprint.". "Low income" includes only individuals with the lowest possible net income option (under €700 a month). "Price sensitive" is measured using participants' reports of the importance of meal price in a typical meal decision. Standard errors are clustered at the individual level.

Table 1.B.26. Heterogeneity analysis using same items I use in the correlation analysis on WTP determinants (Table 1.7)

	Change in WTP compared to baseline					
	(1)	(2)	(3)	(4)	(5)	(6)
	All	Strong norms	In favor	Use info	Own knowledge	High self-control
High EM x Shown label	−0.31***	−0.36***	−0.42***	−0.47***	−0.29***	−0.33***
	(0.05)	(0.06)	(0.07)	(0.08)	(0.05)	(0.07)
Low EM x Shown label	0.14***	0.16***	0.13**	0.21***	0.11**	0.22***
	(0.04)	(0.06)	(0.07)	(0.07)	(0.05)	(0.06)
High emission meal	0.01	−0.01	0.01	−0.01	0.02	0.00
	(0.02)	(0.03)	(0.03)	(0.03)	(0.02)	(0.03)
Low emission meal	−0.06*	−0.04	−0.05	−0.07*	−0.08*	−0.10**
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.05)
Control for third round	0.01	−0.00	0.03	0.02	0.04	0.01
	(0.03)	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)
Participants control	140	71	78	76	69	70
Participants treated	218	107	105	106	135	105
Observations	1,716	880	916	912	932	844

Standard errors in parentheses

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

Notes: Dependent variable: within-subject change in willingness to pay for a meal, compared to baseline. Specifications correspond to Equation 1.1 and do not include a constant, because “Low emissions meal” and “High emissions meal” are mutually exclusive. Regressions include only individuals who report above-average values for the respective items. “In favor of labels in student canteen” is measuring using approval of the statement “I would appreciate if the student canteen would introduce such a measure”. “Self-reported willingness to use info” is measured using approval of the statement “I would include this information in my decision.”. “Self-reported confidence in own knowledge” is measured with two questions: (1) approval of the statement “I already know without labels which emissions are caused by different meals.”, and (2) “I think this information will partially surprise me.” The perceived strength of social norms is measured using the procedure developed by Krupka and Weber (2013). Eating self-control is measured using the questions developed by Haws, Davis, and Dholakia (2016). Standard errors are clustered at the individual level.

1.B.11 Exp. 1: Effect on calorie guesses

Table 1.B.27. Effects of the treatment on calories guessed in Experiment 1

	Guess of calories in				
	(1) Chicken-rice	(2) Courgettes-fries	(3) Beef-potatoes	(4) Cheese sandwich	(5) Veg. pasta
Sees carbon labels	81.31	131.47	3.15	24.13	85.67
	(185.95)	(113.90)	(58.00)	(23.36)	(109.30)
Constant	639.36***	506.27***	732.98***	272.62***	518.82***
	(164.59)	(98.92)	(51.33)	(20.29)	(94.93)
Participants control	71	71	71	71	71
Participants treated	218	218	218	218	218
Observations	217	289	217	289	289

Standard errors in parentheses

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

Notes: To test whether participants conclude other meal characteristics when seeing carbon labels, I ask participants to guess the calories of different meals towards the end of the experiment. Participants in the Treatment see carbon labels during the guess, while participants in the Control group do not. There is no significant effect of seeing the labels on calorie guesses.

1.B.12 Exp. 3: Descriptives on under- and over-estimation

Table 1.B.28. Under- and over-estimation of meal emissions

Meal	Relative emissions	No. underestimated	No. overestimated	No. correct	Total
Vegetable pasta	-0.2 kg	31	249	13	293
Chicken w. rice	0.7 kg	47	163	17	227
Courgettes w. fries	0.7 kg	249	33	11	293
Cheese pasta	0.5 kg	31	24	11	66
Beef w. potatoes	2.7 kg	193	32	2	227
Stir-fried veg.	-0.3 kg	4	61	1	66
Total	654	459	59	55	1.172

Notes: Based on participants in the ATTENT+LABEL treatment. I show under- and overestimation of the emissions caused by those meals that are also used in the experiment decisions. Relative emissions are emissions relative to the cheese sandwich (0.7 kg). I classify a participant as underestimating this amount if their guess for the meal's emissions minus their guess for the cheese sandwich is lower than the actual relative emissions. I classify a participant as overestimating this amount if their guess for the meal's emissions minus their guess for the cheese sandwich is higher than the actual relative emissions.

Table 1.B.29. Number of under- and over-estimations per participant

No. overestimated	0	1	2	3	4	Total
No. underestimated						
0	0	0	0	2	10	12
1	0	1	21	54	0	76
2	1	24	128	0	0	153
3	4	31	0	0	0	35
4	17	0	0	0	0	17
Total	22	56	149	56	10	293

Notes: Relative emissions are emissions relative to the cheese sandwich (0.7 kg). I classify a participant as underestimating this amount if their guess for the meal's emissions minus their guess for the cheese sandwich is lower than the actual relative emissions. I classify a participant as overestimating this amount if their guess for the meal's emissions minus their guess for the cheese sandwich is higher than the actual relative emissions. Each cell shows the number of participants with the respective number of under- or over-estimations.

Table 1.B.30. Number of participants who correctly guessed how the four decision meals rank relative to each other

No. of correctly ranked meals	No. participants
0	11
2	88
3	188
4	6
Total	293

Notes: If a participant indicated emission values for the four decision meals such that the value he indicates for the lowest-ranking meal is the lowest in his ranking, the second-lowest-ranking meal is the second-lowest in his ranking, the third-lowest-ranking meal is the third-lowest, etc. I count him as getting all four relative ranks right. This is true for six participants. 188 participants got three relative ranks right, and 88 got two relative ranks right (i.e. two meals stood in the correct relationship to each other).

1.B.13 Exp. 3: Results split by guess accuracy

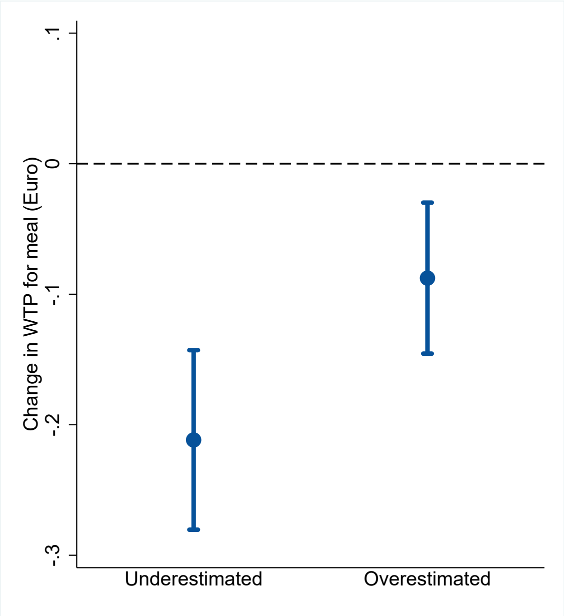


Figure 1.B.5. Replication of Figure 1.15 including only individuals with at least three correct ranks

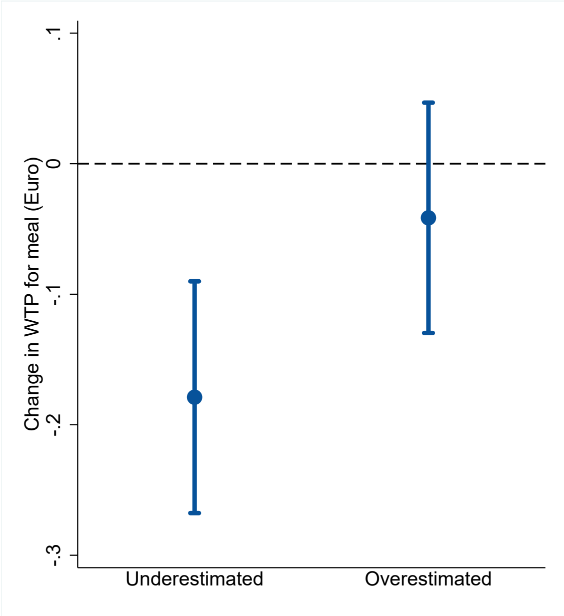


Figure 1.B.6. Replication of Figure 1.15 including only individuals with at most two correct ranks

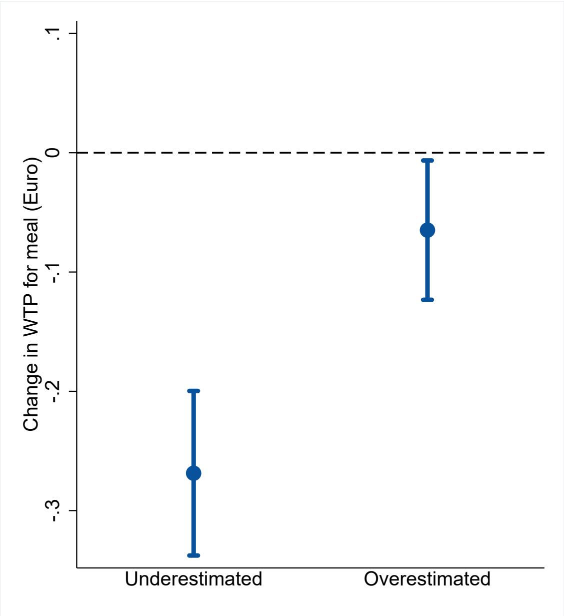


Figure 1.B.7. Replication of Figure 1.15 including only individuals with at least three correctly guessed magnitudes

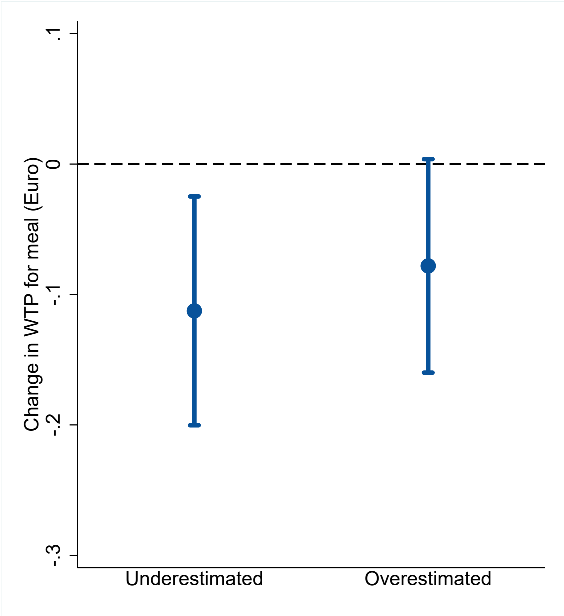


Figure 1.B.8. Replication of Figure 1.15 including only individuals with at most two correctly guessed magnitudes

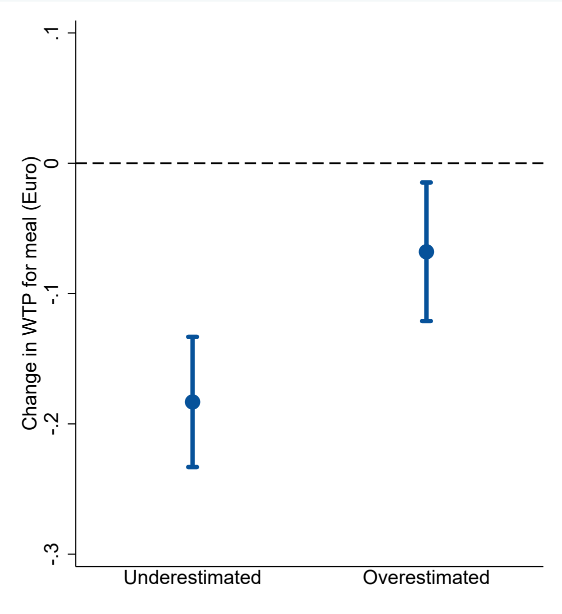


Figure 1.B.9. Replication of Figure 1.15 based on under- or over-estimation of the meal

	Change in WTP compared to baseline	
	(1)	(2)
Underestimated emissions	-0.13***	
	(0.04)	
Underestimation (in kg)		-0.04
		(0.03)
Control for third round	0.05	0.06
	(0.05)	(0.05)
Constant	-0.09***	-0.18***
	(0.03)	(0.03)
Participants	293	267
Obs. underestimate	651	640
Obs. overestimate	471	376
Observations	1,122	1,016

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

Table 1.B.31. Replication of Table 1.3 based on under- or over-estimation of the meal

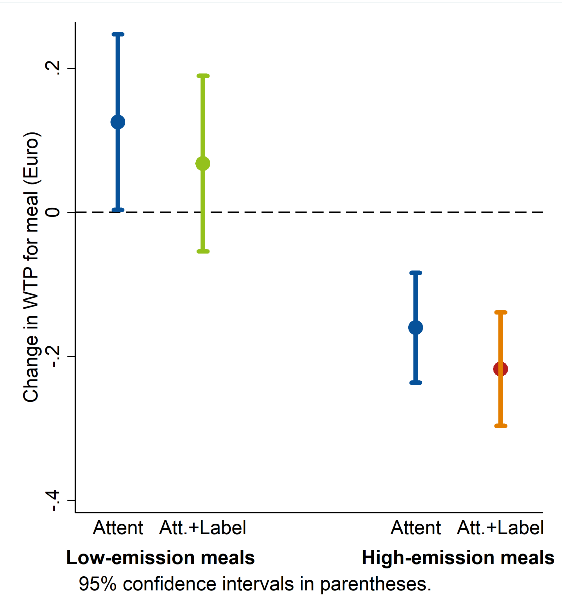


Figure 1.B.10. Replication of Figure 1.16 with only accurate guesses

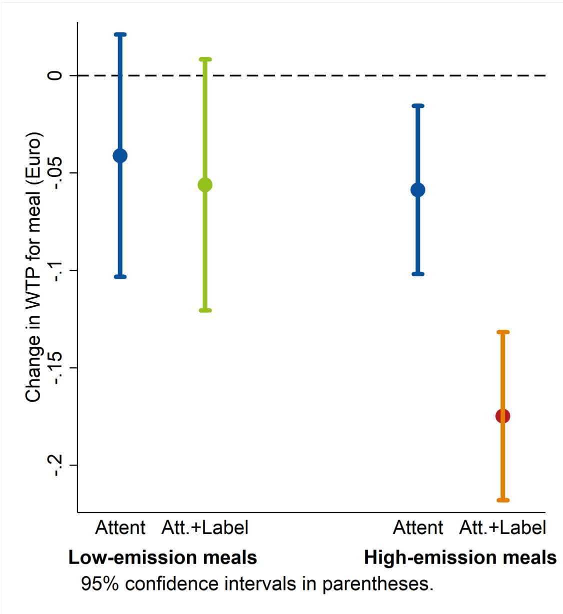


Figure 1.B.11. Replication of Figure 1.16 with only inaccurate guesses

1.B.14 Participants' willingness to pay for the presence of carbon labels

Table 1.B.32. Willingness to pay for seeing carbon labels by treatment group

	(1)
	wtp
Control, then Label	-0.13
	(0.08)
Label, then Offset	-0.11*
	(0.07)
Attent, then Attent	-0.08
	(0.07)
Attent+Label, then Offset	-0.07
	(0.07)
Attent+Offset, then Labels	-0.04
	(0.07)
Control, then Control	0.00
	(.)
Constant	0.28***
	(0.05)
N	731

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

Notes: Average deviation from the average willingness to pay to see emission labels for the final three consumption decisions, by treatment group. "Control, then Control" is the baseline condition.

Table 1.B.33. Correlation between willingness to pay for seeing carbon labels and treatment effect

	(1)
Decrease in WTP for highest-emission meal	-0.21***
	(0.02)
Constant	0.15***
	(0.02)
Observations	397

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

Notes: Dependent variable: Willingness to pay for seeing labels for the final three consumption decisions. Independent variable: The decrease in the participant's willingness to pay for the highest-emission meal when shown emission labels. Regression is restricted to participants who were shown emission values in the experiment. The coefficient signals that participants showing a stronger reaction to carbon labels are also willing to pay a higher amount to be shown the labels.

Appendix 1.C Experiment 2: Additional tables and figures

1.C.1 Time trends

	(1) Meat meal
Treated × Week 1	-2.40
	(1.72)
Treated × Week 2	0.76
	(1.49)
Treated × Week 3	0.24
	(1.98)
Treated × Week 5	-3.25**
	(1.42)
Treated × Week 6	-2.17
	(1.48)
Treated × Week 7	-0.61
	(1.56)
Treated × Week 8	-3.77**
	(1.66)
Treated × Week 9	-1.55
	(1.55)
Treated × Week 10	-6.51***
	(1.85)
Treated × Week 11	-4.23***
	(1.49)
Treated × Week 12	-5.29***
	(1.53)
Treated × Week 13	-5.98***
	(1.59)
Treated × Week 14	-8.35***
	(1.43)
Treatment restaurant	-11.12***
	(1.18)
Other offer: special dish	0.16
	(0.48)
Other offer: special veg. dish	-1.29
	(0.79)
Guests control	7,018
Guests treated	2,857
Observations	130,139

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

	(1) Meat meal
Other offer: pasta	-4.49***
	(0.63)
Other offer: pizza	-4.68***
	(0.80)
Other offer: add. vegan dish	2.07***
	(0.49)
Other offer: grill	-0.46
	(0.77)
Other offer: stew/soup	-2.40**
	(0.99)
Other offer: pan	2.36***
	(0.59)
Second veg. main	-1.93***
	(0.38)
Second meat main	1.54***
	(0.41)
Main components sold in ’000s	-0.73***
	(0.08)
Guests control	7,018
Guests treated	2,857
Observations	130,139

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

Figure 1.C.1. Regression table for Figure 1.11

Note: Regression table split into two columns for readability. Dependent variable: 0/1 indicator for consumption of the meat option, multiplied by 100 to enable the interpretation of coefficients as percentage points. Regression additionally includes weekly controls and day-of-the-week controls.

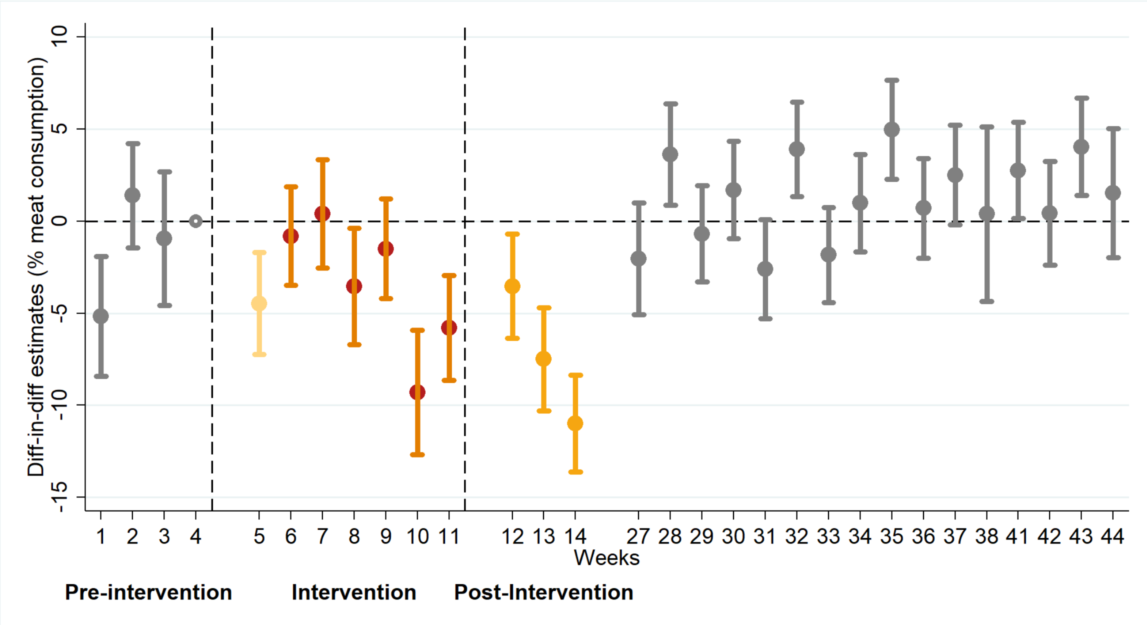


Figure 1.C.2. Event study including data from the following semester

Note: Difference in difference estimates of the likelihood of consuming the meat option (in percentage points), using week 4 of the pre-intervention phase as a baseline. Weeks 1–4 constitute the pre-intervention phase, while weeks 6–11 constitute the intervention phase, and weeks 12–14 the post-intervention phase. Weeks 27 onwards are the new semester. The regression specification closely follows specification (1) in Table 1.2, estimating weekly treatment effects and including weekly time controls and day-of-the-week controls. Bars indicate 95% confidence intervals.

Table 1.C.1. Post-intervention effects throughout the following semester

	Likelihood of consuming meat			
	(1)	(2)	(3)	(4)
Treatment restaurant x Label period	-0.04***	-0.04***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.01)
Treatment restaurant x Post period	-0.02***	-0.02***		
	(0.00)	(0.00)		
Treatment restaurant x June			-0.06***	-0.06***
			(0.01)	(0.01)
Treatment restaurant x July			-0.08***	-0.09***
			(0.01)	(0.01)
Treatment restaurant x October			0.01*	0.01*
			(0.01)	(0.01)
Treatment restaurant x November			0.02**	0.02***
			(0.01)	(0.01)
Treatment restaurant x December			0.02***	0.02***
			(0.01)	(0.01)
Treatment restaurant x January			0.03***	0.03***
			(0.01)	(0.01)
June			0.00	
			(0.01)	
July			0.02***	
			(0.01)	
October			0.01	
			(0.00)	
November			-0.02***	
			(0.00)	
December			0.01***	
			(0.00)	
January			-0.01**	
			(0.00)	
Treatment restaurant	-0.08***	-0.08***	-0.10***	-0.10***
	(0.00)	(0.00)	(0.01)	(0.01)
Label period	0.01***		0.01	
	(0.00)		(0.00)	
Post period	0.01**			
	(0.00)			
Constant	0.51***	0.48***	0.51***	0.48***
	(0.00)	(0.01)	(0.00)	(0.01)
Date effects	No	Yes	No	Yes
Fixed effects	No	No	No	No
Guests control	12,387	12,387	12,387	12,387
Guests treated	5,401	5,401	5,401	5,401
Observations	300,241	300,241	300,241	300,241

Standard errors in parentheses

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

1.C.2 Effect magnitude relative to carbon tax and Exp. 1

It is interesting to see how the effect magnitude I observe in Experiment 1—€120 per tonne—relates to the effect sizes I observe in the natural field experiment. I thus draw on additional student canteen data to estimate the effect a carbon tax would have in the student canteen context to put my natural field experiment result into perspective. I estimate that the carbon labels in the field produce a similar effect as would be expected from a carbon tax of €80 per tonne to €120 per tonne in the same field setting.

For this analysis, I use a larger set of student canteen purchase data to include the time frame from April 2022 to March 2023. Student canteen prices for the same main meal components vary throughout this period due to a price increase in October 2023, which not only increased the general price level but also the price differential between the meat and the vegetarian main meal components. While this difference was on average around €0.33 from April to June 2022 (around 20% of the price of a veg. main meal component sold then), it increased to around €0.50 from October to December 2023(around 25% of the price of a veg. main meal component sold then) and remained at this higher level.

Since the price increase affected all student canteens, I cannot identify the causal effect of the price increase in a difference in difference framework. Instead, Table 1.C.2 takes a descriptive approach, running a simple linear regression controlling for factors other than prices that might affect purchasing behavior, and identifying how the residual variation in purchasing behavior correlates with the price difference between the meat and vegetarian main meal component.

I include the following controls: To control for time trends, I include week and day-of-the-week effects. To control for the main meal components on offer differing in their attractiveness, I include over 100 binary meal-specific control variables controlling for the two most sold main meal components offered in a given canteen on a given day.

Spec. (1) in Table 1.C.2 includes all student[57] purchases of main meal components in this period. Spec. (2) restricts the analysis to purchases of the two main meal components that are most sold in a given canteen on a given day. This is my preferred specification since I can here control for the attractiveness of every meal in the sample. I find that a €0.01 increase in the price difference between vegetarian and meat main components correlates with a 0.25 percentage point decrease in demand for the meat main component and a corresponding increase in demand for the vegetarian main component. An increase in the price difference between the meat and the vegetarian main component of €0.01 can roughly be understood as a carbon tax of €0.01 per kg (or €10 per tonne). This is because the average emissions difference between the meat and the vegetarian main meal component offered in the student canteen is around 1 kg.

The effect I identify for such a carbon tax of €10per tonne—a 0.25 percentage point decrease in demand for the meat main component—is one eighth of the effect of carbon labels identified in the causal analysis shown in the main text (implying rough equivalence of this effect to a carbon tax of €80 per tonne), and one twelfth of the effect of the carbon labels identified within the regression analysis shown in Table 1.C.2 (implying rough equivalence of this effect to

57. I only include purchases made by students in the analysis since employees and guests face a different price structure.

a carbon tax of €120 per tonne). In Experiment 1, the framed field experiment discussed in Section 1.2, I estimate that carbon labels produce a similar effect as a carbon tax of €120 per tonne. My field results can second this estimate: In the student canteen setting, I estimate the effect of the carbon labels to be similar to that of a carbon tax of €80 to €120 per tonne introduced in the same setting.

To provide further context to my results and compare with previous literature, I calculate price elasticities: €0.1 is around 5.0% of the meat meal price at baseline, and 2.5 percentage points is around 5.6% of the baseline demand for the meat meal. This would imply an own-price elasticity of around −1,1 for the meat meal. Since a 2.5 percentage point decrease in demand for the meat main component corresponds to a corresponding increase in demand for the vegetarian main component, figures can also be used to calculate a cross-price elasticity for the vegetarian meal: A 2.5 percentage point increase in demand for the vegetarian meal is around 4.5% of the baseline demand for the vegetarian meal, implying a cross-price elasticity of 0.9.

Compared to estimates of Wirsenius, Hedenus, and Mohlin (2011), these estimates are rather on the higher end. To calculate the impact of an EU-wide carbon tax on animal products, they assume an own price elasticity of −0.5 for eggs, −0.5 for dairy products, −1 for poultry, −0.8 for pork and −1.3 for ruminant meat for food demand in the EU. Further, they estimate a slightly negative cross-price elasticity for cereals (−0.01) following a price increase for meat products, and a zero cross-price elasticity for other vegetarian products.

One reason for the differences in estimates is most likely the vastly different context: Increasing the price of the meat option in the student canteen is different from imposing an EU-wide tax on meat. Students might be especially price-sensitive and might also substitute their meat consumption intertemporally. Since the price of the meat and the vegetarian component offered in the student canteen differs across specific meals offered, the price difference between meat and vegetarian components fluctuates across days: After the price increase, the price difference is €0.4 in around 40% of cases, and €0.1, €0.6, and €0.9 in around 20% of the remaining cases each. Students might thus respond to a particularly high price difference by moving meat consumption to a day with a lower price difference.

Table 1.C.2. Comparison of effects: labels vs. “carbon tax”

	Likelihood of consuming meat	
	(1)	(2)
Price difference (in Euro)	-0.16***	-0.25***
	(0.05)	(0.05)
Treatment restaurant x Label period	-0.04***	-0.03***
	(0.00)	(0.01)
Treatment restaurant	-0.05***	-0.04***
	(0.00)	(0.00)
Constant	0.34***	0.41***
	(0.07)	(0.08)
Week and Day of the week controls	Yes	Yes
Meal-specific controls	Yes	Yes
Guests control	12,053	11,239
Guests treated	5,496	4,878
Observations	384,767	343,891

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 drawing on student canteen data from April 2022–March 2023. Both specifications include week and day-of-the-week effects, as well as over 100 binary controls for the day and student-canteen specific most-sold vegetarian and meat meal. The variable “Price difference” describes the price difference between these two most-sold options. Spec. (2) keeps only sales of the two most popular options.

1.C.3 Effect on carbon footprint

For an analysis of the impact on average greenhouse gas emissions per meal, I restrict the main sample such that it only includes days in the intervention period for which there is a "gastronomic twin" in the pre-intervention period: a day in the pre-intervention period where the same two main meal components were served. Further, I drop any sales not related to the two main components shared between treatment and control restaurants. The reason for this restriction is that the average emissions per meal vary a lot between days due to a changing offer (see Figures 1.E.7 and 1.E.6 for a comparison of daily variations in meat consumption vs. daily variation in average emissions). As vegetarian consumption is, at baseline, higher in the treated than in the control restaurants, a less restricted analysis might falsely attribute changes in the carbon footprint of the meals offered in the pre-intervention vs. in the intervention period to the label. [58] The restricted sample contains 33,427 observations. As shown in Table 1.C.3 in the Appendix, I estimate that labels reduce average emissions per meal by 25 grams or around 3% of the emissions of a baseline meal.

Table 1.C.3. Effect of labels on average emissions per meal

	(1) GHGE (g)	(2) GHGE (g)
Treatment restaurant x Label period	-17.31	-25.39**
	(11.26)	(10.25)
Treatment restaurant	-49.14***	-44.34***
	(7.44)	(6.74)
Label period	5.12	
	(6.26)	
Date effects	No	Yes
Fixed effects	No	No
Guests control	5,075	5,075
Guests treated	1,977	1,977
Observations	33,427	33,427

Standard errors in parentheses

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

Notes: Dependent variable: Emissions caused by main meal component, in gram. The sample is restricted to days in the intervention period for which there is a "gastronomic twin" in the pre-intervention period. Regression follows Spec. (1) and (2) in Table 1.2, using greenhouse gas emissions instead of the choice of the meat meal as the outcome variable.

58. As a simple illustration of why this is necessary: Imagine there is only one pre-intervention and one intervention day. On the pre-intervention day, the offer is a vegetarian meal with emissions of 0.3 kg and a meat meal with 1 kg of emissions per meal. In the treated restaurant, 59% of visitors consume vegetarian at baseline, so average emissions are 0.59 kg. In the control restaurant, 50% consume vegetarian at baseline, so average emissions are 0,65 kg. On the intervention day, the vegetarian offer still has 0.3 kg, but the meat meal now has 1.2 kg. Assuming no change in behavior, average emissions in the treated restaurant are 0.67 kg and 0.75 kg in the control restaurant. A naive analysis would then identify a differential 0,02 decrease in emissions in the treated restaurant compared to the control restaurant, although consumer behavior did not change. Thus, for the emissions analysis, I restrict the sample to establish an identical offer between the pre-intervention and intervention periods.

1.C.4 Heterogeneity in treatment effects

Table 1.C.4 examines treatment effects in different subsamples, using Spec. (2) of Table 1.2. Treatment effects are similar when restricting the sample to only employees (col. 2), to off-peak visit hours (col. 3), to purchases made with an individual payment card (col. 4) and to restaurant guests paying by individual card and visiting the student canteen rather frequently (at least ten times during the 13 weeks, col. 5). Table 1.C.5 shows analyses restricting the sample to guests who pay by individual payment card and for whom I have demographic information (around 1,400 guests). This suggestive analysis indicates that treatment effects are stronger for females, canteen guests below 24 of age, and those who report environmental aspects playing an important role in consumption choice.

Table 1.C.4. Effect of labels on meat consumption, different subsamples

	Likelihood of consuming meat				
	(1)	(2)	(3)	(4)	(5)
	All	Employees	Non-busy time	Card payment	Frequent
Treatment restaurant x Label period	-0.02^{***}	-0.05^{*}	-0.03^{***}	-0.02^{**}	-0.02^{*}
	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)
Treatment restaurant x Post period	-0.07^{***}	-0.11^{***}	-0.06^{***}	-0.09^{***}	-0.08^{***}
	(0.01)	(0.03)	(0.01)	(0.01)	(0.02)
Treatment restaurant	-0.10^{***}	-0.03	-0.10^{***}	-0.05^{***}	-0.05^{***}
	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
Constant	0.48^{***}	0.64^{***}	0.49^{***}	0.39^{***}	0.41^{***}
	(0.01)	(0.03)	(0.02)	(0.02)	(0.02)
Date fixed effects	Yes	Yes	Yes	Yes	Yes
Guest fixed effects	No	No	No	No	No
Guests control	6,924	883	3,784	5,972	1,820
Guests treated	2,810	263	1,678	2,509	721
Observations	120,082	20,836	67,635	70,109	46,988

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 model regression following spec. (2) in Table 1.2. Col.(1) includes all data, Col.(2) only university employees, Col.(3) excludes peak hours, Col.(4) to payments made by individual payment card, and Col.(5) includes only guests who visited the canteen at least ten times during the 13-week sample period. Spec. (1)-(3) estimate robust standard errors, and Spec. (4)-(5) cluster standard errors at the individual level.

Table 1.C.5. Effect of labels on meat consumption, different subsamples

	Likelihood of consuming meat					
	(1) All	(2) Survey	(3) Female	(4) < 24	(5) +Env.	(6) Decide online
Treatment x Label period	−0.02**	−0.03	−0.05*	−0.04*	−0.04*	−0.08
	(0.01)	(0.02)	(0.03)	(0.02)	(0.02)	(0.05)
Treatment x Post period	−0.09***	−0.09***	−0.08***	−0.09***	−0.08***	−0.08
	(0.01)	(0.02)	(0.03)	(0.02)	(0.03)	(0.05)
Treatment	−0.05***	0.00	0.06**	0.01	−0.02	−0.03
	(0.01)	(0.02)	(0.03)	(0.03)	(0.03)	(0.06)
Constant	0.39***	0.25***	0.16***	0.23***	0.11***	0.25***
	(0.02)	(0.03)	(0.04)	(0.04)	(0.03)	(0.08)
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Guest fixed effects	No	No	No	No	No	No
Guests control	5,972	850	465	572	446	152
Guests treated	2,509	522	271	355	234	118
Observations	70,109	15,038	6,945	10,370	7,149	2,693

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 model regression following spec. (2) in Table 1.2. Col.(1) includes all data, Col.(2) only student canteen guests who participated in the pre-intervention field survey. Col.(3) includes, of these, only females. Col.(3) includes only under 24-year olds. Col.(5) includes only survey participants who report that environmental aspects play an important role in their food consumption decisions. Table 1.B.24 reports evidence from Experiment 1 for the same heterogeneity factors. Spec. (1) estimates robust standard errors, and Spec. (2)-(5) cluster standard errors at the individual level.

1.C.5 Field survey results

Below I describe the results of surveys conducted in the control and treatment canteens pre- and post- intervention, as described in section 1.E.6.

Did canteen guests see the labels? Of the post-survey respondents, 340 went to the treated student canteen at least once during the intervention period. 72% of these report having seen the labels. 516 respondents did not go to the treated canteen during the intervention period, according to their individual student canteen cards. However, they might have in fact still gone, but not paid with their individual cards. Of these respondents, 13% report having seen the labels. 177 respondents went to the treated restaurant at least four times during the intervention period. 80% of these guests report having seen the labels.

Do canteen guests feel they reacted to the labels? Of the post-survey respondents who noticed the labels and visited the treated student canteen at least once during the intervention period, 18% report having incorporated the labels in their decisions (agreement of 4 or 5 on a 5-point scale asking how strongly participants incorporated the labels in their choices). Of those who visited the canteen more frequently (146 participants). 16% report having incorporated the labels in their decisions.

How do canteen guests make their consumption choices? 34% of guests report making their choice mainly using the information given on the canteen website. 29% mainly use the digital billboards. 36% report mainly deciding by looking at the food counters. Figure 1.9 shows how the carbon labels were shown in each of these decision contexts. Of the three decision contexts, the carbon labels were most salient on the canteen website. Table 1.C.5 shows how treatment effects differ for guests making their decisions online. Results suggest that effects are stronger for this group.

Do the carbon labels affect other attitudes? I do not find any evidence of the carbon labels affecting my measure of support for a carbon tax or for command-and-control measures (see Table 1.C.6). I also do not find any evidence that they significantly affect students' experience of eating in the canteen.

Table 1.C.6. Effect of the labels on other attitudes

	Approval of...					
	(1) tax	(2) c-c	(3) experience	(4) tax	(5) c-c	(6) experience
Treatment × Post period	−0.03	−0.15	−0.14	0.01	−0.18	−0.15
	(0.08)	(0.10)	(0.09)	(0.10)	(0.12)	(0.11)
Post period	−0.04	0.23***	0.08	−0.04	0.23***	0.08
	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)
Constant	4.31***	4.39***	4.43***	4.33***	4.37***	4.43***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Guest fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Guests control	472	472	472	472	472	472
Guests treated	359	359	359	208	208	208
Observations	1,662	1,662	1,662	1,360	1,360	1,360

Standard errors in parentheses

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

Notes: Dependent variable: Spec. (1) and (4): Agreement with the statement "Flying should be more expensive since it is bad for the environment." Spec. (2) and (5): "It should be prohibited to build new houses not adhering to cur- rent environmental standards." Spec. (3) and (6): ""Eating in the student canteen is a nice experience for me." All are measured on a 7-point scale. I only include individuals who participated in the pre-and the post- survey and who made at least one student canteen purchase during the 14-week study period. In Spec. (1)-(3), I classify an individual as treated if he visits the treated canteen at least once during the intervention period. In Spec. (4)-(6) I only include those treated individuals who visit the treated canteen at least three times during the intervention period. I only include individuals in the analysis who responded to both surveys and include guest fixed effects to control for initial attitudes. Standard errors are clustered at the individual level.

Appendix 1.D Experiments 1 and 3: Details on the experimental set-up

1.D.1 Meals used for elicitation

In the purchasing decisions in experiments 1 and 3, participants make decisions on the same four student canteen meals. These are all meals which are regularly offered in the student canteen. Participants who indicate that they are not vegetarian decide on two vegetarian and two meat meals: Filled courgettes with potato croquettes (1.4 kg of emissions, colored yellow in the labels), Italian vegetable ragout with pasta (0.5 kg of emissions, colored green in the labels), Chicken Schnitzel with rice (1.4 kg of emissions, colored yellow in the labels), and beef ragout with potatoes (3.4 kg of emissions, colored red in the labels). Participants who indicate they are vegetarian decide on four vegetarian meals: Filled courgettes with potato croquettes (1.4 kg of emissions, colored yellow in the labels), Italian vegetable ragout with pasta (0.5 kg of emissions, colored green in the labels), Cheese "Spätzle" with mushrooms (1.2 kg of emissions, colored yellow in the labels), and stir-fried vegetables with rice (0.4 kg of emissions, colored green in the labels). The cheese sandwich is the outside option to every choice and causes 0.7 kg of emissions and is colored green on the labels.

I randomized the order in which meals appear (both in the decision and the emission estimating screens) to avoid order effects. Further, I changed the left-right positioning of the warm meal vs. the cheese roll to right-left for half of the experiment sessions to avoid positioning effects.

1.D.2 Incentivization of elicitations

The elicitation of participants' **willingness to pay for consuming the meals** is incentivized with an adapted BDM mechanism: There is a 50% probability that the specific meal and a 50% probability that the cheese sandwich is randomly drawn as the default meal. If the default meal and the preferred meal indicated in the first part of the decision (e.g. Figure 1.2) coincide, the participant is given the preferred meal at zero price. If the two do not coincide, a price is randomly drawn at which the two options can be exchanged. Each value between €0.00 and €3.00 can be drawn with equal probability, in five-cent steps. If the willingness to pay indicated by the participant in the second part of the decision (e.g. Figure 1.3) is equal to or above the price drawn, the price is deducted from the participants' payment and participants are provided with the preferred option. If willingness to pay is below the price drawn, participants are provided with the less preferred option, and no amount is deducted from participants' payments. The outcome lunch is provided to participants directly after the experiment, together with participants' payment in cash. For this purpose, experiment participants are required to travel to the university campus immediately after completing the experiment. Less than 4% did not pick up their cash payment and meal. The incentivization structure was explained to participants and they were required to pass an extensive comprehension check, which less than 4% of participants did not pass.

This **willingness to pay for seeing labels elicitation** is incentivized with a similar BDM mechanism. There is a 50% probability that the default option is that choices are displayed with, and a 50% probability that the default option is that choices are displayed without labels. If the default display option and the preferred display option coincide, the preferred display option is implemented at zero price. If the two do not coincide, a price is randomly drawn at which the display option can be changed. Each value between €0.00 and €3.00 can be drawn with equal probability, in five-cent steps. If the willingness to pay indicated by the participant in the second part of the decision (similar to Figure 1.3, with display options instead of meals) is equal to or higher than the price drawn, the preferred display option is implemented. The price drawn is only deducted from participants' payment if one of the final three meals is relevant for pay-out. If the willingness to pay is lower than the price drawn, the less-preferred display option is implemented.

1.D.3 Decisions under carbon offsetting

In the Attention+Offset condition in Experiment 3 and the Offset condition in Experiment 1, participants are informed that, if one of the decisions made in this treatment is implemented, the emissions of the meal provided to them (regardless of whether it is the warm meal or the cheese sandwich) are offset by the experimenter with a donation to Atmosfair. The example screens in Subsection 1.D.4 show how this is communicated to experiment participants.

Towards the end of the experiment, after participants have completed all meal decisions, I elicit participants' attitudes towards the effectiveness of carbon offsetting and ask for participants' prior experiences with carbon offsetting. Tables 1.D.1 and 1.D.2 show descriptives pooled across Experiments 1 and 3. Table 1.D.1 shows that 75% of participants had heard of carbon offsetting previously, while 34% have used carbon offsetting themselves.

Table 1.D.2 shows that participants broadly agree with carbon offsetting being effective (Measured as agreement to the statement "Voluntary carbon offsetting is an effective climate protection measure"). They disagree with them replacing other climate protection measures (Measured as agreement to the statement "If I offset emissions for a carbon-intensive activity such as a flight, it is okay to book another flight."). They agree with carbon offsetting not replacing other climate protection activities (Measured as agreement to the statement "Carbon offsetting cannot replace personal efforts to protect the climate."). Interestingly, having experienced the Attention+Offset or the Offset condition earlier in the experiment increases support for the second and decreases support for the third statement.

These descriptive statistics convey that carbon offsetting likely removes a part of environmental guilt, but may not be removing it entirely.

Table 1.D.1. Familiarity with carbon offsetting

	Familiarity with offsetting	
	(1) Heard of	(2) Have used
In offset condition earlier in exp.	-0.04	-0.01
	(0.03)	(0.04)
Constant	0.75***	0.34***
	(0.03)	(0.03)
Observations	732	732

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

Table 1.D.2. Beliefs on carbon offsetting effectiveness

	Familiarity with offsetting		
	(1)	(2)	(3)
	Effective	Can replace	Cannot replace
In offset condition earlier in exp.	0.15	0.45***	-0.50***
	(0.18)	(0.16)	(0.17)
Constant	5.55***	2.86***	8.14***
	(0.14)	(0.12)	(0.13)
Observations	732	732	732

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

1.D.4 Experiment screens (English translation)

Survey start screen

Welcome to the BonnEconLab online study. Please note that you may only take part in this study once. Furthermore, you may only take part if you have registered for the study in our participation database. Please complete this survey on your computer. Participation with mobile devices such as smartphones or tablets is not possible. The payout for this experiment will be done using your personal participant code: 12pI2q5vh Please write down your code! You will need approximately 45 minutes to process this survey. After fully completing the survey, you can collect your payout at our location at the Hofgartenwiese (see map below) until 2 p.m. today. You will not be able to receive your payout at any other time! In this experiment, your payout consists of several components:

- You receive exactly one dish (your lunch).
- You receive an expense compensation of €9.00 in cash.
- You may receive an additional payout of up to €1.60 in addition to the expense compensation. This depends on your answers in the marked part of the study.
- In addition, chance determines whether, depending on your answers in another (also clearly marked) part of the study, you will receive another additional payout of up to €1.10.

Payment will be made in the BonnEconLab pavilion on the Hofgartenwiese (Regina-Pacis-Weg). You will find us at the place marked with a blue arrow under a pavilion.

Decision description screen - Screenshot 1/4

Description of upcoming decisions | Comprehension questions

The second part of the study is about to begin. Your decisions in this part of the study will affect your expense compensation and the dish you receive.

On this page you will find explanations and examples. On the following page we will check your understanding of these explanations. By clicking on the tab above you can switch between the two pages.

Once the comprehension questions have been answered correctly, you can proceed with further work on the survey.

How do your decisions affect your payout?

- In this experiment, your payout consists of three components:
  - You receive exactly one dish (your lunch).
  - You receive an expense compensation. At the moment, the expense compensation is €9.00. You will make a total of 15 decisions over the course of this study. For each of these decisions, you have the option of waiving part of the expense compensation (maximum €3.00). For that, you will receive a court you prefer.
  - In two other parts of the study, you may receive an additional amount of up to €1.60 in addition to the expense compensation, depending on your answers. In addition, depending on your answers in a third part of the study, chance will determine whether you will receive an additional amount of up to €1.10. The relevant parts of the study are clearly marked.
- For each of the 15 decisions, indicate which of the two courts you prefer. Then specify the maximum amount of your expense compensation you would like to forgo in order to receive the preferred court.

The decision that is implemented shall be subject to the following:

- Chance decides whether you will receive your favourite dish for free:
  - Case 1 (50% probability): You will receive your favourite dish for free.
  - Case 2 (50% probability): You will be assigned the non-preferred dish first. In this case, specify the maximum amount of your expense compensation you would like to forgo in order to receive your favourite dish instead.

- If case 2 occurs, it is again a matter of chance:
  - A **surcharge** is determined at random. Any value between €0 and €3 (in 5 cent increments) is equally probable.
  - If the amount you have declared is more than the surcharge, you will receive your preferred dish. For this, the surcharge will be deducted from your expense compensation.
  - If the amount you specify is less than the surcharge, you will receive the non-preferred dish free of charge.

For the other 14 decisions which are not being implemented, the following rules apply:

- These decisions have no effect on the dish you receive.
- These decisions have no effect on your compensation.

You will not know which of the 15 decisions will be implemented until the end of the study. It is therefore in your best interest to make every decision carefully.

Example decision

You can receive either a cheese roll or the 'Baked Feta Cheese with Rice' dish.

Which dish do you prefer? Click on one of the two buttons. Try it!

Baked Feta Cheese with Rice vegetarian	oder	Cheese Roll Details: vegetarian
[Baked Feta Cheese with Rice]		[Cheese Roll]

Example scenario 1

Assuming you made the following decision:

Which dish do you prefer? Click on one of the two buttons.

