Do People Respond to the Climate Impact of their Behavior? The Effect of Carbon Footprint Information on Grocery Purchases

Toke R. Fosgaard[1,4] · Alice Pizzo[2,5] · Sally Sadoff[3]

Abstract

Food production is a primary contributor to climate change, and one way to mitigate its effect is through consumption choices. Finding the most effective way to achieve emissions reduction via consumers behavioral change has recently raised policy-makers' interest but experimental evidence about this is still scarce. In this study, we examine the impact of individualized information about greenhouse gas emissions on grocery purchases. Using a randomized field experiment, we compare the effects of individualized information on the carbon footprint of grocery purchases to individualized information on grocery spending provided through a smartphone app. Compared to the spending information, the carbon footprint information decreases emissions from groceries by 27% in the first month of treatment, with 45% reductions in emissions from beef, the highest emissions food group. Treatment effects fade in the longer run along with app engagement but persist among those engaged with the app. Our results suggest that the provision of emissions information, in particular when paired with sufficient engagement, is a promising avenue for policies to turn food consumption greener.

✉ Alice Pizzo
api.msc@cbs.dk

Toke R. Fosgaard
trefo@dtu.dk

[1] Department of Food and Resources Economics, University of Copenhagen, Copenhagen, Denmark

[2] Department of Management, Society and Communication, Copenhagen Business School, Copenhagen, Denmark

[3] Rady School of Management, University of California San Diego, San Diego, CA, USA

[4] Department of Technology, Management and Economics, The Technical University of Denmark, Kongens Lyngby, Denmark

[5] Present Address: Dalgas Have 15, 2000 Frederiksberg, Denmark

1 Introduction

It is widely agreed among the scientific and institutional communities that there is an urgent need to reduce greenhouse gas (GHG) emissions that contribute to climate change (IPCC 2018). Given that political constraints and the influence of industrial lobbying may limit the extent to which GHG emissions can be addressed using price mechanisms, there is growing interest in behavioral and informational interventions aimed at shifting individual consumption (Arora and Mishra 2021; IPCC 2022). According to Cafaro (2011), consumers have the potential to reduce their carbon footprint by 15 billion tons by 2060 through the adoption of new consumption practices.

Prior work has successfully examined field informational interventions to reduce energy usage and increase demand for more energy-efficient technologies at the household level, such as cars, appliances, and lightbulbs (Guthrie et al. 2015; Hummel and Maedche 2019; Wynes et al. 2018). Indeed, household consumption is estimated to be responsible for up to 72% of global greenhouse gas emissions, which are mostly attributed to transport, housing, and food (Hertwich and Peters 2009; UNEP 2017). Of this, an estimated 20–30% of all GHG emissions originate from food production, making it a critical target for GHG reductions (Commission 2006; IPCC 2018; Vermeulen et al. 2012; Willett et al. 2019). The share of emissions due to food production is similar, if not higher, than household energy usage, which has been a primary focus of prior informational interventions (Goldstein et al. 2020). Informational interventions could potentially also be effective for food consumption, as investigated by recent literature (Dannenberg and Weingärtner 2023; Kanay et al. 2021; Lohmann et al. 2022; Muller et al. 2017; Panzone et al. 2021a; Panzone et al. 2021b; Perino et al. 2014; Suchier et al. 2023).

First, while there is growing awareness that the food production process contributes to climate change, information on the carbon footprint of particular food groups is not readily available and people generally underestimate the impacts (Camilleri et al. 2019; Macdiarmid et al. 2016). For example, despite recent attention to beef as a high emissions food group, many may not be aware of the magnitude: producing a single serving of beef (100 g, 3.53 oz, or 0.22 pounds) generates GHG emissions equivalent to driving 49.86 km (30.98 miles), about the average daily commute in the U.S. (Federal Highway Administration 2017). Second, because emissions vary greatly by food group, shifts in composition can have a large impact (Garnett 2011). For example, emissions related to the production of ground beef are ten times higher than those for chicken (Poore and Nemecek 2018). More generally, interventions that encourage individuals to switch to lower-carbon substitutes or reduce their meat consumption may contribute significantly to the reduction of greenhouse gas emissions: eating fewer animal products can reduce individuals' total carbon footprint by an estimated 22% (Lacroix 2018). Therefore, there is significant potential to reduce carbon emissions by promoting more sustainable food choices among consumers. Furthermore, research shows that consumers are willing to make changes to their dietary choices to reduce their environmental impact, particularly if supported by informative labeling or educational campaigns (Visschers and Siegrist 2015). Nevertheless, food consumption behaviors are very difficult to shift (Liu et al. 2014) and so may not be responsive to light-touch informational interventions.

The goal of this study is to examine whether it is possible to shift actual consumer behavior in response to the environmental impact of food choices. We implement a randomized field experiment among a national sample of the Danish population to test the impact of providing individuals with information about the GHG emissions of their grocery

purchases. We provide the information through a novel smartphone application (app) that collects data on individuals' grocery purchases both before and during our 19-week intervention. We compare the impact of a "Carbon" app that provides item-level carbon-equivalent emissions of individuals' grocery purchases to a "Spending" app that provides item-level cost information. The Carbon app applies GHG emissions converted into kilograms of Carbon Dioxide equivalents (kg of CO_2 e). Both apps provide real-time individualized weekly, monthly, and yearly feedback on purchases, broken down by categories and single items.[1] To measure revealed preferences for the two types of information, we also include a treatment group that makes both apps available to participants. Indeed, while we see growing interest in the use of smart technology to provide personalized feedback and shape individual behavior, little is known about the use or impact of such mobile apps.[2]

We observe 175,146 item-level grocery purchases for 258 participants over a 19-week baseline period and a 19-week treatment period. Our primary outcome of interest is the carbon-equivalent emissions of participants' weekly grocery purchases. We estimate the impact of the Carbon app on carbon