Baked Feta Cheese with Rice

vegetarian

oder

Cheese Roll

Details: vegetarian

Baked Feta Cheese with Rice

Cheese Roll

If you are given the cheese roll: What is the maximum amount of your expense compensation you would be willing to give up in exchange for stir-fry sweet and sour with rice?

(Click on the grey bar to make the slider visible).

0.00€ 1.00€ 2.00€ 3.00€

You would like to give up a maximum of **€1.20** of your allowance to receive the dish **Baked Feta Cheese with Rice** instead of the cheese roll.

Here's what happens in this example (which you have no control over):

- You are first assigned your **less preferred dish**, the cheese roll.
- A **surcharge of €0.60** is randomly determined.

This means for you:

The surcharge with the amount of 0,60 € is lower than the maximum amount of 1,20 € you specified. You will receive the dish 'Baked feta cheese with rice'. For this, € 0.60 will be deducted from your expense compensation.

Decision description screen - Screenshot 3/4

Example scenario 2

Assuming you made the following decision:

Which dish do you prefer? Click on one of the two buttons.

Baked Feta Cheese with Rice

vegetarian

oder

Cheese Roll

Details: vegetarian

Baked Feta Cheese with Rice

Cheese Roll

If you are given the cheese roll: What is the maximum amount of your expense compensation you would be willing to give up in exchange for stir-fry sweet and sour with rice?

(Click on the grey bar to make the slider visible).

0.00€ 1.00€ 2.00€ 3.00€

You would like to give up a maximum of **€1.20** of your allowance to receive the dish **Baked Feta Cheese with Rice** instead of the cheese roll.

Here's what happens in this example (which you have no control over):

- You are first assigned your **less preferred dish**, the cheese roll.
- A **surcharge of 2.00 €** is randomly determined.

This means for you:

The surcharge with the amount of 2.00 € is higher than the maximum amount of 1,20 € you specified. You will receive the cheese roll. Therefore, nothing will be deducted from your expense compensation.

Decision description screen - Screenshot 4/4

Example scenario 3

Assuming you made the following decision:

Which dish do you prefer? Click on one of the two buttons.

Baked Feta Cheese with Rice

vegetarian

oder

Cheese Roll

Details: vegetarian

Baked Feta Cheese with Rice

Cheese Roll

If you are given the cheese roll: What is the maximum amount of your expense compensation you would be willing to give up in exchange for stir-fry sweet and sour with rice?

(Click on the grey bar to make the slider visible).

0.00€ 1.00€ 2.00€ 3.00€

You would like to give up a maximum of **€1.20** of your allowance to receive the dish **Baked Feta Cheese with Rice** instead of the cheese roll. u

Here's what happens in this example (which you have no control over):

- You are assigned your **preferred dish**, 'Baked feta cheese with rice', for free.

This means for you:

You receive the dish 'Baked feta cheese with rice'. Nothing will be deducted from your expense compensation.

Continue to the questions

You can always return to this page while answering the questions.

Comprehension questions - Screenshot 1/3

Description of upcoming decisions | Comprehension questions

Comprehension questions

Please answer the following comprehension questions. If you want to look at the description of the survey again, you can switch back and forth between this page and the previous page by clicking on the tab at the top.

After correctly answering the comprehension questions, you can continue with the further processing of the survey.

Question 1

Assuming you made the following decision:

Which dish do you prefer? Click on one of the two buttons.

Baked Feta Cheese with Rice	oder	Cheese Roll
vegetarian		Details: vegetarian
Baked Feta Cheese with Rice		Cheese Roll

If you are given the cheese roll: What is the maximum amount of your expense compensation you would be willing to give up in exchange for stir-fry sweet and sour with rice?

(Click on the grey bar to make the slider visible).

0.00€ 1.00€ 2.00€ 3.00€

You would like to give up a maximum of **€1.30** of your allowance to receive the dish **Cheese Roll** instead of the Baked Feta Cheese with Rice.

Here's what happens in this example (which you have no control over):

- The decision was carried out.
- You are first assigned your **less preferred dish**, the Baked Feta Cheese with Rice.
- A **surcharge of 0.70 €** is randomly determined.

What do you receive?

- ○ The baked feta cheese with rice and your full expense compensation.
- ○ The baked feta cheese with rice and 0.70 euros will be deducted from your expense compensation.
- ○ The cheese roll and 0.70 euros will be deducted from your expense compensation.
- ○ The cheese roll and your full expense compensation.

Question 2

Assuming you made the following decision:

Which dish do you prefer? Click on one of the two buttons.

Baked Feta Cheese with Rice

vegetarian

oder

Cheese Roll

Details: vegetarian

Baked Feta Cheese with Rice

Cheese Roll

If you are given the cheese roll: What is the maximum amount of your expense compensation you would be willing to give up in exchange for stir-fry sweet and sour with rice?

(Click on the grey bar to make the slider visible).

0.00€ 1.00€ 2.00€ 3.00€

You would like to give up a maximum of **€1.30** of your allowance to receive the dish **Cheese Roll** instead of the Baked Feta Cheese with Rice.

Here's what happens in this example (which you have no control over):

- The decision was carried out.
- You are assigned your **preferred dish**, the cheese roll.

What do you receive?

- The baked feta cheese with rice and your full expense compensation.
- The baked feta cheese with rice and 0.70 euros will be deducted from your expense compensation.
- The cheese roll and 0.70 euros will be deducted from your expense compensation.
- The cheese roll and your full expense compensation.

Question 3

Assuming you made the following decision:

Which dish do you prefer? Click on one of the two buttons.

Baked Feta Cheese with Rice	or	Cheese Roll
vegetarian		Details: vegetarian
Baked Feta Cheese with Rice		Cheese Roll

If you are given the cheese roll: What is the maximum amount of your expense compensation you would be willing to give up in exchange for stir-fry sweet and sour with rice?

(Click on the grey bar to make the slider visible).

0.00€ 1.00€ 2.00€ 3.00€

You would like to give up a maximum of **€1.30** of your allowance to receive the dish **Cheese Roll** instead of the Baked Feta Cheese with Rice.

Here's what happens in this example (which you have no control over):

- The decision was carried out.
- You are first assigned your **less preferred dish**, the Baked Feta Cheese with Rice.
- A **surcharge of 2.70 €** is randomly determined.

What do you receive?

- The baked feta cheese with rice and your full expense compensation.
- The baked feta cheese with rice and 2.70 euros will be deducted from your expense compensation.
- The cheese roll and 2.70 euros will be deducted from your expense compensation.
- The cheese roll and your full expense compensation.

Question 4

How many of the 15 decisions actually have an impact on the dish you are handed and your expense compensation?

- All the 15 decisions have an impact.
- Five of the 15 decisions have an impact.
- One of the 15 decisions has an impact.
- One of the 15 decisions has an influence.

Back to the explanation | Continue with the rest of the survey

Example baseline decision

You can receive either a cheese roll or the dish ‘Stir-fry sweet and sour with rice’ with your payout.

Which dish do you prefer? Click on one of the two buttons.

Stir-fry sweet and sour with rice

vegetarian

or

Cheese Roll

vegetarian

Stir-fry sweet and sour with rice

Cheese Roll

If you are given the Stir-fry sweet and sour with rice: What is the **maximum** amount of your expense compensation you would be willing to give up in exchange for the cheese roll?

(Click on the grey bar to make the slider visible).

0.00€ 1.00€ 2.00€ 3.00€

You would like to give up a maximum of **€0.75** of your allowance to receive the dish **cheese roll** instead of the Stir-fry sweet and sour with rice.

Continue

Introduction carbon footprint estimation

You will now guess for a total of eleven meals how high the CO2 emissions are which are caused by the respective meal.

- You have **60 seconds** to answer each question.
- For each question in which your guess does not deviate from the correct value by more than 30%, **0.10 Euro is added to your payout**.

During each guessing question you will be shown the emissions caused by the meal “Red Thai Curry with Pork and Rice” **as a reference value**.

Your reference value:

Red Thai-Curry with Pork and Rice

CO_2 Causes ± 1,7 kg CO_2
≈ 8,5 km Car drive

Pork

Which assumptions should be taken for the guessing questions?

For the following questions you will not be shown any ingredient lists or a description of the origin of the ingredients. This is because we only want to give you the information which you would normally find in a restaurant. We would like to know how you, based only on the name of the meal on the menu, guess the magnitude of the emissions caused by a meal.

Of course, the emissions of a seemingly identical meal can differ, e.g., depending on the exact ingredients and depending on whether the ingredients were produced in an ecologically sustainable or in a conventional manner. Please assume a conventional production and a conventional meal preparation – just like you would expect it, if you are offered such a meal without any further information in a restaurant.

Please take into account all emissions caused in the agricultural production and in food processing, packaging, conservation and transport of ingredients, up until an ingredient can be purchased in the store. You do not need to take into account emissions which are caused by the transport of ingredients from store to restaurant

Continue

Example carbon footprint estimation

Remaining time on this page. **0:54**

What do you estimate: How high are the greenhouse gas emissions (in CO2-equivalents), which are caused by the meal “Stuffed Zucchini with croquettes”?

Guess the emissions:	As a reference:
Stuffed Zucchini with croquettes CO_2 Causes ? kg CO_2 Vegetarian	**Red Thai-Curry with Pork and Rice** CO_2 Causes 1,7 kg CO_2 ≈ 8,5 km Car drive Pork

I estimate that the meal “Stuffed Zucchini with croquettes” causes emissions of

[ ] kg.

Continue

You will now make four more of the 15 decisions. One of the 15 decisions will be implemented.

You will be shown the greenhouse gas emissions (in CO2 equivalents) of both dishes for the upcoming decisions.

For those interested: More information on the calculation of greenhouse gas emissions:

What assumptions are made in the calculation?

In the calculation, the emissions attributable to a dish are calculated as the sum of the emissions generated in the production of the ingredients. The emissions of each ingredient are calculated "from farm to gate", i.e. all emissions are included that occur during agricultural production and during further processing, packaging, preservation and transport until the ingredient is available for purchase in shops. Not included are the transport from the shop to the restaurant or end consumer and the emissions that arise from any further refrigeration in the restaurant or at the end consumer, as well as the emissions that arise from cooking the dish.

When calculating the values, conventional (i.e. not specifically organically certified) agriculture is assumed. Otherwise, assumptions are made about production that reflect the production of the average product found on our supermarket shelves.

What data is the calculation based on?

The Eaternity database on which the calculations are based is currently the largest and most comprehensive database for calculating the climate-relevant emissions of meals and food products. It includes more than 550 ingredients and other parameters on organic and greenhouse production as well as production, processing, packaging and preservation. The eaternity database is maintained by scientists from the Zurich University of Applied Sciences (ZHAW), the University of Zurich (UZH), the Swiss Federal Institute of Technology Zurich (ETH Zurich), the Research Institute of Organic Agriculture (FiBL), Quantis and other institutions.

Source: eaternity.

Continue

Example decision with labels

You can either get a cheese roll or the dish 'stir-fry sweet and sour with rice' with your payout.

Which dish do you prefer? Click on one of the two buttons.

Stir-fry sweet and sour with rice	or	Cheese Roll
CO_2 Causes 0,4 kg CO_2 ≈ 2,0 km Car drive		CO_2 Causes 0,7 kg CO_2 ≈ 3,5 km Car drive
vegetarian		vegetarian
Stir-fry sweet and sour with rice		Cheese Roll

If you are given the cheese roll: What is the **maximum** amount of your expense compensation you would be willing to give up in exchange for stir-fry sweet and sour with rice?

(Click on the grey bar to make the slider visible).

0.00€ 1.00€ 2.00€ 3.00€

You would like to give up a maximum of **€1.10** of your allowance to receive the dish **stir-fry sweet and sour with rice** instead of the cheese roll.

Continue

For those interested: More information on the calculation of greenhouse gas emissions:

What assumptions are made in the calculation?

In the calculation, the emissions attributable to a dish are calculated as the sum of the emissions generated in the production of the ingredients. The emissions of each ingredient are calculated "from farm to gate", i.e. all emissions are included that occur during agricultural production and during further processing, packaging, preservation and transport until the ingredient is available for purchase in shops. Not included are the transport from the shop to the restaurant or end consumer and the emissions that arise from any further refrigeration in the restaurant or at the end consumer, as well as the emissions that arise from cooking the dish.

When calculating the values, conventional (i.e. not specifically organically certified) agriculture is assumed. Otherwise, assumptions are made about production that reflect the production of the average product found on our supermarket shelves.

What data is the calculation based on?

The Eaternity database on which the calculations are based is currently the largest and most comprehensive database for calculating the climate-relevant emissions of meals and food products. It includes more than 550 ingredients and other parameters on organic and greenhouse production as well as production, processing, packaging and preservation. The eaternity database is maintained by scientists from the Zurich University of Applied Sciences (ZHAW), the University of Zurich (UZH), the Swiss Federal Institute of Technology Zurich (ETH Zurich), the Research Institute of Organic Agriculture (FiBL), Quantis and other institutions.

Source: eaternity.

Introduction decisions with offsetting

You will now make four more of the 15 decisions. One of the 15 decisions will actually be implemented.

If it is one of the now following four choices that is implemented, the greenhouse gas emissions of the dish you have been handed will be offset by a donation to the NGO atmosfair. This happens regardless of whether the dish was originally assigned to you or whether you exchanged it for the other dish by paying a surcharge. Atmosfair uses the donation to support sustainable energy projects so that the emissions are saved elsewhere. In this way, the dish handed out to you becomes emission-neutral / CO2-neutral.

For those interested: Further information on CO2 offsetting:

How does the CO2 offset work?

The donation to atmosfair is used to develop renewable energies in countries where they hardly exist yet, i.e. mainly in developing countries. In this way, atmosfair saves CO2 that would otherwise have been produced by fossil energies in these countries.

Example projects

- Atmosfair uses donations to reduce the selling price of energy-efficient stoves in Nigeria. In Nigeria, 75% of families cook on open fires, and a family of 7 consumes 5 tonnes of wood per year. This enormous consumption of firewood has already led to almost total deforestation and the progressive spread of deserts, especially in the poor north of the country. Energy-efficient stoves use about 80% less wood.
- Atmosfair uses donations to make small-scale biogas plants more affordable in Nepal. This project targets families living in rural areas who previously used wood as an energy source for cooking. In this way, the increasing deforestation of Nepal's forests can be counteracted.
- Atmosfair uses donations to support a small hydropower plant in Honduras. In this way, four villages that previously used wood and diesel generators for energy supply could be connected to the electricity grid for the first time. In addition, electricity can be fed into the national grid, replacing electricity from gas-fired power plants.

Source: atmosfair

Continue

Example decision with offsetting

You can either receive a cheese roll or the dish ‘Italian Vegetable ragout with pasta’ with your payout.

The emissions attributable to each dish are offset by a donation to the NGO atmosfair. Atmosfair supports sustainable energy projects with the donation, so that the emissions are saved elsewhere.

Which dish do you prefer? Click on one of the two buttons.

Italian Vegetable ragout with pasta		Cheese Roll
CO2 neutral Emission offset vegetarian	or	CO2 neutral Emission offset vegetarian
Italian Vegetable ragout with pasta		Cheese Roll

If you are assigned the cheese roll: What is the **maximum** amount of your expense compensation that you would be willing to give up in exchange for Italian Vegetable ragout with pasta? (Click on the grey bar to make the slider visible).

0.00€ 1.00€ 2.00€ 3.00€

You would like to give up a maximum of **0.75 €** of your expense compensation to receive the **Italian Vegetable ragout with pasta** instead of the cheese roll.

Continue

For those interested: Further information on CO2 offsetting:

How does the CO2 offset work?

The donation to atmosfair is used to develop renewable energies in countries where they hardly exist yet, i.e. mainly in developing countries. In this way, atmosfair saves CO2 that would otherwise have been produced by fossil energies in these countries.

Example projects

- Atmosfair uses donations to reduce the selling price of energy-efficient stoves in Nigeria. In Nigeria, 75% of families cook on open fires, and a family of 7 consumes 5 tonnes of wood per year. This enormous consumption of firewood has already led to almost total deforestation and the progressive spread of deserts, especially in the poor north of the country. Energy-efficient stoves use about 80% less wood.
- Atmosfair uses donations to make small-scale biogas plants more affordable in Nepal. This project targets families living in rural areas who previously used wood as an energy source for cooking. In this way, the increasing deforestation of Nepal's forests can be counteracted.
- Atmosfair uses donations to support a small hydropower plant in Honduras. In this way, four villages that previously used wood and diesel generators for energy supply could be connected to the electricity grid for the first time. In addition, electricity can be fed into the national grid, replacing electricity from gas-fired power plants.

Source: atmosfair

Introduction calorie estimation

You will now estimate the energy value of each dish in kilocalories (kcal) for a total of five dishes. For each estimation question, the completion time is **limited to 60 seconds**. For each estimation question where your estimate does not deviate from the correct value by more than 30%, **your payout increases by 0.10 euros**.

What assumptions should be made for the estimation?

You will not be presented with ingredient lists for the following estimation questions. This is because we want to give you, as much as possible, only the information that you would find in the restaurant. We want to know how you estimate the energy value of a dish, based solely on the name of the dish in the menu.

Continue

Example calorie estimation

Remaining time on this page. **0:54**

What do you estimate: What is the energy value in kilocalories (kcal) of the dish ‘Beef ragout with potatoes’?

Beef ragout with potatoes

Beef

I estimate that the dish ‘Beef ragout with potatoes’ has

kcal.

Continue

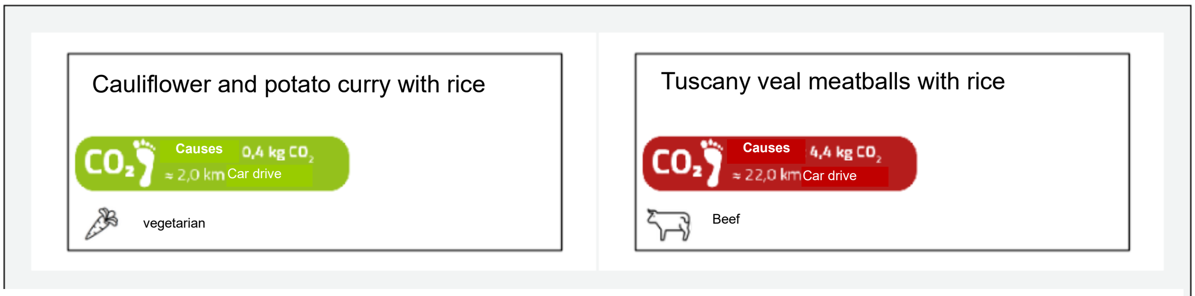
WTP for label presence

You are about to make the last three of the 15 decisions. One of the 15 decisions will actually be implemented.

But now there are two differences:

5. There are now three **new dishes** that you have not seen in your previous decisions.
6. You can see **emission labels** for these three dishes. These labels show the greenhouse gas emissions of the dishes in CO2 equivalents.

For example, two of the labels might look like this:



The display of the labels can either be preset so that:

- The labels are also displayed to you, or that
- The labels are not displayed to you.

Chance decides whether the display setting of the labels corresponds to your wishes without charge.

- Case 1 (probability 50%): We (do not) display the labels according to your wishes.
- Case 2 (probability 50%): The labels are initially preset so that it does not correspond to your wishes. For this case, you specify the **maximum** amount of your expense compensation you would like to give up in order to get your preferred display setting instead.

If **case 2** occurs, chance decides again:

- A **price** is determined randomly. Every value between 0€ and 3€ (in 5 cent steps) is equally probable.
- If the given amount is **higher** than the price, you will still get your **preferred display setting**. For this**, the charge will be deducted from your expense compensation**. However, this will only happen if one of the three dishes shown equally actually determines your payout.
- If the specified amount is **less** than the price, you will receive your **non-preferred display setting for free**.

Which display settings do you prefer? Click on one of the two buttons.

Labels should be **shown** | Labels should **not be shown**

If the display of labels is **not** preset and one of the three choices, you make now actually determines your payout: What is the **maximum** amount of your expense compensation you would like to give up in order to have the labels displayed?

(Click on the gray bar to make the slider visible).

0.00€ 1.00€ 2.00€ 3.00€

You want to give up a maximum of **1.70 €** of your expense compensation to **unlock the display of labels**.

Continue

Appendix 1.E Experiment 2: Details on the experimental set-up

1.E.1 Canteen set-up in Bonn

The natural field experiment was conducted in the student canteens of the University of Bonn from April 2022 to July 2022. The whole of April (four weeks) served as a pre-intervention phase in which baseline consumption decisions were observed. Emission labels were introduced in the treatment student canteen from the beginning of May to mid-June 2022 (seven weeks). From mid-June to mid-July 2022 (three weeks, which ended with the summer closing of the treated student canteen), I continue to observe consumption decisions to examine post-intervention behavior.

There are three student canteens in Bonn: The treatment student canteen, the first control restaurant (located 1.7 km from the treatment restaurant), and the second control restaurant (located 4.7 km from the treatment restaurant and frequented much less than the other two restaurants). Menu planning is centralized among the three student canteens, and there is thus a large overlap in the daily offering. All three student canteens offer two main meal components, which differ daily but are mostly the same across student canteens. In addition, each of the student canteens might offer additional options, which are student-restaurant-specific. The larger control restaurant sometimes offers pizza or pasta in addition, and all student canteens might serve leftover main meal components from the previous day, soup, and side dishes. In the treatment restaurant, only the main meal components were equipped with carbon labels, and sides and leftover main meal components were not labeled. [59] Correspondingly, the dependent variable in my main regression is whether the main meal component a restaurant guest chooses contains meat or is vegetarian.

1.E.2 Canteen visiting patterns

An average student canteen guest visited the student canteen 20 times from April to mid-July. Around 74% visit 10 times or more, and around 45% visit 20 times or more. 90% of guests visited the same student canteen at least 80% of the time. The student canteens offer very cheap meals, with complete meals costing between €1.00 and €3.00. In fast food restaurants located in the surrounding area, meals are priced at €4.00 upward. In a survey I conducted among over 1,000 student canteen guests (survey 2 described in the Appendix), over 40% of students report that they would have difficulty finding an affordable meal if the student canteens did not exist. Switching between student canteens and other gastronomic offers is thus also not frequent. Figure 1.E.1 in the Appendix includes an analysis based on the trackable personal card payments. I classify restaurant guests as "Treatment" or "Control" visitors based on their consumption behavior in the first two weeks. Around 3% of purchases made by "Control" visitors are made in the treated restaurant throughout the entire 14-week period. For "Treatment" visitors, the

59. The main reason for this was that I wanted to test carbon labeling in a manner that was feasible for the student canteen to implement long-term. While main meal components are planned and known beforehand, sides and leftover dishes are decided spontaneously. Further, leftover main meal components only make up a smaller part of daily sales and the emissions caused by side dishes are almost negligible compared to those of the main meal components. Sales of all products are tracked, and label effects in the main sample are conservatively calculated over all main meal components offered, i.e. including main meal components spontaneously added to the menu but not labeled.

percentage fluctuates between 5% and 9%. Figure 1.E.2 further examines which percentage of these non-home visits involve the consumption of a meat main component. There is no clear trend throughout the study period.

Further, an analysis of daily restaurant guests shows that the labeling intervention did not lead to a decrease in student canteen guests, relative to the control restaurant (see Figure 1.10). The introduction of carbon labels in the treatment restaurant was displayed as a measure taken by the student canteens themselves, with no connection presented to the University of Bonn or me specifically as the researcher. The introduction of the emission labels was explained on billboards and leaflets available inside the student canteen, as shown in Figure 1.E.5. I conducted two surveys accompanying the measure, one before the intervention period and one after the intervention period, further described in the Appendix. The surveys and the labeling measures were advertised through different channels, and the survey was advertised as a chance to voice one's opinion on the offer of the student canteen. It is thus unlikely that restaurant guests drew a connection between the initiative and the survey.

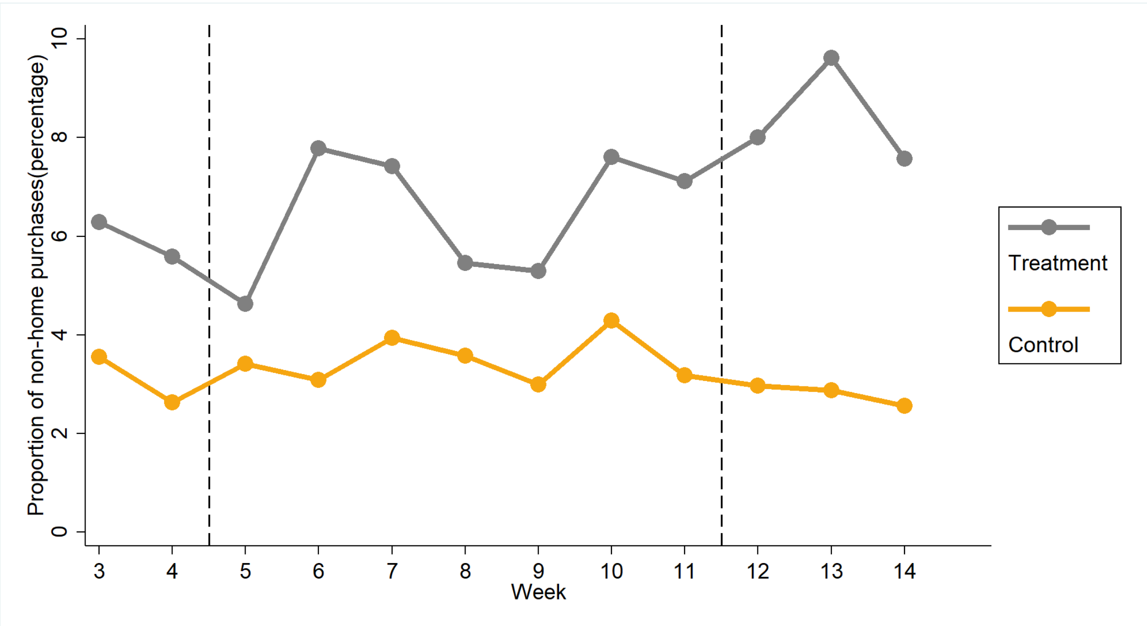


Figure 1.E.1. Visits to the "non-home" canteen

Note: In percentage points relative to total canteen visits. Classification as the "home" canteen based on behavior in the first two weeks. The sample is similar to that in spec. (4) in Table 1.2, but the intention to treat is calculated based entirely on the first two weeks, based on a minimum of two visits during this period. $N = 37,030$

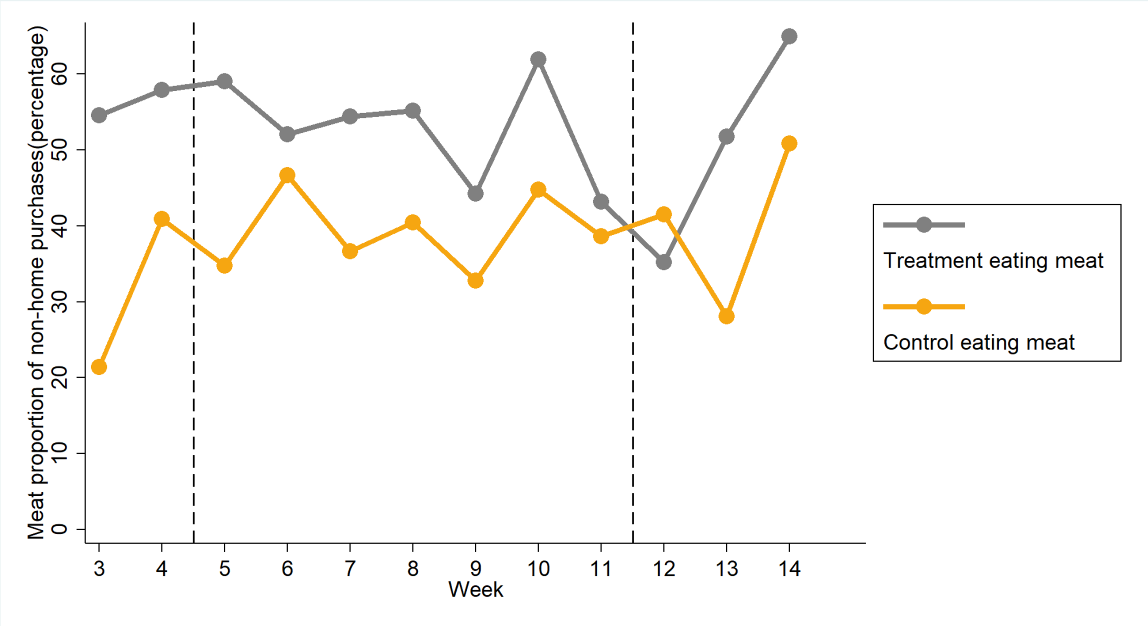


Figure 1.E.2. Meat consumption when visiting the "non-home" canteen

Note: In percentage points relative to total "non-home" canteen consumption. Classification as the "home" canteen based on behavior in the first two weeks. The sample is similar to that in spec. (4) in Table 1.2, but the intention to treat is calculated based entirely on the first two weeks, based on a minimum of two visits during this period. $N = 37,030$

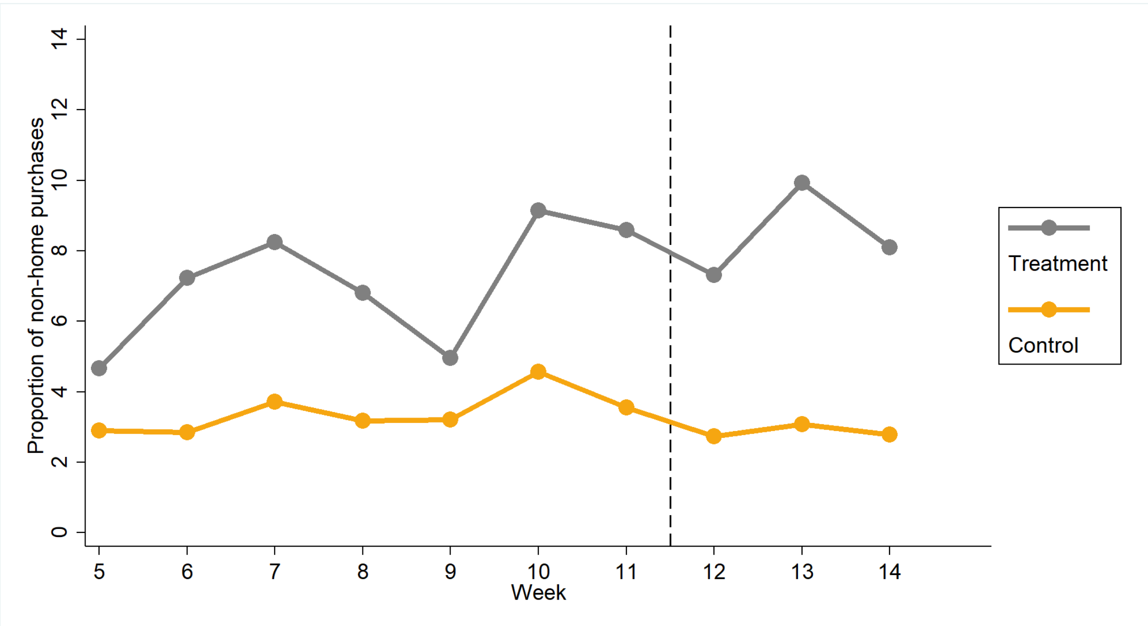


Figure 1.E.3. Visits to the "non-home" canteen, classification based on full pre-intervention phase

Note: In percentage points relative to total canteen visits. Classification as the "home" canteen based on behavior in the first four weeks (pre-intervention phase). The sample is similar to that in spec. (4) in Table 1.2, but also includes observations from week 5 of the sample period. $N = 45,628$

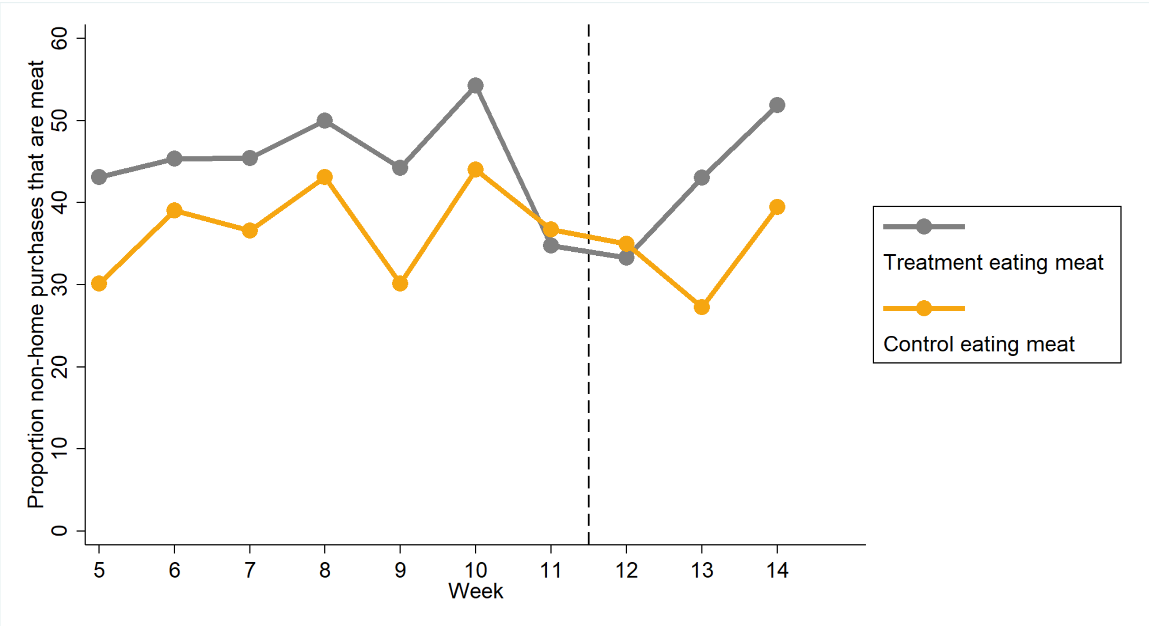


Figure 1.E.4. Meat consumption when visiting the "non-home" canteen, classification based on full pre-intervention phase

Note: In percentage points relative to total "non-home" canteen consumption. Classification as the "home" canteen based on behavior in the first two weeks (pre-intervention phase). The sample is similar to that in spec. (4) in Table 1.2, but also includes observations from week 5 of the sample period. $N = 45,628$

1.E.3 Carbon label calculation

For the carbon labels, I calculated emission values with the application Eaternity Institute (2020), using ingredient lists provided by the student canteen. The design of the carbon labels was proposed by the student canteen, based on what is technically feasible and possibly implementable as a long-run measure. Examples are shown in Figure 1.9. They were coded in a traffic-light system, with thresholds determined such that approximately a third of the main components offered by the student canteen during the study period would be classified as green, one-third as yellow, and one-third as red. This corresponded to thresholds of 0.7 kg and 1 kg. [60]

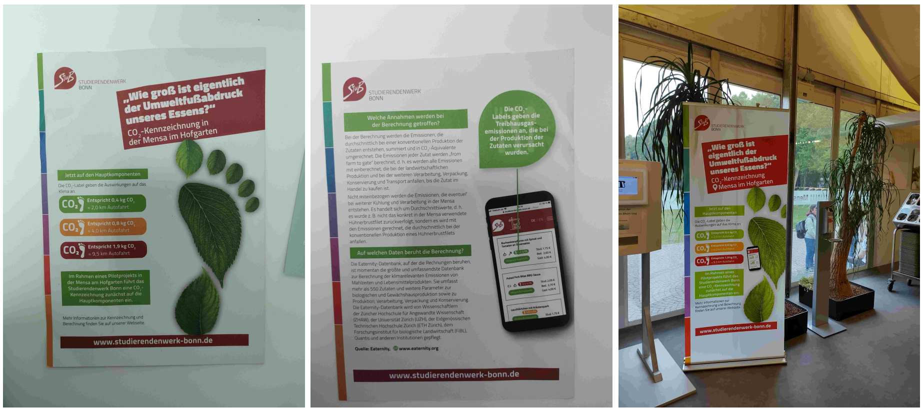


Figure 1.E.5. Explanation of the carbon labeling initiative in the canteen

Note: Leaflets (left and center) and billboards at the entrance of the student canteen (right).

1.E.4 Data set

The main data set covers purchase data from April 1st, 2022 to July 8th, 2022. Spec. (1) in Table 1.E.1 performs the basic analysis shown in the main text in Table 1.2 in Col.(1) on all data before any exclusions.

- Starting from week 9 of the treatment period (May 30th to June 3rd), Ukrainian refugees received free meals in the treated student canteen and the larger control restaurant, using specific student canteen cards. I thus identify these sales and exclude them from all analyses. For the treated restaurant, they make up 12% of total sales in week 9,25% in week 10, and between 14% and 18% for the rest of the observation period. For the control restaurant, they make up between 2% and 7% of total sales. Spec. (2) in Table 1.E.1 shows how this exclusion affects results.
- During the first week of the label period (May 2nd to May 6th), the display was irregular, as the student canteen needed some "trial and error" to get the system running. On

60. Carbon emission labels for a given meal are calculated as the sum of the emissions caused by each of the ingredients. For each ingredient, emission values are calculated "from farm to gate". Hereby, it is assumed that the production process mirrors the average conventional production, e.g. I do not track the specific chicken breast bought by the student canteen but assume average conventional production. Emissions caused by the student canteen cooling, freezing, and cooking ingredients on-site are not included. These calculation details are explained to students on the student canteen website and on leaflets lying out on-site in the student canteen.

some days, the labels were only displayed in the student canteen or online. Further, the student canteen had a special "Healthy Campus" week during the first week of May, during which it offered additional extraordinary meals which were also irregularly labeled. It is thus not clear whether the decrease in meat consumption observed during this week (see Figure 1.11) can be attributed to the carbon labels. To be conservative, I exclude this week from the main analysis. Spec. (3) in Table 1.E.1 additionally excludes week 5 from the sample.

- There are seven days on which the treatment restaurant and the larger control restaurant did not offer the same main meal components: 7th of April, 19th of April, 20th of April, 17 of May, 15th of May, 24th of June, and 27th of June. This is because, although menu planning is centralized, one of the student canteens may not have delivered an ingredient on time or may realize another ingredient is about to expire and independently adjust its meal offer. Any differences in the choice of the main meal component between treatment and control restaurants on these days are likely mainly influenced by differences in offer rather than by differences in label treatment. I thus exclude these days. Spec. (4) in Table 1.E.1 additionally excludes these seven days from the sample (the final sample used in the main text).

For each purchase, I have data on the mode of purchase (student canteen card or debit card), meal category (combined with daily menus, this provides the specific meal name), student canteen card ID (if the purchase is made with the student canteen card), cash register number, date of purchase, time of purchase (exact to the minute), and purchase value.

Table 1.E.1. Field estimates of the effect of carbon labels on meat consumption, testing robustness to different data exclusion criteria

	Likelihood of consuming meat			
	(1) Full data	(2) Excl. Ukr.	(3) +Excl. W5	(4) +Excl. diff. offer
Treatment restaurant x Label period	-0.02**	-0.03***	-0.02***	-0.02***
	(0.01)	(0.01)	(0.01)	(0.01)
Treatment restaurant x Post period	-0.01	-0.07***	-0.07***	-0.07***
	(0.01)	(0.01)	(0.01)	(0.01)
Treatment restaurant	-0.10***	-0.10***	-0.10***	-0.10***
	(0.01)	(0.01)	(0.01)	(0.01)
Label period	0.01***	0.01**	0.01**	0.01
	(0.00)	(0.00)	(0.00)	(0.00)
Post period	0.02***	0.01	0.01	0.01*
	(0.00)	(0.00)	(0.00)	(0.00)
Constant	0.51***	0.51***	0.51***	0.51***
	(0.00)	(0.00)	(0.00)	(0.00)
Date effects	No	No	No	No
Fixed effects	No	No	No	No
Guests control	7,298	7,217	6,798	5,589
Guests treated	3,278	2,939	2,716	2,329
Observations	155,411	150,345	137,962	120,121

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

Note: Spec. (1) includes all data from weeks 1 to week 14. Spec. (2) excludes consumption by Ukrainian refugees. Spec. (3) additionally excludes the first week of the label period (week 5). Spec. (4) additionally excludes seven days on which the offer of the treatment and control canteens strongly differed, resulting in the final sample analyzed in Table 1.2. Specification follows 1.2.

1.E.5 Descriptive statistics on meat consumption and average emissions

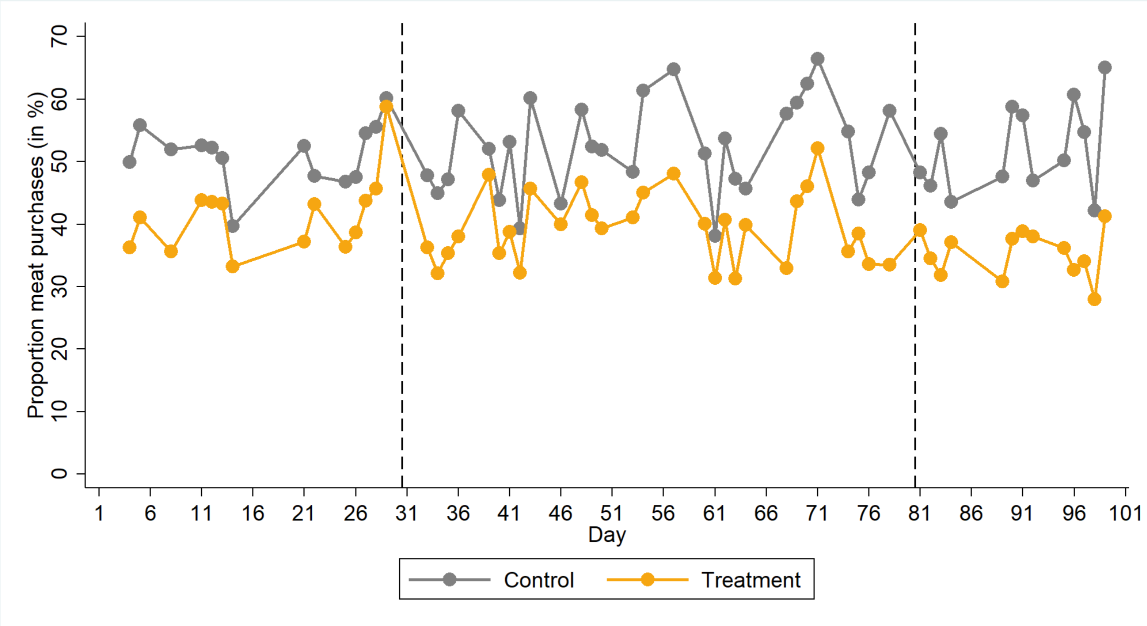


Figure 1.E.6. Proportion of meat meals sold in the canteen

Note: using the final sample but including week 5. $N = 130, 132$

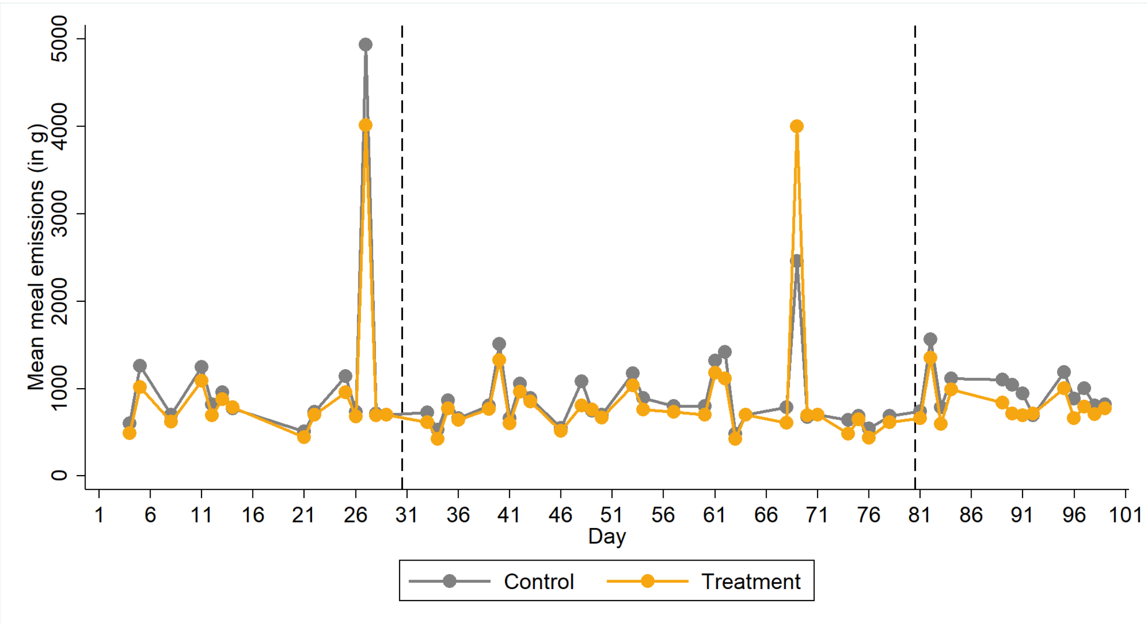


Figure 1.E.7. Average emissions per meal sold in the canteen

Note: Using the final sample but including week 5. $N = 130, 132$

1.E.6 Survey accompanying natural field experiment

Pre-intervention survey: During the second week of April, I conducted a survey among student canteen guests at the treatment student canteen and the first, larger, control restaurant. The survey was advertised as an opportunity to voice one's opinion on the offer of the student canteen, took participants around five minutes, and motivated potential participants with the chance to win one of ten €50 coupons for the student canteen. The survey was advertised through multiple channels. First, I put up posters advertising the survey in many faculties throughout the University of Bonn. Second, I distributed leaflets in front of the treatment restaurant and the larger control restaurant, together with research assistants (see Figure 1.E.8). It is common for students and student groups to advertise surveys, projects, and events in this manner. Finally, the experimental lab at the University of Bonn sent out an e-mail to its entire participant pool advertising participation.

Figure 1.E.8. Leaflet advertising participation in the survey

Note: Leaflet was distributed in front of the student canteen.

In the survey, respondents indicated their student canteen card number and consented to their survey responses being connected to their consumption decisions from April to July. They filled out questions on demographics, environmental attitudes, political preferences, and preferences towards the student canteen offer. Responses to the questions on student canteen offer and participant comments were analyzed, summarized, and presented to the gastronomic manager of the student canteens. Over 1,700 restaurant guests participated in this first survey, 94% of these students.

Post-intervention survey: From the 22nd of June, I started sending out invitations to participate in a second survey. These were sent out by e-mail to those participants of the first survey who indicated their e-mail addresses and consented to be contacted for a second survey. This was the case for 94% of participants in survey 1. Of the 1,558 I invited to the survey, 918 filled out survey 2. I invited participants in a staggered fashion over two weeks and sent a reminder on the 7th of July. Again, survey respondents had the opportunity to win one of ten 50 €coupons for the student canteen.

In survey 2, I repeated some of the questions from survey 1, to assess whether attitudes changed differentially in the treatment student canteen. As pre-registered, the main attitudes of interest were (1) agreement with the statement "Flying should be more expensive, since it is bad for the environment", as a proxy for support for carbon taxes, and (2) agreement to the statement "It should be prohibited to build new houses not adhering to current environmental standards" as a proxy for support for command-and-control policy instruments to cut carbon. The final (3) outcome of interest is the participants' subjective experience of eating in the student canteen, assessed by agreement to the statement "Eating in the student canteen is a nice experience for me". The survey further included some questions of interest to the student canteen following the outcome of the first survey. At the end of the survey, participants could indicate whether and how they had perceived the emission labels, as well as voice their opinion on the initiative.

References

Allcott, Hunt, Daniel Cohen, William Morrison, and Dmitry Taubinsky. 2022. “When do "nudges" increase welfare?” Working Paper, Working Paper Series 30740. National Bureau of Economic Research. https://doi.org/10.3386/w30740. [9]

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. [9, 35]

Allcott, Hunt, and Dmitry Taubinsky. 2015. “Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market.” *American Economic Review* 105 (8): 2501–38. [9, 43]

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. [9]

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. [10]

Barahona, Nano, Cristóbal Otero, and Sebastián Otero. 2023. “Equilibrium effects of food labeling policies.” *Econometrica* 91 (3): 839–68. [9]

Bilén, David. 2022. “Do carbon labels cause consumers to reduce their emissions? Evidence from a large-scale natural experiment.” Working paper. Mimeo, Gothenburg University. [6, 10]

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2022. “Salience.” *Annual Review of Economics* 14 (1): 521–44. [7]

Bouchaud, Jean-Philippe, Philipp Krueger, Augustin Landier, and David Thesmar. 2019. “Sticky expectations and the profitability anomaly.” *Journal of Finance* 74 (2): 639–74. [27]

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. [10]

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. [9, 35]

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. [7, 24, 27]

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. [6, 8–10]

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. [16]

Chetty, Raj. 2009. “Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods.” *Annual Review of Economics* 1 (1): 451–88. [7, 9, 27]

Clark, Michael A., Nina G. Domingo, Kimberly Colgan, Sumil K. Thakrar, David Tilman, John Lynch, Inês L. Azevedo, and Jason D. Hill. 2020. “Global food system emissions could preclude achieving the 1.5°C and 2°C climate change targets.” *Science* 370 (6517): 705–8. [5]

Conlon, John J. 2024. “Attention, Information, and Persuasion.” Working Paper. Stanford. https://johnjconlon17.github.io/website/conlon_attention_persuasion.pdf. [9]

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. [5]

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. [5, 43]

DellaVigna, Stefano. 2009. “Psychology and economics: Evidence from the field.” *Journal of Economic Literature* 47 (2): 315–72. [7, 27]

DellaVigna, Stefano, John A List, and Ulrike Malmendier. 2012. “Testing for altruism and social pressure in charitable giving.” *Quarterly Journal of Economics* 127 (1): 1–56. [9]

DellaVigna, Stefano, John A List, Ulrike Malmendier, and Gautam Rao. 2016. “Voting to tell others.” *Review of Economic Studies* 84 (1): 143–81. [9]

Dohmen, Thomas, and Tomáš Jagelka. 2023. “Accounting for individual-specific reliability of self-assessed measures of economic preferences and personality traits.” Working Paper, Working Paper Series 16027. IZA Discussion Paper. https://docs.iza.org/dp16027.pdf. [18]

Eaternity Institute. 2020. *Eaternity data base.* https://eaternity.org/. Accessed: 2020-08-26. [15, 20, 112]

Epstein, Larry G., Jawwad Noor, and Alvaro Sandroni. 2008. “Non-Bayesian updating: a theoretical framework.” *Theoretical Economics* 3 (2): 193–229. [27]

European Commission. 2023. *Farm to Fork strategy.* https://food.ec.europa.eu/horizontal-topics/farm-fork-strategy_en. Accessed: 2023-01-30. [6]

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. [6]

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. [6]

Goldin, Jacob, and Daniel Reck. 2022. “Optimal defaults with normative ambiguity.” *Review of Economics and Statistics* 104 (1): 17–33. [9]

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. [5]

Harrison, Glenn W, and John A List. 2004. “Field experiments.” *Journal of Economic literature* 42 (4): 1009–55. [6]

Haws, Kelly L., Scott W. Davis, and Utpal M. Dholakia. 2016. “Control over what? Individual differences in general versus eating and spending self-control.” *Journal of Public Policy & Marketing* 35 (1): 37–57. [38, 42, 70]

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. [6, 7, 9, 10]

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. [6, 8–10, 27]

IPCC. 2023. *Climate change 2023: Synthesis report. A report of the Intergovernmental Panel on Climate Change. Contribution of Working Groups I, II, and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]* in press. [5]

Jalil, Andrew J., Joshua Tasoff, and Arturo V. Bustamante. 2020. “Eating to save the planet: Evidence from a randomized controlled trial using individual-level food purchase data.” *Food Policy* 95: 101950. https://doi.org/https://doi.org/10.1016/j.foodpol.2020.101950. [10]

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. [5]

Krupka, Erin L., and Roberto A. Weber. 2013. “Identifying social norms using coordination games: Why does dictator game sharing vary?” *Journal of the European Economic Association* 11 (3): 495–524. [38, 42, 70]

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. [9]

Lohmann, Paul, 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. [6, 9, 10]

Nafziger, Julia. 2020. “Spillover effects of nudges.” *Economics Letters* 190: 109086. [43]

Obama, Barack. 2015. *Executive order – Using Behavioral Science insights to better serve the American people.* Media Release, White House, Washington, DC, September 15, 2015. [6]

OECD. 2019. *Enhancing climate change mitigation through agriculture.* OECD. [5]

Osman, Magda, and Katie Thornton. 2019. "Traffic light labelling of meals to promote sustainable consumption and healthy eating." *Appetite* 138: 60–71. [10]

Panzone, Luca A., Alistair Ulph, Denis Hilton, Ilse Gortemaker, and Ibrahim Adebisi Tajudeen. 2021. "Sustainable by design: choice architecture and the carbon footprint of grocery shopping." *Journal of Public Policy & Marketing* 40 (4): 463–86. [10]

Poore, Joseph, and Thomas Nemecek. 2018. "Reducing food's environmental impacts through producers and consumers." *Science* 360 (6392): 987–92. [5]

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. [7, 12, 15]

Reisch, Lucia A., Cass R. Sunstein, Mark A. Andor, Friederike C. Doebbe, Johanna Meier, and Neal R. Haddaway. 2021. "Mitigating climate change via food consumption and food waste: A systematic map of behavioral interventions." *Journal of Cleaner Production* 279: 123717. [6]

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. [7]

Rodemeier, Matthias. 2021. "Buy baits and consumer sophistication: Theory and field evidence from large-scale rebate promotions." Working paer, Working Paper Series 124. CAWM Discussion Paper. https://www.econstor.eu/handle/10419/234136. [9]

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. [9]

Rondoni, Agnese, and Simona Grasso. 2021. "Consumers behaviour towards carbon footprint labels on food: A review of the literature and discussion of industry implications." *Journal of Cleaner Production* 301: 127031. [10]

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. [5]

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. [18]

Schulze Tilling, Anna. 2021b. *Quantifying the role of greenhouse gas emissions in consumption choice.* Https://doi.org/10.1257/rct.7858. [15, 31]

Shewmake, Sharon, Abigail Okrent, Lanka Thabrew, and Michael Vandenbergh. 2015. "Predicting consumer demand responses to carbon labels." *Ecological Economics* 119: 168–80. [8, 9]

Spaargaren, Gert, CSA Van Koppen, Anke M. Janssen, Astrid Hendriksen, and Corine J. Kolfschoten. 2013. "Consumer responses to the carbon labelling of food: A real life experiment in a canteen practice." *Sociologia Ruralis* 53 (4): 432–53. [10]

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. [9, 10, 43]

Taufique, Khan M.R., Kristian S. Nielsen, Thomas Dietz, Rachael Shwom, Paul C. Stern, and Michael P. Vandenbergh. 2022. "Revisiting the promise of carbon labelling." *Nature Climate Change* 12 (2): 132–40. [7, 12, 15]

Thunström, Linda. 2019. "Welfare effects of nudges: The emotional tax of calorie menu labeling." *Judgment and Decision Making* 14 (1): 11. [9, 38]

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. [9, 43]

Visschers, Vivianne H.M., and Michael Siegrist. 2015. "Does better for the environment mean less tasty? Offering more climate-friendly meals is good for the environment and customer satisfaction." *Appetite* 95: 475–83. [10]

Vlaeminck, Pieter, Ting Jiang, and Liesbet Vranken. 2014. "Food labeling and eco-friendly consumption: Experimental evidence from a Belgian supermarket." *Ecological Economics* 108: 180–90. [10]

Wirsenius, Stefan, Fredrik Hedenus, and Kristina Mohlin. 2011. "Greenhouse gas taxes on animal food products: rationale, tax scheme and climate mitigation effects." *Climatic change* 108 (1-2): 159–84. [80]

Wolfram, Jessica. 2021. "Companies bet carbon labels can help the planet. Will consumers catch on?" Accessed January 30, 2023. https://www.washingtonpost.com/climate-solutions/2021/06/17/carbon-footprint-emissions-label. [6]

Chapter 2

Real-Time Feedback and Social Comparison Impact Resource Use and Welfare: Evidence from a Field Experiment *

Joint with Mark Andor, Lorenz Götte, Michael Price, and Lukas Tomberg

2.1 Introduction

In the last decade, there has been increasing interest in behavioral economic interventions from research, society, and politics. A whole series of empirical studies have investigated the effectiveness of behavioral interventions in a wide range of areas,[1] behavioral insights groups have been installed around the world (Obama, 2015; OECD, 2017), several books on behavioral interventions became bestsellers (Ariely and Jones, 2008; Thaler and Sunstein, 2009; Kahneman, 2011; Halpern, 2015), and Daniel Kahneman, Robert Shiller and Richard Thaler each won a Nobel prize.

However, the welfare effects of such interventions have rarely been studied. From a cost-effectiveness standpoint, many behavioral interventions triumph (Allcott and Mullainathan, 2010; Benartzi et al., 2017). By design, they usually do not impose monetary costs on the

* We are grateful for comments and suggestions by Max Auffhammer, Fabian Dehos, Eugen Dimant, Valeria Fanghella, Stefano DellaVigna, Christian Hönow, Joachim Schleich, Christoph M. Schmidt, and Colin Vance, as well as by participants of the Advances with Field Experiments Conference 2022 (AFE), the Annual Conference of the Association of Environmental and Resource Economists 2022 (AERE), the Annual Conference of the European Economic Association 2023 (EEA), the Annual Conference of the German Committee for Environmental and Resource Economics 2023 (AURÖ), the Mannheim Energy Conference 2023, the 2022 World Economic Science Association Conference (ESA), the CAIS Research Incubator, the 9th Retreat of the CRC TR 224, the Behavioral group at IDOS, and seminars at University of Duisburg-Essen, Grenoble Ecole de Management and the Mercator Research Institute on Global Commons and Climate Change (MCC). Furthermore, we thank Sven Hansteen, Florin Martius and Valerie Peetz for excellent research assistance. In addition, we are indebted to Michael Schild for his tremendous support in the technical implementation of the study. This project was conducted in research partnership with the Center for Advanced Internet Studies (CAIS). We gratefully acknowledge financial support by the Ministry of Culture and Science of the State of North Rhine-Westphalia and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project B07) and Germany's Excellence Strategy – EXC 2126/1-390838866. The experimental design was pre-registered with the AEA RCT Registry as trial AEARCTR-0006747 and IRB approval was obtained from the GfeW (Certificate No. nQ9jc48B).

1. For recent meta-analyses see, for example, Jachimowicz et al. (2019), Nisa et al. (2019), Cadario and Chandon (2020), and Mertens et al. (2022).

consumer, and implementation costs are low. These characteristics make behavioral interventions particularly attractive to policymakers and even interventions with small effects are often assessed as “cost-effective”. However, such cost-effectiveness analyses usually ignore non-monetary costs and benefits, including shame, pride, and social pressure (DellaVigna, List, and Malmendier, 2012; DellaVigna et al., 2016; Butera et al., 2022), (dis-)utility resulting from changes in consumption behavior (Allcott and Kessler, 2019), or the non-instrumental (dis-)utility of information provision (Sharot and Sunstein, 2020).

This paper contributes to this discussion by offering a comprehensive welfare analysis of two popular behavioral interventions affecting consumers along two different dimensions. While one intervention primarily affects consumers by providing information that *motivates* behavioral change, the other primarily provides information that *facilitates* a change in behavior. While many interventions can be classified along similar lines (see Congiu and Moscati, 2020), it is still unclear whether targeting one or the other dimension is more effective in creating behavioral change and a positive welfare impact. To provide a clean evaluation for informational interventions,[2] we compare both interventions in a common context: fostering resource conservation in showering behavior. Showering consumes a considerable amount of resources, both due to the energy necessary to heat the water used and the water use itself.[3] Furthermore, showering behavior is prone to behavioral biases, such as salience bias (Tiefenbeck et al., 2018), and thus to inefficiently high levels of resource consumption. At the same time, it is a daily activity that consumers can directly control, making it an ideal application for behavioral interventions.

We focus on social comparison reports (SC) as a representative and well-studied[4] example of an intervention providing consumers with information that *motivates* a change in behavior. Our SC reports provide treated participants with weekly information on their average water use and CO_2 emissions per shower and compare it to that of other participants. This communicates the social desirability of conserving resources, motivating participants to conserve water and energy.

We focus on real-time feedback (RTF) as a representative and well-studied[5] example of an intervention providing consumers with information that *facilitates* a change in behavior. Our RTF intervention activates participants’ shower heads to glow in different colors when showering, depending on how much water has already been consumed during the current shower. This solves two problems impeding behavioral change in the absence of RTF. First, the quantity of energy and water consumed while showering is usually only imperfectly observable to consumers, leading to consumers either not being able to act upon this aspect at all or making consumers reliant on imperfect heuristics. Consumption can be thought of as a shrouded attribute in this

2. Our focus in this project is on behavioral interventions providing information, which constitute a subset of the behavioral interventions described in, e.g., Congiu and Moscati (2020). Our findings might not apply to other types of interventions, e.g., interventions changing product attributes or product placement.

3. A typical individual in our sample uses 33 liters of hot water per shower, which requires on average 1.6 kWh to heat it up. In comparison, the average household in the European Union uses 1.0 kWh for lighting per day (Faberi et al., 2015), and a modern refrigerator uses 0.63 kWh per day (Michel, Attali, and Bush, 2016).

4. See, for example, Allcott (2011), Ferraro and Price (2013), Allcott and Rogers (2014), Brent, Cook, and Olsen (2015), and Goette et al. (2021).

5. See, for example, Tiefenbeck et al. (2018), Tiefenbeck et al. (2019), Goette, Han, Lim, et al. (2021), Fang et al. (2023), and Byrne et al. (2024).

setting.[6] Second, consumers' attention might not be focused on energy and water conservation at the moment of showering – even if a consumer had developed a perfect heuristic, she would still require an unusually high level of self-control to be able to direct her attention toward this heuristic during every shower. The RTF intervention addresses both of these problems by making resource consumption salient and actionable, facilitating behavioral change.

Our experiment design allows us to draw conclusions on both the behavioral and welfare effects of both interventions SC and RTF, as well as their combination (BOTH). We equip all of our 1,003 field experiment participants in 574 households with smart shower heads to track showering behavior, capturing water use, temperature, and time of every shower. After observing baseline showering behavior for four weeks, participants in the treatment groups experience an information intervention for four weeks based on the experimental group to which they are exogenously assigned. For participants receiving RTF, we remotely activate the real-time feedback function in their shower heads. For participants receiving SC, we configure weekly e-mail reports. The third treatment, BOTH, combines both interventions: Participants receive real-time feedback through their shower head and social comparison reports. Further, we elicit participants' willingness to pay to experience the interventions for four weeks. We elicit willingness to pay twice. We first ask participants for their willingness to pay before having experienced any of the interventions, and then again after they have experienced one of the interventions due to their exogenous assignment. The willingness to pay bids are binding and thus constitute revealed preferences, as the field experiment includes a phase in which the treatments are allocated endogenously based on the elicited willingness to pay.

The willingness to pay elicitation constitutes the basis of our welfare analysis. The willingness to pay subsumes the private benefits from resource savings, the change in shower comfort, the psychological costs and benefits of the interventions, and any further factors relevant to the consumers. Adding the societal benefits from the reduction in carbon emissions due to the reduction in hot water use caused by the interventions and subtracting the program costs, taking into account the marginal cost of public funds, yields estimates of the welfare effects of the interventions.

In terms of behavioral impact, we find that the RTF intervention decreases water and energy use per shower by 28.8%. SC reports decrease water and energy use per shower by 9.4%, and the combination BOTH by 35.0%. RTF is thus around three times as effective as SC reports in our context, and the combination of both interventions is most effective. Our results suggest that providing consumers with information facilitating behavioral change is more effective than providing information motivating consumers to change behavior.

While all interventions are highly valued by participants, RTF and the combination of both interventions are around 10% more highly valued than SC reports alone. We find that the experience of the interventions does not substantially alter the willingness to pay for them. Furthermore, the endogenous allocation of the interventions based on willingness to pay instead of exogenous allocation based on a random draw leads to similar conservation effects. Willingness to pay is, on average, much above and beyond actual and perceived cost savings realized

6. The most prominent examples of shrouded attributes are of a financial nature, such as taxes (Chetty, Looney, and Kroft, 2009), shipping costs (Gabaix and Laibson, 2006; Brown, Hossain, and Morgan, 2010), or marginal prices (Ito, 2014). In the context of resource consumption, which is what we are concerned with, it is consumption itself that is often not observable to consumers (Jessoe and Rapson, 2014).

through the interventions, implying the existence of psychological benefits. For around 25% to 28% of participants, however, willingness to pay is lower than perceived savings, implying that these participants incur psychological costs with the intervention.

In terms of welfare impact, we estimate the highest welfare increase for the combined intervention and the lowest increase for the SC intervention. Recognizing that our sample of volunteers may be particularly motivated to conserve resources and to receive the interventions, we conduct an additional welfare analysis under particularly conservative assumptions with regard to conservation effects and consumer surplus. This conservative analysis yields positive welfare effects for both RTF and the combined intervention, while the welfare effect of SC is close to zero. Our results indicate that providing consumers with information facilitating behavioral change has a slightly more positive effect on consumer welfare than providing information motivating behavioral change.

We contribute to the literature in three ways. First, our study contributes to the relatively young literature on the welfare effects of behavioral interventions. As recent literature has pointed out (Allcott and Kessler, 2019; Butera et al., 2022), policymakers must be cautious of behavioral interventions producing the desired behavioral effect at the cost of imposing disproportionate psychological costs on consumers. In response, a dynamic strand of research has emerged. Essentially, two main approaches to welfare analysis have developed, which differ in their way of quantifying changes in consumer surplus: one that derives consumer surplus from structural models or sufficient statistics (Chetty, Looney, and Kroft, 2009; DellaVigna, List, and Malmendier, 2012; DellaVigna et al., 2016; Rodemeier, 2021; Allcott et al., 2022; Rodemeier and Löschel, 2022; Löschel, Rodemeier, and Werthschulte, 2023), and one that directly measures consumer surplus by eliciting the willingness to pay to receive a behavioral intervention (Allcott and Kessler, 2019; Butera et al., 2022). Importantly, this second approach allows one to capture the psychological costs and benefits of an intervention that are independent of any effect on consumer behavior. For example, an intervention may change a consumer's feelings about his or her choice without affecting the choice itself, or it may affect the consumer's experience while making the choice. Our paper aligns with this second strand of research, providing the first welfare comparison of two interventions affecting consumers along different dimensions. The two interventions and their combination are evaluated within the same experimental context, providing an apples-to-apples comparison so far lacking in the literature.

Another difference to previous literature lies in our double elicitation of consumer willingness to pay, both before and after receiving the intervention. Previous literature on welfare effects differs not only in the consumption context studied but also in the timing of the willingness to pay inquiry. Allcott and Kessler (2019) elicit consumer willingness to pay to continue receiving an intervention, while Butera et al. (2022) elicit willingness to pay for receiving an intervention for the first time. By eliciting both measures within the same consumption context, we provide evidence on whether economically large differences exist between the two measures and provide guidance for future research on how an experiment can be designed to capture both.

Second, our paper relates to conceptual work classifying behavioral interventions. Our focus on the two dimensions of motivating vs. facilitating behavioral change most closely relates to Congiu and Moscati (2020) who classify interventions into a *message* and an *environment* dimension. Our findings can also inform earlier classifications by Johnson et al. (2012), Hansen and Jespersen (2013), and Mongin and Cozic (2018) with empirical evidence. The added value of our paper lies in our ability to provide a clean and comprehensive comparison of two information interventions affecting behavior among different dimensions. Since consumption context, participant sample, and timeline are common, differences in treatment and welfare effects can be cleanly and causally attributed to differences in the nature of the interventions.

Finally, our study contributes to previous literature on the efficacy of behavioral interventions in reducing resource consumption. Social comparison reports have been prominently discussed in previous literature (Allcott, 2011; Ferraro and Price, 2013; Allcott and Rogers, 2014; Brent, Cook, and Olsen, 2015; Goette, Leong, and Qian, 2019; Goette et al., 2021), but are not generally cost-effective across different countries (Andor, Gerster, et al., 2020) and may impose substantial (intangible) costs that reduce the welfare effects of such interventions (Allcott and Kessler, 2019). Real-time feedback has been identified as a promising intervention to influence showering behavior, with previous studies reporting higher effect sizes than for social comparison reports (e.g., Tiefenbeck et al., 2018).[7] However, the existing evidence is derived from very different consumption contexts, e.g., studying the electricity consumption of a specific appliance versus aggregated electricity consumption. Our paper adds to recent research comparing different behavioral interventions within the same consumption context (Brandon et al., 2019; Fang et al., 2023). We advance this research by providing evidence on the welfare effects instead of solely focusing on behavioral effects. Furthermore, we differ from the two studies in that we evaluate two interventions that are inherently different in their nature. In our study, we investigate social comparison reports and real-time feedback, two interventions that have been extensively discussed and investigated over the past decade but whose effectiveness and welfare consequences have never been compared within one experiment. Brandon et al. (2019) investigate the effects of two different types of social nudges on electricity consumption patterns and have no real-time feedback element in their field experiment, while Fang et al. (2023) focus on real-time feedback but combine this with an intervention mainly providing information on environmental impact.

Our findings are relevant for all consumption contexts in which fine-grained data is available with smart sensors and the question is how to best utilize this data to increase the effectiveness of interventions, adding to a broader strand of literature (e.g., Jessoe and Rapson, 2014; Asensio and Delmas, 2015; Grubb and Osborne, 2015; Ito, Ida, and Tanaka, 2018; Brülisauer et al., 2020; Gerster, Andor, and Goette, 2020). They also connect to the literature focusing on shrouded attributes and the importance of making shrouded attributes salient at the moment of purchase (e.g., Gabaix and Laibson, 2006; Chetty, Looney, and Kroft, 2009; DellaVigna and Pollet, 2009; Grubb, 2009; Brown, Hossain, and Morgan, 2010; Taubinsky and Rees-Jones, 2018). The shrouded attribute in our setting is the quantity consumed – making it both difficult for consumers to allocate their attention to this attribute and costly to calculate adequate estimates. Since consumers in this case need to develop heuristics to guide their behavior, and our interventions help develop effective and simple heuristics – for example, to stop showering when real-time feedback gives a specific cue – we also connect to the literature on heuristic thinking (e.g., Lacetera, Pope, and Sydnor, 2012; List et al., 2023).

The paper proceeds as follows: In section 2.2, we describe the experimental design and our study sample. In section 2.3, we present the behavioral impact of the interventions and

7. For more detailed reviews of behavioral interventions applied in the context of resource consumption see, for example, Andor and Fels (2018) and Khanna et al. (2021).

explore the mechanisms underlying treatment differences, specifically focusing on the role of social comparison reports in motivating behavioral change and the role of real-time feedback in facilitating behavioral change. We then present participants' willingness to pay for receiving the interventions, and our analysis of the welfare effects of each of the interventions. Section 2.4 concludes.

2.2 The field experiment

2.2.1 Technical equipment

We provided all study participants with a smart shower head, a WiFi gateway, and installation instructions. Details on the functioning and installation of the shower head are included in Appendix 2.E. The shower head and WiFi gateway, connected to the participants' WiFi network, formed a data infrastructure that allowed the shower head to continuously send data on the water use in liters, date, time, and temperature of every shower to the researchers. The participants did not have access to this data during the study.

2.2.2 Sample

Participants were recruited in collaboration with the German survey institute *forsa*, which maintains a sample of panelists that is representative of German-speaking internet users aged 14 and above. In late August 2020, *forsa* sent out an invitation e-mail to 9,376 panelists, living in the Rhine-Ruhr-Area, Germany's largest metropolitan area.[8]

The invitation email conveyed some general information about the environmental impact of showering, information on the compensation for participation,[9] and information about the tasks as a study participant, which included installing the shower head and data infrastructure and completing several surveys throughout the study. The study was presented to the participants under the name "Project Sustainable Showering". Those who were interested in participating were referred in the email to a project website where more detailed information was provided and where they had the opportunity to register for study participation. The invitation email and the project homepage neither explained that a randomized experiment was to be conducted, nor which interventions were planned. Registration took place via an online form on the project homepage in which we queried contact details, consent to data processing within the study, and initial information on household size and technical circumstances to ensure that participants met the requirements to take part in the study. The content of the invitation email is depicted in Appendix 2.B, and the content of the project website is presented in Appendix 2.C.

8. This area was chosen because two of the research institutions involved in the research project, RWI – Leibniz Institute for Economic Research in Essen and the University of Bonn, are located in this region, thus enabling potential technical support at the premises of the participating households.

9. Participants were compensated for study participation in two ways. First, they were allowed to keep the smart shower head after the study period and could then control it independently using an app. Secondly, the participants received a monetary amount of 35 EUR, which was paid out to them at the end of the study in the form of a voucher of their choice.

1,100 of the invited persons registered for the study, representing 11.7% of those originally invited. We selected the final study participants among the eligible registrations[10] targeting smaller households with fewer showers in the home because we consider it more likely for these households that the information provided to the participants in the course of the experiment would reach all shower users. Since we targeted a total sample size of around 600 households and expected some attrition during the study, we confirmed the participation of a total of 685 participants and sent them an invitation for the baseline survey in October 2020. 647 participants filled out the baseline survey within the given time frame. Those who did not fill out the survey after a reminder email were not considered for study participation anymore. From the outset, we divided the participants into two study waves. The first study wave began immediately, while the second wave began in January 2021. The division into the two study waves was made to logistically simplify the roll-out of the smart shower heads. This division allowed for regional clustering of the respective waves, which in turn minimized travel time for on-site technical support. Apart from the difference in timing, the research design was identical for both study waves. The final study sample consists of 574 households.[11]

Table 2.1. Socio-economic summary statistics

Variable	Explanation	Mean	Std. Dev.	N
Age	Age of respondent	53.85	13.31	567
Female	Dummy: 1 if respondent is a woman	0.39	–	572
Household size	Number of persons living in the household	1.75	0.59	574
Children	Dummy: 1 if children live in the household	0.07	–	574
College degree	Dummy: 1 if respondent has a university degree	0.49	–	565
Number of showers	Number of showers installed in the household	1.25	0.52	574
Income	Monthly net household income in EUR	3489.93	1445.03	541

Notes: This table shows summary statistics for the final sample of study participants, i.e., participants who filled out the baseline survey and successfully installed the shower head infrastructure. The number of observations per variable varies, because participants had the option of not answering some of the questions. Since our main specification relies on fixed-effects models that do not require information on individual time-constant covariates, the lack of information on socioeconomic characteristics of some participants does not limit the size of our estimation sample. Household income was determined using an interval scale and converted into a continuous income measure for the empirical analysis, in which a participant's income was set to the value of the midpoint of the selected income interval.

Table 2.1 provides summary statistics on the study sample. As a result of our selection process, the average household size is smaller than that of the German population (1.75 compared to 2.0; German Federal Institute for Population Research, 2023b) as is the share of households with children (7% compared to 29%; German Federal Statistical Office, 2022). Correspondingly, the monthly net household income is slightly lower than in the population (3,490 EUR compared to 3,612 EUR; German Federal Statistical Office, 2021). The study sample furthermore consists of slightly older participants (54 years compared to 51 years; own calculation of the average age of the German adult population based on German Federal Institute for Population Research, 2023a), a higher share of participants with a college degree (49% compared to 19%;

10. We had to delete 25 registrations because they were either obvious duplicates or the participants provided an address outside the Rhine-Ruhr-Area. Further, 203 registered persons did not meet the technical requirements needed for successful installation of the shower head and the data infrastructure.

11. The remaining 73 households either dropped out after the baseline survey or were unable to install the smart shower heads or the data infrastructure.

German Federal Statistical Office, 2022) and a lower share of female participants (39% females compared to 51%; German Federal Statistical Office, 2022).

2.2.3 Experimental design

Our experiment consists of three phases illustrated in Figure 2.1: A baseline phase, an exogenous treatment phase, and an endogenous treatment phase.[12] In the baseline phase, i.e., the first four weeks of the experiment, we tracked showering behavior without introducing any behavioral intervention. The baseline phase was followed by the exogenous treatment phase, i.e., four weeks in which the behavioral interventions were exogenously assigned to the participants. In the following four weeks, the endogenous treatment phase, we reassigned the experimental groups based on participants' willingness to pay (WTP) to receive the interventions. The primary purpose of this endogenous treatment phase was to make the WTP inquiry consequential and thus incentive-compatible (Carson and Groves, 2007).[13] Furthermore, it allows us to observe the conservation effects in a setting in which participants receive the interventions based on their WTP. The experimental design was pre-registered with the AEA RCT Registry as trial AEARCTR-0006747 (Andor, Goette, et al., 2020) and IRB approval was obtained from the GfeW (Certificate No. nQ9jc48B).

2.2.3.1 Experimental groups

For the exogenous treatment phase, we split the sample into five equally-sized experimental groups. We allocated the experimental groups randomly, with stratification ensuring that they did not significantly differ from each other concerning baseline water use per shower, age, household size, number of showers installed in the household, environmental attitudes, perceived social norms regarding energy use, and education. The five experimental groups are characterized as follows:

Super-control group: This group received no behavioral intervention throughout the study and served two purposes: First, it provided data on showering behavior over time in the absence of any intervention and second, this group was used to form reference groups needed for the behavioral interventions: Each participant who was not in the super-control group was randomly assigned to a reference group consisting of nine members of the super-control group. Participants who received the social comparison reports were compared to their reference group in the comparison reports. Similarly, the thresholds in the real-time feedback intervention were based on the behavior of the reference group. Because the super-control group was not eligible for an intervention in either phase of the study, the WTP inquiries filled out by these households were purely hypothetical.[14] Yet, to ensure that these participants stayed at a similar attention level regarding their study participation as the other participants, they received weekly emails

12. Baseline phase of wave 1 (wave 2): From October 15, 2020 to November 20, 2020 (from January 1, 2021 to March 11, 2021). Exogenous treatment phase of wave 1 (wave 2): From November 21, 2020 to December 21, 2020 (from March 12, 2021 to April 14, 2021). Endogenous treatment phase of wave 1 (wave 2): From December 22, 2020 to January 21, 2021 (from April 15, 2022 to May 14, 2021).

13. Details on the treatment assignment mechanism in the endogenous treatment phase are provided in Appendix 2.A.1.3.

14. This concerns only participants in this group - all other participants are incentivized to indicate their true WTP values since these co-determine the interventions these households receive during the endogenous treatment phase.

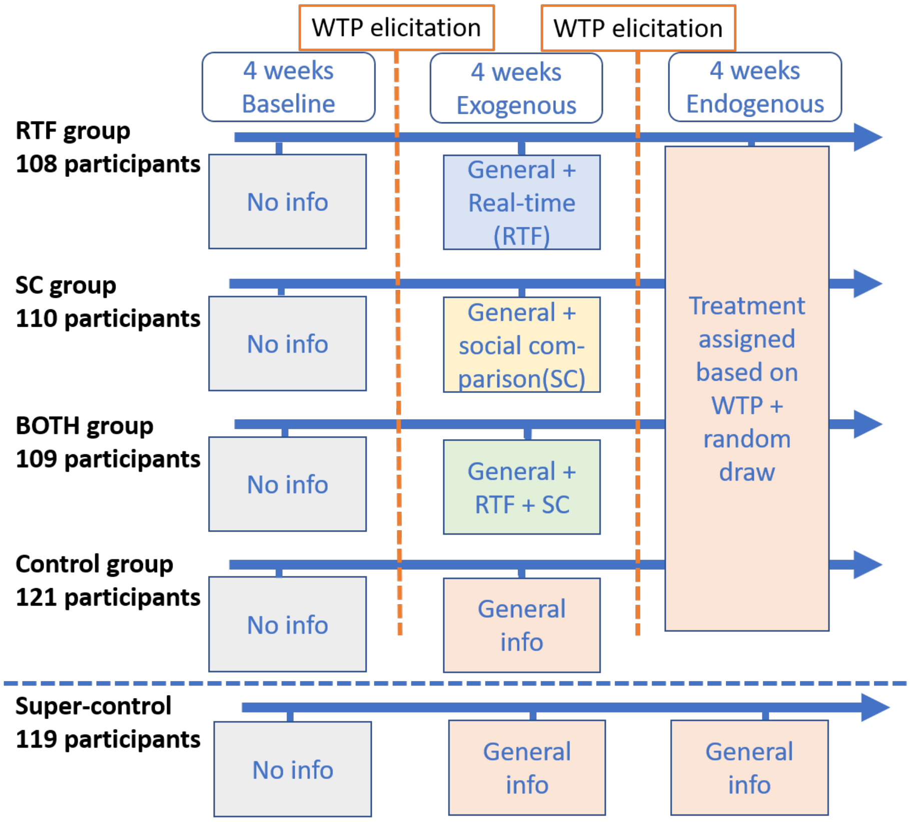


Figure 2.1. Overview of the experimental design

Note: This figure shows the experimental design. Participant treatment during the exogenous phase depended on treatment group assignment. Participant treatment during the endogenous phase depended on WTP indicated during the two elicitations (before and after the exogenous phase) and a random draw.

("newsletters") with general information about showering and its energy intensity throughout the study; as did all other study participants. The upper panel of the exemplary newsletter in Appendix 2.D contains one example of the general information texts.

Control group: Like the super control-group, this group received weekly newsletters but no intervention in the exogenous treatment phase. However, just as the SC, RTF and BOTH group, this group was eligible for receiving an intervention in the endogenous treatment phase, making this group's WTP inquiry binding and thus incentive compatible.

Social comparison (SC) group: During the exogenous treatment phase, participants assigned to this group received a social comparison section in their weekly newsletters in addition to the general information section. In this social comparison section, the average water use per shower (and corresponding CO_2 emissions, given the average water temperature) was reported and compared to the average of the participant's reference group constituting of untreated households from the super-control group. Furthermore, weekly averages of the water use per shower of both the participant household and its reference group were plotted in a time diagram. In addition, participants were shown how their water consumption per shower ranks compared to the nine reference households, i.e., they were assigned a rank between one (lowest use) and ten (highest use). For each rank, an additional emoticon was displayed: A smiling face for comparatively low consumption, a neutral face for medium consumption, and a frowning face for high consumption. To ensure that enough data was available to calculate the social comparison, the information was updated every two weeks, so that the second newsletter repeated the information of the first newsletter, the fourth newsletter repeated the information of the third newsletter, and so on. The middle panel of the exemplary newsletter in Appendix 2.D is an example of the social comparison information.

Real-time feedback (RTF) group: For participants assigned to this intervention, the real-time feedback function was activated in participants' shower heads. The LEDs built into the shower head signaled the water consumption of the shower in real time. Shower heads were configured to light up green for the first 15 liters of a shower. Thereafter they turned blue, purple, red and eventually started to flash in the red color. The thresholds at which the colors changed were determined by the behavior of the participant's untreated reference group. In detail, the shower head would start to flash red once the reference group's average water use per shower over the past two weeks was exceeded. The thresholds of the remaining colors were allocated in regular intervals between 15 liters and the average water use per shower of the reference group. For these participants, a RTF section was added to the weekly newsletter. This section contained information on the water consumption corresponding to each of the color thresholds, but participants were not informed about the origin of the color thresholds. For each threshold in liters, the corresponding average CO_2 emissions caused by heating the water to the usual water temperature of the household were also detailed. To ensure that enough data was available to reliably construct the thresholds, the thresholds were updated every second week, as they were for the SC group so that the newsletter in between was a repetition of the newsletter of the week before. The lower panel of the exemplary newsletter in Appendix 2.D is an example of how the real-time feedback functionality was communicated.

BOTH group: For participants assigned to this intervention, the SC section was added to the weekly newsletters, and the RTF function in the smart shower heads was activated. Participants thus received both the interventions of the SC group and the RTF group. The weekly newsletters of this group thus contained a SC section as well as a RTF section, and the full newsletter in Appendix 2.D is a representative newsletter of a participant that received the BOTH intervention. Having the information from both interventions, the participants in this group were able to infer that their shower thresholds depended on their reference group's average water use per shower.

2.2.3.2 Elicitation of willingness to pay for the interventions

Towards the end of the baseline phase and the end of the exogenous treatment phase, participants filled out a willingness to pay (WTP) inquiry, in which they indicated how much they are willing to pay to receive each of the interventions (SC, RTF, and BOTH) in the endogenous treatment phase.

The WTP inquiry was conducted in the form of a survey which consisted of several multiple price lists (MPL), with one MPL for each of the three interventions. In each MPL, the participants were asked to decide if they preferred to receive four weeks of the intervention in question plus a varying monetary amount or whether they preferred to forego the intervention and receive an amount of 15 EUR. An example of an MPL eliciting WTP for the social comparison reports is provided in Figure 2.2. The participants had to fill out an MPL for each treatment twice, once before and once towards the end of the exogenous treatment phase, totaling six MPLs. Except for the super-control group, these inquiries were binding in that one of the decisions made, i.e., the decision of a randomly drawn row (of 15) in one randomly selected MPL (out of six), was actually implemented in the endogenous treatment phase. So the participant would, depending

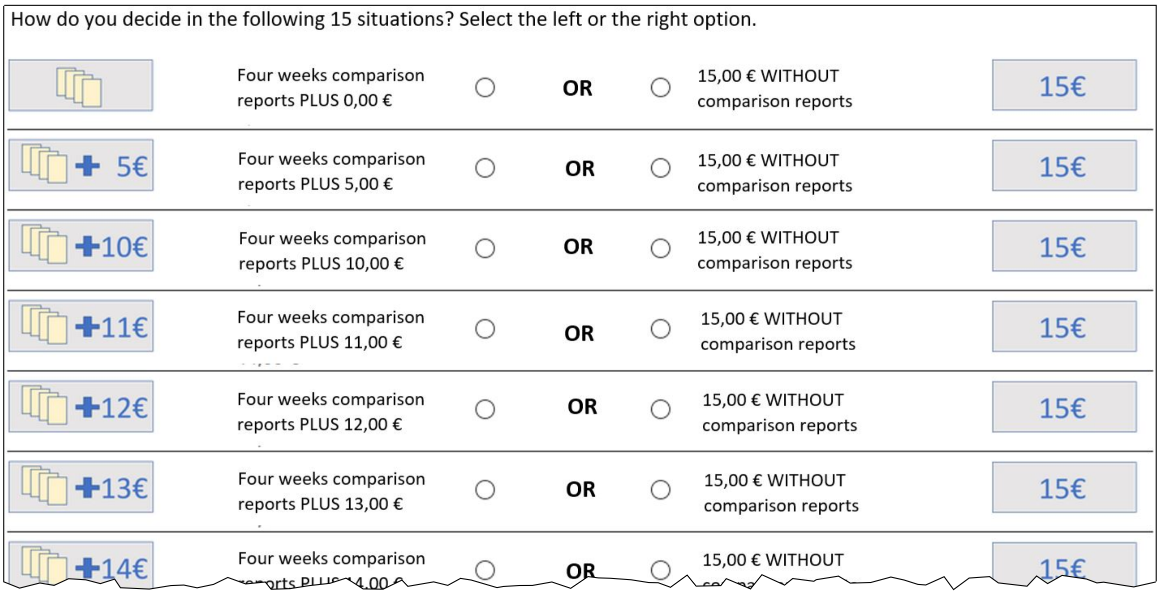


Figure 2.2. Illustration of the multiple price list for SC

Note: This figure is abridged - left side options continue with 15 €, 16 €, 17 €, 18 €, 19 €, 20 €, 25 €, 30 €. Multiple price lists for RTF and BOTH were identical apart from the left-hand-side illustrations.

on her decision, receive the treatment and the corresponding monetary amount or the fixed amount of 15 EUR.

Our WTP inquires thus allow us to elicit, in an incentivized manner, how much consumers are willing to pay to receive any of the interventions for a four-week period. For example, if in a given choice a participant would prefer to receive RTF and 12 Euro to receiving 15 Euro, we interpret this as the consumer having at least a WTP of 3 Euro for receiving RTF for 4 weeks. If this choice was selected for implementation, we would activate the RTF feature of this participant's shower head during the exogenous treatment phase, without this requiring any effort by the participant. Since participants' decisions are consequential, it is costly for participants to under- or overstate their WTP, and we thus avoid hypothetical response bias. Also, note that all participants kept the shower head at the end of the study as part of their study compensation (see 2.B for the study invitation and information on participant compensation), and it was made clear to participants that this would not be affected by the WTP elicitation. Before eliciting participants' WTP, we explained the structure in detail and provided examples.

2.3 Results

2.3.1 Data

As a first step, we present the shower data from the baseline phase in Table 2.2.[15] For the baseline phase as well as the exogenous treatment phase, we pool the super-control group and the control group to maximize power and refer to this pooled group as the "(super-)control group". We see a slight difference in baseline water use and baseline shower temperature in the (super-)control group compared to the three treatment groups. This difference is only statistically significant in one case (water temperature of the SC group) and will not affect our regression results, since we include household fixed effects in all regression specifications.

Table 2.2. Shower statistics in the baseline phase by experimental group

Variable	Unit	(Super-) control	SC	RTF	BOTH
Baseline water use	Liters per shower	34.07	32.16	31.89	32.31
			(0.84)	(0.94)	(0.76)
Baseline shower water temperature	°C	38.12	37.17	36.88	37.81
			(2.43)	(1.33)	(0.73)
Baseline water flow	Liters per Minute	9.18	9.21	9.05	8.87
			(0.13)	(0.54)	(1.21)
Baseline Shower frequency	Number of showers per day	0.96	1.12	1.08	0.99
			(1.13)	(1.31)	(0.38)
No. of households		236	110	109	109
No. of observations		8,416	4,207	4,360	4,056

Notes: This table shows baseline averages. The test statistics in parentheses are t-values for comparison to the (super-)control group obtained from OLS regressions with either baseline water use, water temperature, water flow, or shower frequency as the outcome variable and group membership as the explanatory variable with standard errors clustered at the household level and controlling for time fixed effects (not in the case of shower frequency as this is a cross-sectional variable). 10 households in the (super-)control group were only able to properly connect their shower heads after the baseline phase, so they are excluded from these statistics and all analyses in this section.

2.3.2 Impact of the interventions on behavior

Figure 2.3 depicts the daily average water use per shower in the different treatment groups in the baseline phase and in the exogenous treatment phase.[16] During the first four weeks, average water use per shower is not distinguishable between the (super-)control group and the treatment groups. The start of the treatment phase, as indicated by the blue vertical line in Figure 2.3 leads to a sharp and immediate drop in water use per shower for all groups that received treatment, while the behavior of the control group does not change systematically. Moreover, visual inspection suggests that RTF and BOTH lead to stronger reductions in water use per shower than SC.

15. We cleaned the raw shower data, in that we remove showers with water consumption above 200 liters or below 5 liters. This is necessary so that water withdrawals that are very likely not showers do not add additional noise to the results. A similar approach is taken by Fang et al. (2023). It should also be noted that showers taken while the WiFi gateway was not connected to the Internet were, for technical reasons, recorded with the time stamp of the next successful data transmission from the WiFi gateway and not the actual time stamp of the shower. This is the case for 17.2% of showers taken throughout the entire study period. To correct this, we redefine the time stamps of such offline showers and allocate them evenly over the period between the last and the next successful data transmission.

16. This graph pools data from the two study waves, such that the onset of the treatment phase was on November 20, 2020 for the first study wave and on March 11, 2021 for the second study wave.

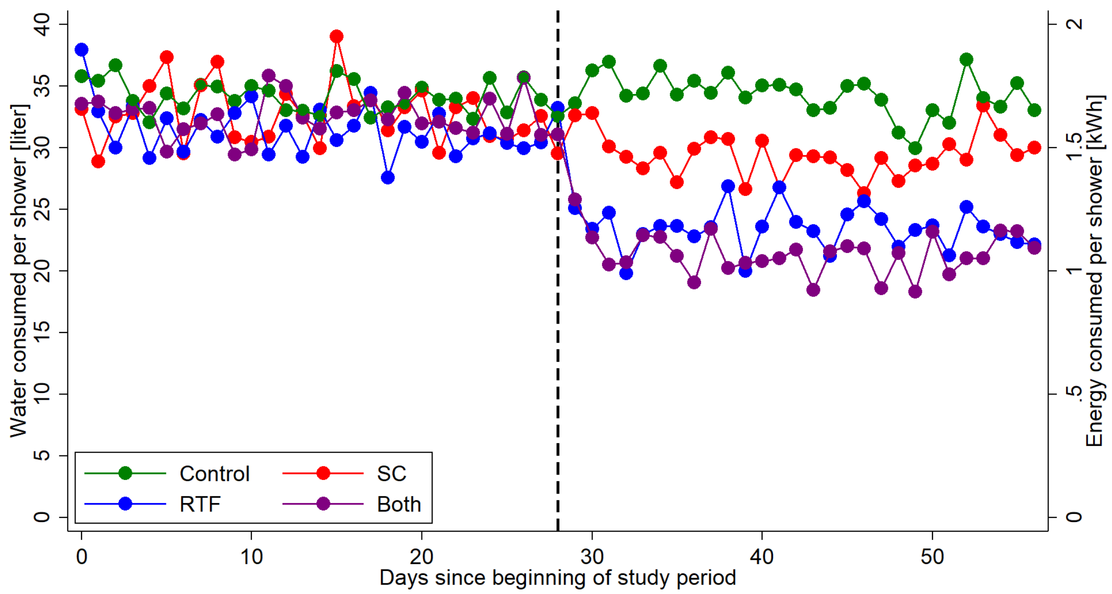


Figure 2.3. Exogenous treatment phase - average water use per shower by day and treatment group

Note: This figure compares average water use per shower between treatment groups during the exogenous treatment phase, shown on the left vertical axis. The right vertical axis translates water consumption into energy consumption required to heat the water, given a water temperature of 38 °C (the average water temperature in our sample).

In a next step, we estimate the treatment effects using a difference in differences regression model of the following form:

$$y_{jit} = \alpha_i + \beta_1 SC_{it} + \beta_2 RTF_{it} + \beta_3 BOTH_{it} + \tau_t + \epsilon_{jit}, \tag{2.1}$$

where y_{jit} represents the outcome variable, i.e., water used during shower j by household i on day t. α_i represents household-specific fixed-effects. SC_{it}, RTF_{it} and $BOTH_{it}$ are indicators of household i receiving the respective treatment at time t. τ_t represents day-specific fixed effects and ϵ_{jit} is the error term that is clustered at the household level.

The estimates of the average treatment effects in the exogenous treatment phase are depicted in Table 2.3. The SC treatment reduced water use per shower by 3.19 liters, the RTF treatment by 9.79 liters, and the BOTH treatment by 11.91 liters. Relative to the average water use of the (super-)control group in the exogenous treatment phase, which is 34.1 liters, these effects correspond to a 9.4%, 28.8%, and 35.0% reduction in water use per shower. All effects are statistically significant at the 1% level. The effect of SC is significantly smaller than the effects of RTF and BOTH (p-values of the differences: <0.00) and the effect of BOTH is somewhat larger than that of RTF (p-value of the difference: 0.13). The robustness checks presented in Appendix 2.A.1.1 indicate that the estimates of the effects remain largely unchanged if household or time fixed effects are excluded from the regression model. In addition, the interventions have only a negligible effect on shower temperature and shower frequency. We can thus interpret a decrease in water use per shower as an equivalent decrease in energy consumption. Effects are similar when restricting the sample to only one- or multi-person households, although Table 2.A.3 suggests the effect of BOTH is higher in one-person households and 2.A.4 suggests men react more strongly to SC than women.

The effect of 28.8% that we find for RTF is slightly larger than that found in Tiefenbeck et al. (2018), who estimate that RTF reduces water use per shower by 22%. The 9.4% reduction

Table 2.3. Average treatment effects on water use per shower in the exogenous phase

	(I) Volume (Liters)	(II) Relative to control group
SC	-3.19***	-9.4%
	(0.66)	
RTF	-9.79***	-28.8%
	(0.88)	
BOTH	-11.91***	-35.0%
	(1.22)	
(super-)control group average	34	
No. of households	564	
No. of observations	38,453	

Notes: OLS estimates with household and time fixed effects included. Standard errors are clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. The (super-)control group average corresponds to the average water use per shower of the control and super-control group in the exogenous treatment phase. Column (II) contains manual calculations obtained by dividing the treatment effects in column (I) by the (super-)control group average. These percentages are indistinguishable from those obtained when the dependent variable is divided by the average of the (super-)control group before performing the regression analysis.

in water use per shower caused by the SC intervention is substantially larger than that reported by Allcott (2011), who finds that SC reports decrease aggregate household electricity use by 2%. One reason for this difference could be that our SC reports are appliance-specific and thus could make behavioral adjustments easier.

Previous studies (e.g., Allcott, 2011; Tiefenbeck et al., 2018; Andor, Gerster, et al., 2020) have shown that the effectiveness of behavioral interventions on resource conservation in absolute terms increases in baseline resource use. For example, this may be because there is a higher potential for savings when initial resource use is high. We find a similar pattern with our sample. As shown in Figure 2.4, this pattern is particularly pronounced for RTF and BOTH, where the treatment effects increase from below 5 liters per shower in the second decile of baseline water use to around 15 liters in the 10th decile. By contrast, the effect of SC is close to zero in the first six deciles of baseline water use and increases to about 5 liters thereafter.

2.3.2.1 Including the endogenous treatment phase

Next, we include data from the endogenous treatment phase. This analysis is insightful from several angles: First, we can investigate the treatment effects when the interventions are allocated based on individual preferences for receiving them. This is an important use case as such interventions, as long as they are not mandated by the government, will be rolled out on an opt-in basis. To avoid incorrectly attributing potential long-run effects of the interventions from the exogenous treatment phase to the endogenous treatment phase, we control for the interventions received in the exogenous treatment phase in the regression. The results of this analysis, presented in column (I) of Table 2.4, reveal that treatment effects are slightly lower in the endogenous treatment phase compared to the results from the exogenous treatment phase, which represent estimates of the average treatment effects (ATE) in absence of self-selection into treatment. Specifically, we find that in the endogenous treatment phase SC reduces water consumption per shower by 2.83 liters, RTF by 6.73 liters, and BOTH by 9.07 liters. These ef-

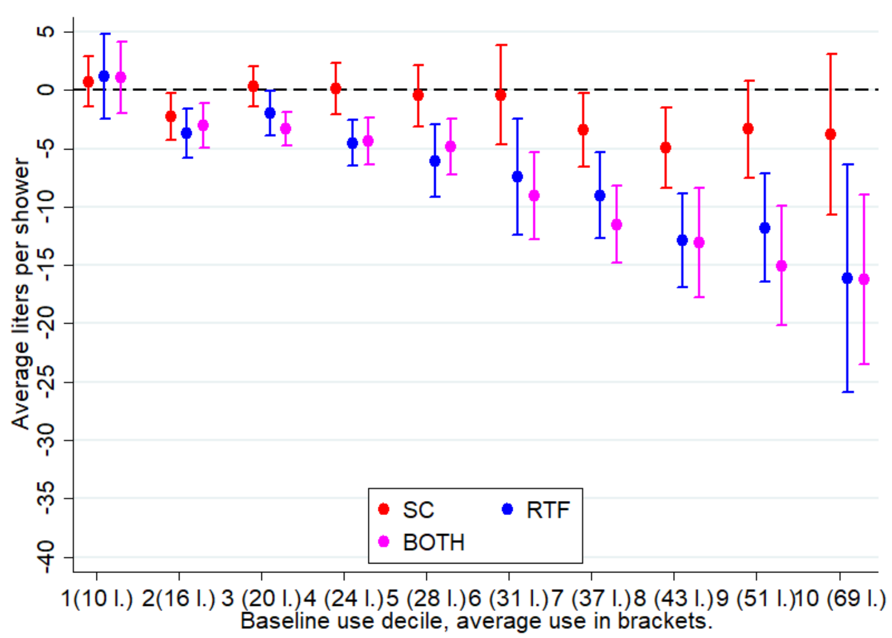


Figure 2.4. Treatment effect by baseline deciles

Note: The markers visualize effect sizes by baseline use decile, with tails indicating 95% confidence intervals.

fects can still be considered quite large, showing that the interventions can lead to considerable conservation effects even with endogenous allocation.

Table 2.4. Average treatment effects on water use per shower in both treatment phases

	(I) Only endo. phase	(II) Pooled exo. & endo. phase	(III) Pooled exo. & endo. phase (IV approach)
SC	-2.83***	-3.10***	-3.97***
	(1.04)	(0.54)	(0.56)
RTF	-6.73***	-8.34***	-9.05***
	(1.23)	(0.73)	(0.78)
BOTH	-9.07***	-10.50***	-11.60***
	(1.09)	(0.74)	(0.77)
Controls for treatment in the exo. phase	Yes	Yes	Yes
Instrumental variables	No	No	Yes
No. of households	564	564	564
Observations	37,113	54,527	54,527

Notes: OLS/IV estimates with household and time fixed effects included. Standard errors are clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. Standard errors in column III are calculated by bootstrap. In column III, we instrument treatment receipt in the endogenous treatment phase by the factors governing which of the 90 MPL questions was chosen for implementation. These are 1) whether the first or second elicitation was chosen for implementation, 2) which of the three MPL lists was chosen, and 3) which of the 15 questions on the MPL was chosen.

As an additional step, we pool the data from the exogenous and endogenous treatment phases. This maximizes the size of our estimation sample. These results are depicted in column (II) of Table 2.4. The estimated coefficients are, as expected in between the coefficients from the endogenous phase presented in column (I) and those from the exogenous phase presented in Table 2.3. The standard errors of the estimated coefficients decrease substantially due to the larger sample. As a result, the effects of all three interventions are now significantly different from each other on the 5%-level, substantiating that SC has a complementary effect to RTF, as the joint provision of SC and RTF in the BOTH group leads to a significant increase in the effect compared to the RTF group.

In column (III) of Table 2.4, we correct for the endogeneity of treatment assignment in the endogenous phase by using an instrumental variables (IV) approach.[17] We find that the treatment effects increase slightly compared to the ones from the pooled model in column (II) and are similar to the ones from the exogenous phase in Table 2.3.

2.3.2.2 Exploring the mechanisms underlying treatment differences

Having established that RTF induces greater behavior change than SC and that the effect of RTF is enhanced when combined with SC, it is natural to explore how we can explain these differences in effectiveness in terms of the mechanisms through which the different interventions may operate.

Regarding the mechanisms that SC could operate through, we adopt the intuition by Allcott (2011) who argues that SC could work because "if households are uncertain about some part of their production function, the social comparisons may facilitate social learning about their privately-optimal level of energy use" (p. 1084). In our case this means that households may not have thought about how much water they would need to use to have a satisfactory shower experience, but begin to optimize when they learn that others use much less. Second, he argues that "the treatment may directly affect the 'moral cost' of energy use" (p. 1084), i.e., in our case, excessive water use while showering becomes associated with (increased) moral costs once a household receives SC reports. Both mechanisms boil down to the idea that the SC reports increase the motivation and thus the intention to conserve water and energy while showering.

However, it is well known that people tend to exhibit an "intention-behavior gap", meaning that the right intentions are not always followed by changes in behavior that are in line with those intentions. Two of the key problems that drive the intention-behavior gap, as summarized by Sheeran and Webb (2016, p. 507), are "fail to get started", e.g., not remembering or ignoring the intention to conserve when getting into the shower, and "fail to keep goal pursuit on track", e.g., failing to put the intention to take shorter showers into practice while enjoying a pleasant shower, or simply not knowing when to stop showering in order to achieve the personal conservation goals.[18] Combining the SC reports with the RTF intervention has the potential to overcome both of these challenges, thus bridging the intention-behavior gap: Because the light signals are directly emitted by the shower heads and could not be turned off during the study, they reminded participants of their conservation intentions each time they began to shower, and the color changes reminded them continuously while they were showering, making it unlikely to be distracted from those intentions.

Even if participants were able to fully bridge the intention-behavior gap without RTF, they would still have difficulties achieving their conservation goals because the exact amount of water

17. For details on this approach see Appendix 2.A.1.3.

18. Note that we use the term "conservation goal" here to describe whatever abstract goal the participants have in their mind when they start changing their behavior. This is not to be confused with experimentally induced goal setting as for example investigated in Goette, Han, Lim, et al. (2021).

and energy they use while showering would remain a "shrouded attribute" (Gabaix and Laibson, 2006). This is because the standard deviation of water flow (liters per minute) within a household is quite high in our sample, being 0.70 at the median.[19] Thus, even participants with perfect self-control and the use of heuristics such as using a stopwatch to monitor shower time would not be able to accurately control their water use per shower. For example, a person who uses an average of 35 liters of water per shower, which takes about 3.5 minutes at 10 liters per minute, would normally be about 2.5 liters off if she sets a stopwatch to 3.5 minutes while showering. RTF addresses this problem by providing information on the amount of water used while showering, which should allow the person showering to directly relate shower duration to the amount of water and energy used, thus facilitating the precise pursuit of her conservation goals. This reasoning is in line with Jessoe and Rapson's (2014) finding that high-frequency electricity consumption information allows households to better understand electricity use quantities and thereby to more accurately respond to price signals.

Table 2.5. Average treatment effect on the standard deviation of water use per shower within households

	(I) Volume (Liters)	(II) Relative to control group
SC	-2.23***	-13.1%
	(0.68)	
RTF	-4.29***	-25.2%
	(0.71)	
BOTH	-5.83***	-34.3%
	(0.74)	
(super-)control group average	17	
No. of households	564	
No. of observations	1,127	

Notes: OLS estimates with household fixed effects included. Standard errors are clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. The (super-)control group average corresponds to the standard deviation of water use per shower of the control and super-control group in the exogenous treatment phase. Column (II) contains manual calculations obtained by dividing the treatment effects in column (I) by the (super-)control group average.

Table 2.5 shows the effect of the interventions on variation in water use per shower, comparing within-household variation in water use in the pre-intervention period with within-household variation in the intervention period. Treatment effects are significantly stronger for RTF and BOTH compared to SC. While all three interventions reduce this variation, RTF and BOTH reduce it more than twice as much as SC. This suggests that RTF makes it significantly easier for participants to achieve a certain target behavior in terms of water use per shower than SC alone.

Similarly, we show the treatment effects on the proportion of showers ended within a given color threshold in Table 2.6. These results show that, on average, participants who received SC decreased the proportion of showers that ended while the shower head was flashing red by 8 percentage points and, in turn, increased the proportion of showers that ended while the shower

19. This variation cannot be explained by different people using a shower, as the median standard deviation is still 0.64 even for one-person households.

Table 2.6. Average treatment effect on the proportion of showers ended within a certain (hypothetical) color threshold

	(I)	(II)	(III)	(IV)	(V)
	Effect on the proportion of showers ended while the shower head was / would have been...				
	...green	...blue	...purple	...red	...flashing red
SC	0.04*	0.03	-0.02	0.03*	-0.08***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
RTF	0.12***	0.05**	0.02	-0.00	-0.18***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
BOTH	0.22***	0.05*	-0.03	0.01	-0.25***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
Control group average	0.26	0.11	0.15	0.12	0.35
No. of households			446		
No. of observations			882		

Notes: OLS estimates with household fixed effects included. Standard errors are clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. The dependent variable is the share of showers per shower head ended within a given color threshold. For the baseline phase and for the treatment groups that did not receive real-time feedback, i.e., SC and the control group, the color thresholds represent only hypothetical values, that is, the thresholds indicate the colors that a given shower head would have signaled if real-time feedback had been enabled. We calculated these hypothetical color thresholds as we did for the groups receiving real-time feedback, i.e., we based them on the average water consumption per shower of a participant's reference group, which consisted of participants from the super-control group. Therefore, participants in the super-control group are not included in these analyses because there are no valid hypothetical color thresholds available for them. See Table 2.A.5 in the Appendix for an overview of the individual transitions between color thresholds.

head was still green by 4 percentage points. In contrast, participants that received RTF (BOTH) decreased their share of showers ended while the shower head was flashing red to a much larger extent, i.e., 18 (25) percentage points. In turn, they increased the share of showers ended while the shower head was still green by 12 (22) percentage points and the share of showers ended while the shower head was blue by 5 (5) percentage points. Again, these results suggest that RTF allowed participants to condition their behavior much more closely on actual water use during showering than participants in the SC group were able to do.

2.3.3 Impact of the interventions on willingness to pay

As a further step, we analyze the WTP data, which we expect to subsume the utility loss and the perceived monetary savings due to shorter showers as well as any psychological costs and benefits caused by the interventions. As the participants were informed that they could keep the shower heads after the end of the study, we do not believe that the participants included their evaluation of the shower head itself in their WTP. Our study is unique in that we elicit participants' WTP both before and after utilizing the interventions. By doing so, we gather information on the anticipated utility of the interventions for the participants, or in other words, how much they are willing to pay in the hopes of gaining from the interventions. By comparing the WTP data collected before and after the participants have used the interventions for a month, we can ascertain if the participants' perception of the interventions has shifted.

Since WTP was collected with an MPL that provides interval-censored data, we take a conservative approach and set the WTP values to the respective lower limit of the respective interval, providing lower-bound values of WTP. Thus, those who always opt for the intervention in an MPL would receive a WTP value of 15, those who indicated WTP between 10 and 15 EUR received a WTP value of 10, and so forth. Since the lower end of the MLP is an open interval limit for those who always chose the voucher without an intervention, we set the WTP to -20 in this case.

Table 2.7 depicts the average WTP values and the shares of participants with negative WTP for the full sample (panels I and II) as well as separately for the respective treatment groups (panels III and IV) and for the control group (panels V and IV). For each panel, the first column depicts WTP elicited in the baseline phase and the second column depicts WTP elicited in the exogenous treatment phase. From panel I, we find that WTP for SC is significantly lower than that for RTF (10%-level) and BOTH (1%-level), but all average WTP figures are at around 10 EUR. In addition, we find that WTP is slightly lower in the exogenous treatment phase than in the baseline phase in all panels. However, as indicated by the small and non-significant regression results in panel VII, this decrease does not appear to be due to treatment experience.

Table 2.7. Average willingness to pay for the interventions

		WTP elicited in the baseline phase		WTP elicited in the exogenous treatment phase		
		Mean	95%-Conf. int.	Mean	95%-Conf. int.	Obs.
Full sample						
	Average WTP amounts					
I	WTP for SC	9.58	[8.90 – 10.27]	8.92	[8.22 – 9.61]	439
I	WTP for RTF	10.12	[9.42 – 10.83]	9.53	[8.82 – 10.25]	439
I	WTP for BOTH	10.31	[9.63 – 10.98]	9.76	[9.06 – 10.45]	439
	Share of participants with negative WTP					
II	WTP for SC	6.8%	[4.5% – 9.2%]	5.9%	[3.7% – 8.1%]	439
II	WTP for RTF	5.9%	[3.7% – 8.1%]	5.7%	[3.5% – 7.9%]	439
II	WTP for BOTH	6.2%	[3.9% – 8.4%]	5.5%	[3.3% – 7.6%]	439
WTP in the respective treatment groups						
	Average WTP amounts					
III	WTP for SC	8.93	[7.48 – 10.38]	8.59	[7.21 – 9.97]	109
III	WTP for RTF	9.63	[7.97 – 11.29]	9.13	[7.54 – 10.72]	105
III	WTP for BOTH	11.13	[9.91 – 12.35]	9.67	[8.34 – 11.01]	107
	Share of participants with negative WTP					
IV	WTP for SC	6.4%	[1.7% – 11.1%]	9.2%	[3.7% – 14.7%]	109
IV	WTP for RTF	8.6%	[3.1% – 14.0%]	6.7%	[1.8% – 11.5%]	105
IV	WTP for BOTH	3.7%	[0.1% – 7.4%]	3.7%	[0.1% – 7.4%]	107

		WTP elicited in the baseline phase		WTP elicited in the exogenous treatment phase		
		Mean	95%-Conf. int.	Mean	95%-Conf. int.	Obs.
WTP in the control group						
	Average WTP amounts					
V	WTP for SC	9.64	[8.26 – 11.03]	9.44	[8.18 – 10.70]	118
V	WTP for RTF	10.80	[9.63 – 11.97]	10.25	[9.01 – 11.50]	118
V	WTP for BOTH	10.58	[9.29 – 11.88]	10.14	[8.90 – 11.39]	118
	Share of participants with negative WTP					
VI	WTP for SC	7.6%	[2.8% – 12.5%]	4.2%	[0.5% – 7.9%]	118
VI	WTP for RTF	3.4%	[0.1% – 6.7%]	4.2%	[0.5% – 7.9%]	118
VI	WTP for BOTH	5.9%	[1.6% – 10.3%]	5.9%	[1.6% – 10.3%]	118
Difference in differences analysis of treatment experience on WTP						
	Average WTP amounts					
VII	WTP for SC			-0.14	[-2.07 – 1.79]	878
VII	WTP for RTF			0.05	[-1.79 – 1.88]	878
VII	WTP for BOTH			-1.02	[-2.51 – 0.47]	878
	Share of participants with negative WTP					
VIII	WTP for SC			6.1%	[-2.9% – 15.1%]	878
VIII	WTP for RTF			-2.8%	[-10.5% – 5.0%]	878
VIII	WTP for BOTH			0.0%	[-7.0% – 7.0%]	878

Notes: The sample in panels I and II consists of all participants who filled out both WTP inquiries and who were not part of the super-control group, as the WTP inquiries for the super-control group were purely hypothetical. Six participants did not fill out any of the WTP inquiries, and three of the remaining participants did not fill out the second WTP inquiry. The sample in panels III and IV consists of those who filled out both WTP inquiries and who received the treatment for which the WTP is reported in the respective row in the exogenous treatment phase. The sample in panels V and VI consists of all participants who filled out both WTP inquiries and who were part of the control group (not the super-control group) in the exogenous treatment phase. The sample in panels VII and VIII is the same as in panels I and II, but it combines the data from the first and the second WTP elicitation. The depicted coefficients are estimates using a difference in differences regression, controlling for individual fixed effects.

Focusing on the proportion of participants who indicated a negative WTP for the interventions (panels II, IV, VI, and VIII), i.e., participants who were willing to forgo money to avoid receiving the interventions, we find that these shares vary between 3.4% and 9.2%, with hardly any systematic differences between the interventions.

An important aspect of the WTP is what the participants think they can save financially through the interventions. In the next step, we therefore leverage a survey question about the perceived monetary savings that we elicited at the end of the exogenous treatment period.[20] On average, participants in the SC group reported perceived savings of 4.12 EUR in the exogenous treatment month. Average perceived savings in the RTF group were 5.48 EUR and perceived savings in the BOTH group were 6.23 EUR. Using the average number of 31 showers per household in the exogenous treatment phase, the treatment effects in liters per shower from Table 2.3, and an estimated cost of about 1 Eurocent per liter of warm shower water, we estimate the actual monetary savings for one month of treatment to be around 1 EUR for SC, 3 EUR for RTF, and 3.7 EUR for BOTH. This indicates that participants tended to overestimate the monetary savings from the interventions.

Table 2.8. Comparison of willingness to pay with perceived savings for those experiencing treatment

Treat. (N)	Post-treatment WTP		Perceived savings		Net non-monetary value		Share negative net non-monetary value	
SC (89)	8.59	[7.21 – 9.97]	4.12	[2.73 – 5.52]	4.34	[2.24 – 6.43]	28.1%	[18.6% – 37.6%]
RTF (92)	9.13	[7.54 – 10.72]	5.48	[3.90 – 7.05]	3.59	[1.35 – 5.83]	25.0%	[16.0% – 34.0%]
BOTH (93)	9.67	[8.34 – 11.01]	6.23	[4.85 – 7.61]	3.38	[1.60 – 5.17]	26.9%	[17.7% – 36.1%]

Notes: In contrast to Table 2.7, only participants who answered the survey question eliciting the perceived savings are included in this table.

In Table 2.8, we contrast the perceived savings with WTP stated in the second WTP elicitation by those who experience the respective treatment. Subtracting the perceived savings from WTP provides us with an estimate of the net non-monetary value of the interventions (third column of Table 2.8). This term can be understood as the net psychological impact of the intervention, including potential psychological benefits consumers experienced by adjusting their behavior (e.g., warm glow from consuming less resources, time savings) and potential psychological costs consumers experienced by adjusting their behavior (e.g., loss of comfort due to a shorter shower). It also includes psychological benefits to consumers that are independent of the interventions' impact on showering behavior (e.g., continuing to take short showers and feeling even better about it, enjoying showering with a glowing light, enjoying discussions with friends and family resulting from the intervention), as well as psychological costs to consumers that are independent of the interventions' impact on showering behavior (e.g., continuing to take long showers but feeling bad about it, being bothered by the light, disliking receiving emails). While this net non-monetary value is significantly positive on average, the share of participants with negative net non-monetary value is between 25% and 28% (fourth column), with no significant difference between the interventions.

20. The question read: "Do you think INTERVENTION has helped you reduce your electric and water bills?" and if yes "Please estimate: "By how many euros did your energy and water costs decrease last month thanks to the INTERVENTION?"; where INTERVENTION was the specific intervention experienced by the respondent during the exogenous treatment phase.

2.3.4 Impact of the interventions on welfare

To estimate the effect of the interventions on welfare, we need to consider not only the change in consumer surplus but also the changes in externalities and the implementation costs of the interventions. In doing so, we lean on the approach by Allcott and Kessler (2019) and calculate the change in welfare ΔW as follows:

$$\Delta W_j = \Delta V_j - \phi_R \times \Delta R_j - (MCF - MU_j) \times c_j, \tag{2.2}$$

where ΔV_j is the experienced change in consumer surplus, which we measure as WTP elicited after treatment experience of intervention j.[21] ΔR_j is the change in resource use while showering, i.e., the treatment effect of intervention j. ϕ_R is the external effect resulting from the carbon emissions associated with heating one liter of shower water and c_j is the cost associated with the intervention, i.e., the cost of the smart shower head. Assuming that these costs are borne by the government, we multiply the cost of the intervention with the marginal costs of public funds (MCF). Since the revenues of the firms providing the interventions do not represent a welfare loss, the unit cost of the interventions must be reduced by the per-unit markup (MU_j) on the retail price of the interventions.

We use three scenarios to parameterize the welfare equation. The first is the "base scenario", which relates directly to the results of our experiment. The second scenario is a "conservative" scenario, in which we make pessimistic assumptions to obtain a lower bound estimate of the welfare effects. The third scenario is an "opt-in" scenario, in which we assume that only those receive the interventions who have a positive WTP for the interventions and who are therefore likely to opt-in for an intervention even in a completely voluntary setting.

For the base scenario, the values for ΔV_j are provided in Table 2.7 (WTP elicited in the exogenous treatment phase). With regard to the external effects, we focus on the effects of carbon emissions associated with the energy use required to heat water for showering. Based on the average shower temperature of 38 °C, averages for the energy sources used to heat water in Germany and (marginal) emission factors for the respective energy sources, we assume that one liter of water consumption is, on average, associated with 15 grams of CO_2 (see Appendix 2.F for details). To obtain a monetary value, we multiply the 15 grams of CO_2 by the social cost of carbon estimated in Rennert et al. (2022) of 160 EUR per ton of CO_2, resulting in external costs of 0.0024 EUR/liter.[22] ΔR_j is the effect of the respective intervention in liters per shower taken from Table 2.3 and multiplied by 31, which is the average number of showers taken in the exogenous treatment month. The costs associated with the intervention are not easily quantified. To approximate these costs, we assume the retail costs of the smart shower heads of 80 EUR and a lifetime per shower head of 5 years. Thus, we assume costs per month of $80/(5 \times 12) = 1.33$ EUR. Operational costs are abstracted from, as they are expected to be low on a large scale since the interventions can be fully automated. For the marginal costs of public funds (MCF),

21. Consumer surplus is defined as the difference between maximum WTP and the price paid. We assume in our welfare analysis that the cost of the interventions are borne by the government, i.e., considered in c_j. Therefore, consumers do not bear a direct cost, which is why WTP is an appropriate measure of consumer surplus.

22. Specifically, Rennert et al. (2022) estimate the social cost of carbon to be 185 Dollars. We transform this value to EUR assuming the long-term average exchange rate of Dollars to Euros over the past 10 years, which is 0.8627 (European Central Bank, 2023).

we assume the factor of 1.85 found by Kleven and Kreiner (2006).[23] We assume a 35% per-unit markup on the retail price, which is the value that De Loecker and Eeckhout (2018) estimate as the average markup for Germany in 2016.

The results of the welfare change due to the interventions (ΔW) in the base scenario are depicted in the upper panel of Table 2.9 and indicate that the average welfare effect of SC is 6.83 EUR per month, the effect of RTF is 7.86 EUR per month and the effect of BOTH is largest at 8.56 EUR per month. This means that all three interventions lead to substantial welfare gains.

Table 2.9. Welfare effects per month of intervention in EUR

	ΔV	$\phi_R \times \Delta R$	$(MCF - MU) \times c$	ΔW	$MVPF$
Base scenario					
SC	8.59	−0.24	2.00	6.83 EUR	6.62
RTF	9.13	−0.73	2.00	7.86 EUR	7.39
BOTH	9.67	−0.89	2.00	8.56 EUR	7.92
Conservative scenario					
SC	2.42	−0.08	2.47	0.03 EUR	1.87
RTF	2.57	−0.24	2.47	0.35 EUR	2.11
BOTH	2.72	−0.30	2.47	0.55 EUR	2.27
Opt-in scenario					
SC	12.01	−0.26	2.00	10.27 EUR	9.20
RTF	12.21	−0.73	2.00	10.94 EUR	9.70
BOTH	11.89	−0.87	2.00	10.76 EUR	9.57

Notes: Welfare changes and MVPF are calculated as follows: $\Delta W = \Delta V - \phi_R \times \Delta R - (MCF - MU) \times c$, $MVPF = \frac{\Delta V - \phi_R \times \Delta R}{c}$.

For the conservative scenario, we scale our WTP values and treatment effects down, recognizing that our sample of volunteers may be particularly motivated to conserve resources and to receive the interventions. To this end, we compare our average WTP estimate for the SC intervention, which is 8.59 EUR, to the average WTP in Allcott and Kessler (2019), which to our knowledge is the only other study that has elicited WTP for a social comparison intervention. They find that the average WTP for receiving four home energy reports on residential natural gas consumption is 2.81 Dollars, or 2.42 EUR,[24] which is 72% lower than the WTP for the SC intervention in our study. Since there are no comparable WTP estimates available for the RTF and BOTH interventions, we reduce all of our WTP values in the conservative scenario by 72%. To scale down the treatment effects, we draw on the results by Tiefenbeck et al. (2019), who conducted a study on the effectiveness of RTF while showering among hotel guests, i.e., in a population without volunteer selection bias and no monetary incentive to conserve water. They find a conservation effect of RTF of 11.4%. Given the control group's water use of 34.1 liters during the exogenous treatment period, an 11.4% reduction would correspond to water savings of 3.89 liters per shower, which is 60.3% less than the effect we found in the exogenous treatment phase (Table 2.3). In our conservative scenario, we thus assume 3.89 liters as the treatment

23. Kleven and Kreiner (2006) estimate the MCF based on several labor supply elasticity scenarios. We use the result of Scenario 6, which they denote as the "natural baseline scenario".

24. Assuming the long-term average exchange rate of Dollars to Euros over the past 10 years, which is 0.8627 (European Central Bank, 2023).

effect of RTF and also reduce the effects of the other interventions by 60.3%, resulting in a SC effect of 1.27 liters and a BOTH effect of 4.73 liters per shower. Furthermore, we assume that markups (MU_j) are zero. In addition, we reduce our emission factor for one liter of shower water by 16%, which is the share of energy used to heat water in private households that comes from renewable energy or district heating and that thus is associated with low or zero carbon emissions (see Appendix 2.F for details). The calculation in the conservative scenario results in external costs of 0.002 EUR/liter.

The results of the conservative scenario, which we interpret as a lower bound on the average welfare effects of the interventions, are shown in the middle panel of Table 2.9 and indicate that all three interventions still yield modest welfare gains ranging from 3 cents (SC) to 55 cents (BOTH) per month even under conservative assumptions.

For the opt-in scenario, we use the same parameters as in the base scenario. However, we prune the sample so that only individuals with a positive WTP after treatment experience remain in the sample.[25] The results of the opt-in scenario are shown in the bottom panel of Table 2.9 and can be interpreted as an upper bound for the welfare effects. As expected, these are much larger than in the base scenario, ranging from 10.27 (SC) to 10.94 EUR (RTF) per month. However, these welfare effects are not average effects for the entire sample, but only for those with positive WTP and thus for a smaller overall group.

Encouraged by Hendren and Sprung-Keyser (2020), we also report the marginal value of public funds (MVPF) in Table 2.9. To calculate the MVPF, we use the same input values as for the calculated welfare effects of the interventions (ΔW), but standardize the welfare gains, or in Hendren and Sprung-Keyser's (2020) terms, the societal willingness to pay, by the cost of the policy. Thus, the MVPF provides us with an estimate of the social benefits per unit of public funds, which has the advantage of allowing a unified analysis of the welfare effects of different policies and making them directly comparable across domains.

We find that, with the exception of the values in the conservative scenario, all MVPFs exceed the value of 5. A MVPF of over 5 implies that the interventions would be among the most welfare-efficient social policy measures (c.f. Hendren and Sprung-Keyser, 2020) and in the order of magnitude of several measures targeting children's health and education outcomes. Concluding, the MVPF analysis underlines that the interventions in our experiment promise considerable welfare gains.

2.4 Conclusion

This paper provides evidence comparing popular behavioral economic interventions in terms of resource conservation and welfare impacts. In a field experiment, we compare two information interventions aimed at reducing energy and water consumption while showering and their combination. One intervention provides social comparison information motivating consumers to conserve, while the other provides real-time information while showering, facilitating resource conservation for participants. Both interventions are highly effective in reducing resource use, but the facilitating intervention and the combination of both interventions lead to significantly greater reductions than the motivating intervention alone. Participants' revealed willingness to

25. Table 2.A.6 in the Appendix reports the treatment effects for this subsample, which we use to calculate ΔR in the opt-in scenario.

pay for the interventions and our welfare analysis show that both intervention types improve welfare, but that the facilitating and the combined intervention do so to a greater extent. Thus, our findings from a resource conservation setting suggest that information that facilitates behavior change is more effective and welfare enhancing than information that motivates behavior change.

Our results are in line with previous findings highlighting the importance of making shrouded attributes salient to consumers (e.g., Gabaix and Laibson, 2006; Chetty, Looney, and Kroft, 2009; DellaVigna and Pollet, 2009; Grubb, 2009; Brown, Hossain, and Morgan, 2010; Taubinsky and Rees-Jones, 2018). Our findings corroborate that in environments such as showering and energy use in the home, quantity consumed is a shrouded attribute. Making quantities salient to consumers removes barriers to behavioral change, facilitating reactions to prices (Jessoe and Rapson, 2014) and knowledge about environmental impact (Fang et al., 2023).

Our findings also showcase the potential for behavioral change in consumption settings in which fine-grained data can be made available. While information interventions motivating behavioral change, such as social comparison reports (Allcott and Mullainathan, 2010; Benartzi et al., 2017), can achieve effects in some settings even in the absence of fine-grained data, appliance-specific data has the potential to generate much larger effects. Increasing technological development and technological diffusion is making low-cost solutions to assessing real-time information more available than ever before. For example, a large German retailer has recently started selling a RTF shower head at very affordable prices (Tchibo, 2023). Policymakers should be aware of these developments and the potential they hold for behavioral interventions made possible through such technology. From a practical point of view, RTF has the advantage over SC reports that it can be used with little information on the consumers or it could even work without the need to process or collect data, for example by hardcoding the light signals in the shower head based on pilot studies like ours. This could help to overcome privacy-related objections to behavioral interventions.

While the application we examine in detail – showering behavior – is in itself an everyday and energy-intensive behavior and thus important to consider in energy conservation efforts, there is potential for future research to assess whether the pattern we observe also holds for other water- and energy-intense behaviors. For example, future research could compare motivating and facilitating information interventions to address consumers' room heating and cooling behaviors. We would expect patterns to be similar since the energy used for room heating and cooling is – similar to the energy used for heating water when showering – a shrouded, non-salient attribute to the consumer. We thus expect a similar energy-saving potential for information interventions making this attribute salient.

In contrast, the distinction between motivating and facilitating information interventions might be less pronounced in an environment where salience bias is less of an issue. This would apply to environments in which consumers are better able to develop heuristics for their consumption or where the consumption itself is more salient (e.g., taking a plane). Comparing motivating and facilitating information interventions in such applications would be an interesting avenue for future research.

More generally, our findings contribute to the still very limited literature assessing the welfare effects of (behavioral) interventions by directly eliciting consumers' willingness to pay for experiencing the interventions (Allcott and Kessler, 2019; Allcott et al., 2022; Butera et al., 2022). In our setting, the behavioral interventions on average create a win-win situation: They reduce environmental externalities and simultaneously create a psychological benefit to consumers. We also show that the intervention experience did not change willingness to pay.

Appendix 2.A Supplementary analyses

2.A.1 Further analyses on the impact of the interventions behavior

2.A.1.1 Robustness checks

Table 2.A.1. Treatment effects robust to different specifications

	Volume (Liters)		
SC	-3.79**	-4.66**	-3.19***
	(1.55)	(1.98)	(0.66)
RTF	-9.91***	-10.79***	-9.79***
	(1.26)	(1.68)	(0.88)
BOTH	-12.01***	-12.88***	-11.91***
	(1.33)	(1.74)	(1.22)
Time FE	No	Yes	Yes
Household FE	No	No	Yes
No. of households	564	564	564
No. of observations	38,453	38,453	38,453

Notes: Standard errors are clustered at the household level and are reported in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 2.A.2. Negligible effect on shower temperature and frequency

	Shower temperature	Showers per day
SC	-0.17	0.05
	(0.15)	(0.04)
RTF	-0.19	-0.03
	(0.18)	(0.03)
BOTH	-0.29*	-0.02
	(0.15)	(0.03)
No. of households	564	564
No. of observations	38,560	32,148

Notes: OLS estimates with household and time fixed effects included. For Spec. (2), we construct a data set including the number of showers per household per day during the baseline phase and exogenous phase of the study. Standard errors are clustered at the household level and are reported in parentheses. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Table 2.A.3. Average treatment effects on water use per shower in the exogenous phase split by household size

	(I) Full sample	(II) 1-person households	(III) 2- and 3-person households
SC	-3.19***	-3.39***	-3.14***
	(0.66)	(1.09)	(0.79)
RTF	-9.79***	-8.36***	-10.14***
	(0.88)	(1.60)	(1.03)
BOTH	-11.91***	-15.34***	-10.93***
	(1.22)	(3.50)	(1.08)
(super-)control group average	34	31	35
No. of households	564	188	376
No. of observations	38,453	9,081	29,372

Notes: Dependent variable is shower volume in Liters. OLS estimates with household and time fixed effects included. Standard errors are clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. The (super-)control group average corresponds to the average water use per shower of the control and super-control group in the exogenous treatment phase, restricting to the respective households. Column (II) restricts the sample only to one-person households, and column (III) restricts the sample to only 2- or 3-person households. Of these, 329 are two-person and 47 are three-person.

Table 2.A.4. Average treatment effects on water use per shower in the exogenous phase split by gender (1-person households)

	(I) All 1-person households	(II) Male 1-person households	(III) Female 1-person households
SC	-3.39***	-4.18***	-1.88
	(1.09)	(1.52)	(1.42)
RTF	-8.36***	-7.19***	-8.85***
	(1.60)	(2.39)	(2.01)
BOTH	-15.34***	-16.78***	-12.75**
	(3.50)	(4.93)	(4.93)
(super-)control group average	31	32	29
No. of households	188	108	79
No. of observations	9,081	5,598	3,428

Notes: Dependent variable is shower volume in Liters. OLS estimates with household and time fixed effects included. Standard errors are clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. The (super-)control group average corresponds to the average water use per shower of the control and super-control group in the exogenous treatment phase, restricting to the respective households. Column (I) includes all one-person households, (II) restricts the sample only to male one-person households, and column (III) restricts the sample to only female 1-person households.

2.A.1.2 Analysis of the transition between color thresholds due to the interventions

Table 2.A.5. Transition matrix: Changes in the color threshold in which a participant's showers ended most often

			to					
			green	blue	purple	red	flashing red	N
Cont	**from**	green	84%				16%	25
		blue	25%	33%	8%		33%	12
		purple	8%	8%	46%	8%	31%	13
		red	14%		21%	21%	43%	14
		flashing red	4%	2%	4%	2%	89%	54
SC	**from**	green	95%			3%	3%	38
		blue	86%	14%				7
		purple	15%	46%	23%		15%	13
		red			100%			1
		flashing red	8%	12%	4%	2%	75%	51
RTF	**from**	green	97%				3%	31
		blue	89%	11%				9
		purple	20%	60%		10%	10%	10
		red	30%	40%	20%	10%		10
		flashing red	29%	4%	18%	6%	43%	49
BOTH	**from**	green	100%					25
		blue	87%	13%				15
		purple	54%	38%	8%			13
		red	100%					1
		flashing red	29%	18%	9%	5%	38%	55

Notes: We present the changes between the baseline phase and the exogenous treatment phase. The sum of the individual rows is 100% (if not, this is due to rounding error). As an example, we can interpret the last row of the table as follows: Of the 55 participants in the BOTH group who finished most of their showers in the baseline phase while the shower head would have flashed red, 29% (18%, 9%, 5%, 38%) finished most showers in the exogenous treatment phase while the shower head was green (blue, purple, red, flashing red). We calculated the hypothetical color thresholds in the baseline phase as we did in the treatment phases, that is, we based them on the average baseline water consumption per shower of a participant's reference group, which consisted of participants from the super-control group. Therefore, participants in the super-control group are not included in this table because there are no valid hypothetical color thresholds available for them.

2.A.1.3 Details on treatment assignment in the endogenous phase and the IV approach

To explain the functioning of our instrumental variables, we first explain the details of treatment assignment in the endogenous treatment phase:

(1) It was randomly drawn which of the two WTP inquiries was relevant for the allocation.
(2) From the relevant WTP inquiry, it was randomly drawn which of the three MPLs was relevant for the allocation, i.e., whether the MPL concerning WTP for SC, RTF or BOTH was the relevant MPL.
(3) One of the 15 rows in the relevant MPL, i.e., one decision, was randomly selected. If the participant chose the intervention in this row the participant received the intervention in the endogenous treatment phase and if the participant did not choose the intervention, she was allocated to the control group in the endogenous treatment phase.

Given this treatment assignment mechanism, we can divide our participants into two groups: The first group is the group of "compliers", i.e., those who received the intervention randomly drawn in step 2 of the treatment assignment mechanism. This group is composed of two subgroups, the first consisting of those participants who reported the maximum WTP for all three interventions in both WTP surveys, which is the case for 43% of our participants, and the second consisting of those who reported different WTP across interventions or WTP inquiries but whose decision in the specific MPL row drawn in step 3 of the assignment mechanism was in favor of the intervention. The second group consists of the so-called "never takers", who are those who were assigned a specific intervention in step 2 of the allocation mechanism, but decided against receiving the intervention in the MPL row drawn in step 3. Our approach of using the random draws from all three steps of the assignment mechanism as instrumental variables allows us to estimate the causal local average treatment effect (LATE) for the group of compliers (Angrist, Imbens, and Rubin, 1996).

2.A.1.4 Treatment effect for the opt-in scenario (sub sample of those with WTP>0)

Table 2.A.6. Average treatment effects on water use per shower in the opt-in scenario (sub sample of those with WTP>0)

	(I) Volume (Liters)
SC	-3.43***
	(0.85)
RTF	-9.78***
	(1.04)
BOTH	-11.74***
	(1.43)
Control group average	33
No. of households	358
No. of observations	25,416

Notes: OLS estimates with household and time fixed effects included. The results show treatment effects in the exogenous treatment phase for households who would choose to take up an intervention in an opt-in scenario, i.e., the estimation sample consists of those who stated a positive WTP after treatment experience or who are part of the control group in the exogenous treatment phase and who stated a positive WTP for the SC intervention in the second WTP inquiry. The super-control group is not part of the estimation sample. Standard errors clustered at the household level. ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Willingness to pay analysis correcting for censored data

Table 2.A.7. Average Willingness to pay for interventions – interval regression

	Baseline WTP			WTP after experience		
	Mean	95%-Conf. int.	Obs	Mean	95%-Conf. int.	Obs
Average WTP amounts						
WTP for SC	17.46	[15.45 – 19.47]	437	14.14	[10.85 – 17.42]	109
WTP for RTF	19.80	[17.47 – 22.14]	438	17.48	[13.09 – 21.87]	105
WTP for BOTH	19.73	[17.52 – 21.93]	436	16.69	[13.00 – 20.38]	108

Notes: Willingness to pay data included from those households who were not part of the super-control group, i.e. filled out incentivized willingness to pay inquiries.

Table 2.A.8. Difference in differences-analysis of WTP – interval regression

	WTP SC	WTP RTF	WTP DUAL
Exo Phase	-0.94	-1.10	-1.54
	(1.30)	(1.53)	(1.20)
Exo Phase x SC	-0.36	0.87	1.73
	(1.89)	(2.24)	(2.01)
Exo Phase x RTF	-1.18	-1.17	-0.01
	(1.70)	(2.14)	(1.84)
Exo Phase x DUAL	-1.84	-1.43	-2.46
	(1.71)	(2.08)	(1.82)
Constant	17.08***	20.48***	20.06***
	(1.62)	(1.77)	(1.70)
lnsigma			
Constant	2.60***	2.72***	2.67***
	(0.04)	(0.05)	(0.05)
No. of participants	440	440	440
Observations	876	876	875

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

Notes: Willingness to pay data included from those households who were not part of the Super-Control group, i.e. filled out incentivized willingness to pay inquiries. Household fixed effects included.

Appendix 2.B Invitation for study participation

Dear XYZ!

With this email we would like to invite you to participate in a new study by our long-time scientific partner, RWI - Leibniz Institute for Economic Research*. The study is financed by the Ministry of Culture and Science of the State of North Rhine-Westphalia and examines ways to reduce resource consumption. In a first step, only our panelists have the opportunity to participate.

WHAT EXACTLY IS THE STUDY ABOUT?
As a core element of the study, participants will be provided with a high-quality shower head

that measures water consumption per shower. During the four-month study phase, this information will be automatically shared with RWI. Additionally, you will be provided with various information about your showering behavior during the course of the study. At the end of the study, you will be able to keep the shower head, which then ceases to share information. This will allow you to keep track of your water consumption yourself on a permanent basis.

HOW CAN I SIGN UP?
You can find all further information and registration form on the project website: [URL]

YOUR BENEFITS
As an **expense allowance**, you will receive vouchers or additional resource saving tools worth a total of € 35 in addition to the shower head before the end of the study. You will receive these vouchers as part of your participation in short, monthly surveys conducted by RWI to accompany the study. At the end of the study, we will offer you a selection of popular suppliers where you can redeem your vouchers. Furthermore, you will of course be informed about the results of the study in the form of a short report. So again in a nutshell:

- Modern shower head with innovative functions
- Vouchers as an expense allowance - choose your preferred provider from our list of providers
- Report with the results of the study
- Participation in a study, the findings of which will make the world a little bit more sustainable

WHY SHOWERS?
What's not commonly known is that showering involves a relatively high consumption of resources, as not only water but also significant amounts of energy are used to heat the water. For example, with the energy consumed by an ordinary shower, you could light an average household for two and a half days, or watch TV for 26 hours.

HOW IS MY DATA PROTECTED?
The study is subject to strict data protection regulations. More detailed information can be found on the project website: [LINK].

DON'T FORGET TO REGISTER!
You can register on the project homepage [LINK].

If you have any questions about the study, please contact the project team at [E-MAIL ADDRESS].

Your forsa.omninet Team

* RWI - Leibniz Institute for Economic Research in Essen is an independent, non-profit scientific research institution. As one of the leading economic research institutes in Germany, the institute is a member of the Leibniz Association.

Appendix 2.C Website to register for the study

The project website aimed at participants included the following subpages:

- **Home page:** Welcome message and general introduction to the study
- **Project goals:** Short overview of the experiment's objectives
- **Your contribution:** Information on environmental potential of the study
- **FAQ:** Answers to typical questions
- **Who we are:** Short overview of the research team
- **Sign up:** Overview of the general terms & conditions and signup form

The following subsection features a translated version of selected pages.

2.C.1 Website text content

Home Page
Welcome to the "sustainable showering" project homepage!

We are pleased that you are interested in our research project! On this website, we would like to introduce the project to you and hope to win you as a study participant.

This project is a scientific research project of RWI - Leibniz Institute for Economic Research (Essen) in research partnership with the Center for Advanced Internet Studies (CAIS).

The project is funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia under the title "Digitization of Sustainable Behavior." [Click here to go to the project page on the RWI homepage. Please note that you can only register on the current page].

As a core element of your study participation, you will be provided with a high-quality shower head that measures water consumption per shower. During the four-month study phase, this information will be automatically shared with RWI. Additionally, you will be provided with various information about your showering behavior during the course of the study. At the end of the study, you will be able to keep the shower head, which then ceases to share information. This will allow you to keep track of your water consumption yourself on a permanent basis.

As an expense allowance, in addition to the showerhead, you will receive optional vouchers and/or additional resource saving tools worth a total of €35 until the end of the study. You will receive the vouchers as part of your participation in short, monthly surveys that will be conducted to accompany the study.

Feel free to take a further look at the project page and sign up!

Project goals

"Sustainable showering":
Of course, showering is an indispensable part of personal hygiene. Simultaneously however,

showering is a very resource-intensive activity and hardly perceived as such. Many people underestimate the associated energy consumption, caused in particular by heating water. For example, an average shower consumes about 2.6 kilowatt hours. The same amount of energy could be used to light an average household for two and a half days, or to watch television for 26 hours. So the question is whether it is possible to shower more "sustainably."

The goals of the project:

- Gather information on the exact amount of energy and water used in showering.
- Development of helpful measures to reduce resource consumption.

By saving energy consumption while showering, an individual's CO2 emissions can be significantly reduced. An average shower taken causes annual emissions of 325 kg of CO2. It takes 26 trees to reabsorb this amount of CO2 within one year. Or put another way: Just 14 showers emit as much CO2 as a tree can absorb within a year.

Your contribution

Your contribution to climate protection:
By participating in this study, you are making a concrete contribution to climate protection. This happens in two ways:

- You gain exclusive access to the savings measures investigated in this study, which can be used to reduce your personal resource consumption and the associated energy costs and CO2 emissions.
- The knowledge generated in the study can be used to formulate concrete recommendations for resource conservation in the context showering and other energy-intensive activities. In this way, policy measures can be developed to effectively save CO2 emissions.

Every reduction in CO2 emissions makes a valuable contribution to combating climate change. Further information on the causes and consequences of climate change can be found, for example, on the website of the Federal Agency for Civic Education: https://www.bpb.de/gesellschaft/umwelt/klimawandel/.

Our project is primarily focused on saving energy, i.e., measures are being developed to save hot water. However, water is also a valuable resource in itself and is scarce in many parts of the world, as well as regionally in Germany, at least at times. See, for example, here: https://www.umweltbundesamt.de/presse/pressemitteilungen/wassersparen-sinnvoll-ausgereizt-uebertrieben. Thus, with the study, we will identify scientific findings that are also of great interest to other parts of the world.

FAQ

Do I receive an expense allowance for my participation?
Yes. You will be compensated in two ways for the time and effort you have incurred. First, you get to keep the high-quality shower head used in the study after completion. With this shower head, you can continue to keep track of your water consumption by using a smartphone app. Second, you will receive optional vouchers and/or additional resource saving tools worth €35 until the end of the study. You will receive the vouchers as part of your participation in short, monthly surveys that accompany the study. At the end of the study, we will offer you a selection of popular suppliers where you can redeem your vouchers.

How much effort will it take for me to participate?
At the beginning of the study, we will send you a shower head and a WiFi gateway, which you can easily install yourself with the help of the enclosed instructions. If you encounter any difficulties, we will gladly assist you with the installation. During the course of the study, we will send you monthly online surveys by e-mail, each of which can be completed in about 15 minutes. In addition, you will receive newsletters from us at regular intervals, which you only need to read. The study involves no further effort on your part and is completed after four months.

How is my data collected?
The study is subject to strict data protection regulations. For details, please read the project's privacy policy.

I am going on vacation within the study phase. Is this a problem?
No. However, it would be nice if you let us know when you register. You have the option to do this on the registration form under "Notify the Project Team" or by emailing [E-MAIL ADDRESS].

There is more than one shower in my household. Can I still participate?
Of course. In this case, we ask you to install the new shower head at the shower that you personally use most often.

My shower is not compatible with conventional shower heads.
If you for example have a Rain Dance shower that requires a special shower head, participation is unfortunately not possible. If you are unsure about your shower head model, please feel free to contact us. You can find the contact information here.

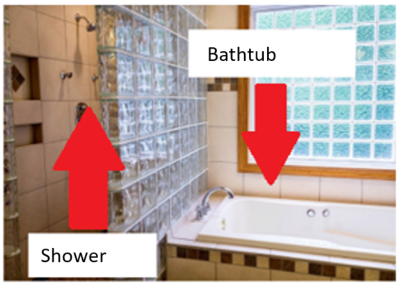
2.C.2 Signup form content

1. We kindly ask you to provide information about the number of persons living in your household permanently:
Household size

- Total number of persons (adults and children)
- Number of children (age below 18)

2. We kindly ask you to provide information about your shower setup:

Picture 1: Shower and bathtub combined



Picture 2: Shower and bathtub separate

How many separate or combined showers and bath tubs are there in your home?

- Number of bath tubs with shower (combination, see picture 1)
- Number of separate showers (not in bath tub, as seen in picture 2)
- Number of bath tubs (without shower or not used as shower, as seen in picture 2)

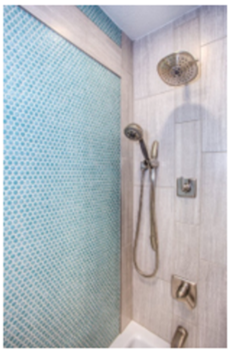
3a. This question is about the shower you use most frequently:

How many household members use this shower at least once a week?

- Number of shower users

Picture 1: The shower head is screwed onto a shower hose.

Picture 2: The shower head protrudes from the wall.



Picture 3: There are two shower heads. One is screwed onto a shower hose, and the other protrudes from the wall.

Which of these descriptions fits the shower in question?

- No selection made
- Number of separate showers (not in bath tub, as seen in picture 2)
- Number of bath tubs (without shower or not used as shower, as seen in picture 2)
- Number of bath tubs (without shower or not used as shower, as seen in picture 2)

Visually, the shower head we are going to provide you with resembles an ordinary hand shower (see picture below). It can be screwed onto ordinary 1/2 inch shower hoses, which can be found in most households.

3b. We kindly ask for your assessment about the shower you personally use the most

- No selection made
- I think my shower is compatible with this shower head
- I am unsure whether my shower is compatible with this shower head or not
- I do not think my shower is compatible with this shower head

4. Technical requirements

To set up data transmission, you need to connect the shower head to your WiFi network using a smartphone app. This step can be completed in a minute and is necessary for study participation. To perform this step, you will need your WiFi password and a smartphone. The WiFi password will not be shared with us.

- No selection made
- I own a smartphone and have access to my WiFi password
- I own a smartphone and am unsure about my WiFi password
- I do not own a smartphone and/or do not have access to my WiFi password

In addition to the smart shower head, we will provide you with a WiFi gateway, which is required to connect the shower head to your network. This device is very small (dimensions: 4.5 x 4.5 x 1cm, see picture below). It should be located within a radius of about 5 meters around your shower (ideally in the same room) and should be continuously powered.

- I have access to an electrical outlet within a radius of 5 meters around my shower

Appendix 2.D Exemplary newsletter for the BOTH group

Showering sustainably
Your personal newsletter

Dear Mr/Ms XYZ,

do you remember how much energy is consumed in a five-minute shower at a water temperature of 38 degrees Celsius? That's right, it's 2.2 kWh!
With these 2.2 kWh you could also

- toast 132 slices of bread
- boil 16.5 litres of water at 100 degrees Celsius with an electric kettle

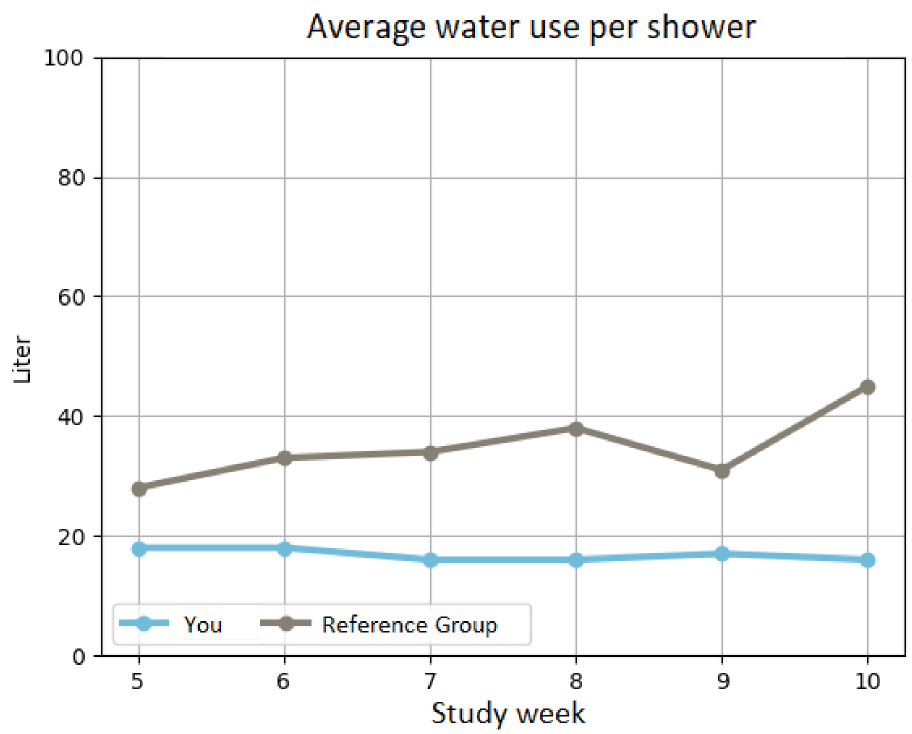
The generation of 2.2 kWh of energy emits 650 g of CO2. It takes 19 beech trees to reabsorb this amount of CO2 within one day.

Your water consumption in comparison

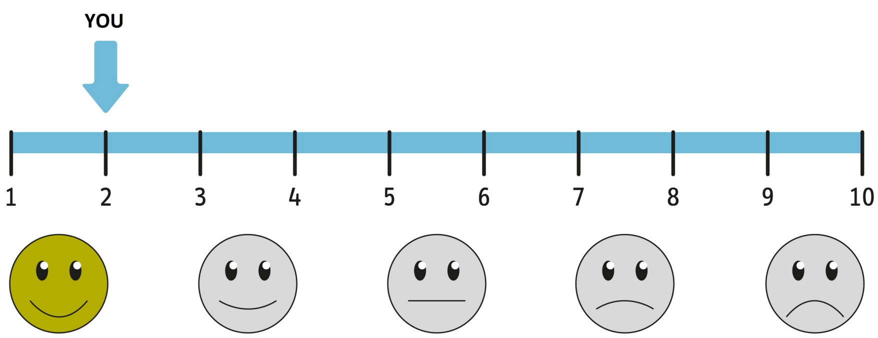
In the last two weeks, you have used an average of **17 litres of water** per shower. At your average water temperature and the type of water heating used by an average German citizen, this corresponds to emissions of **244.5 g of CO2 per shower**. It takes 7.1 beeches to reabsorb this amount of CO2 within one day.

The average consumption of 9 other randomly selected participating households was **36 litres**. The average CO2 emission per shower was **537.0 g CO2 per shower**. It takes 15.7 beeches to reabsorb this amount of CO2 within one day.

In the following figure you can see how your average consumption and the average consumption of your comparison group developed in the last weeks. The last two weeks were weeks 9 and 10 of the study.



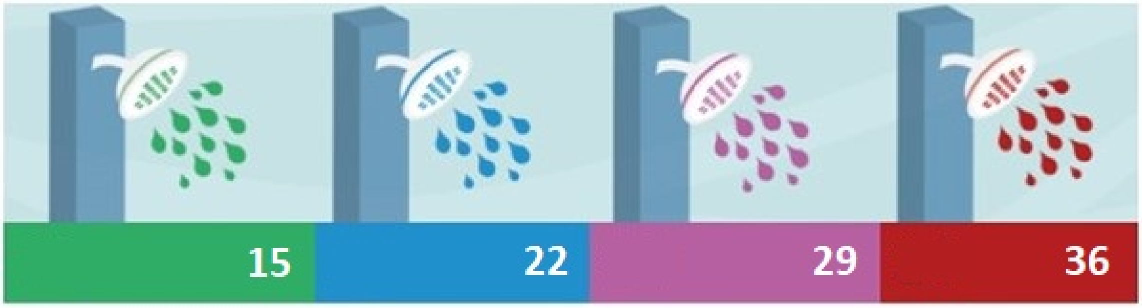
Compared to 9 other participating households, your rank is 2nd place:



One of these households had lower average water use per shower than you. Eight of these households had a higher average water consumption.

Your smart showerhead

Your shower head will help you keep track of your water consumption while showering. It will light up in color and as your water consumption increases, the color will change. Finally, once your water consumption exceeds 36 litres, it will start flashing red. 36 litres of water consumption corresponds to 537 g of CO_2 emissions based on your average water temperature and the energy consumption of an average German citizen. It takes 15.7 beech trees to reabsorb this amount of CO2 within one day. The exact sequence of colours is as follows:



During the first 15 litres of your shower, the shower head will glow green.
After **15 litres** (225 g CO_2 emissions), your shower head will begin to glow blue.
After **22 litres** (330 g CO_2 emissions), your shower head will begin to glow purple.
After **29 litres** (435 g CO_2 emissions), your shower head will start to glow red.
Finally, after **36 litres** (537 g CO_2 emissions), your shower head will begin to flash red.

Sincerely,
Your "nachhaltiges Duschen" team

Please do not reply to this mail. For questions and comments, please contact us at [E-MAIL ADDRESS].

Appendix 2.E Technical details and installation of the smart shower heads

Participants received a package with the smart shower head and the WiFi gateways. In parallel, participants received the installation instructions in digital form via email. Upon request, a printed version of the instructions was enclosed in the package.

The installation of the smart shower heads and the infrastructure for data transmission took place in the following steps.

The existing shower head had to be unscrewed from the shower hose and the smart shower head had to be screwed on.

The proper functioning of the shower head could be confirmed by brief light signals, which the shower head emits during start-up. Apart from this short signal, the shower head to the participants functioned like any regular shower head during the Baseline Phase.

The WiFi gateway had to be plugged into a power outlet near the shower head. Participants then had to download an app to connect the WiFi gateway to their home network. This was done selecting the correct WiFi network and entering the network key. A successful connection of the gateway was indicated by light signals.

If problems arose, participants could reach the project team by e-mail. The problems were then solved either by e-mail, phone call or personal visit. Only for a minority of the study participants were the problems not solvable. In most cases, the issues resulted from WiFi configuration errors or too much distance between the gateway and the smart shower head, which could be easily solved.

After successful installation, the shower head transmitted information for every shower taken (time stamp, amount of water used, average water temperature, water flow, length of shower breaks) via Bluetooth to the WiFi gateway, which then transmitted the information to the research team via the Internet. Participants had no way to access this information during the study.

If the WiFi gateway was not plugged in or had no internet connection during a shower, the data from that shower was stored in the shower head (up to 200 showers can be stored) and transmitted with the next successful connection. Interruptions of showers of less than 3 minutes were interpreted by the smart shower head as shower breaks, while an interruption of more than 3 minutes signaled the start of a new shower.

Appendix 2.F Details on the calculation of shower costs and carbon intensity

In the following, we describe our approach and assumptions to calculate (1) the costs and (2) the carbon intensity per liter of shower water. We draw on calculations by the Consumer Organisation of the German state of Rhineland-Palatinate (Consumer Protection Agency RLP, 2020) on the average cost and energy consumption per shower. The assumed temperature in their calculations corresponds exactly to the average baseline temperature in our control group of 38°C (Table 2.2). In the next step, we calculate the cost and energy demand per liter of shower water, distinguishing between water heating by electricity, gas and oil. We then multiply the energy demand per liter of shower water with the CO_2 intensity of one kilowatt-hour of heat supply with the respective energy source. For gas- and oil-fired water heating, we use the estimates of the CO_2 intensity provided by the German Federal Environment Agency (German Environment Agency, 2019; Table 62 "Emission factors for the provision of heat from fossil fuels in private households, the tertiary sector and industry"). For electric water heating, we use an estimate for the marginal emission factor of electricity generation in Germany provided by Huber et al. (2021). The use of a marginal emission factor compared to average emissions per kWh is appropriate because a reduction in electricity demand leads to a reduction in electricity generation at the marginal power plant, which is typically fossil fuel-fired and therefore more carbon-intensive than the average electricity mix (see, e.g., Holland et al., 2022). We then use data from the German Federal Ministry for Economic Affairs and Energy (2020) on the distribution of energy consumption for water heating in private households to calculate a representative CO_2-factor per liter of shower water. In the base scenario, we make the assumption that shower water in Germany is heated only by electricity, gas or oil. This is a simplification, but it includes a large 84% of the final energy consumption for water heating in German households in 2018. The remaining households obtained district heating or heated water with renewable energy. As a representative CO_2-factor, this calculation yields a value of 0.015 kg CO_2 per liter of shower water.

In the conservative scenario, we reduce our emission factor for one liter of hot water consumption by 16%, assuming that the share of energy used to heat water in private households that comes from renewable energy or district heating is associated with zero carbon emissions. While district heating does in principle entail carbon emissions, there may be cases where it is fed entirely by industrial waste heat, so that the use of this waste heat does not result in further emissions, which is what we assume in the conservative scenario. The calculation in the conservative scenario results in external costs of 0.002 EUR/liter.

Table 2.F.1. Estimation of CO_2 per liter

	(I) Energy cost per shower (40 L at 38°C)	(II) Cost per kWh	(III) kWh per L at 38°C	(IV) kg CO_2 per kWh	(V) kg CO_2 per L	(VI) PJ used for water heating in 2018	(VII) Share	(VIII) kg CO_2 per L (weighted avg.)
Elec.	0.40 EUR	0.28 EUR	0.036	0.550	0.020	55.307	17.9%	
Gas	0.10 EUR	0.05 EUR	0.050	0.246	0.012	190.384	61.6%	0.015
Oil	0.15 EUR	0.07 EUR	0.054	0.318	0.017	63.307	20.5%	
			$\frac{(I)/40}{(II)}$		$(III) \times (IV)$		$\frac{(VI)}{\sum(VI)}$	$\sum[(V) \times (VII)]$

Notes: Sources: First column: Consumer Protection Agency RLP (2020), Second column: Consumer Protection Agency RLP (2020), Fourth column: Huber et al. (2021) for Elec. & German Environment Agency (2019) for Gas/Oil, Sixth column: German Federal Ministry for Economic Affairs and Energy (2020)

References

Allcott, Hunt. 2011. "Social norms and energy conservation." *Journal of Public Economics* 95 (9-10): 1082–95. [124, 127, 136, 138]

Allcott, Hunt, Daniel Cohen, William Morrison, and Dmitry Taubinsky. 2022. "When do "nudges" increase welfare?" Working Paper, Working Paper Series 30740. National Bureau of Economic Research. https://doi.org/10.3386/w30740. [126, 147]

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. [124, 126, 127, 144, 145, 147]

Allcott, Hunt, and Sendhil Mullainathan. 2010. "Behavior and energy policy." *Science* 327 (5970): 1204–5. [123, 147]

Allcott, Hunt, and Todd Rogers. 2014. "The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation." *American Economic Review* 104 (10): 3003–37. [124, 127]

Andor, Mark A, and Katja M Fels. 2018. "Behavioral economics and energy conservation – a systematic review of non-price interventions and their causal effects." *Ecological Economics* 148: 178–210. [127]

Andor, Mark A, Andreas Gerster, Jörg Peters, and Christoph M Schmidt. 2020. "Social norms and energy conservation beyond the US." *Journal of Environmental Economics and Management* 103: 102351. [127, 136]

Andor, Mark A, Lorenz Goette, Mike K Price, Anna Schulze Tilling, and Lukas Tomberg. 2020. *The effects of information on consumer surplus in resource consumption.* AEA RCT Registry. November 18. https://doi.org/10.1257/rct.6747-1.0. [130]

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of causal effects using instrumental variables." *Journal of the American Statistical Association* 91 (434): 444–55. [151]

Ariely, Dan, and Simon Jones. 2008. *Predictably irrational.* HarperCollins New York. [123]

Asensio, Omar I, and Magali A Delmas. 2015. "Nonprice incentives and energy conservation." *Proceedings of the National Academy of Sciences* 112 (6): E510–E515. [127]

Benartzi, Shlomo, John Beshears, Katherine L. Milkman, Cass R. Sunstein, Richard H. Thaler, Maya Shankar, Will Tucker-Ray, William J. Congdon, and Steven Galing. 2017. "Should governments invest more in nudging?" *Psychological science* 28 (8): 1041–55. [123, 147]

Brandon, Alec, John A List, Robert D Metcalfe, Michael K Price, and Florian Rundhammer. 2019. "Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity." *Proceedings of the National Academy of Sciences* 116 (12): 5293–98. [127]

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. [124, 127]

Brown, Jennifer, Tanjim Hossain, and John Morgan. 2010. "Shrouded attributes and information suppression: Evidence from the field." *Quarterly Journal of Economics* 125 (2): 859–76. [125, 127, 147]

Brülisauer, Marcel, Lorenz Goette, Zhengyi Jiang, Jan Schmitz, and Renate Schubert. 2020. "Appliance-specific feedback and social comparisons: Evidence from a field experiment on energy conservation." *Energy Policy* 145: 111742. [127]

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. [124, 126, 147]

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. [124]

Cadario, Romain, and Pierre Chandon. 2020. "Which healthy eating nudges work best? A meta-analysis of field experiments." *Marketing Science* 39 (3): 465–86. [123]

Carson, Richard T, and Theodore Groves. 2007. "Incentive and informational properties of preference questions." *Environmental and Resource Economics* 37 (1): 181–210. [130]

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and taxation: Theory and evidence." *American Economic Review* 99 (4): 1145–77. [125–127, 147]

Congiu, Luca, and Ivan Moscati. 2020. "Message and Environment: A framework for nudges and choice architecture." *Behavioural Public Policy* 4 (1): 71–87. [124, 126]

Consumer Protection Agency RLP. 2020. *Warmwasser: Komfortables Sparen – So geht's!* As of: February 2020. Verbraucherzentrale Rheinland-Pfalz. https://www.verbraucherzentrale-rlp.de/sites/default/files/2020-03/VZ_WarmWasser_2020_12_web.pdf, accessed on September 22, 2020. [164, 165]

De Loecker, Jan, and Jan Eeckhout. 2018. "Global Market Power." Working Paper, Working Paper Series 24768. National Bureau of Economic Research. https://doi.org/10.3386/w24768. [145]

DellaVigna, Stefano, John A List, and Ulrike Malmendier. 2012. "Testing for altruism and social pressure in charitable giving." *Quarterly Journal of Economics* 127 (1): 1–56. [124, 126]

DellaVigna, Stefano, John A List, Ulrike Malmendier, and Gautam Rao. 2016. "Voting to tell others." *Review of Economic Studies* 84 (1): 143–81. [124, 126]

DellaVigna, Stefano, and Joshua M Pollet. 2009. "Investor inattention and Friday earnings announcements." *Journal of Finance* 64 (2): 709–49. [127, 147]

European Central Bank. 2023. *Euro foreign exchange reference rates: US dollar (USD).* Change from 24 January 2013 to 24 January 2023. European Central Bank (ECB). https://www.ecb.europa.eu/stats/policy_and_exchange_rates/euro_reference_exchange_rates/html/eurofxref-graph-usd.en.html, accessed on January 24, 2023. [144, 145]

Faberi, Stefano, Loriana Paolucci, Bruno Lapillonne, and Karine Pollier. 2015. *Trends and policies for energy savings and emissions in transport.* Report, ODYSSEE-MURE project. https://www.odyssee-mure.eu/publications/archives/energy-efficiency-trends-policies-transport.pdf, accessed on October 12, 2023. [124]

Fang, Ximeng, Lorenz Goette, Bettina Rockenbach, Matthias Sutter, Verena Tiefenbeck, Samuel Schoeb, and Thorsten Staake. 2023. "Complementarities in behavioral interventions: Evidence from a field experiment on resource conservation." *Journal of Public Economics* 228: 105028. [124, 127, 134, 147]

Ferraro, Paul J, and Michael K Price. 2013. "Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment." *Review of Economics and Statistics* 95 (1): 64–73. [124, 127]

Gabaix, Xavier, and David Laibson. 2006. "Shrouded attributes, consumer myopia, and information suppression in competitive markets." *Quarterly Journal of Economics* 121 (2): 505–40. [125, 127, 139, 147]

German Environment Agency. 2019. *Emissionsbilanz erneuerbarer Energieträger: Bestimmung der vermiedenen Emissionen im Jahr 2018.* Table 62: Emission factors of heat supply from fossil fuels in private households, the tertiary sector and industry. Umweltbundesamt. https://www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/2019-11-07_cc-37-2019_emissionsbilanz-erneuerbarer-energien_2018.pdf, accessed on September 22, 2020. [164, 165]

German Federal Institute for Population Research. 2023a. *Altersstruktur der Bevölkerung.* Demografieportal. Bundesinstitut für Bevölkerungsforschung. https://www.demografie-portal.de/DE/Service/Impressum/Impressum.html, accessed on November 11, 2023. [129]

German Federal Institute for Population Research. 2023b. *Zahl der Privathaushalte und durchschnittliche Haushaltsgröße in Deutschland (1871-2021).* Bundesinstitut für Bevölkerungsforschung. https://www.bib.bund.de/DE/Fakten/Fakt/L49-Privathaushalte-Haushaltsgroesse-ab-1871.html, accessed on February 2, 2023. [129]

German Federal Ministry for Economic Affairs and Energy. 2020. *Energiedaten.* As of: May 2020. Table 7b: Endenergieverbrauch nach Anwendungsbereichen II. Bundesministerium für Wirtschaft und Energie. https://www.bmwk.de/Redaktion/DE/Artikel/Energie/energiedaten-gesamtausgabe.html, originally retrieved on September 22, 2020; the URL was last accessed on January 26, 2023 and contains a newer version of the document. [164, 165]

German Federal Statistical Office. 2021. *Laufende Wirtschaftsrechnungen: Einkommen, Einnahmen und Ausgaben privater Haushalte - 2020 - Fachserie 15 Reihe 1.* Published 15. Dezember 2021. Article number: 2150100207004. Statistisches Bundesamt (Destatis). [129]

German Federal Statistical Office. 2022. *Bevölkerung und Erwerbstätigkeit - 2020 (Endergebnisse) - Fachserie 1 Reihe 3.* Published 31. Januar 2022. Article number: 2010300207004. Statistisches Bundesamt (Destatis). [129, 130]

Gerster, Andreas, Mark A Andor, and Lorenz Goette. 2020. "Disaggregate Consumption Feedback and Energy Conservation." Working Paper, Working Paper Series 182. CRC TR 224 Discussion Paper Series University of Bonn and University of Mannheim. https://www.wiwi.uni-bonn.de/bgsepapers/boncrc/CRCTR224_2020_182.pdf. [127]

Goette, Lorenz, Hua-Jing Han, Zhi Hao Lim, et al. 2021. “The dynamics of goal setting: Evidence from a field experiment on resource conservation.” Working Paper, Working Paper Series 283. CRC TR 224 Discussion Paper Series University of Bonn and University of Mannheim. https://www.wiwi.uni-bonn.de/bgsepapers/boncrc/CRCTR224_2021_283.pdf. [124, 138]

Goette, Lorenz, Zhengyi Jiang, Jan Schmitz, and Renate Schubert. 2021. “The effects of upward and downward social comparisons on energy consumption behavior: Evidence from a field study on air-conditioning usage.” *Sustainability and Environmental Decision Making,* 409–40. [124, 127]

Goette, Lorenz, Ching Leong, and Neng Qian. 2019. “Motivating household water conservation: A field experiment in Singapore.” *PLOS ONE* 14 (3): e0211891. [127]

Grubb, Michael D. 2009. “Selling to overconfident consumers.” *American Economic Review* 99 (5): 1770–807. [127, 147]

Grubb, Michael D, and Matthew Osborne. 2015. “Cellular service demand: Biased beliefs, learning, and bill shock.” *American Economic Review* 105 (1): 234–71. [127]

Halpern, David. 2015. *Inside the nudge unit: How small changes can make a big difference.* Random House. [123]

Hansen, Pelle Guldborg, and Andreas Maaløe Jespersen. 2013. “Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy.” *European Journal of Risk Regulation* 4 (1): 3–28. [126]

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. “A unified welfare analysis of government policies.” *Quarterly Journal of Economics* 135 (3): 1209–318. [146]

Holland, Stephen P, Matthew J Kotchen, Erin T Mansur, and Andrew J Yates. 2022. “Why marginal CO2 emissions are not decreasing for US electricity: Estimates and implications for climate policy.” *Proceedings of the National Academy of Sciences* 119 (8): e2116632119. [164]

Huber, Julian, Kai Lohmann, Marc Schmidt, and Christof Weinhardt. 2021. “Carbon efficient smart charging using forecasts of marginal emission factors.” *Journal of Cleaner Production* 284: 124766. [164, 165]

Ito, Koichiro. 2014. “Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing.” *American Economic Review* 104 (2): 537–63. [125]

Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2018. “Moral suasion and economic incentives: Field experimental evidence from energy demand.” *American Economic Journal: Economic Policy* 10 (1): 240–67. [127]

Jachimowicz, Jon M, Shannon Duncan, Elke U Weber, and Eric J Johnson. 2019. “When and why defaults influence decisions: A meta-analysis of default effects.” *Behavioural Public Policy* 3 (2): 159–86. [123]

Jessoe, Katrina, and David Rapson. 2014. “Knowledge is (less) power: Experimental evidence from residential energy use.” *American Economic Review* 104 (4): 1417–38. [125, 127, 139, 147]

Johnson, Eric J, Suzanne B Shu, Benedict GC Dellaert, Craig Fox, Daniel G Goldstein, Gerald Häubl, Richard P Larrick, John W Payne, Ellen Peters, David Schkade, et al. 2012. “Beyond nudges: Tools of a choice architecture.” *Marketing Letters* 23: 487–504. [126]

Kahneman, Daniel. 2011. *Thinking, fast and slow.* Macmillan. [123]

Khanna, Tarun M, Giovanni Baiocchi, Max Callaghan, Felix Creutzig, Horia Guias, Neal R Haddaway, Lion Hirth, et al. 2021. “A multi-country meta-analysis on the role of behavioural change in reducing energy consumption and CO2 emissions in residential buildings.” *Nature Energy* 6 (9): 925–32. [127]

Kleven, Henrik Jacobsen, and Claus Thustrup Kreiner. 2006. “The marginal cost of public funds: Hours of work versus labor force participation.” *Journal of Public Economics* 90 (10-11): 1955–73. [145]

Lacetera, Nicola, Devin G Pope, and Justin R Sydnor. 2012. “Heuristic thinking and limited attention in the car market.” *American Economic Review* 102 (5): 2206–36. [127]

List, John A, Ian Muir, Devin Pope, and Gregory Sun. 2023. “Left-digit bias at lyft.” *Review of Economic Studies* 90 (6): 3186–237. [127]

Löschel, Andreas, Matthias Rodemeier, and Madeline Werthschulte. 2023. “Can self-set goals encourage resource conservation? Field experimental evidence from a smartphone app.” *European Economic Review* 160: 104612. [126]

Mertens, Stephanie, Mario Herberz, Ulf JJ Hahnel, and Tobias Brosch. 2022. “The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains.” *Proceedings of the National Academy of Sciences* 119 (1): e2107346118. [123]

Michel, Anette, Sophie Attali, and Eric Bush. 2016. *Energy efficiency of white goods in Europe: monitoring the market with sales data.* Final report. Study realised on behalf of ADEME by SOWATT and Bush Energie. https://storage.topten.eu/source/files/Market-Monitoring-2016-EN-Topten.eu.pdf, accessed on October 12, 2023. [124]

Mongin, Philippe, and Mikaël Cozic. 2018. "Rethinking nudge: Not one but three concepts." *Behavioural Public Policy* 2 (1): 107–24. [126]

Nisa, Claudia F, Jocelyn J Bélanger, Birga M Schumpe, and Daiane G Faller. 2019. "Meta-analysis of randomised controlled trials testing behavioural interventions to promote household action on climate change." *Nature Communications* 10 (1): 1–13. [123]

Obama, Barack. 2015. *Executive order – Using Behavioral Science insights to better serve the American people.* Media Release, White House, Washington, DC, September 15, 2015. [123]

OECD. 2017. *Behavioural insights and public policy: Lessons from around the world.* OECD Publishing, Paris. [123]

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. [144]

Rodemeier, Matthias. 2021. "Buy baits and consumer sophistication: Theory and field evidence from large-scale rebate promotions." Working Paper, Working Paper Series 124. CAWM Discussion Paper Series. https://www.econstor.eu/handle/10419/234136. [126]

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. [126]

Sharot, Tali, and Cass R Sunstein. 2020. "How people decide what they want to know." *Nature Human Behaviour* 4 (1): 14–19. [124]

Sheeran, Paschal, and Thomas L Webb. 2016. "The intention–behavior gap." *Social and Personality Psychology Compass* 10 (9): 503–18. [138]

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. [127, 147]

Tchibo. 2023. *LED shower head with water consumption indicator.* Tchibo GmbH online store. https://www.tchibo.de/products/402100280/led-duschkopf-mit-wasserverbrauchsanzeige, accessed on September 19, 2023. [147]

Thaler, Richard H., and Cass R. Sunstein. 2009. *Nudge: Improving decisions about health, wealth, and happiness.* Penguin. [123]

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. [124, 127, 135, 136]

Tiefenbeck, Verena, Anselma Wörner, Samuel Schöb, Elgar Fleisch, and Thorsten Staake. 2019. "Real-time feedback promotes energy conservation in the absence of volunteer selection bias and monetary incentives." *Nature Energy* 4 (1): 35–41. [124, 145]

Chapter 3

Tastes better than expected: Post-intervention effects of a vegetarian month in the student canteen *

Joint with Charlotte Klatt

3.1 Introduction

While meat consumption plays an important role in our diets, providing essential nutrients and energy, its environmental and health impacts are significant (Springmann et al., 2018; Bonnet et al., 2020).[1] Reducing meat consumption has thus been the target of a variety of interventions. Previous research has found that meat intake can be reduced by changing the arrangement of food counters and menus (e.g. Kurz, 2018; Garnett et al., 2020; Lohmann et al., 2024), installing carbon labels (e.g. Bilén, 2022; Lohmann et al., 2022; Ho and Page, 2023; Schulze-Tilling, 2023), changing prices (e.g. Garnett et al., 2021), and introducing green defaults (see Meier et al., 2022, for a meta-review).

However, interventions aimed at decreasing meat consumption are often short-lived.[2] Even if a canteen or supermarket were to persistently implement a policy, an individual would only be exposed to it for a limited time, until they graduate, change jobs, or move house. A crucial question is thus whether we can expect such interventions to have an impact post-intervention, and if so, why. One possible reason might be that an intervention changes beliefs. For example, Jalil, Tasoff, and Bustamante (2023) still see the effects of an information intervention to

* We are grateful to the Studierendenwerk Bonn for their support and cooperation. We are thankful to Astrid Dannenberg, Thomas Dohmen, Ximeng Fang, Hanna Fuhrmann-Riebel, and Lorenz Götte for their comments and suggestions. We gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 (Project B07) and Germany's Excellence Strategy – EXC 2126/1-390838866. The experimental design was pre-registered (AsPredicted #127656) and IRB approval was obtained from the GfeW (Certificate No. BRJCfLS6).

1. Poore and Nemecek (2018) estimate that meat and dairy provide only 18 % of calories consumed, while using 83% of global farmland and producing 60% of agriculture's greenhouse gas emissions. Further, individuals with high intakes of red and processed meat generally show modestly higher mortality rates in high-income Western countries (Godfray et al., 2018).

2. For example, a German supermarket charged the "true costs" of its food products, including environmental and social costs, but did so only for a week (tagesschau.de, 2023). Student canteens often operate with special weeks focusing on the environment and/or health and then return to normal operations.

decrease meat consumption after three years. Interventions might also persistently affect attitudes or the perception of social norms, as argued by Gravert and Shreedhar (2022). Both of these channels entail the intervention directly affecting future consumption. Additionally, interventions might impact long-run consumption behavior indirectly, mediated by the change in consumption behavior they induce in the short run. For example, an intervention's initial effect on consumption could lead to the build-up of habit stock, which in turn affects consumption more permanently (Stigler and Becker, 1977). Alternatively, the initial change in consumption behavior might enable individuals to learn about their preferences, and this might subsequently affect behavior.[3]

This paper studies the intervention and post-intervention effects of a one-month intervention to decrease meat consumption in the student canteen, and offers suggestive evidence on the channels driving post-intervention effects. The intervention affected only one of the University's three student canteens and consisted of the complete removal of all meat options from the menu. The canteen instead offered a higher variety of vegetarian and vegan options.[4] To study the intervention's effects on canteen guests, we combine individual-level canteen consumption data with survey data capturing changes in perceived social norms and canteen guests' perception of the canteen and their own consumption behavior. Our combined data allows us to evaluate whether the intervention led to a change in individuals' behavior during and in the two months following the intervention, and to provide suggestive evidence on the likely drivers of these post-intervention effects.

To examine the effect of the intervention on the relative share of meat meals sold, we first analyze the full canteen consumption data (over 270, 000 purchases made by over 4, 500 guests over six months). We use a difference-in-difference framework comparing which percentage of meals sold in the treatment canteen contained meat, both over time and relative to the control canteens. While the intervention by design led to a 100% decrease in the proportion of meat meals sold in the treatment canteen during the intervention month, it also led to a decrease of 7 to 12% in the two months following the intervention, relative to baseline.

To examine in how far this effect is attributable to a change in canteen guests' behavior, rather than merely a change in canteen frequenting patterns, we additionally perform an intent-to-treat (ITT) analysis at the guest level, classifying regular canteen guests as treatment or control based on their pre-intervention behavior. In our main specification, we include guest-level fixed effects to control for baseline meat consumption. We estimate that the intervention on average led to guests usually visiting the treatment canteen pre-intervention being less likely to choose a meat meal when they visit one of the canteens in the two months following the intervention period. Specifically, the share of meat meals they consume decreases by 4% relative to baseline. On average, the intervention did not significantly impact treated canteen guests' likelihood of frequenting the student canteens in general or the treated canteen in particular, neither during the intervention period nor post-intervention.

We next consider evidence for the relevance of different channels driving these post-intervention effects. One possible channel could be a change in social norms towards meat con-

3. For example, Charness, Chemaya, and Trujano-Ochoa (2023) shows this to be the case concerning risk preferences, with individuals making lottery choices differently after having experienced making risk choices.

4. This received attention from regional and national news (Die Welt, 2023; Kölner Stadt Anzeiger, 2023; t-online, 2023).

sumption, but we find little evidence for this: We track potential changes in perceived social norms by conducting surveys in the treatment and control canteens pre-and post-intervention. Around 400 canteen guests responded to both surveys, allowing us to perform a difference-in-difference analysis for changes in perceived social norms, while controlling for guests' characteristics with guest-level fixed effects. We find little to no evidence of the intervention changing the perceived descriptive norm towards meat consumption (elicited by asking respondents to guess which percentage of canteen guests chooses a vegetarian/vegan meal) or the perceived injunctive norm (elicited by asking respondents to rate the social appropriateness of different meat consumption behaviors and to then indicate what they believe to be the most common response among other respondents).[5]

Our suggestive evidence rather points towards learning and habit formation being the most relevant channels: To shed light on other possibly relevant mechanisms we ask respondents of the post-intervention survey to self-report the reasons for decreasing their meat consumption post-intervention (conditional on respondents indicating their behavior had changed post-intervention). The most frequently cited reason was learning about the taste of previously untried vegetarian options, followed by a perceived improvement in the vegetarian offerings in the cafeteria (although there was in fact no difference in offerings before and after the intervention), and making vegetarian eating more of a habit. Suggestive evidence from the post-intervention period further supports that having previously experienced a meal impacts meal choices: Canteen guests are generally more likely to choose a meal if they had already experienced it in a previous visit. A slight majority of canteen guests indicated approval for repeating such an intervention month annually (52% of the control and 58% of treatment guests).

We mainly contribute to two strands of literature. The first concerns the evaluation of interventions to decrease meat consumption in a student canteen context. Lohmann et al. (2022) and Schulze-Tilling (2023) conduct natural field experiments in the student canteen showing that carbon labels decrease guests' meat consumption, and Garnett et al. (2020) and Kurz (2018) show that the order of canteen food counters and menu ordering can influence choices. Notably, Kurz (2018) finds evidence of effects persisting in a 13-week post-intervention period. Garnett et al. (2021) further finds evidence of price changes in the canteen affecting meat consumption. This paper studies a more drastic type of intervention, which sharply decreased meat sales while not reducing the number of student canteen guests. Post-intervention effects are, in comparison, much more modest and comparable in magnitude with those found by Kurz (2018).

The second strand of literature we contribute to examines the possible drivers of such post-intervention effects. Evidence in the food consumption domain is scant in this regard. However, different possible drivers have been examined in the resource consumption domain. Byrne et al. (2021) provide experimental evidence that the post-intervention effects of an intervention to reduce shower length are driven by consumers forming a habit of paying attention to consumption. Goetz, Mayr, and Schubert (2022) explain persistent spillover effects of an intervention to save hot water using a theoretical framework in which households strive to be consistent with their environmental self-image. Castillo and Petrie (2023) explain the persistent effects of high-frequency information and monetary incentives on gas usage with households

5. Following Cialdini and Trost (1998), the "descriptive norm" an individual perceives refers to his impression of how others behave, while his perception of the "injunctive norm" refers to his impression of what others think one should behave like. Our procedure for identifying the perceived injunctive norm follows Krupka and Weber (2013).

in different treatment groups experimenting with different room temperatures, learning about differences in comfort, and adjusting behavior accordingly. In other domains such as blood donations (e.g. Bruhin et al., 2021) and gym attendance (Acland and Levy, 2015), persistence effects seem well-explicable with the Stigler and Becker (1977) habit-formation model. We contribute to this literature by combining our observations of post-intervention effects with survey data assessing the relevance of possible channels, providing the first evidence of possible drivers of post-intervention effects in the food consumption domain. Our findings suggest that even short-lived interventions can affect food consumption behavior in the longer run, by helping individuals learn about their food preferences.

The rest of this paper proceeds as follows. Section 3.2 describes the experiment setting and data, as well as the surveys we conducted pre- and post- intervention. Section 3.3 analyzes the effect of the intervention on canteen sales. Section 3.4 analyzes the effect of the intervention on guests' behavior in an intent-to-treat analysis. Section 3.5 provides suggestive evidence on the channels for post-intervention effects, and examines whether the intervention led to a change in the perception of social norms. Section 3.6 discusses the interventions' popularity among canteen guests. Finally, section 3.7 discusses our findings.

3.2 Experiment setting and data

3.2.1 The canteen intervention and canteen data

The intervention we study was implemented in one of the student canteens of the University of Bonn in May 2023. The student canteen named the intervention a "vegan-vegetarian month", during which all meat options were removed from the menu of one of the student canteens and replaced with vegan or vegetarian alternatives. In the following, we will refer to the intervention as the vegetarian month and the canteen in which it was implemented as the treatment canteen. The vegetarian month was initiated jointly by the operators of the canteen, student representatives, and other local student organizations involved in the organization of the student canteen.[6] The student canteens offer very cheap meals, with complete meals costing between €1.00 and €4.00 (prices as of 2023/2024). In fast food restaurants located in the surrounding area, meals are priced upward of €4.00.

We observe consumption decisions in all student canteens and cafes of the University of Bonn. Besides the treatment canteen, there are two other main student canteens in Bonn, which we use as control canteens in our analysis.[7] The first larger control canteen is located 1.7 km from the treatment canteen, and the second smaller control canteen is located 4.7 km from the treatment canteen and frequented much less than the other two canteens. All three student canteens usually offer one vegetarian main meal component and one main meal component containing meat. The options offered as main meal components differ daily, but meal planning is

6. Collaboration and exchange between these parties was fostered by the initiative "Nachhaltige Ernährung im Studienalltag" (Sustainable consumption in daily student life) funded by the consumer protection agency of the state of North-Rhine-Westphalia. We were involved in the planning of the vegetarian month to the extent that we made recommendations on how to adjust its design to allow for the cleanest scientific evaluation possible. This mainly involved making recommendations on the timing of the implementation.

7. We also have data from several University-run cafes throughout the city of Bonn. These cafes have much fewer sales and a different offer than the main canteens and are thus not included in the analysis. We do not see an effect of the vegetarian month on the sales of these cafes.

centralized across the three canteens such that usually the same main meal components are offered across canteens on a given day. During the intervention month in May 2023, the treatment canteen deviated from centralized meal planning: It eliminated the meat-containing main meal component from its menu and instead offered two vegetarian main meal components which it chose independently from the coordinated menu. After the vegetarian month, it once again adhered to the centralized menu. The vegetarian month was announced less than a week before it was implemented (See Instagram announcement in Figure 3.D.1). In addition to main meal components, all canteens offer side dishes, desserts, and a vegetarian stew which can be supplemented with a sausage.[8] Further, the larger control canteen sometimes offers pizza or pasta in addition, and student canteens might serve leftover main meal components from the previous day.

Our main analysis focuses on whether canteen guests purchase meat-containing or vegetarian main meal components, as these make up the bulk of lunch purchases in the canteens and were mainly re-designed during the vegetarian month. We focus on purchases made between February 1st (three months before the intervention month) and July 31st (two months after the intervention month). For each purchase, we observe which meal is purchased, the price paid, and the location, day, and time of the purchase. From February to July 2023, a total of 276,673 main meal components were purchased in the student canteens, with 69% of these made in the control and 31% of these made in the treatment canteen. 62% of all purchases were made with a personalized payment card, allowing us to track the consumption decisions of individual guests across time.

3.2.2 Survey design and data

To examine the channels through which the intervention might have persistently affected consumer behavior, we conducted a pre-intervention survey in the beginning of April and a post-intervention survey in mid-July 2023. 839 participants completely filled out the first survey, 902 the second, and 396 participated in both. Demographic characteristics of our survey respondents are shown in Table 3.1. Section 3.D.2 in the Appendix details how we recruited survey participants.

At the beginning of both surveys, participants provided their respective student canteen payment card identifiers. This allows us to link survey responses to canteen consumption decisions.[9] Survey participants were provided with information on this linkage and consented to the procedure.

Both surveys collected participants' demographic information (gender, age, and study program).[10] Further, participants were asked to estimate various figures concerning student life in Bonn. Specifically, we asked for estimates of the percentage of students spending a semester abroad, engaging in university politics, and eating meat or fish on a regular day at the university. We effectively asked respondents to provide an estimate for each item twice, by asking once in a positive framing and again in a negative framing (e.g. asking which percentage of students spend a semester abroad and which percentage of students does not spend a semester

8. This was replaced by a vegetarian sausage during the vegetarian month in the treatment canteen.
9. For individuals who did not participate in our survey, we can still track an individual's consumption decisions over time, but do not have the auxiliary information provided by survey participants.
10. We make use of this data in the suggestive heterogeneity analyses shown in Tables 3.A.7 and 3.A.6.

abroad).[11] As pre-registered, we are only interested in the questions on meat consumption in our main analysis and use these to elicit participants' perceived descriptive norms. We included the other questions to obfuscate our main interest, and to minimize possible experimenter demand effects following Dannenberg, Klatt, and Weingärtner (2024). To additionally elicit respondents' personal norm towards meat consumption and their perception of the injunctive norm, we follow the procedure developed by Krupka and Weber (2013). We first asked respondents as how socially appropriate they perceive different meat consumption behaviors and then asked them to guess what most other respondents answered to this question.[12] For obfuscation purposes, we also asked similar questions to elicit the perceived injunctive norm towards spending a semester abroad and engaging in university politics. Translated survey screens are shown in section 3.D.2.

While we generally took care to obfuscate the purpose of the surveys, we did include questions focusing on the student canteen at the end of the second survey. As we placed these at the end of the last survey, they are unlikely to have influenced the answers respondents provided previously. Specifically, we asked participants whether they believe that they opt for a vegetarian canteen meal more often after the vegetarian month, and if so why. Participants could agree or disagree with possible reasons we provided, and indicate further reasons in an open comment box. Finally, we asked all participants how much they would support different canteen policies: offering a vegetarian day a week, offering more vegetarian and vegan meals, or offering only vegetarian meals for a month.

Table 3.1. Comparison of the respondents of the two surveys

	Pre-intervention	Post-intervention	Both
% Female	63%	59%	64%
Age	22.5	22.6	22.7
% Full-time students	96%	93%	96%
% Treatment canteen	34%	36%	39%
N	835	902	392

Notes: The first column shows descriptive statistics for participants of the survey we conducted pre-intervention at the beginning of April, the second column shows statistics for participants of the survey we conducted post-intervention mid-July, and the third column shows statistics for individuals who responded to both of our surveys. Age is approximated, as respondents indicated their age on an interval. For the purpose of these descriptive statistics, we assume age to be equal to the midpoint of the indicated interval. The percentage of respondents frequenting the treatment canteen we report here is based on respondents' self-report of which canteen they visit most frequently. For a part of survey respondents (556 in total), we can deduce whether survey respondents mainly frequent the treatment canteen or the control canteens based on linked consumption data. We base our heterogeneity analyses in Table 3.A.6 and 3.A.7 on this classification, attributing 36% of individuals to the treatment canteen. Data-based and self-reported canteen classification are identical for 98% of these respondents.

3.3 Canteen-level analyses

In this section, we analyze the effect of the intervention on canteen-level sales. This analysis can be understood as a first step in assessing the effects of the vegetarian month and does not take a stance on the reasons why the intervention led to a change in sales. Section 3.4.2 will assess whether these changes in sales are indeed caused by a change in student canteen guests' consumption behavior or if they are merely attributable to changes in canteen frequenting patterns.

11. We follow this procedure for a higher reliability of participant responses, as pre-registered.
12. These guesses were not financially incentivized.

3.3.1 Descriptive statistics

Figure 3.1 shows the number of main meal components sold in the control and treatment canteens during the pre-intervention, intervention, and post-intervention periods, as well as the respective trends in revenue made with the sale of main meal components. While sales in the control canteens are constantly higher than in the treatment canteen, sale trends pre-intervention look quite parallel. During the first eight weeks of the data period, the University was on semester break, with classes resuming from week 12 to week 24. In weeks 25 and 26, the University was again on semester break.[13] During the intervention and post-intervention phase, sales in the canteens continue to follow similar trends, with no indication of the vegetarian month leading to a decrease in average sales. Figure 3.2 shows the same figures for the sale of meat main components. Meat sales and revenue made with meat sales drop to zero in the treatment canteen during the intervention month, while meat sales roughly remain at pre-intervention levels in the control canteens.

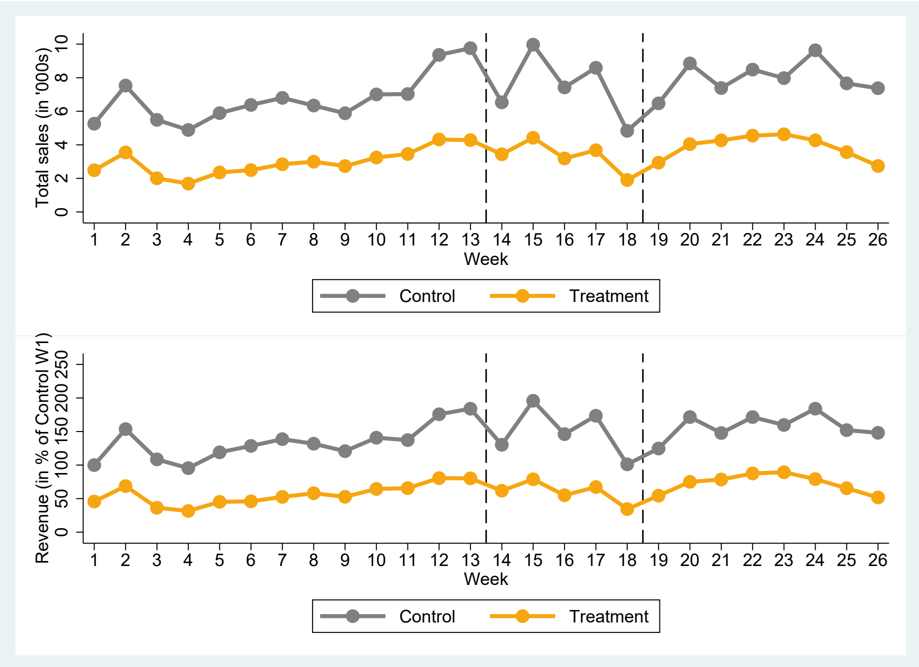


Figure 3.1. Main meal components sold February 23-July 23

Note: Upper figure shows the weekly number of main meal components sold in the treatment and control canteens, across the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Lower figure shows the respective weekly revenue made with the sale of main meal components, normalized relative to revenue made in the control canteens in week 1. During the first eight weeks of the data period, the University was on semester break, with classes resuming from week 12 to week 24, with a Pentacost break in week 18. In weeks 25 and 26, the University was again on semester break.

13. During the semester breaks, there are still exams taking place, and students submit homework essays, etc., so there are still activities on campus and sales accordingly do not drop to zero.

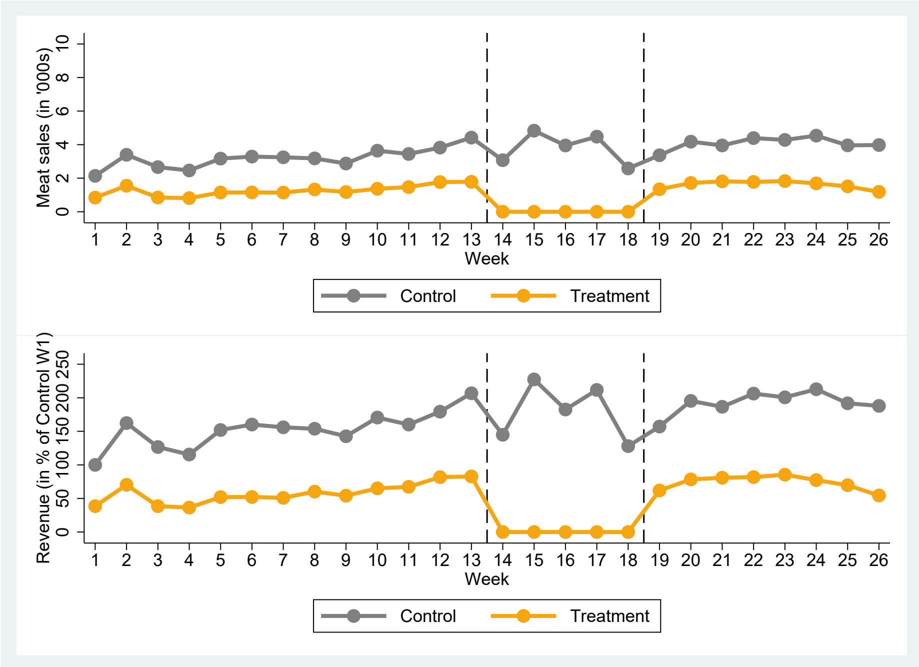


Figure 3.2. Meat main meal components sold February 23-July 23

Note: Upper figure shows the weekly number of meat main meal components sold in the treatment and control canteens, across the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Lower figure shows the respective weekly revenue made with the sale of meat main meal components, normalized relative to revenue in the control canteen in week 1. During the first eight weeks of the data period, the University was on semester break, with classes resuming from week 12 to week 24, with a Pentacost break in week 18. In weeks 25 and 26, the University was again on semester break.

3.3.2 Effect of the intervention on the proportion of meat sales

We first analyze whether the intervention led to a change in the proportion of meat meals purchased in the treatment canteen relative to the control canteen. The main variable of interest in this analysis is whether canteen guests choose the meat or the vegetarian main meal component, with the most basic difference-in-difference specification being:

$$\begin{aligned} Meat_{pt} = \alpha + \beta_1 InterPeriod_t + \beta_2 PostPeriod_t + \gamma Treat_p \\ + \delta_1 (Treat_p \times InterPeriod_t) + \delta_2 (Treat_p \times PostPeriod_t) + \epsilon_{it} \end{aligned} \tag{3.1}$$

The variable $Meat_{pt}$ is a binary outcome describing whether the main meal component purchased in purchase p on day t is meat-based, i.e. $Meat_{pt}$ equals 1 if a meat-based main meal component is purchased, and 0 if a vegetarian main meal component is purchased. $InterPeriod_t$ is an indicator for whether this purchase occurred during the intervention period (May 2023), and $PostPeriod_t$ is an indicator for whether this purchase occurred in the nine weeks following the intervention period (June/July 2023). $Treat_p$ is an indicator of whether the purchase is made in the treatment canteen. $(Treat_p \times InterPeriod_t)$ identifies differential changes in purchasing behavior during the vegetarian month in the treatment canteen. $(Treat_p \times PostPeriod_t)$ identifies differential changes in purchasing behavior after the vegetarian month in the treatment canteen.

Table 3.2 shows regression results. Col. (1) follows Equation 3.1, while Col. (2) exchanges the time indicator dummy variables for daily controls. Since meal planning is centralized across canteens outside of the intervention period, including daily fixed effects controls more precisely for changes in the attractiveness of the main meal components. Col. (3) additionally includes canteen-level controls for additional options on offer. Col. (4) additionally includes canteen-level sales. A full table including coefficients estimated on control variables is shown in Table 3.A.1 in the Appendix. Figure 3.A.1 shows an event plot.

Spec. (1)-(4) estimate that the vegetarian month led to a decrease of 45 to 46 percentage points in the proportion of meat main component purchased in the treatment canteen. This roughly corresponds to 100% of the proportion of meat main component sales at baseline and is not a particularly surprising result, since the intervention by design eliminated all meat sales. Differing point estimates across Col. (1)-(4) are merely attributable to the specifications estimating slightly differing counterfactual meat sales. The coefficient “Treat x PostPeriod” examines whether the vegetarian month led to a change in the proportion of meat meals sold post-intervention. Across Col. (1)-(4), we estimate that the proportion of meat main components sold decreased by 3 to 5 percentage points, or 7.5% to 12% of the baseline level (42.7 percentage points). We thus find that the intervention led to a significantly lower proportion of meat main components sold in the treatment canteen in the two months following the intervention. This drop in relative meat sales might be attributable to canteen guests changing their meat consumption behavior, but might also be due to possible changes in guests’ frequenting patterns (e.g. changing preferences for one or the other canteen, or for frequenting the canteen in general). The analysis in section 3.4 is better equipped to isolate a change in meat consumption behavior.

Table 3.2. Canteen-level estimates of effect on meat sales

	Likelihood of consuming meat (in pp.)			
	Base	Date FE	+Controls	+Sales
Treat x Inter period	-45.66***	-45.81***	-45.26***	-45.06***
	(0.40)	(0.40)	(0.46)	(0.46)
Treat x Post period	-4.84***	-5.10***	-3.78***	-3.18***
	(0.46)	(0.46)	(0.49)	(0.50)
Treat	-4.98***	-4.63***	-2.85***	-6.11***
	(0.30)	(0.30)	(0.54)	(0.74)
Inter period	2.99***			
	(0.31)			
Post period	3.73***			
	(0.26)			
Constant	47.65***	34.49***	33.71***	42.13***
	(0.17)	(0.94)	(1.07)	(1.68)
Date fixed effects	No	Yes	Yes	Yes
Guest fixed effects	No	No	No	No
Control for other offer	No	No	Yes	Yes
Guests control	8,353	8,353	8,353	8,353
Guests treated	3,575	3,575	3,575	3,575
Observations	276,673	276,673	276,673	276,673

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 when visiting the canteen. Col. (1) follows Equation 3.1. The constant describes the proportion of meat main meal components sold in the control canteens pre-intervention. Specifications (2) and (3) include daily date-fixed effects to control for the daily changing offer of main meal components, which is common across canteens pre- and post-intervention. The "PostPeriod" and "Inter period" indicators are thus dropped due to collinearity. Specification (3) includes controls for changes in other elements of the canteens' offers. These include controls for additional offers of the control canteens (special meals, pasta, pizza, other additional meals), as well as a control for whether a second vegetarian main is on offer. This is sometimes the case in all canteens if there are left-overs from the previous day. Specification (4) additionally includes canteen-level sales as a control. Guest numbers are lower-bound estimates since they only include guests paying with an individual payment card. The full table including the coefficients estimated on control variables is shown in Table 3.A.1. Standard errors are robust.

3.4 Guest-level analyses

While section 3.3 identifies the effect of the vegetarian month on canteen-level sales, it is not clear from this analysis alone in how far effects are attributable to a change in meat consumption behavior as opposed to a mere change in the composition of guests frequenting the canteen. Canteen-level effects are likely in part driven by changes in the composition of canteen guests. This section will thus provide a guest-level analysis, examining whether usual treatment canteen guests on average purchase less meat meals in the canteens during and after the intervention. For this purpose, we construct an intent-to-treat (ITT) sample of canteen guests to whom we can associate a treatment group based on their pre-intervention purchasing behavior.

3.4.1 Construction of the intent-to-treat sample

62% of the purchases made in the canteens are paid with a personalized payment card. For the guest-level analysis, we restrict the sample to these purchases.[14] Further, we drop instances of a canteen guest purchasing more than one main meal component in one visit.[15] For the main analysis, we further restrict the sample to purchases made by canteen guests who visited a student canteen at least five times in the three months preceding the vegetarian month (80% of the remaining sample) and who either spent at least 80% of these visits at the treatment canteen or at least 80% of these visits at one of the control canteens (92% of the remaining sample). These restrictions allow us to categorize individuals as treatment or control based on intention to treat. Specifically, the treatment group in the IIT analysis consists of canteen guests who primarily visited the treatment canteen and the control group consists of canteen guests who primarily visited the control canteens pre-intervention. Guests' classification as treated or control is thus based entirely on pre-intervention data and stays constant throughout the data period. The resulting sample consists of 117,642 purchases made by a total of 4,513 guests. 84% of these purchases are made by university students, 16% by university employees, and less than 0.5% by guests not affiliated with the university.

3.4.2 Intent-to-treat effect of the intervention on guest behavior

Using the sample constructed in section 3.4.1, we analyze changes in the consumption behavior of guests who mainly frequent the treatment canteen pre-intervention compared with guests who mainly frequent the control canteen pre-intervention. The main variable of interest is whether canteen guests choose the meat or the vegetarian main meal component when they visit the canteen, with the most basic difference-in-difference specification being:

$$\begin{aligned} Meat_{it} = \alpha + \beta_1 InterPeriod_t + \beta_2 PostPeriod_t + \gamma Treat_i + \\ + \delta_1 (Treat_i \times InterPeriod_t) + \delta_2 (Treat_i \times PostPeriod_t) + \epsilon_{it} \end{aligned} \tag{3.2}$$

The variable $Meat_{it}$ is a binary outcome describing whether the main meal component purchased by individual i on day t is meat-based, i.e. $Meat_{it}$ equals 1 if a meat-based main meal component is purchased, and 0 if a vegetarian main meal component is purchased. $InterPeriod_t$ is an indicator for whether this purchase occurred during the intervention period (May 2023), and $PostPeriod_t$ is an indicator for whether this purchase occurred in the nine weeks following the intervention period (June/July 2023). $Treat_i$ is an indicator for whether the purchase is made by an individual classified as treated based on pre-intervention purchase patterns.[16] ($Treat_i \times InterPeriod_t$) identifies intent-to-treat effects of the vegetarian month. ($Treat_i \times PostPeriod_t$) identifies intent-to-treat post-intervention effects of the vegetarian month. Standard errors are clustered at the individual level.

14. Of the remaining purchases, 85% are made with a debit or credit card, and the remaining amount with cash. For an analysis at the canteen level including all sales data, please see section 3.3.

15. Canteen guests might purchase multiple main meal components because they are very hungry or because they are inviting a friend. Since we cannot distinguish between the two, we drop all instances of multiple main meal components being purchased. These are 5% of the remaining purchases.

16. As described in section 3.4.1, we classify canteen guests as treated or control guests based on consumption behavior in the three months preceding the vegetarian month. The $Treat_i$ indicator is thus independent of whether the specific purchase $Meat_{it}$ occurred in the treatment or control canteen.

Regression results are shown in Table 3.3. Col. (1) performs the regression described in Equation 3.2. For canteen guests in the treated group, the likelihood of choosing a meat main meal component when visiting one of the canteens is decreased by 42 percentage points during the intervention period. After the intervention period, the likelihood of choosing a meat main meal component is decreased by 4 percentage points, relative to baseline. One possible factor explaining that estimates are slightly smaller than in our canteen-level analysis might be guests choosing different student canteens during the intervention and post-intervention period than pre-intervention.[17] Col. (2) exchanges the “InterPeriod” and “PostPeriod” indicators for daily fixed effects that capture changes in the attractiveness of the daily-changing meals on offer in the student canteens pre- and post-intervention. This does not change the estimated coefficients.

While the estimates in Col. (1) and (2) are by design not impacted by a possible increase in regular treatment canteen guests visiting the control canteens and vice versa, they might still be driven by changes in guests’ decision to visit a student canteen in general. Specifically, the vegetarian month might have led to guests with a taste for meat avoiding the student canteens and guests with a taste for vegetarian options increasingly frequenting the student canteens. Col. (3) thus additionally includes guest fixed effects. In this manner, we can control for individual canteen guests’ taste for meat. This is our preferred specification to assess the impact of the intervention on guests’ canteen consumption behavior. We find that the intervention on average led to a decrease of 35 percentage points in the proportion of meat meals purchased by the treated group, i.e. the likelihood of an average treated guest to consume meat when in the canteen is reduced by 35 percentage points during the intervention. Post-intervention, we estimate that the proportion of meat meals purchased decreased by 2 percentage points, i.e. the likelihood of an average treated guest to consume meat when in the canteen is reduced by 2 percentage points in the two months following the intervention period. This translates into a 4% decrease in meat consumption relative to baseline meat consumption in the treated group (42%).

Figure 3.3 estimates the time trend based on equation 3.2, exchanging the “InterPeriod” and “PostPeriod” indicators for weekly indicators. The coefficients estimated for the weeks preceding the intervention period move around 0, supporting the validity of the parallel trend assumption for our difference-in-difference analysis. Post-intervention, the coefficients move around 0 in the two weeks immediately following the intervention. We estimate negative coefficients for the following six weeks. [18]

To further investigate a possible effect of the intervention on canteen frequenting patterns, Table 3.4 examines the effect of the intervention without conditioning on guests’ decision to visit the student canteen. For this purpose, we expand the data set to resemble a balanced panel, inserting zeros for days on which a regular canteen guest did not frequent one of the canteens.[19] Col. (1) examines the effect of the intervention on guests’ likelihood of visiting one of the student

17. Section 3.B provides further statistics on the intervention influencing guests’ decision to visit the treatment or one of the control canteens. Especially for guests with high baseline meat consumption, the proportion of visits to the control canteen relative to the treatment canteen seems to have increased during the intervention period. The estimates identified in the intent-to-treat analysis are not affected by such changes in canteen frequenting patterns.

18. Week 27 consists of just one day of data, since this is the last day in our data period.

19. Note that the analysis shown in Table 3.3 uses only observations of guests visiting one of the student canteens on a given day and then examines guests’ choice of main meal component. This decision is thus made conditional on guests’ previous decision to visit the canteen. The analysis in Table 3.4 artificially expands the data set to include a zero observation for each day a regular student canteen guest could have visited one of the canteens. In

Table 3.3. ITT estimates of effect on meat consumption

	Likelihood of consuming meat (in pp)		
	Base	Date FE	Date+Guest FE
Treat x Inter period	-0.42***	-0.42***	-0.35***
	(0.01)	(0.01)	(0.01)
Treat x Post period	-0.04***	-0.04***	-0.02**
	(0.01)	(0.01)	(0.01)
Treat	-0.03*	-0.02	
	(0.02)	(0.02)	
Inter period	0.03***		
	(0.01)		
Post period	0.04***		
	(0.01)		
Constant	0.45***	0.32***	0.37***
	(0.01)	(0.01)	(0.01)
Date fixed effects	No	Yes	Yes
Guest fixed effects	No	No	Yes
Guests control	3,371	3,371	3,371
Guests treated	1,142	1,142	1,142
Observations	117,642	117,642	117,642

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. Estimates show the change in the likelihood of consuming meat in percentage points. Col. (1) corresponds to Equation 3.2. The Constant term describes the proportion of meat meals sold to the control group pre-intervention. Specifications (2) and (3) include date-fixed effects to control for the daily changing offer of main meal components, which is common across canteens pre- and post-intervention. The "PostPeriod" and "Inter period" indicators are thus dropped due to collinearity. Specification (3) includes individual guest-level effects. Standard errors are clustered at the individual level.

canteens on a given day, performing a similar analysis as shown in Col. (3) of Table 3.3, but exchanging the dependent variable with a binary variable equalling 1 if the guest visited one of the canteens on a given day, and 0 if not. The baseline likelihood with which treatment guests visit one of the student canteens at lunchtime pre-intervention is 28 pp. (see Col. (1) in Table 3.C.1). Treated canteen guests' likelihood of visiting one of the canteens does not seem to be affected by the intervention. Table 3.C.2 repeats this analysis focusing specifically on guests' likelihood of visiting their "home" canteen rather than visiting any student canteen. Again, we do not find a significant effect of the intervention.

contrast to the data set used in Table 3.3, we here classify guests as treatment or control canteen guests based only on weeks 1-11 of our pre-intervention period, and then drop these from the analysis and instead use weeks 12-13 as a shorter pre-intervention phase. The reason for this is that in this analysis guests' decision of whether or not to visit the student canteen is part of the analysis outcome, but at the same time also part of the criteria defining the ITT samples. To avoid introducing endogeneity, we use part of the pre-intervention phase (weeks 1-11) to assign ITT groups, and the remaining part (weeks 12-13), to include pre-intervention behavior in the analysis. Tables 3.C.3 and 3.C.1 repeat the analyses in Table 3.4 with all specifications shown in Table 3.3 and find similar results across specifications.

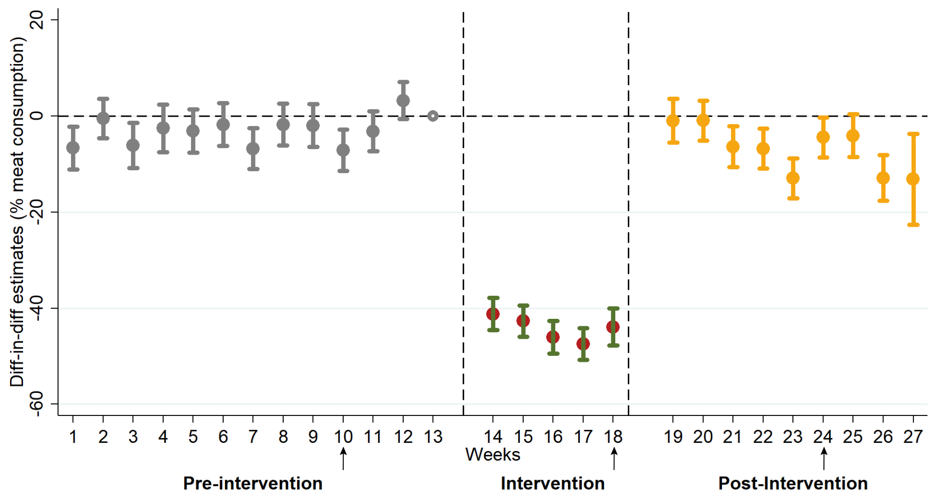


Figure 3.3. ITT event plot

Note: Figure shows coefficients estimated in a regression analysis following equation 3.2, but including weekly interaction terms and time controls. We additionally control for day of the week. Coefficients show the estimated change in the likelihood of consuming the meat main meal component, in percentage points. Our data for week 27 includes only one day, since it was the last day of July and our sample period. The first arrow indicates Easter, the second Pentecost, and the third the beginning of the semester break. Bars indicate 95% confidence intervals.

Col. (2) examines guests' likelihood of visiting one of the canteens at lunchtime and then consuming the meat main component. The baseline likelihood for treatment guests to do so is 12 pp. (see Col. (1) in Table 3.C.3), which is decreased by 10.3 pp.(86%) during the intervention phase and 1.8 pp. (17%) post-intervention. Col. (3) examines guests' likelihood of visiting one of the canteens at lunchtime and then consuming the vegetarian main meal component. The baseline likelihood for treatment guests is 17 pp. (see Col. (4) in Table 3.C.3), which is increased by 9.8 pp (58%) during the intervention and 1.2 pp. (7%) post-intervention. Figures 3.C.1, 3.C.2, and 3.C.3 show event plots for each of the three analyses. Additionally, time trends on regular canteen guests' decision to visit one of the student canteens at lunchtime are shown for all canteen guests in Figure 3.B.3, and separated by previous meat consumption in Figures 3.B.6, 3.B.9, and 3.B.12. Table 3.C.4 additionally replicates Table 3.C.1 conditioning on previous meat consumption. While there is on average no effect of the intervention on visits, there seem to be considerable heterogeneities, with canteen visits increasing during and post-intervention for guests with low previous meat consumption and decreasing for guests with high previous meat consumption. This suggests that changes in guest frequenting patterns are likely a relevant factor in explaining the change in sales identified in section 3.3, and supports the importance of including individual fixed effects when assessing the effect of the intervention on guest consumption behavior.

Table 3.4. ITT estimates without conditioning on the decision to visit one of the canteens

	Visit(in pp)	Visit+Meat(in pp)	Visit+Veg(in pp)
	Date+Guest FE	Date+Guest FE	Date+Guest FE
Treat x Inter period	-0.44	-10.26***	9.82***
	(0.74)	(0.63)	(0.67)
Treat x Post period	-0.57	-1.82***	1.24**
	(0.76)	(0.49)	(0.57)
Constant	27.72***	11.00***	16.71***
	(0.68)	(0.46)	(0.56)
Date fixed effects	Yes	Yes	Yes
Guest fixed effects	Yes	Yes	Yes
Guests control	2,722	2,722	2,722
Guests treated	922	922	922
Observations	262,368	262,368	262,368

Standard errors in parentheses

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

Notes: For the purpose of this analysis, we expand the main data set to resemble a balanced panel, i.e. inserting zeros for day and student canteen guest combinations on which a regular canteen guest did not visit the canteen. The construction of the data set is explained more in detail in the main text. The dependent variable in Col. (1) is a 0/1 indicator for visiting one of the student canteens and consuming any main meal component. Col. (2) is a 0/1 indicator for visiting one of the student canteens and then consuming the meat main component. Col. (3) is a 0/1 indicator for visiting one of the student canteens and then consuming the vegetarian main component. Regression specification is as in Col. (3) of Table 3.3. Standard errors are clustered at the individual level. The number of guests is a bit lower than in the main analysis in Table 3.2 because treatment and control group assignment criteria are applied using only data from weeks 1 to 11. Weeks 1 to 11 are used exclusively for this purpose and then dropped from the analysis, as explained in the main text. Figures 3.C.1, 3.C.2, and 3.C.3 show event plots corresponding to each of the three columns. Tables 3.C.1 and 3.C.3 repeat the analysis with different specifications, following all specifications of Table 3.3. Table 3.C.2 looks specifically at guests' likelihood of visiting their "home" canteen rather than one of the canteens in general.

3.4.3 Heterogeneity analysis of post-intervention effects

We now examine how estimated treatment effects differ depending on canteen guests' consumption behavior. Table 3.5 splits the sample by the frequency with which guests visited one of the student canteens during the intervention period, and repeats the main specification shown in Col. (3) of Table 3.3 on the restricted samples. Col. (1) includes the full sample, while Col. (2) restricts the sample to control guests, and treatment guests who visited the treatment canteen at least once during the intervention period. Col. (3) restricts the sample to control guests, and treatment guests for whom we did not register a visit to the treatment canteen during the intervention period. Note that these guests might have still visited the canteen but used a different payment method than their individual payment cards. Col. (4) - Col. (6) perform further sample splits depending on the number of visits during the intervention period. Treatment effects are more pronounced if guests visit the treatment canteen at least once. However, there is limited evidence of a clear increasing relationship between number of visits and effect sizes. Specifically, we estimate the largest treatment effects for guests visiting the canteen between 3 and 6 times, and estimate smaller effects for guests visiting more frequently. One reason for this could be that guests who come to the student canteen often during the intervention period are also more likely to have primarily consumed vegetarian meals pre-intervention, as shown in section 3.B.

Table 3.5. ITT estimates by visits in intervention month

	All	At least 1 visit?		Number visits		
		yes	no	1-2	3-6	over 6
Treat x Inter period	-34.84***	-35.05***		-43.79***	-39.78***	-31.42***
	(1.38)	(1.39)		(2.61)	(1.96)	(1.98)
Treat x Post period	-1.62**	-1.59**	-2.88	-0.47	-3.25***	-0.68
	(0.68)	(0.70)	(2.48)	(1.50)	(1.20)	(0.90)
Constant	36.93***	37.14***	39.43***	39.98***	39.12***	37.25***
	(0.92)	(0.93)	(1.05)	(1.04)	(1.02)	(0.99)
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Guest fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Guests control	3,371	3,371	3,371	3,371	3,371	3,371
Guests treated	1,142	915	227	232	364	319
Observations	117,642	114,967	90,524	91,553	96,518	102,594

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 when visiting the canteen. Col. (1) corresponds to Col. (3) in Table 3.3 and includes the full sample. Col. (2) restricts the sample to control guests and treatment guests who made at least one purchase with their personalized card during the intervention period. Col. (3) includes only control guests and those treatment guests for whom we did not register such a purchase – Note, however, that they might have still visited the canteens during the time frame, but used a different payment method than their personalized card. Col. (4) - Col. (6) restrict the sample of treatment guests by the number of visits registered during the intervention period. Numeric thresholds are chosen such that each category corresponds to roughly one third of regular student canteen guests. Standard errors are clustered at the individual level. Table 3.C.5 repeats the analyses without conditioning on canteen guests visiting the canteen, i.e. examining guests decision to visit the canteen and then consume meat.

For canteen guests already consuming close to no meat pre-intervention, the vegetarian month can — mechanically — not lead to a decrease in meat meals consumed post-intervention. Patterns are similar when examining student canteen meat consumption without conditioning on an individual visiting the student canteen (i.e. repeating the analysis from Col. (2) in Table 3.4 on the respective sub-samples), as examined in Table 3.C.5.

Table 3.6 analyzes treatment effects splitting the sample by meat consumption pre-intervention. The coefficients estimated for "Treat x Post period" suggest that the post-intervention effects of the vegetarian month are strongest for canteen guests with high previous meat consumption. Patterns are similar when examining the respective sub-samples without conditioning on guests' decision to visit one of the student canteens, as examined in Table 3.C.6.

Tables 3.A.6 and 3.A.7 further examine heterogeneity by demographic characteristics. For this analysis, we restrict the sample to those canteen guests who took part in our surveys and provided demographic information. Correspondingly, the sample size is decreased to a total of 570 canteen guests, making results more of a suggestive nature. Results indicate that younger guests (21 and younger), female guests, and those who study Law (rather than Culture, Economics, or Social Studies) show larger post-intervention effects. Table 3.A.8 examines effects separately for university employees (around 16% of the ITT sample). The coefficient on post-intervention effect sizes for university employees is negative, but insignificant.

Table 3.6. ITT estimates by pre-intervention meat consumption

	All	By percentage of meat meals pre-intervention		
		0-10%	10%-68%	over 68%
Treat x Inter period	-34.84***	-2.68***	-42.67***	-76.73***
	(1.38)	(0.35)	(1.44)	(1.70)
Treat x Post period	-1.62**	-1.02	-2.03	-2.22**
	(0.68)	(0.80)	(1.44)	(1.11)
Constant	36.93***	0.01	26.80***	82.22***
	(0.92)	(0.32)	(2.07)	(2.07)
Date fixed effects	Yes	Yes	Yes	Yes
Guest fixed effects	Yes	Yes	Yes	Yes
Guests control	3,371	1,117	1,122	1,132
Guests treated	1,142	400	375	367
Observations	117,642	38,644	38,731	40,267

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 when visiting the canteen. Col. (1) corresponds to Col. (3) in Table 3.3 and includes the full sample. Col. (2) - (4) restrict the sample based on canteen guests' purchasing behavior pre-intervention, as registered by their personalized payment cards. Each column corresponds to around one-third of canteen guests. Col. (2) restricts the sample to guests who consumed meat in 0% to 10% of their meals pre-intervention, Col. (3) to guests who consumed meat in 10% to 68% of their meals pre-intervention, and Col. (4) to guests who consumed meat in over 68% of their meals pre-intervention. Percentage thresholds are chosen such that each category corresponds to roughly one third of regular student canteen guests. Standard errors are clustered at the individual level. Table 3.C.6 repeats the analyses without conditioning on canteen guests visiting the canteen, i.e. examining guests' decision to visit the canteen and then consume meat.

3.5 Channels for post-intervention effects

Why did the vegetarian month lead to a post-intervention change in consumption behavior? Prominent reasons why effects might last could be (1) habit formation (as modeled by Stigler and Becker (1977)), (2) learning about one's preferences for vegetarian options (similar to individuals learning about their risk preferences in Charness, Chemaya, and Trujano-Ochoa (2023)), or (3) a change in perceived social norms with respect to meat consumption (as suggested by Gravert and Shreedhar (2022)[20]). This section will first present survey evidence on the possible relevance of each of these channels, and then present evidence from additional analyses of canteen data.

3.5.1 Survey evidence

Guests' self-reported motives

As a first step, we examine the reasons that the respondents of our post-intervention survey self-report as motives to change their behavior after the vegetarian month. Of the respondents

20. Nyborg et al. (2016) argue that policy can help shift social norms towards new self-sustaining social norms.

of the post-intervention survey, 325 (36%) report mainly going to the treatment canteen[21], of which 287 report having visited the treatment canteen at least once during the intervention period. 54 of the respondents (19%) report that they believe they are choosing the vegetarian option more often after the vegetarian month than before.[22] Table 3.7 shows the reasons which the respondents of the post-intervention survey selected for this being the case. 67% believe the change is driven by learning about the taste of different vegetarian options, 39% believe it is driven by an alleged improvement in the vegetarian and vegan offer of the canteen, and 31% believe it is due to a change in habits.

One might also think of the following reasoning to explain post-intervention effects: The vegetarian month might lead to students obtaining their desired level of meat consumption by consuming meat outside of the student canteen, e.g. always at dinner time or by frequenting fast food restaurants instead. This then might become a habit that persists even after the vegetarian month, e.g. persistently eating vegetarian at lunchtime and eating meat for dinner. As 11% of survey respondents agree to this item, we believe that the above explanation might play a smaller role in driving post-intervention effects. [23]

A further channel one might think of to explain treatment effects is a possible change in guests' perception of the social norm towards meat consumption. We asked respondents whether a change in students' general consumption behavior or their friend group's behavior drives their change in consumption. 9% report this to be the case. Section 3.5.1 investigates this possible channel more systematically.

The patterns reported here are similar when we designate guests as treatment or control guests based not on their self-reported group, but on their consumption data as described in section 3.4.1, and also when restricting the sample only to guests for whom we see a reduction in meat consumption in the consumption data, as shown in Table 3.A.5. They are also similar for usual control canteen guests who frequented the treatment canteen at least once during the intervention period.[24]

21. Participants' indication of their main canteen seem to be quite trustworthy: For those respondents survey respondents for which we have a self-reported main canteen and can additionally deduce canteen frequenting behavior based on the consumption data (556 participants), there is a 98% correlation between the two values.

22. Of the remaining 287 respondents, 74 (26%) report that they do not believe that they are choosing the vegetarian option more often after the vegetarian month than before, and 159 (55%) report this cannot be the case for them because they already only consumed vegetarian meals in the canteen pre-intervention. For those participants for whom we can also deduce from the consumption data whether they visited the treatment canteen during the intervention period, self-report and data match in 79% of cases. In almost all of the remaining cases, the respondent reports having visited the intervention canteen but we do not observe a payment with the respondent's individual payment card in the data. This would be explicable with the respondent using a different payment method in this visit. Self-report and consumption data diverge more widely concerning the perceived reduction in meat consumption. The patterns reported here are similar when we designate guests as treatment or control guests based not on their self-reported group, but on their consumption data as described in section 3.4.1, and also when restricting the sample only to guests for whom we see a reduction in meat consumption in the consumption data, as shown in Table 3.A.5.

23. The item reads: "I started consuming meal rather outside of the canteen during the vegetarian month and stuck to that after the month was over." Unfortunately, we do not have data on students' consumption outside of the institutions run by the university, and correspondingly, we cannot directly assess whether meat consumption increased in these settings. However, we do have data for the university-run cafes surrounding the student canteens. There, we do not see the vegetarian month causing a change in consumption behavior.

24. For those mainly frequenting the control canteen, 37% report visiting the treatment canteen at least once during the intervention period, and 24% of these report that the vegetarian month made them consume more veg-

Table 3.7. Guests' self-reported motives for decreasing meat consumption

Motive	Count	Percentage
Norm1	5	9.26%
Norm2	5	9.26%
Taste	36	66.67%
Habit	17	31.48%
Offer	21	38.89%
Spillover	6	11.11%
None of above	6	9.80%
Total respondents	54	100%

Notes: Table shows reasons cited by guests who report mainly going to the treatment canteen and that they consume the vegetarian option more frequently after the vegetarian month. Multiple options were selectable. The statements read: I now consume the vegetarian option more frequently, because ... (1) more students eat vegetarian/vegan meals since the month. (2) my friends eat more vegetarian/vegan meals since the month. (3) I got to know vegetarian vegan meals in the month which were new for me and which I find tasty. (4) It's becoming a habit for me to eat vegan/vegetarian. (5) The vegetarian/vegan offer has improved since the month. (6) I started consuming meat rather outside of the canteen during the month and stuck to that after the month was over.

Changes in perceived social norms

We additionally evaluate possible changes in perceived social norms more systematically by comparing the norm perceptions we elicited in our pre-intervention and post-intervention surveys. We elicited norms in both the treatment and control canteens, and both before and after the intervention, so our data allows for a difference-in-difference analysis of possible changes in the perceived social norm. To avoid that our results are confounded by changes in the composition of survey respondents between the first and the second survey, our main analysis is restricted to survey respondents who filled out both surveys. Further, we include individual fixed effects to control for respondents' norm perception at baseline.

Table 3.8 investigates possible changes in the perceived descriptive norm. Col. (1) shows possible changes in respondents' guess for the percentage of canteen guests NOT eating meat in the canteen, while Col. (2) investigates possible changes in respondents' guess for the percentage of canteen guests eating meat. We elicited both items to increase the reliability of our estimates, as detailed in section 3.2. Section 3.D.2 shows translated screenshots of the elicitation. We find no evidence for a sizeable overall change in the perception of the descriptive norm over time, nor a sizeable differential effect for regular guests of the treatment canteen.

Table 3.9 examines possible changes in the perceived injunctive norm. We follow the procedure developed by Krupka and Weber (2013) to identify changes in the perceived injunctive norms. For Col. (1) - (3), respondents were asked as how socially appropriate they personally perceive the consumption behavior of a student who consumes a meat-containing lunch on one out of five typical days (Col. 1), or who does so on three out of five days (Col. 2), or five out of five days (Col. 3). Col. (4) - (6) then identify the perceived injunctive norm by asking respondents to guess what most other respondents answered to the previous questions.

etarian after than before the vegetarian month. Patterns of reported reasons are similar as in the treatment group, and are shown in detail in Table 3.A.4.

We find no evidence for a differential change in perceived injunctive norms for respondents frequenting the treatment canteen, i.e. there is no change in the personal norm or the perception of the injunctive norm attributable to the experience of the vegetarian month.

Overall, the analysis, together with the self-reported motives reported above, suggests that a change in perceived social norms does not seem to be a major driver of the treatment effects identified for the treatment canteen.

Our elicitation of the personal norm towards meat consumption shown in Col. (1) - (3) of Table 3.9 might also be used as an indication of the relevance of a change in beliefs towards the negative consequences of meat consumption driving treatment effects: It could be argued that the vegetarian month may have led to increased discussion about animal welfare and the environmental impact of meat consumption, which might have changed guests' beliefs towards meat consumption, and in turn influenced treatment effects. Such a change in beliefs would have arguably affected guests' perceived personal norms. However, Table 3.9 does not show any evidence of this being the case, as we find no differential change in the personal norm for guests visiting the treatment canteen. [25]

25. At the same time, there seems to be an overall trend towards a stricter personal norm towards meat consumption in both the treatment and the control canteens, and it is of course possible that the vegetarian month affected norms across canteens and contributed to this trend. Any change in consumption behavior occurring due to such a common trend is not causally identified in our difference-in-difference analysis, and would lead to our analysis underestimating treatment effects.

Table 3.8. Estimates of the intervention possibly changing perceived descriptive norms

	Perceived descriptive norm	
	% not eating meat	% eating meat
Treat x Post period	1.37	-0.72
	(1.67)	(1.52)
Post period	0.02	-0.07
	(1.12)	(1.04)
Constant	52.50***	48.21***
	(0.42)	(0.38)
Guest fixed effects	Yes	Yes
Guests control	236	236
Guests treated	156	156
Observations	784	784

Standard errors in parentheses

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

Notes: Regression includes only survey data from respondents who filled out both surveys, allowing for the inclusion of guest fixed effects. Dependent variables differ by column. Col. (1) assesses changes in respondents' guess for which percentage of students of the University of Bonn do NOT consume a fish- or meat-containing meal on a typical university day, between surveys one and two. Col. (2) assesses changes in respondents' guess for which percentage of students of the University of Bonn do consume a fish- or meat-containing meal on a typical university day, between surveys one and two. Answers were not incentivized, but the purpose of the study was obfuscated as described in section 3.2. Guest fixed effects control for each respondents' norm perception at baseline. The number of control guests refers to the number of respondents who filled out both surveys and usually frequent the control canteens, and the number of treatment guests refers to the number of respondents who filled out both surveys and usually frequent the treatment canteen. Translated screen shots of the survey questions are shown in section 3.D.2. Standard errors are clustered at the individual level.

Table 3.9. Estimates of the intervention possibly changing perceived injunctive norms

	Personal norm			Perceived injunctive norm		
	1/5	3/5	5/5	1/5	3/5	5/5
Treat x Post period	0.13	0.14	0.25	-0.03	-0.15	-0.04
	(0.12)	(0.12)	(0.16)	(0.13)	(0.15)	(0.17)
Post period	0.16**	0.02	-0.19**	0.06	0.09	0.04
	(0.08)	(0.07)	(0.09)	(0.09)	(0.10)	(0.11)
Constant	4.67***	3.51***	2.43***	4.81***	3.58***	2.42***
	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)
Guest fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Guests control	236	236	236	236	236	236
Guests treated	156	156	156	156	156	156
Observations	784	784	784	784	784	784

Standard errors in parentheses

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

Notes: Regression includes only survey data from respondents who filled out both surveys, allowing for the inclusion of guest fixed effects. Dependent variables differ by column. Col. (1) - (3) assess changes in respondents' personal norm towards meat consumption, between surveys one and two. Specifically, they assess changes in respondents' perception of how socially appropriate a student of the University of Bonn is behaving if they consume meat on 1 out of 5 typical university days (Col. 1), on 3 out of 5 typical university days (Col. 2) or on 3 out of 5 typical university days (Col. 3). Approval is indicated on a 7-point scale ranging from 0 (not socially appropriate) to 6 (very socially appropriate). Col. (4) - Col. (6) elicit the perceived injunctive norm by asking respondents' what they believe is the most common answer to these questions among other respondents. Answers were not incentivized, but the purpose of the study was obfuscated as described in section 3.2. Guest fixed effects control for each respondents' norm perception at baseline. The number of control guests refers to the number of respondents who filled out both surveys and usually frequent the control canteens, and the number of treatment guests refers to the number of respondents who filled out both surveys and usually frequent the treatment canteen. Translated screen shots of the survey questions are shown in section 3.D.2. Standard errors are clustered at the individual level.

3.5.2 Additional analysis of canteen data

Effect of being familiar with a meal on meal choice

The most popular reason for changing consumption behavior reported in Table 3.7 is that respondents got to know vegetarian and vegan meals which were new for them in the course of the vegetarian month, and that they found them tasty. To gather further evidence on whether this might play a role in driving effects, we investigate in how far canteen guests' decision to consume the meat option in the post-intervention period correlates with whether they have consumed the meat or the vegetarian meal on offer previously. Table 3.10 regresses two dummy indicators of whether a canteen guest had consumed the meat/vegetarian option on offer already in the (pre-)intervention period on the guest's choice of the meat option. Columns (2) and (3) include individual fixed effects to control for differences in an individual's general taste for meat or vegetarian meals. Column (3) additionally includes date fixed effects to control for how attractive canteen guests on average perceive the meat and vegetarian option on offer on a given day to be. Results indicate that having tasted the meat meal on offer previously correlates with an 11 percentage point increase in the likelihood of choosing the meat meal, while having tasted the vegetarian meal on offer correlates with a 6 percentage point decrease in the likelihood of choosing the meat option.[26] Importantly, these estimates cannot be interpreted as causal, as they might still correlate with individual-specific tastes (e.g. an individual who loves chicken will be more likely to have consumed a chicken-including meal in the past and will also be more likely to consume it in the present without effects being driven by previous exposure to the meal). However, the date-fixed effects in the regression do control for differences in tastiness canteen guests would on average agree upon, and the individual-fixed effects control for an individual's general inclination to consuming meat. We thus interpret the results as suggestive evidence for previous exposure to a meal influencing an individual's likelihood of choosing the meat meal.

26. On some days, there are multiple meat or vegetarian main meal components on offer, most often because there are main meal components left over from the previous day. In these cases, the "Know meat meal" dummy turns one if a guest has tasted one of the meat components on offer previously, and the "Know veg. meal" dummy turns one if a guest has tasted one of the vegetarian components on offer previously.

Table 3.10. Correlation between previous experience of meal options and meat consumption

	Meat consumption Post-period		
	Base	Date FE	Date+Guest FE
Treat	2.28*		
	(1.30)		
Know meat meal	41.85***	11.39***	11.38***
	(0.72)	(0.63)	(0.66)
Know veg. meal	-26.21***	-6.12***	-5.35***
	(0.65)	(0.49)	(0.51)
Constant	42.29***	45.41***	41.05***
	(0.77)	(0.24)	(0.97)
Date fixed effects	No	Yes	Yes
Guest fixed effects	No	No	Yes
Guests control	2,813	2,813	2,813
Guests treated	929	929	929
Observations	35,879	35,879	35,879

Standard errors in parentheses

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

Notes: We use data from the pre-intervention and intervention period to create indicators, for each canteen guest and each meal on offer, of whether the guest had already consumed the meal previously. We restrict the analysis of consumption choices to the post-intervention data, and “Know meat meal” and “Know veg. meal” are indicators for whether the guest making the consumption choice had already consumed the meat meal or the vegetarian meal on offer in the pre-intervention or intervention period. Col. (2) additionally includes daily fixed effects to control for changes in daily meal offer. They control for differences in the attractiveness of the vegetarian and meat option offered on a given day, as on average perceived by canteen guests. Col. (3) additionally includes guest fixed effects to control for a guests’ general inclination to consume meat in the student canteen. Standard errors are clustered at the individual level.

Effect of the intervention on the canteen menu

Among the reasons survey respondents report for a change in behavior, an apparent change in the offer of the student canteen is one popular reason. A comparison of meals offered before and after the intervention, however, shows that the student canteen offered similar meals pre- and post-intervention. Of the vegetarian meals offered in the post-intervention period, 84% had already been offered at least once pre-intervention (78% of the meat meals). We thus interpret these results as evidence of canteen guests’ perception of canteen meals having changed. Canteen guests might have learned about the canteens’ vegetarian meals on offer, which, although not new to the menu, might well be “new” to the guests who have not experienced them previously.

Note also that meal planning is centralized across student canteens pre- and post- intervention, so if the vegetarian month was to have led to a change in the offer, this would affect both the treatment and control canteens. It would thus not be a factor explaining the effects identified in the ITT regressions, since it would affect groups equally.

3.6 Canteen guests' approval of the intervention

Of the 902 canteen guests who took part in the post-intervention survey, 75% would like a vegetarian day a week, 80% would like more vegan and vegetarian meals, and 54% are in favor of a vegetarian month every year. Figure 3.4 depicts approval for the different policies grouping respondents by whether they pre-dominantly visit the control or treatment canteens. Approval ratings do not significantly differ between the two groups.

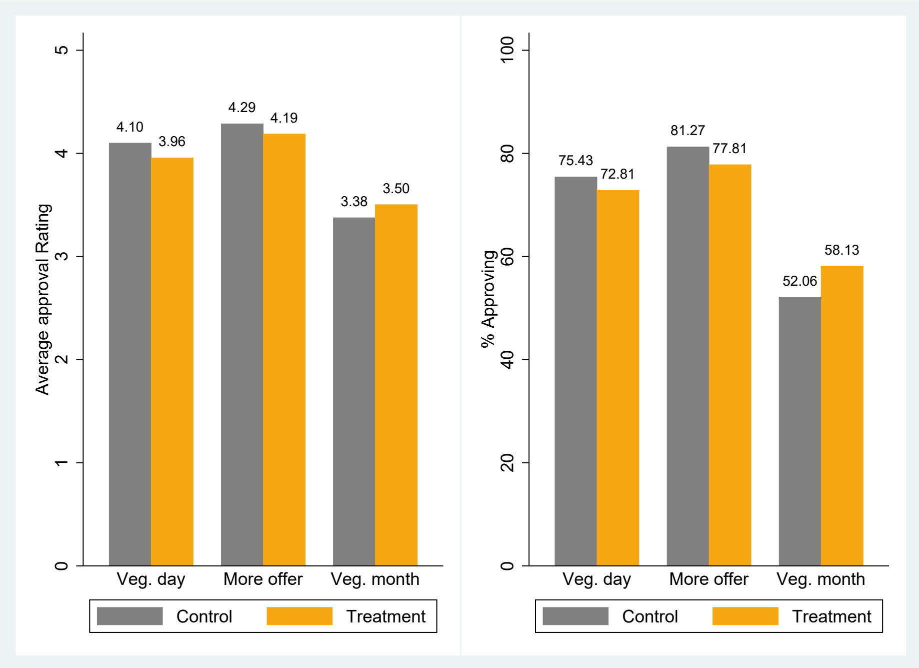


Figure 3.4. Canteen guests' approval of different policies

Note: We use survey data from 582 control and 320 treatment guests (based on self-reported main student canteen). The left figure shows the average approval for different policies, as indicated on a scale from 1 (does not apply at all) to 5 (completely applies). The right figure shows the percentage of respondents indicating approval of 4 or 5. The questions were phrased as: "For more climate and animal protection, I wish that the student canteens in Bonn would... (1) ... offer one vegetarian day a week. (2) ... offer more vegan/ vegetarian meals. (3) ... only offer exclusively vegan and vegetarian meals for one month every year."

The differences in policy approval are greater when examining guests usually frequenting the treatment canteen separately by their perceived reduction in meat consumption, as shown in Figure 3.5. Policy approval is highest among survey participants who already ate exclusively vegetarian in the canteen before the intervention, with over 90% in favor of a vegetarian day and an increase in vegetarian options. Among respondents who said they had reduced their meat consumption after the vegetarian month, 76% are in favor of a vegetarian day a week and 63% would like to have a vegetarian month every year. In contrast, of those who stated that they had not reduced their meat consumption, only around 38% are in favor of a vegetarian day a week, and only around 14% support having a vegetarian month every year.

3.7 Discussion

We study the effects of a somewhat radical intervention to decrease meat consumption in the student canteen: One of the University's student canteens completely eliminated meat options from

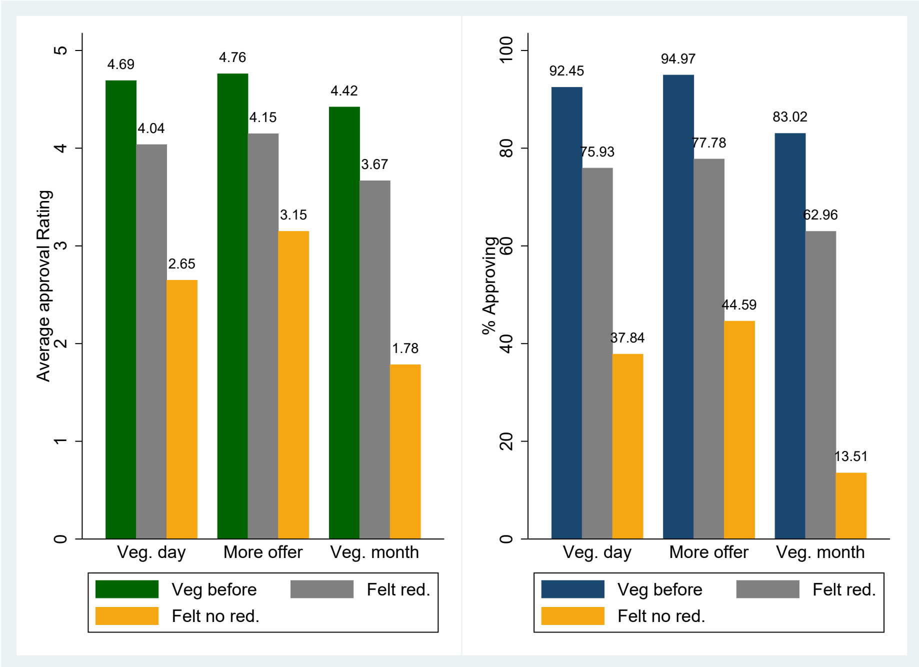


Figure 3.5. Canteen guests' approval of different policies by self-reported reduction in meat consumption

Note: We use survey data from those 287 guests who indicated mainly going to the treatment canteen and indicated how the intervention changed their consumption behavior: 159 guests report already eating exclusively vegetarian at the canteen pre-intervention, 54 guests indicate to have reduced meat consumption post-intervention and 74 guests indicate to not have reduced meat consumption post-intervention. The left figure shows the average approval for different policies, as indicated on a scale from 1 (does not apply at all) to 5 (completely applies). The right figure shows the percentage of respondents indicating approval of 4 or 5. The questions were phrased as: "For more climate and animal protection, I wish that the student canteens in Bonn would... (1) ... offer one vegetarian day a week. (2) ... offer more vegan/ vegetarian meals. (3) ... only offer exclusively vegan and vegetarian meals for one month every year."

its menu for a month and instead offered more vegetarian and vegan options. In a difference-in-difference analysis using the usual guests of the other two student canteens as a control group, we estimate that the intervention did not lead guests usually frequenting the treatment canteen to reduce their visits during or after the intervention. We estimate that the intervention decreased the proportion of meat meals sold in the treatment canteen by 7 to 12% in the two months following the intervention, relative to the baseline period. In an intent-to-treat analysis including guest fixed effects, we estimate that guests who usually frequented the treatment canteen pre-intervention were 4% less likely to choose a meat meal in the canteen in the post-intervention period, relative to baseline levels.

Post-intervention effects are thus noticeable but modest. Effect sizes are comparable to the post-intervention effects Kurz (2018) identifies for an intervention changing choice architecture in the canteen. It remains immensely difficult to permanently shift consumers' meat consumption behavior, and even a radical intervention such as that examined here is no "silver bullet". A one-month intervention period also does not seem to be long enough or perhaps the intervention not strong enough to shift guests' perception of the social norm towards meat consumption.

Our evidence suggests that guests' learning about their preferences for different vegetarian and vegan options mainly drives post-intervention effects. We expect this channel to play a role in any intervention successfully shifting meat consumption behavior, even if it is only implemented for a short time. Other types of interventions targeting food consumption behavior might additionally impact consumers through other channels such as a change in social norms or attitudes. For the intervention we study, a possible change in social norms does not seem to be a major driver of post-intervention effects, with auxiliary survey evidence offering little to no evidence of the intervention producing a change in perceived social norms.

The experience made with the vegetarian month might also hold interesting insights in the sense that it is an example of a rather radical policy not sparking as much discontent as some opinion leaders seem to have expected pre-intervention. At the beginning of the intervention month, the regional agricultural minister received harsh criticism from within her own party for supporting the policy (Die Welt, 2023; Kölner Stadt Anzeiger, 2023; t-online, 2023), and the intervention sparked considerable online discussions on Twitter/X and Reddit, with some people expressing concerns about the intervention drastically decreasing student canteen guest numbers. Our results suggest that there might not have been that much reason for concern: We do not see any evidence of the intervention decreasing the number of student canteen guests, a majority of guests are in favor of the intervention even after the intervention month, and the green student party who co-initiated the policy measure was re-elected as the strongest student party. Of course, the social dynamics, environmental awareness, and baseline consumption of vegetarian dishes among university students likely differ from those of the general population, and this might have led to a more favorable outcome than had the policy been implemented in other groups of the population. However, the concerns expressed before the intervention month were also context- and group-specific, i.e., people expected this specific segment of the population to react more negatively to the intervention than it in fact did. In this sense, our paper connects to Andre et al. (2021), who show evidence of people underestimating the prevalence of climate-friendly behaviors among fellow citizens.

One reason the intervention was, to a reasonable extent, accepted by students might have been that the intervention was developed jointly by student representatives, canteen operators, and local student organizations. Collaboration and exchange between these parties was fostered by a state-funded initiative.[27] Student involvement in the development of the measure may have impacted its acceptance in different ways, e.g. by leading to a greater sense of ownership and commitment among students, or by influencing the implementation of the vegetarian month such that it is better accepted by students. In a back-of-the-envelope calculation, we estimate that the canteen avoided around seven tonnes of CO_2 emissions during the intervention period.[28]

Importantly, the student canteen environment cannot be understood as representative of the population as such, and effect sizes will likely differ among other groups of the population. Further, our results focus on meat consumption *in the canteen*. We do not observe students' meat consumption outside of the canteen. It is of course possible that students decreased meat consumption in the canteen while simultaneously increasing their meat consumption at home or in other dining settings. While such potential displacement effects seem plausible during the intervention period, it seems less plausible for them to play a major role post-intervention when meat meals are again available in the canteen. This intuition is backed by our survey evidence.

27. This is the initiative "Nachhaltige Ernährung im Studienalltag" (Sustainable consumption in daily student life) funded by the consumer protection agency of the state of North-Rhine-Westphalia.

28. 16,634 purchases were made in the treatment canteen during the intervention period. Pre-intervention, 42% of purchases were meat meals. Schulze-Tilling (2023) estimates that the difference in emissions between a meat meal and a vegetarian meal is on average around 1 kg in the student canteen in Bonn. We thus calculate: 16,634 times 0.42 = 6,986 kg.

Our findings offer valuable insights into how and why interventions to decrease meat consumption may have an impact post-intervention, as well as on the behavioral frictions that might make consumers avoid vegetarian or vegan dishes. Future research might further investigate the potential of interventions to help individuals to update their beliefs and perceptions about vegetarian options, leading to more informed and potentially lasting changes in their dietary habits.

Appendix 3.A Additional tables and figures

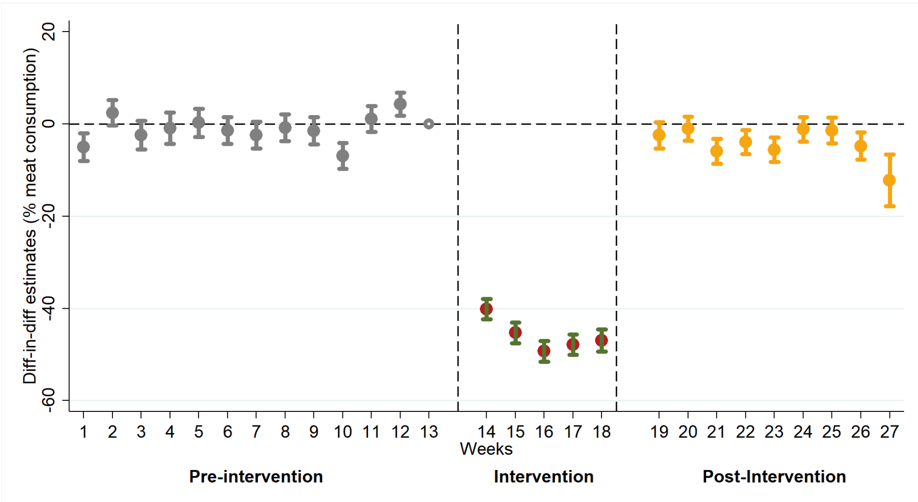


Figure 3.A.1. Canteen-level event plot

Note: Figure shows coefficients estimated in a regression analysis following equation 3.1, but including weekly interaction terms and time controls. We additionally control for day of the week and the controls added in Col. (4) of Table 3.2. Coefficients show the estimated change in the likelihood of consuming the meat main meal component, in percentage points. Our data for week 27 includes only one day, since it was the last day of July and our sample period. Bars indicate 95% confidence intervals.

Table 3.A.1. Canteen-level estimates including controls

	Likelihood of consuming meat (in pp.)			
	Base	Date FE	+Controls	+Sales
Treat x Inter period	-45.66***	-45.81***	-45.26***	-45.06***
	(0.40)	(0.40)	(0.46)	(0.46)
Treat x Post period	-4.84***	-5.10***	-3.78***	-3.18***
	(0.46)	(0.46)	(0.49)	(0.50)
Treat	-4.98***	-4.63***	-2.85***	-6.11***
	(0.30)	(0.30)	(0.54)	(0.74)
Inter period	2.99***			
	(0.31)			
Post period	3.73***			
	(0.26)			
Sales				-0.01***
				(0.00)
Constant	47.65***	34.49***	33.71***	42.13***
	(0.17)	(0.94)	(1.07)	(1.68)
Special			-2.06***	-2.05***
			(0.40)	(0.40)
Pasta			1.12**	0.56
			(0.45)	(0.46)
Pizza			-1.62***	-2.15***
			(0.36)	(0.37)
Extra meal type 1			-1.98***	-1.78***
			(0.49)	(0.49)
Extra meal type 2			-0.91	-0.50
			(0.59)	(0.59)
2nd veg main			-2.42***	-2.26***
			(0.28)	(0.28)
Extra meal type 3			-1.15*	-0.83
			(0.61)	(0.61)
Control res. 2			10.16***	5.63***
			(0.63)	(0.94)
Date fixed effects	No	Yes	Yes	Yes
Guest fixed effects	No	No	No	No
Control for other offer	No	No	Yes	Yes
Guests control	8,353	8,353	8,353	8,353
Guests treated	3,575	3,575	3,575	3,575
Observations	276,673	276,673	276,673	276,673

Standard errors in parentheses

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

Notes: This is a full version of Table 3.2 including coefficients estimated for controls. These include indicators for additional offers (special dish, pasta dish, pizza, and other extra meals) as well as for a 2nd vegetarian main on offer and a separate intercept for the second control canteen. Standard errors are robust.

Table 3.A.2. Estimates of the intervention possibly changing perceived descriptive norms, assigning treatment group based on consumption data

	Perceived descriptive norm	
	% not eating meat	% eating meat
Treat x Post period	1.03	-1.19
	(2.55)	(2.33)
Post period	1.13	0.13
	(1.72)	(1.58)
Constant	50.89***	49.44***
	(0.64)	(0.59)
Guest fixed effects	Yes	Yes
Guests control	107	107
Guests treated	70	70
Observations	354	354

Standard errors in parentheses

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

Notes: Regression includes only survey data from respondents who filled out both surveys, allowing for the inclusion of guest fixed effects. Additionally, we restrict the sample to those responses which we can link to individual consumption data, and assign treatment group to each individual based on previous consumption behavior, using the same rule as in our main ITT analysis. Dependent variables differ by column. Col. (1) shows respondents' guess for which percentage of students of the University of Bonn do NOT consume a fish- or meat-containing meal on a typical university day. Col. (2) shows respondents' guess for which percentage of students of the University of Bonn do consume a fish- or meat-containing meal on a typical university day. Answers were not incentivized, but the purpose of the study was obfuscated as described in section 3.2. Standard errors are clustered at the individual level.

Table 3.A.3. Estimates of the intervention possibly changing perceived injunctive norms, assigning treatment group based on consumption data

	Personal norm			Perceived injunctive norm		
	1/5	3/5	5/5	1/5	3/5	5/5
Treat x Post period	-0.01	0.02	-0.10	-0.10	-0.10	0.11
	(0.18)	(0.17)	(0.22)	(0.19)	(0.20)	(0.25)
Post period	0.27**	0.14	0.08	0.25**	0.21*	0.16
	(0.12)	(0.10)	(0.11)	(0.12)	(0.12)	(0.15)
Constant	4.69***	3.63***	2.65***	4.68***	3.59***	2.44***
	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)
Guest fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Guests control	107	107	107	107	107	107
Guests treated	70	70	70	70	70	70
Observations	354	354	354	354	354	354

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

Notes: Regression includes only survey data from respondents who filled out both surveys, allowing for the inclusion of guest fixed effects. Additionally, we restrict the sample to those responses which we can link to individual consumption data, and assign treatment group to each individual based on previous consumption behavior, using the same rule as in our main ITT analysis. Dependent variables differ by column. Col. (1) - (3) show responses for the question of how ethically correct a student of the University of Bonn is behaving if they consume meat on 1 out of 5 typical university days (Col. 1), on 3 out of 5 typical university days (Col. 2) or on 3 out of 5 typical university days (Col. 3). Approval is indicated on a 7-point scale ranging from 0 (not socially appropriate) to 6 (socially appropriate). Col. (4) - Col. (6) elicit the perceived norm by asking respondents' what they believe is the most common answer to these questions among other respondents. Answers were not incentivized, but the purpose of the study was obfuscated as described in section 3.2. Standard errors are clustered at the individual level.

Table 3.A.4. Self-reported motives for decreasing meat consumption among guests mainly frequenting the control canteen

Motive	Count	Percentage
Norm1	6	12.5%
Norm2	7	14.58%
Taste	27	56.25%
Habit	26	54.17%
Offer	17	35.42%
Spillover	1	2.08%
None of above	4	8.33%
Total respondents	48	100%

Notes: Table shows reasons cited by guests who report mainly going to the control canteen and that they consume the vegetarian option more frequently after the vegetarian month. 203 (37%) of respondents who reported mainly going to the control canteen report going to the treatment canteen at least once during the intervention period, and of these, 48 (24%) report consuming the vegetarian option more frequently after the vegetarian month. Multiple options were selectable. The statements read: I now consume the vegetarian option more frequently, because ... (1) more students eat vegetarian/vegan meals since the month. (2) my friends eat more vegetarian/vegan meals since the month. (3) I got to know vegetarian vegan meals in the month which were new for me and which I find tasty. (4) It's becoming a habit for me to eat vegan/vegetarian. (5) The vegetarian/vegan offer has improved since the month. (6) I started consuming meat rather outside of the canteen during the month and stuck to that after the month was over.

Table 3.A.5. Self-reported motives for decreasing meat consumption among guests with a behavioral change in the consumption data

Motive	Count	Percentage
Norm1	1	7.69%
Norm2	1	7.69%
Taste	7	53.85%
Habit	4	30.77%
Offer	6	46.15%
Spillover	0	0%
None of above	0	0%
Total respondents	13	100%

Notes: Table shows reasons cited by guests who we classify as Treatment following the procedure described in section 3.4.1, report that they consume the vegetarian option more frequently after the vegetarian month, and consume a higher proportion of vegetarian main meal components after than before the vegetarian month according to their consumption data. Multiple options were selectable. The statements read: I now consume the vegetarian option more frequently, because ... (1) more students eat vegetarian/vegan meals since the month. (2) my friends eat more vegetarian/vegan meals since the month. (3) I got to know vegetarian vegan meals in the month which were new for me and which I find tasty. (4) It's becoming a habit for me to eat vegan/vegetarian. (5) The vegetarian/vegan offer has improved since the month. (6) I started consuming meat rather outside of the canteen during the month and stuck to that after the month was over.

Table 3.A.6. ITT estimates by gender and study

	All	Gender split		Split treatment group + Full control	
		male	female	Culture, Economics, Society	Law
Treat x Inter period	-28.33***	-32.46***	-23.55***	-23.49***	-35.73***
	(2.83)	(4.03)	(4.06)	(3.31)	(4.96)
Treat x Post period	-1.18	-0.24	-2.20	-0.91	-1.88
	(1.60)	(2.41)	(2.00)	(2.23)	(1.86)
Constant	27.01***	35.47***	16.95***	25.56***	29.78***
	(2.21)	(3.13)	(3.10)	(2.22)	(2.63)
Date fixed effects	Yes	Yes	Yes	Yes	Yes
Guest fixed effects	Yes	Yes	Yes	Yes	Yes
Guests control	357	177	176	357	357
Guests treated	199	102	94	111	70
Observations	15,961	8,792	7,018	13,404	11,891

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. Col. (1) corresponds to Col. (3) in Table 3.3, but includes only canteen guests for whom we have demographic information. Col. (2) - (3) restrict the sample based on canteen guests' gender, as indicated in the surveys. Col. (4) - Col. (5) restricts the treatment group based on field of study. Both include the full control sample, but restrict the treated group to only Humanities and Arts, Economics, and Social Sciences in Col. (4) and to only Law students in Col. (5). Standard errors are clustered at the individual level.

Table 3.A.7. ITT estimates by age

	All	Age split		
		under 22	22-23	over 23
Treat x Inter period	-28.33***	-32.30***	-26.10***	-24.91***
	(2.83)	(4.78)	(4.98)	(4.98)
Treat x Post period	-1.18	-5.16**	-1.24	1.92
	(1.60)	(2.32)	(1.84)	(3.61)
Constant	27.01***	30.34***	33.08***	13.99***
	(2.21)	(3.17)	(3.76)	(4.39)
Date fixed effects	Yes	Yes	Yes	Yes
Guest fixed effects	Yes	Yes	Yes	Yes
Guests control	357	156	99	101
Guests treated	199	71	53	74
Observations	15,961	6,212	4,806	4,895

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. Col. (1) corresponds to Col. (3) in Table 3.3, but includes only canteen guests for whom we have demographic information. Col. (2) - (4) restrict the sample based on canteen guests' age, as indicated in the surveys. Standard errors are clustered at the individual level.

Table 3.A.8. ITT estimates for university employees

	All	Group
		Only employees
Treat x Inter period	-34.84***	-57.45***
	(1.38)	(4.75)
Treat x Post period	-1.62**	-2.08
	(0.68)	(2.08)
Constant	36.93***	59.28***
	(0.92)	(2.69)
Date fixed effects	Yes	Yes
Guest fixed effects	Yes	Yes
Guests control	3,371	554
Guests treated	1,142	101
Observations	117,642	18,746

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.

Appendix 3.B Effect of the intervention on canteen visiting patterns

To examine whether the vegetarian month led to guests changing their canteen visiting patterns, we classify guests as treated or control guests based on their pre-intervention behavior, similarly to the procedure described in section 3.4.1, but with one difference: We base classification only on the first 11 weeks instead of the entire 13-week pre-intervention period, to also be able to present a short pre-trend for each of the analyses.[29]

We first examine how frequently guests we classified as treated or control visit the respective other canteen. As can be seen in Figure 3.B.1, around 4% to 8% of the purchases made by the group we classify as treated are made in the control canteen, while between 1% and 5% of the purchases made by the group we classify as control are made in the treatment canteen. The percentage fluctuates across weeks, but there is no clear pattern of the vegetarian month differentially leading to increases in "non-home" visits in the treated group. Figure 3.B.2 additionally examines whether those in the treatment group visiting the control canteen consume a higher proportion of meat meals during their visits to the control canteen. Also here, there is if at all only a slight increase, with guests eating the meat meal on 40% of these visits before the intervention, and between 40% and 60% during the intervention period.

To examine whether the vegetarian month led to guests usually frequenting the treatment canteen visiting the student canteens less in general, we analyze which proportion of the usual canteen guests visited one of the canteens on a given day on which the student canteens are operating. In the first two weeks following the weeks used for classification, a typical regular student canteen guest would visit the canteen on 30% of the days on which it is open. This percentage consistently decreases – among both groups – throughout the sample period. This is likely due to canteen frequenting patterns changing in both groups across time: Someone who meets our criteria for being a regular canteen guest in the first weeks of the intervention period might have, for example, already left the university by week 26. To examine whether the intervention led to a change in the canteen frequenting behavior of the treated, we examine the intervention period for differential trends between control and treatment groups. There is no evidence of this being the case. Thus, overall, the intervention period does not seem to have had an effect on overall sales on canteen frequenting behavior.

However, the composition of guests might have changed during the intervention period. Figures 3.B.6, 3.B.9, and 3.B.12 reproduce Figure 3.B.3 restricting the sample based on previous meat consumption. Figure 3.B.6 shows that for canteen guests in the lowest tercile of meat consumption pre-intervention (0-10% of their pre-intervention meals contained meat), the intervention led to a visible increase in canteen visits in the treatment group relative to the control group. For guests in the second tercile (Figure 3.B.9), trends are very similar to overall trends. For guests in the third tercile (over 68% of pre-intervention meals contained meat), there is a visible decrease in canteen visits in the treatment group relative to the control group. Figures 3.B.4

29. The reasoning for this is as follows: We are looking at outcome variables in this section that also form part of our treatment and control group definitions: quantity of visits and whether the guest eats in the treatment or control canteen. There is thus little value to be gained from considering the pre-intervention behavior of these outcomes since they are influenced by the nature of our treatment group definition. To still be able to examine a pre-trend, we base classification in this section only on the first 11 weeks. In the main analysis presented in section 3.4, we show the full pre-intervention trend since the dependent variable we examine in the main analysis – meat consumption – plays no role in our sample definitions.

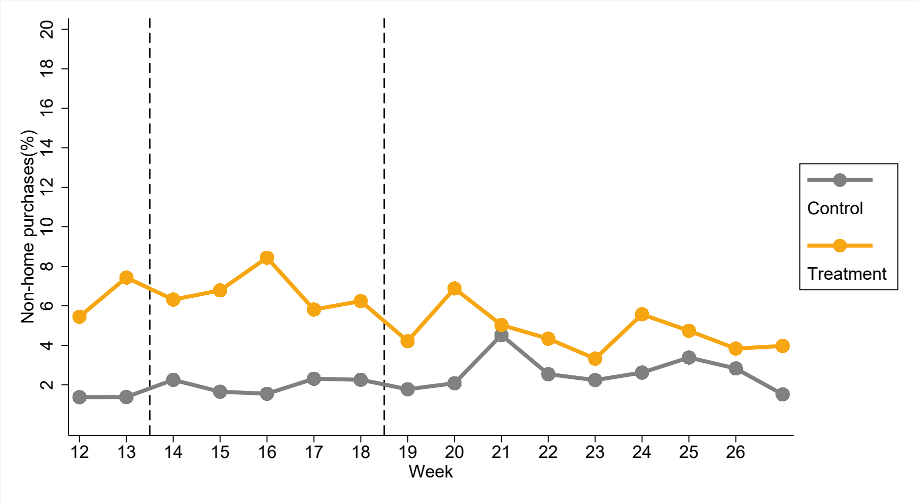


Figure 3.B.1. Visits to the "non-home" canteen

Note: This figure shows the weekly percentage of meals which guests classified as treated consumed in the control canteen and vice-versa. Weeks 1 to 11 are excluded from the graph, since classification as control or treated is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on purchases made by 2,722 control and 922 treatment guests. The number of guests is a bit lower than in the main analysis in Table 3.2 because treatment and control group assignment criteria are applied using only data from weeks 1 to 11.

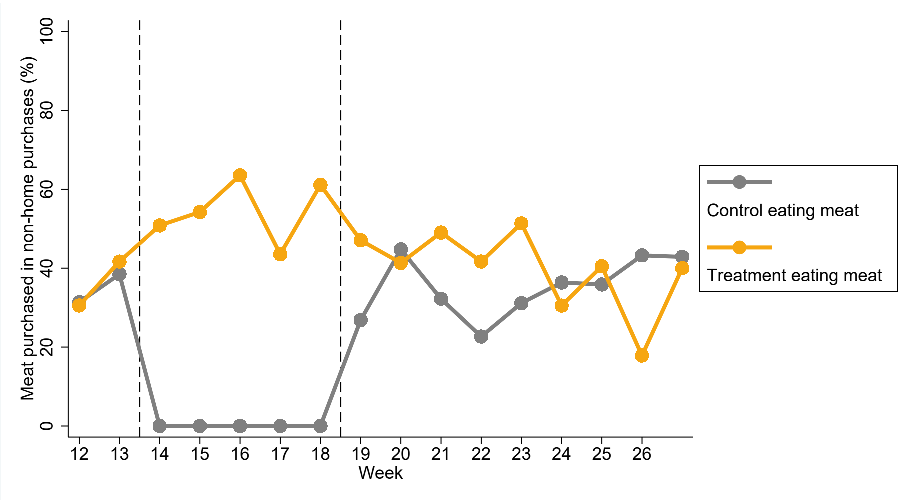


Figure 3.B.2. Meat consumption when visiting the "non-home" canteen

Note: This figure shows the percentage of meat meals consumed among guests eating at their "non-home" canteen. Weeks 1 to 11 are excluded from the graph, since classification as control or treated is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on purchases made by 2,722 control and 922 treatment guests. The number of guests is a bit lower than in the main analysis in Table 3.2 because treatment and control group assignment criteria are applied using only data from weeks 1 to 11.

to 3.B.11 recreate Figures 3.B.1 and 3.B.2 splitting the sample by baseline meat consumption and finds a similar pattern: treatment guests with low baseline meat consumption do not seem to frequent the control canteen more frequently during the intervention period, while guests in the highest tercile of previous meat consumption seem to be visiting the canteen more fre-

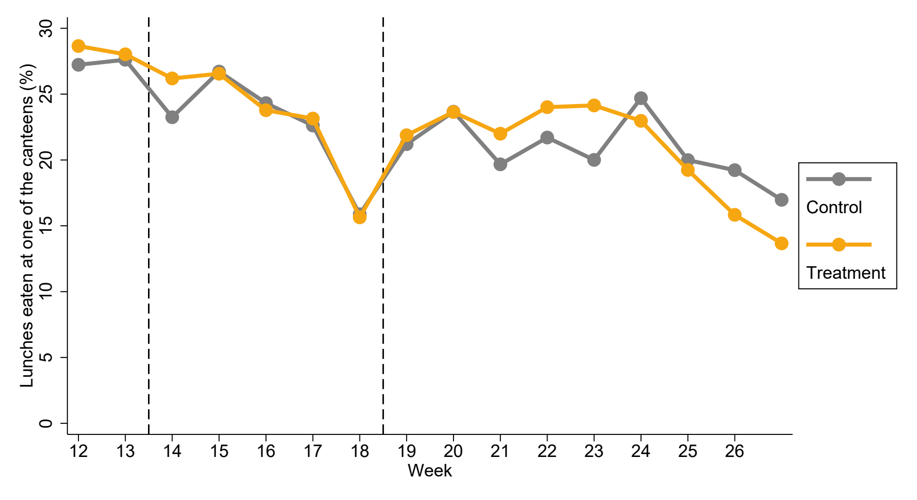


Figure 3.B.3. Percentage of usual canteen guests eating lunch at the student canteen

Note: This figure shows the percentage of lunch meals consumed at one of the student canteens for guests classified as regular treatment or control canteen guests. Weeks 1 to 11 are excluded from the graph, since classification as a regular student canteen guest is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on purchases made by 2,722 control and 922 treatment guests. The number of guests is a bit lower than in the main analysis in Table 3.2 because treatment and control group assignment criteria are applied using only data from weeks 1 to 11.

quently. For these guests, the proportion of meat meals consumed during these visits also seems to increase.

Thus, simply examining changes in average canteen frequenting patterns seems to mask important heterogeneities: While the vegetarian month did not lead to a change in average consumption patterns, the composition of guests in the canteen seems to have changed during the intervention period.

3.B.1 Restricting to guests with low meat consumption at baseline

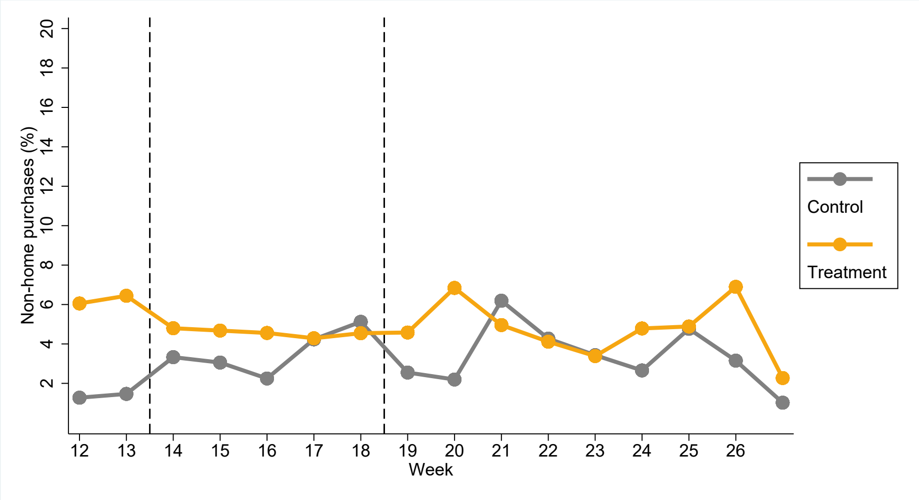


Figure 3.B.4. Visits to the "non-home" canteen (sub-sample low meat at baseline)

Note: This figure shows the weekly percentage of meals which guests classified as treated consumed in the control canteen and vice-versa. Weeks 1 to 11 are excluded from the graph, since classification as control or treated is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on 32,845 purchases made by 852 control and 315 treatment guests. These are guests consuming meat in less than 10% of their meals in weeks 1 to 11. We split guests into three terciles based on previous meat consumption, this is the lowest tercile.

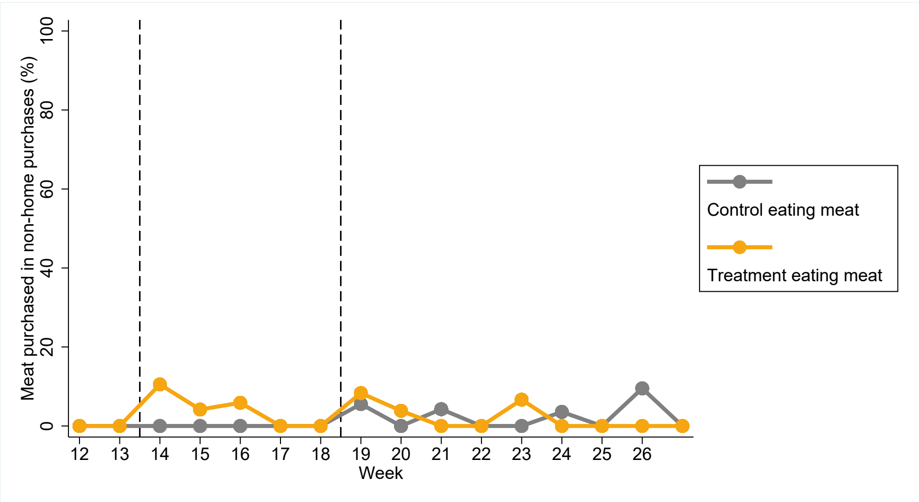


Figure 3.B.5. Meat consumption when visiting the "non-home" canteen (sub-sample low meat at baseline)

Note: This figure shows the percentage of meat meals consumed among guests eating at their non-home canteen. Weeks 1 to 11 are excluded from the graph, since classification as control or treated is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on 32,845 purchases made by 852 control and 315 treatment guests. These are guests consuming meat in less than 10% of their meals in weeks 1 to 11. We split guests into three terciles based on previous meat consumption, this is the lowest tercile.

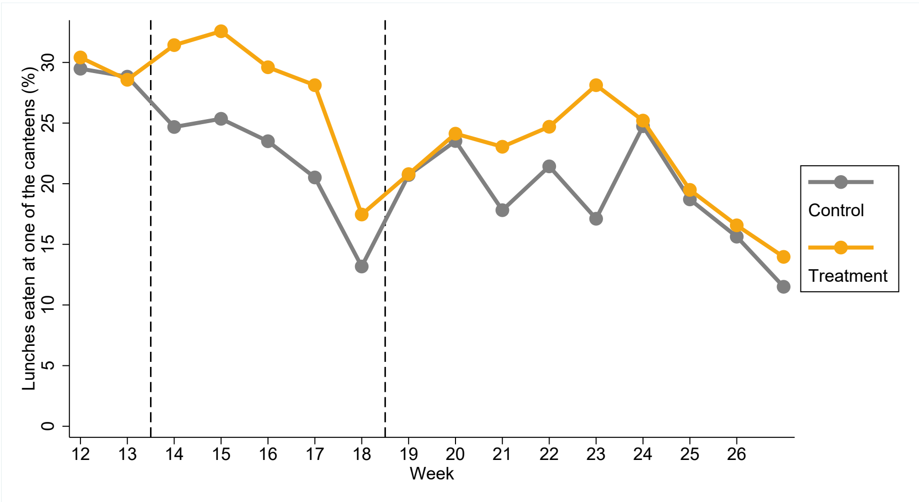


Figure 3.B.6. Percentage of usual canteen guests eating lunch at the student canteen (sub-sample low meat at baseline)

Note: This figure shows the percentage of meals consumed at one of the student canteens for guests classified as treatment or control canteen guests. Weeks 1 to 11 are excluded from the graph, since classification as a regular student canteen guest is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on purchases made by 852 control and 315 treatment guests. These are guests consuming meat in less than 10% of their meals in weeks 1 to 11. We split guests into three terciles based on previous meat consumption, this is the lowest tercile.

3.B.2 Restricting to guests with medium meat consumption at baseline

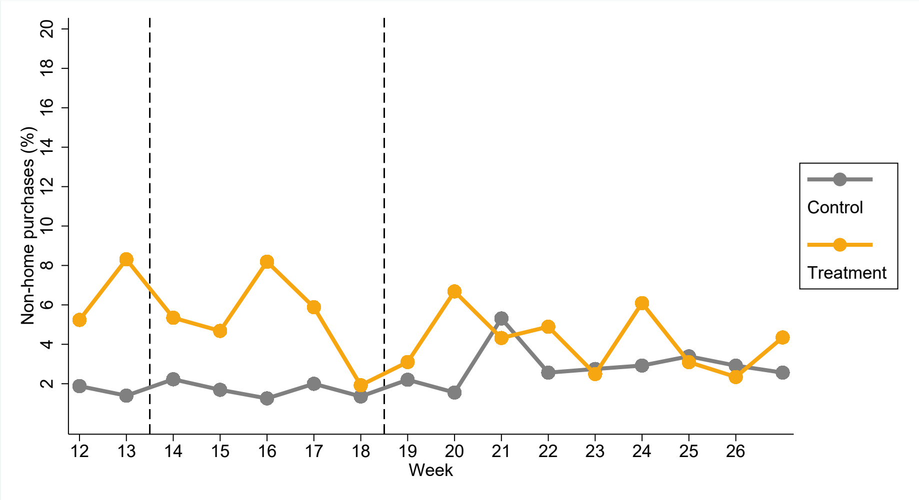


Figure 3.B.7. Visits to the “non-home” canteen (sub-sample med. meat at baseline)

Note: This figure shows the weekly percentage of meals which guests classified as treated consumed in the control canteen and vice-versa. Weeks 1 to 11 are excluded from the graph, since classification as control or treated is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on 34,175 purchases made by 909 control and 306 treatment guests. These are guests consuming meat in over 10%, but less than 68% of their meals in weeks 1 to 11. We split guests into three terciles based on previous meat consumption, this is the medium tercile.

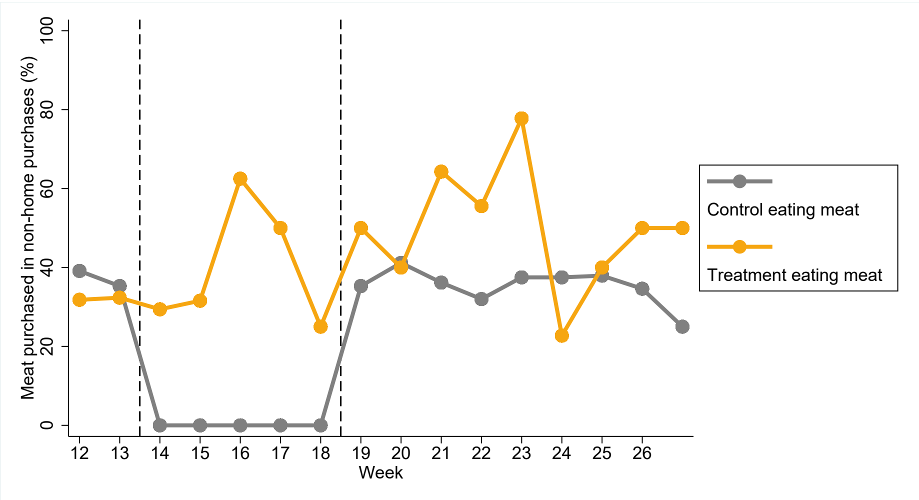


Figure 3.B.8. Meat consumption when visiting the “non-home” canteen (sub-sample med. meat at baseline)

Note: This figure shows the percentage of meat meals consumed among guests eating at their non-home canteen. Weeks 1 to 11 are excluded from the graph, since classification as control or treated is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on 34,175 purchases made by 909 control and 306 treatment guests. These are guests consuming meat in over 10%, but less than 68% of their meals in weeks 1 to 11. We split guests into three terciles based on previous meat consumption, this is the medium tercile.

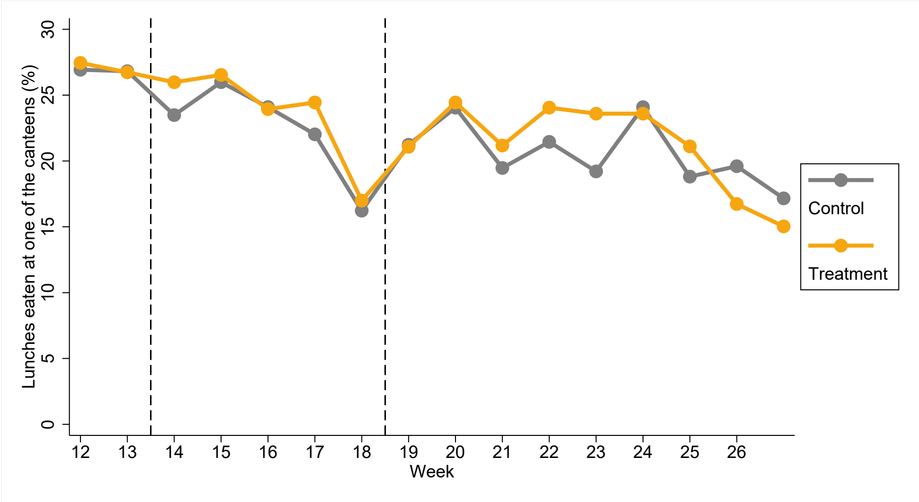


Figure 3.B.9. Percentage of usual canteen guests eating lunch at the student canteen (sub-sample med. meat at baseline)

Note: This figure shows the percentage of meals consumed at one of the student canteens for guests classified as treatment or control canteen guests. Weeks 1 to 11 are excluded from the graph, since classification as a regular student canteen guest is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on purchases made by 909 control and 306 treatment guests. These are guests consuming meat in over 10%, but less than 68% of their meals in weeks 1 to 11. We split guests into three terciles based on previous meat consumption, this is the medium tercile.

3.B.3 Restricting to guests with high meat consumption at baseline

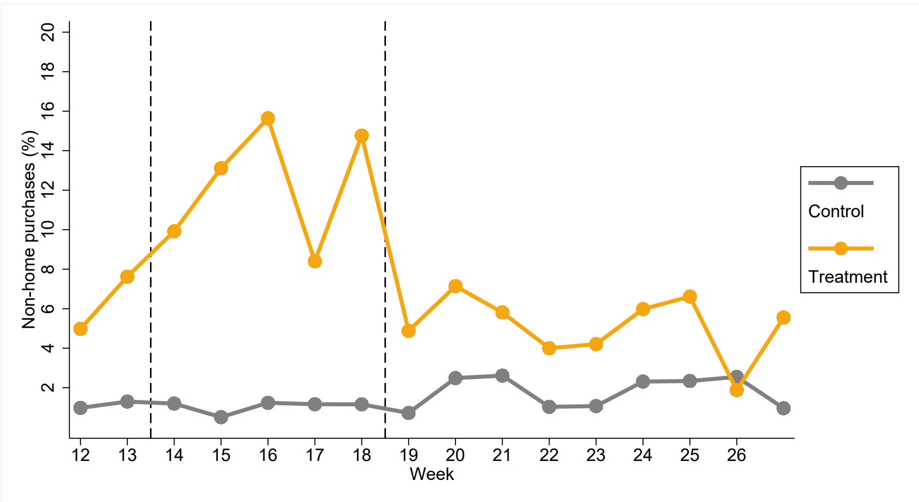


Figure 3.B.10. Visits to the "non-home" canteen (sub-sample high meat at baseline)

Note: This figure shows the weekly percentage of meals which guests classified as treated consumed in the control canteen and vice-versa. Weeks 1 to 11 are excluded from the graph, since classification as control or treated is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on 36,296 purchases made by 961 control and 301 treatment guests. These are guests consuming meat in 68% and over of their meals in weeks 1 to 11. We split guests into three terciles based on previous meat consumption, this is the highest tercile.

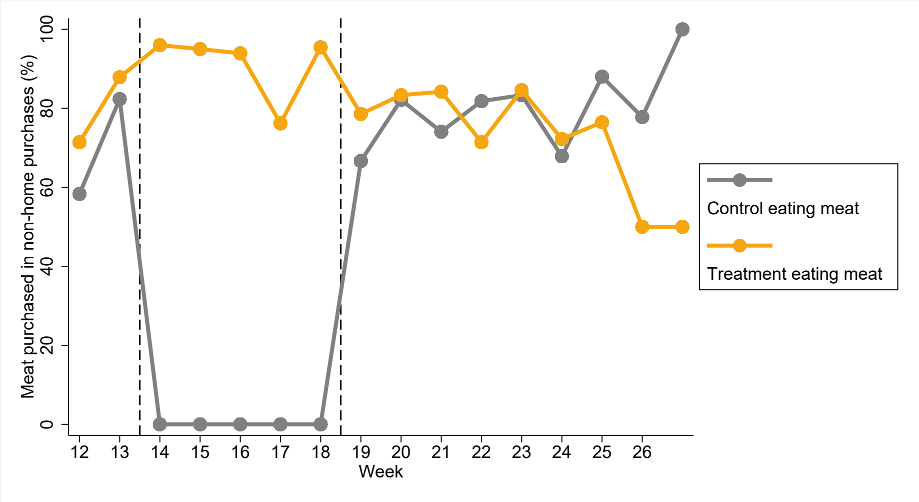


Figure 3.B.11. Meat consumption when visiting the “non-home” canteen (sub-sample high meat at baseline)

Note: This figure shows the percentage of meat meals consumed among guests eating at their non-home canteen. Weeks 1 to 11 are excluded from the graph, since classification as control or treated is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on 36,296 purchases made by 961 control and 301 treatment guests. These are guests consuming meat in 68% and over of their meals in weeks 1 to 11. We split guests into three terciles based on previous meat consumption, this is the highest tercile.

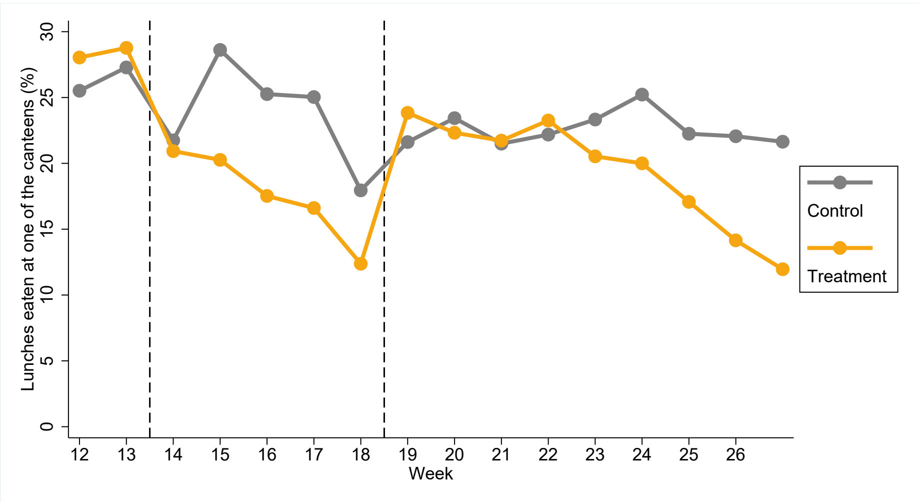


Figure 3.B.12. Percentage of usual canteen guests eating lunch at the student canteen (sub-sample high meat at baseline)

Note: This figure shows the percentage of meals consumed at one of the student canteens for guests classified as treatment or control canteen guests. Weeks 1 to 11 are excluded from the graph, since classification as a regular student canteen guest is determined based on these weeks, following the procedure described in section 3.4.1. Graph shows the final two weeks of the pre-intervention phase (weeks 1-13, February to April 2023), intervention phase (weeks 14 - 18, May 2023), and post-intervention phase (weeks 19 - 25, June to July 2023). Based on purchases made by 961 control and 301 treatment guests. These are guests consuming meat in 68% and over of their meals in weeks 1 to 11. We split guests into three terciles based on previous meat consumption, this is the highest tercile.

Appendix 3.C Additional intent-to-treat estimates

Table 3.C.1. ITT estimates of the effect of the intervention on canteen visits

	Visit(in pp)		
	Base	Date FE	Date+Guest FE
Treat x Inter period	-0.44	-0.44	-0.44
	(0.74)	(0.74)	(0.74)
Treat x Post period	-0.57	-0.57	-0.57
	(0.76)	(0.76)	(0.76)
Treat	0.92	0.92	
	(0.92)	(0.92)	
Inter period	-4.67***		
	(0.36)		
Post period	-6.26***		
	(0.38)		
Constant	27.42***	27.48***	27.72***
	(0.46)	(0.77)	(0.68)
Date fixed effects	No	Yes	Yes
Guest fixed effects	No	No	Yes
Guests control	2,722	2,722	2,722
Guests treated	922	922	922
Observations	262,368	262,368	262,368

Standard errors in parentheses

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

Notes: Regressions analyze the likelihood of consuming a main meal component at one of the student canteens for guests classified as treatment or control canteen guests. Weeks 1 to 11 are excluded from the analysis, since classification as a regular student canteen guest is determined based on these weeks, following the procedure described in section 3.4.1. Dependent variable is is a 0/1 indicator for visiting one of the student canteens. The number of guests is a bit lower than in the main analysis in Table 3.2 because treatment and control group assignment criteria are applied using only data from weeks 1 to 11. Weeks 1 to 11 are dropped from the analysis as explained in the main text around Table 3.4. Col. (1) corresponds to Equation 3.2 apart from the dependent variable. The Constant term describes the likelihood of the control group to visit one of the student canteens pre-intervention. Specifications (2) and (3) include date-fixed effects to control for the daily changing offer of main meal components, which is common across canteens. The “PostPeriod” and “Inter period” indicators are thus dropped due to collinearity. Specification (3) includes individual fixed effects. Standard errors are clustered at the individual level.

Table 3.C.2. ITT estimates of the effect of the intervention on visits specifically to the home canteen

	Visit(in pp)		
	Base	Date FE	Date+Guest FE
Treat x Inter period	-0.11	-0.11	-0.11
	(0.73)	(0.73)	(0.73)
Treat x Post period	0.42	0.42	0.42
	(0.75)	(0.75)	(0.75)
Treat	-0.52	-0.52	
	(0.91)	(0.91)	
Inter period	-4.74***		
	(0.36)		
Post period	-6.45***		
	(0.38)		
Constant	27.04***	27.27***	27.14***
	(0.45)	(0.76)	(0.68)
Date fixed effects	No	Yes	Yes
Guest fixed effects	No	No	Yes
Guests control	2,722	2,722	2,722
Guests treated	922	922	922
Observations	262,368	262,368	262,368

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

Notes: Analysis repeats that in Table 3.C.1, but counts only those visits in line with the ITT group guests are assigned to, i.e. for guests classified as treated it only counts visits to the treatment canteen, and for guests classified as control it only counts visits to the control canteen. This is the case for 97.7% of visits. Standard errors are clustered at the individual level.

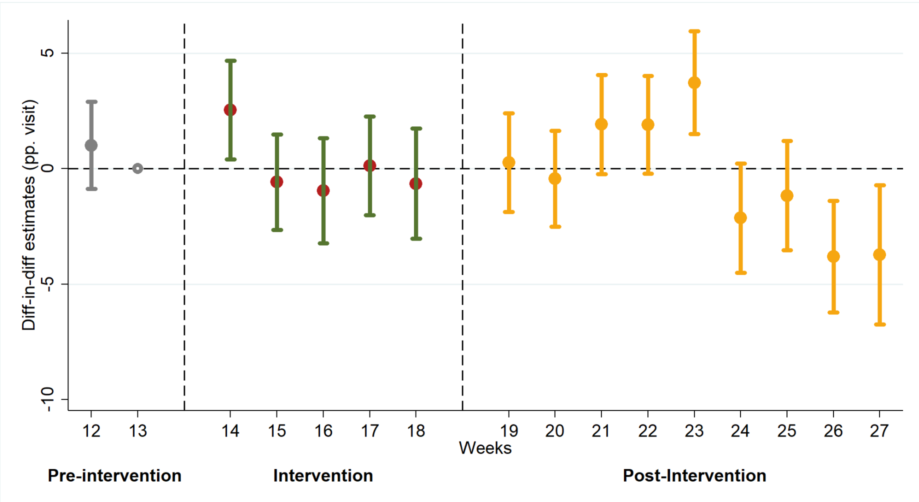


Figure 3.C.1. ITT event plot on visits, corresponding to Table 3.4 Spec. (1)

Note: Figure shows coefficients estimated in a regression analysis following Table 3.4 Spec. (1), but including weekly interaction terms. Coefficients show the estimated change in the likelihood of visiting one of the student canteens, in percentage points. Our data for week 27 includes only one day, since it was the last day of July and our sample period. Bars indicate 95% confidence intervals.

Table 3.C.3. ITT estimates of the effect of the intervention on canteen visits split by choice of main meal component

	Visit+Meat(in pp)			Visit+Veg(in pp)		
	Base	Date FE	Date+Guest FE	Base	Date FE	Date+Guest FE
Treat x Inter period	-10.26***	-10.26***	-10.26***	9.82***	9.82***	9.82***
	(0.63)	(0.63)	(0.63)	(0.67)	(0.67)	(0.67)
Treat x Post period	-1.82***	-1.82***	-1.82***	1.24**	1.24**	1.24**
	(0.49)	(0.49)	(0.49)	(0.57)	(0.57)	(0.57)
Treat	0.14	0.14		0.78	0.78	
	(0.68)	(0.68)		(0.76)	(0.76)	
Inter period	-0.60**			-4.08***		
	(0.24)			(0.27)		
Post period	-1.01***			-5.25***		
	(0.25)			(0.28)		
Constant	11.57***	10.97***	11.00***	15.86***	16.51***	16.71***
	(0.34)	(0.54)	(0.46)	(0.39)	(0.64)	(0.56)
Date fixed effects	No	Yes	Yes	No	Yes	Yes
Guest fixed effects	No	No	Yes	No	No	Yes
Guests control	2,722	2,722	2,722	2,722	2,722	2,722
Guests treated	922	922	922	922	922	922
Observations	262,368	262,368	262,368	262,368	262,368	262,368

Standard errors in parentheses

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

Notes: Regressions analyze the likelihood of visiting one of the canteens and then consuming a meat or vegetarian main component, for guests classified as treatment or control canteen guests. Weeks 1 to 11 are excluded from the analysis, since classification as a regular student canteen guest is determined based on these weeks, following the procedure described in section 3.4.1. The dependent variable in Col. (1)-(3) is a 0/1 indicator for visiting one of the student canteens and then consuming the meat option. The dependent variable in Col. (4)-(6) is a 0/1 indicator for visiting one of the student canteens and then consuming the vegetarian option. The number of guests is a bit lower than in the main analysis in Table 3.2 because treatment and control group assignment criteria are applied using only data from weeks 1 to 11. Weeks 1 to 11 are dropped from the analysis as explained in the main text as explained in the main text around Table 3.4. Col. (1) and (4) correspond to Equation 3.2 apart from the dependent variable. The Constant terms describe the likelihood of the control group visiting one of the student canteens pre-intervention and then consuming the meat meal/ the vegetarian meal. Specifications (2),(3),(5), and (6) include date-fixed effects to control for the daily changing offer of main meal components, which is common across canteens. The “Post period” and “Inter period” indicators are thus dropped due to collinearity. Specifications (4) and (6) include individual fixed effects. Standard errors are clustered at the individual level.

Table 3.C.4. ITT estimates of the effect of the intervention on canteen visits split by previous meat consumption levels

	All	By percentage of meat meals pre-intervention		
		0-10%	10%-68%	over 68%
Treat x Inter period	-0.44	6.15***	1.02	-8.39***
	(0.74)	(1.21)	(1.23)	(1.33)
Treat x Post period	-0.57	2.52*	0.72	-4.64***
	(0.76)	(1.32)	(1.32)	(1.31)
Constant	27.72***	27.93***	27.00***	28.21***
	(0.68)	(1.21)	(1.17)	(1.17)
Date fixed effects	Yes	Yes	Yes	Yes
Guest fixed effects	Yes	Yes	Yes	Yes
Guests control	2,722	852	909	961
Guests treated	922	315	306	301
Observations	262,368	84,024	87,480	90,864

Standard errors in parentheses

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

Notes: Regressions analyze the likelihood of visiting one of the canteens and then consuming a meat or vegetarian main component, for guests classified as treatment or control canteen guests. Weeks 1 to 11 are excluded from the analysis, since classification as a regular student canteen guest is determined based on these weeks, following the procedure described in section 3.4.1. Weeks 1 to 11 are dropped from the analysis as explained in the main text as explained in the main text around Table 3.4. All columns follow the same specification as in 3.2, but use guests' decision to visit one of the student canteens as the outcome variable. Col. (2) includes only guests in the lower tercile of previous meat consumption, col. (3) includes guests in the medium tercile and col. (4) includes guests in the highest tercile. Standard errors are clustered at the individual level.

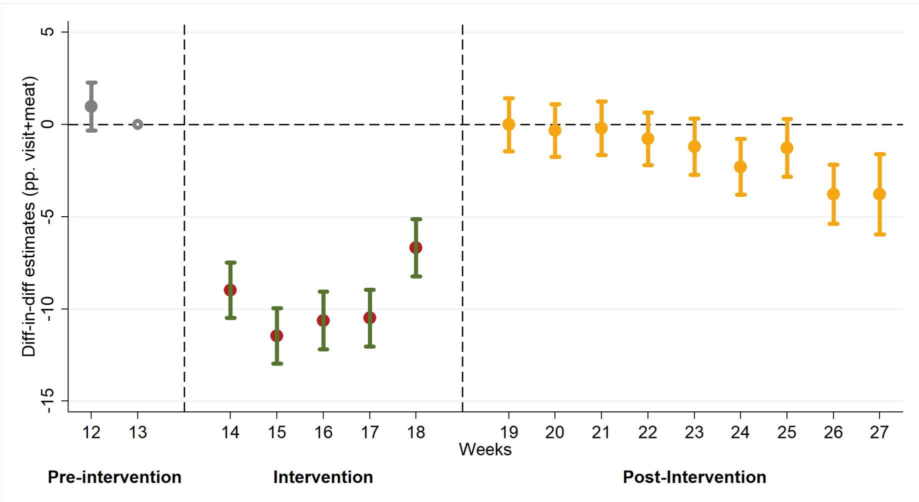


Figure 3.C.2. ITT event plot on visit+meat, corresponding to Table 3.4 Spec. (2)

Note: Figure shows coefficients estimated in a regression analysis following Table 3.4 Spec. (2), but including weekly interaction terms. Coefficients show the estimated change in the likelihood of visiting one of the student canteens and then consuming the meat main, in percentage points. Our data for week 27 includes only one day, since it was the last day of July and our sample period. Bars indicate 95% confidence intervals.

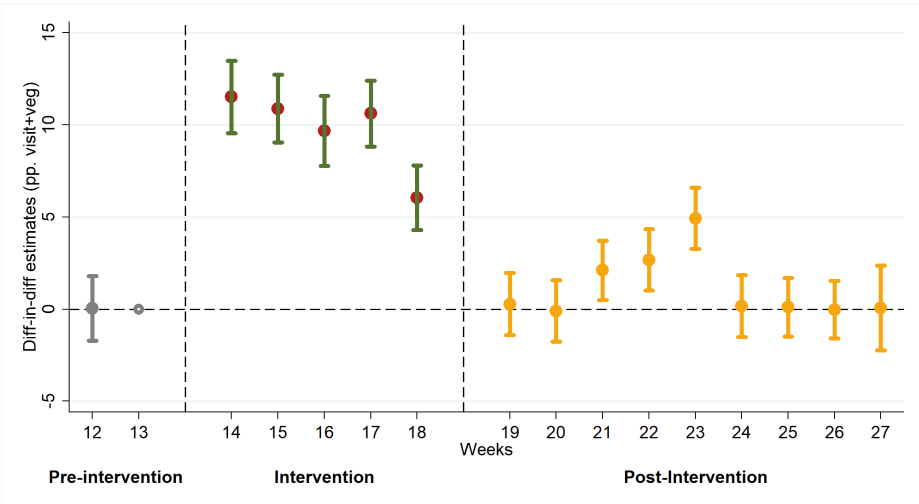


Figure 3.C.3. ITT event plot on visit+veg, corresponding to Table 3.4 Spec. (3)

Note: Figure shows coefficients estimated in a regression analysis following Table 3.5 Spec. (3), but including weekly interaction terms. Coefficients show the estimated change in the likelihood of visiting one of the student canteens and then consuming the vegetarian main, in percentage points. Our data for week 27 includes only one day, since it was the last day of July and our sample period. Bars indicate 95% confidence intervals.

Table 3.C.5. ITT estimates by visits in the intervention month, unconditional on visit

	All	At least 1 visit?		Number visits		
		yes	no	0-2	3-6	over 6
Treat x InterPeriod	-10.26***	-11.72***		-6.91***	-11.20***	-11.51***
	(0.63)	(0.72)		(0.81)	(1.02)	(1.44)
Treat x Post period	-1.82***	-1.93***	1.69***	-1.94***	-2.80***	-0.98
	(0.49)	(0.55)	(0.45)	(0.67)	(0.91)	(0.95)
Constant	11.00***	11.25***	10.39***	10.29***	10.98***	10.97***
	(0.46)	(0.48)	(0.51)	(0.49)	(0.51)	(0.52)
Date fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Guest fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Guests control	2,722	2,722	2,722	2,722	2,722	2,722
Guests treated	922	737	185	350	266	196
Observations	262,368	249,048	209,304	221,184	215,136	210,096

Standard errors in parentheses

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

Notes: Regressions analyze the likelihood of visiting one of the canteens and then consuming the meat main component, for guests classified as treatment or control canteen guests. Weeks 1 to 11 are excluded from the analysis, since classification as a regular student canteen guest is determined based on these weeks, following the procedure described in section 3.4.1. The dependent variable is a 0/1 indicator for visiting one of the student canteens and then consuming the meat option. The number of guests is a bit lower than in the main analysis in Table 3.5 because treatment and control group assignment criteria are applied using only data from weeks 1 to 11. Weeks 1 to 11 are dropped from the analysis as explained in the main text as explained in the main text around Table 3.4. Col. (1) corresponds to Col. (2) in Table 3.4 and includes the full sample. Col. (2) restricts the sample to control guests and treatment guests who made at least one purchase with their personalized card during the intervention period. Col. (3) includes only control guests and those treatment guests for whom we did not register such a purchase – Note, however, that they might have still visited the canteens during the time frame, but used a different payment method than their personalized card. Col. (4) - Col. (6) restricts the sample of treatment guests by number of visits registered during the intervention period. Standard errors are clustered at the individual level.

Table 3.C.6. ITT estimates by previous meat consumption, unconditional on visit

	All	By percentage of meat meals pre-intervention		
		0-10%	10%-68%	over 68%
Treat x InterPeriod	-10.26***	-0.86***	-9.97***	-20.59***
	(0.63)	(0.17)	(0.83)	(1.45)
Treat x Post period	-1.82***	-0.40*	-0.77	-4.67***
	(0.49)	(0.21)	(0.78)	(1.18)
Constant	11.00***	0.00	8.07***	24.01***
	(0.46)	(0.06)	(0.75)	(1.11)
Date fixed effects	Yes	Yes	Yes	Yes
Guest fixed effects	Yes	Yes	Yes	Yes
Guests control	2,722	852	909	961
Guests treated	922	315	306	301
Observations	262,368	84,024	87,480	90,864

Standard errors in parentheses

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

Notes: Regressions analyze the likelihood of visiting one of the canteens and then consuming the meat main component, for guests classified as treatment or control canteen guests. Weeks 1 to 11 are excluded from the analysis, since classification as a regular student canteen guest is determined based on these weeks, following the procedure described in section 3.4.1. The dependent variable is a 0/1 indicator for visiting one of the student canteens and then consuming the meat option. The number of guests is a bit lower than in the main analysis in Table 3.5 because treatment and control group assignment criteria are applied using only data from weeks 1 to 11. Weeks 1 to 11 are dropped from the analysis as explained in the main text as explained in the main text around Table 3.4. Col. (1) corresponds to Col. (2) in Table 3.4 and includes the full sample. Col. (2) restricts the sample to guests who consumed meat in 0% to 10% of their meals pre-intervention, Col. (3) to guests who consumed meat in 10% to 68% of their meals pre-intervention, and Col. (4) to guests who consumed meat in over 68% of their meals pre-intervention. Standard errors are clustered at the individual level.

Appendix 3.D Details on the experimental set-up

3.D.1 Materials from the canteen

Figure 3.D.1. Instagram Post announcing the vegetarian month

Note: Published on the 26th of April. Translation: Announcement: This May is going to be green! Vegan-vegetarian month from the 2nd of May until the 2nd of June in the canteen "am Hofgarten". For five weeks, the canteen "am Hofgarten" will serve only vegan and vegetarian dishes — from breakfast to dessert. This month is part of a comprehensive re-strategizing for a sustainable canteen of the future, which the student canteens in Bonn are currently engaging in.

3.D.2 Surveys

To make survey participation attractive, survey participants had the chance to win one of 20 50€coupons for the student canteen by participating in one of the surveys.[30] We advertised the survey by distributing leaflets in front of the treatment and the larger control student canteen (see figures 3.D.2 and 3.D.3. It is common for students and student groups to advertise surveys, projects, and events in this manner. Further, the experimental lab at the University of Bonn sent out an e-mail to its entire participant pool advertising survey participation. The e-mail texts are shown below. Finally, respondents of the first survey could indicate their e-mail address at the end of the first survey and agree to be contacted directly by e-mail for the second survey. We advertised the survey as a survey on life as a student in Bonn.

Advertisement - Leaflets

30. Per survey ten vouchers were randomly distributed among survey respondents.

Figure 3.D.2. Leaflet advertising participation in the first survey

Note: Translation: Your opinion is wanted! Answer a short survey about daily student life in Bonn! Among all participants we are raffling 10 x 50 Euro student canteen credit! Scan the QR-Code to participate now. (*) The raffle will include all participants who have completely filled out the survey by the 18th of April.

Figure 3.D.3. Leaflet advertising participation in the second survey

Note: Translation: New survey on daily student life in Bonn! Your opinion is wanted! Answer a short survey! Among all participants we are raffling 10 x 50 Euro student canteen credit! Scan the QR-Code to participate now. (*) The raffle will include all participants who have completely filled out the survey by the 16th of July.

Advertisement - Emails

Text sent to the participant pool of the BonnEconLab to advertise the first survey (translated from German)

* Study on Daily Student Life in Bonn: Participate and Win Credit for Your Canteen Card!*

Hello XX,

We would like to bring to your attention a study currently being conducted by a doctoral student at the Bonn Graduate School of Economics:

Participate and Win Credit for Your Cafeteria Card!

Dear Students,

How do you perceive your daily life as a student in Bonn? What are your views on your fellow students?

This is being explored in a current study by the Bonn Graduate School of Economics at the University of Bonn. Filling out the online survey takes five minutes, and there are ten €50 credits for the canteen card up for grabs! All questionnaires completed in full by April 16, 2023, will be entered into the draw. Participation is only possible if you own a canteen card.

Here is the link to the survey: https://studienalltag.econ.uni-bonn.de/room/umfrage

Please Note:

- The study is not being conducted by BonnEconLab, so registration is not through our participation database. You can access the survey directly via the link above.
- You are, of course, welcome to share the survey invitation with your fellow students.

Best regards,

Your BonnEconLab

Text sent to the participant pool of the BonnEconLab to advertise the second survey (translated from German)

* New Study on Daily Student Life in Bonn: Participate and Win Credit for Your Canteen Card!

Hello XX,

We would like to bring to your attention a study currently being conducted by a doctoral student at the Bonn Graduate School of Economics:

Participate and Win Credit for Your Cafeteria Card!

Dear Students,

How do you perceive your daily life as a student in Bonn? This is being explored in a current study by the Bonn Graduate School of Economics at the University of Bonn. Specifically, it looks at your views on studying abroad as part of your degree, your attitude towards involvement in university politics, and your opinions on some of the services offered by the Student Services Organization.

This online survey is the second of two surveys being conducted in the context of the study. It doesn't matter whether you have participated in the first survey or not: you are welcome to participate in this survey and the associated prize draw. We look forward to your participation!

Filling out the survey takes five minutes, and there are ten €50 credits for the canteen card up for grabs! All questionnaires completed in full by July 16, 2023, will be entered into the draw. Participation is only possible if you own a canteen card.

Here is the link to the survey: https://studienalltag.econ.uni-bonn.de/room/umfrage

Please Note:

- The study is not being conducted by BonnEconLab, so registration is not through our participation database. You can access the survey directly via the link above.
- You are, of course, welcome to share the survey invitation with your fellow students.

Best regards,
Your BonnEconLab

Core part of the survey: Elicitation of the perceived social norm

Questions on studying abroad and involvement in student politics were added to obfuscate the purpose of the survey. In the elicitation of the perceived descriptive norm questions, the slider only appeared once participants clicked on the bar, to avoid giving survey participants a reference point.

In the following questions, we are interested in your perception of the University of Bonn and its students.

Next

Please make a guess: Which percentage of the students of the University of Bonn spends at least one semester abroad during their studies?

Please click on the blue bar to make the slider visible.

0% 100%

Next

Please make a guess: Which percentage of the students of the University of Bonn eats a fish or meat containing meal for lunch on a typical University day?

Please click on the blue bar to make the slider visible.

0% 100%

Next

Please make a guess: Which percentage of the students of the University of Bonn eats a fish or meat containing meal for lunch on a typical University day?

Please click on the blue bar to make the slider visible.

0% 100%

You guess: **44%** eat a fish or meat containing meal for lunch on a typical University day.

Next

Please make a guess: Which percentage of the students of the University of Bonn is **not** involved in university politics or student associations?

Please click on the blue bar to make the slider visible.

0% 100%

Next

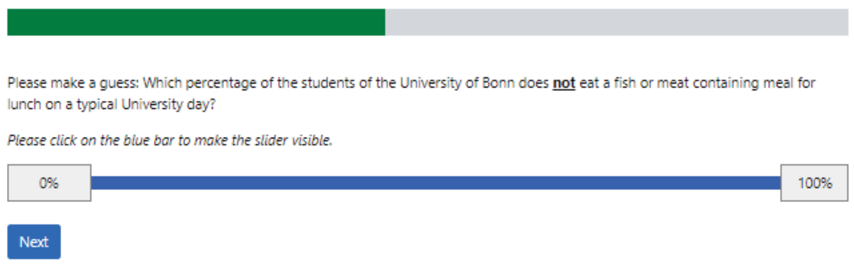
Please make a guess: Which percentage of the students of the University of Bonn does **not** spend at least one semester abroad during their studies?

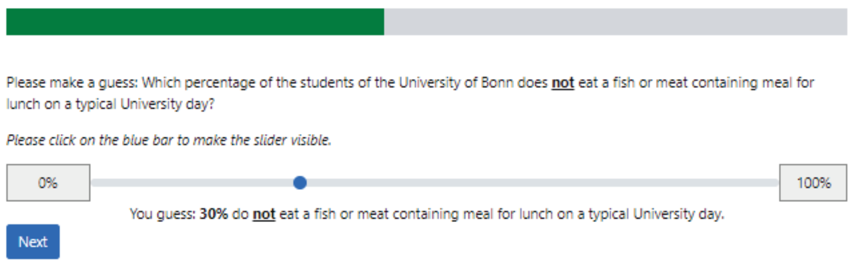
Please click on the blue bar to make the slider visible.

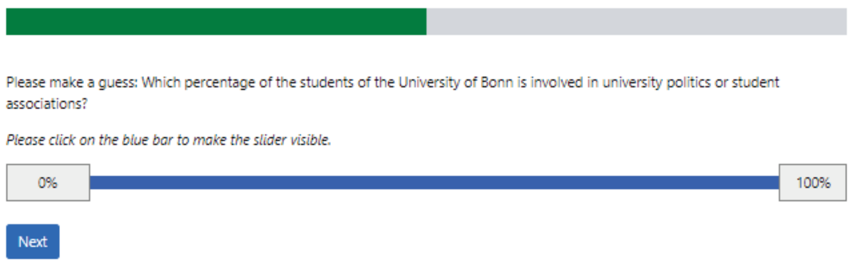
0% 100%

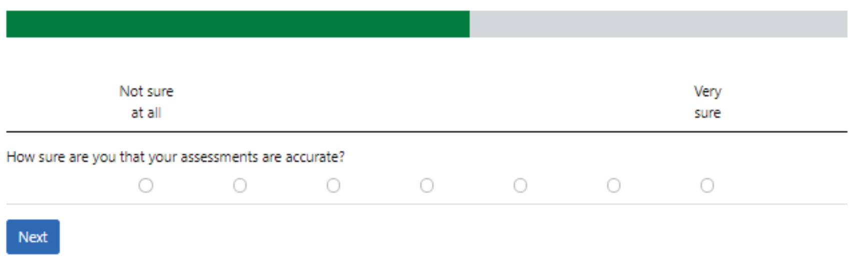
Next

Elicitation of the perceived descriptive norm









Elicitation of the personal injunctive norm and perceived social injunctive norm

In the following questions, we are interested in your **personal assessments.**

Next

How socially appropriate do you find it if a student of the University of Bonn ...

	Not socially appropriate						Socially appropriate
... does <u>**not**</u> spend a semester abroad during their studies.							

Next

What do you think was the most common answer **from the other survey participants** to the previous question? That is, how socially appropriate do most other participants find it if a student at the University of Bonn...

	Not socially appropriate						Socially appropriate
...<u>**does not**</u> spend a semester abroad during their studies.							

Next

How socially appropriate do you find it if a student of the University of Bonn ...

	Not socially appropriate						Socially appropriate
... eats a fish or meat containing meal for lunch on 1 out of 5 typical University days.							
... eats a fish or meat containing meal for lunch on 3 out of 5 typical University days.							
... eats a fish or meat containing meal for lunch on 5 out of 5 typical University days.							

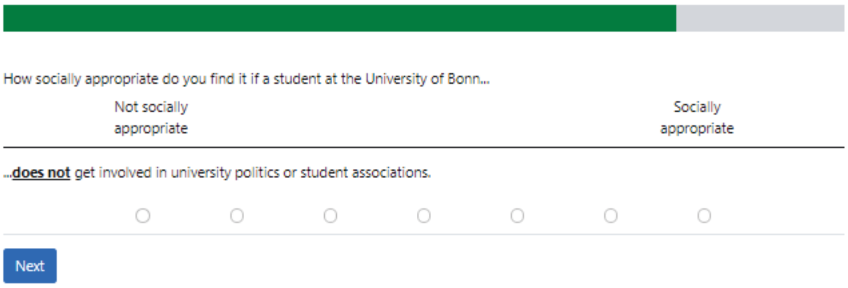
Next

What do you think was the most common answer **from the other survey participants** to the previous question? That is, how socially appropriate do most other participants find it if a student at the University of Bonn ...

	Not socially appropriate						Socially appropriate
... eats a fish or meat containing meal for lunch on 1 out of 5 typical University days.							
... eats a fish or meat containing meal for lunch on 3 out of 5 typical University days.							
... eats a fish or meat containing meal for lunch on 5 out of 5 typical University days.							

Next

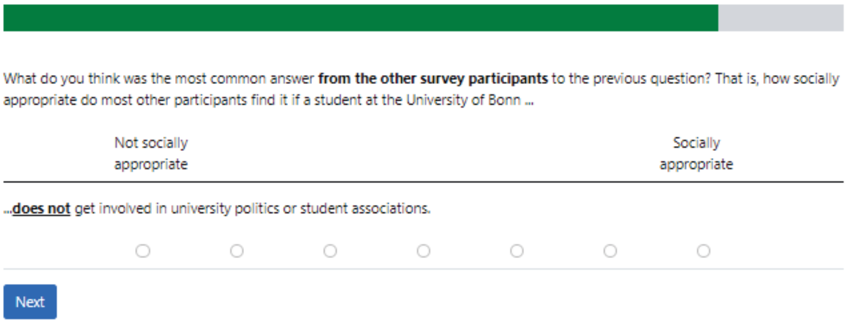
Elicitation of the personal injunctive norm and perceived social injunctive norm



How socially appropriate do you find it if a student at the University of Bonn...

	Not socially appropriate						Socially appropriate
...**does not** get involved in university politics or student associations.	○	○	○	○	○	○	○

Next



What do you think was the most common answer **from the other survey participants** to the previous question? That is, how socially appropriate do most other participants find it if a student at the University of Bonn ...

	Not socially appropriate						Socially appropriate
...**does not** get involved in university politics or student associations.	○	○	○	○	○	○	○

Next

Elicitation of the perception of the intervention

In the following, we are interested in your perception of the vegan-vegetarian month at the Mensa am Hofgarten, which likely represented a change in the daily routine of some students.

Did you visit the Mensa at Hofgarten during the vegan-vegetarian month at least once?

- ○ Yes
- ○ No

Next

If YES:

Do you believe that after the vegan-vegetarian month you opt for the vegetarian main component more often than before the vegan-vegetarian month?

- ○ Yes
- ○ No
- ○ No, because I always chose the vegetarian main component even before the vegan-vegetarian month.

	Not sure at all						Very sure
How sure are you of this answer?	○	○	○	○	○	○	○

Next

If YES:

What do you think are the reasons you opt for the vegan-vegetarian main component more frequently after the intervention month? Please check all options that apply to **you personally**.

I opt for the vegan-vegetarian main component more frequently because...

- ☐ ...more students have been eating more vegetarian/vegan dishes since the intervention month.
- ☐ ...my friends have been eating more vegetarian/vegan dishes since the intervention month.
- ☐ ...I discovered new vegan/vegetarian dishes during the intervention month that I like.
- ☐ ...it has become more of a habit for me to eat vegan/vegetarian.
- ☐ ...the vegan/vegetarian options have improved since the intervention month.
- ☐ ...I started to consume meat more outside the cafeteria during the intervention month and have continued to do so after the month ended.
- ☐ None applicable
- ☐ Prefer not to respond

Next

Space for comments, in case you have additional reasons or if the month has affected you in other ways:

Next

All participants regardless of previous answers:

For more climate and animal protection, I wish the student canteens in Bonn would ...

	Does not apply at all				Completely applies
... offer one vegetarian day a week.	○	○	○	○	○
... offer more vegan/vegetarian meals.	○	○	○	○	○
... offer exclusively vegan and vegetarian meals for one month every year.	○	○	○	○	○

Next

References

Acland, Dan, and Matthew R Levy. 2015. “Naiveté, projection bias, and habit formation in gym attendance.” *Management Science* 61 (1): 146–60. [174]

Andre, Peter, Teodora Boneva, Felix Chopra, and Armin Falk. 2021. “Fighting climate change: The role of norms, preferences, and moral values.” Working Paper, Working Paper Series 14518. IZA Discussion Paper. https://docs.iza.org/dp14518.pdf. [197]

Bilén, David. 2022. “Do carbon labels cause consumers to reduce their emissions? Evidence from a large-scale natural experiment.” Working paper. Mimeo, Gothenburg University. [171]

Bonnet, Céline, Zohra Bouamra-Mechemache, Vincent Réquillart, and Nicolas Treich. 2020. “Regulating meat consumption to improve health, the environment and animal welfare.” *Food Policy* 97: 101847. [171]

Bruhin, Adrian, Lorenz Goette, Simon Haenni, and Lingqing Jiang. 2021. “Oops!... I did it again: Understanding mechanisms of persistence in prosocial behavior.” Working Paper, Working Paper Series 15642. CEPR Discussion Papers. https://doi.org/10.3386/w28707. [174]

Byrne, David P, Lorenz Goette, Leslie A Martin, Lucy Delahey, Alana Jones, Amy Miles, Samuel Schob, Thorsten Staake, Verena Tiefenbeck, et al. 2021. “The habit-forming effects of feedback: Evidence from a large-scale field experiment.” Working Paper, Working Paper Series 285. CRC TR 224 Discussion Paper Series University of Bonn and University of Mannheim. https://www.crctr224.de/research/discussion-papers/archive/dp285. [173]

Castillo, Marco, and Ragan Petrie. 2023. “Incentivized learning and attention-driven treatment effects.” Working Paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4261298. [173]

Charness, Gary, Nir Chemaya, and Dario Trujano-Ochoa. 2023. “Learning your own risk preferences.” *Journal of Risk and Uncertainty* 67 (1): 1–19. [172, 187]

Cialdini, Robert B., and Michael R. Trost. 1998. “Social influence: Social norms, conformity and compliance.” In *The handbook of social psychology,* 4th ed., edited by Daniel T. Gilbert and Susan T. Fiske, 2: 151–92. New York, NY: McGraw Hill. [173]

Dannenberg, Astrid, Charlotte Klatt, and Eva Weingärtner. 2024. “The effects of social norms and observability on food choice.” *Food Policy*. [176]

Die Welt. 2023. “Großer Ärger für CDU-Ministerin nach Teilnahme am Veggie-Monat.” Accessed March 18, 2024. https://www.welt.de/politik/deutschland/plus245384312/Ernaehrung-Grosser-Aerger-fuer-CDU-Ministerin-wegen-Teilnahme-am-Veggie-Monat.html. [172, 197]

Garnett, Emma E, Andrew Balmford, Theresa M Marteau, Mark A Pilling, and Chris Sandbrook. 2021. “Price of change: Does a small alteration to the price of meat and vegetarian options affect their sales?” *Journal of Environmental Psychology* 75: 101589. [171, 173]

Garnett, Emma E, Theresa M Marteau, Chris Sandbrook, Mark A Pilling, and Andrew Balmford. 2020. “Order of meals at the counter and distance between options affect student cafeteria vegetarian sales.” *Nature Food* 1 (8): 485–88. [171, 173]

Godfray, H Charles J, Paul Aveyard, Tara Garnett, Jim W Hall, Timothy J Key, Jamie Lorimer, Ray T Pierrehumbert, Peter Scarborough, Marco Springmann, and Susan A Jebb. 2018. “Meat consumption, health, and the environment.” *Science* 361 (6399): eaam5324. [171]

Goetz, Alexander, Harald Mayr, and Renate Schubert. 2022. “Beware of side effects? Spillover evidence from a hot water intervention.” Working Paper. https://congress-files.s3.amazonaws.com/2022-07/BewareOfSideEffects.pdf. [173]

Gravert, Christina, and Ganga Shreedhar. 2022. “Effective carbon taxes need green nudges.” *Nature Climate Change* 12 (12): 1073–74. [172, 187]

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. [171]

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. [171]

Kölner Stadt Anzeiger. 2023. “NRW-Ministerin rechtfertigt sich nach Teilnahme an Veggie-Monat.” Accessed March 18, 2024. https://www.ksta.de/politik/nrw-politik/nrw-ministerin-nimmt-an-veggie-monat-teil-kritik-aus-der-cdu-575117. [172, 197]

Krupka, Erin L., and Roberto A. Weber. 2013. "Identifying social norms using coordination games: Why does dictator game sharing vary?" *Journal of the European Economic Association* 11 (3): 495–524. [173, 176, 189]

Kurz, Verena. 2018. "Nudging to reduce meat consumption: Immediate and persistent effects of an intervention at a university restaurant." *Journal of Environmental Economics and management* 90: 317–41. [171, 173, 196]

Lohmann, Paul, 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. [171, 173]

Lohmann, Paul, Elisabeth Gsottbauer, James Farrington, Steve Human, and Lucia Reisch. 2024. "An online randomised controlled trial of price and non-price interventions to promote sustainable food choices on food delivery platforms." Working Paper. %7Bhttps://pmlohmann.com/project/delivery-apps/delivery-apps.pdf%7D. [171]

Meier, Johanna, Mark A Andor, Friederike C Doebbe, Neal R Haddaway, and Lucia A Reisch. 2022. "Do green defaults reduce meat consumption?" *Food Policy* 110: 102298. [171]

Nyborg, Karine, John M Anderies, Astrid Dannenberg, Therese Lindahl, Caroline Schill, Maja Schlüter, W Neil Adger, Kenneth J Arrow, Scott Barrett, Stephen Carpenter, et al. 2016. "Social norms as solutions." *Science* 354 (6308): 42–43. [187]

Poore, Joseph, and Thomas Nemecek. 2018. "Reducing food's environmental impacts through producers and consumers." *Science* 360 (6392): 987–92. [171]

Schulze-Tilling. 2023. "Changing consumption behavior with carbon labels: Causal evidence on behavioral channels and effectiveness." Working Paper. https://anna-schulze-tilling.github.io/papers/Schulze_Tilling_JMP.pdf. [171, 173, 197]

Springmann, Marco, Michael Clark, Daniel Mason-D'Croz, Keith Wiebe, Benjamin Leon Bodirsky, Luis Lassaletta, Wim De Vries, Sonja J Vermeulen, Mario Herrero, Kimberly M Carlson, et al. 2018. "Options for keeping the food system within environmental limits." *Nature* 562 (7728): 519–25. [171]

Stigler, George J, and Gary S Becker. 1977. "De gustibus non est disputandum." *American Economic Review* 67 (2): 76–90. [172, 174, 187]

t-online. 2023. "CDU-Ministerin nimmt an Veggie-Tag teil – und muss sich rechtfertigen." Accessed March 18, 2024. https://www.t-online.de/region/koeln/id_100177642/nrw-cdu-ministerin-isst-veggie-gericht-und-muss-sich-dafuer-rechtfertigen.html. [172, 197]

tagesschau.de. 2023. "Würstchen zum "wahren Preis" 88 Prozent teurer." Accessed March 18, 2024. https://www.tagesschau.de/wirtschaft/verbraucher/penny-umweltfolgekosten-102.html. [171]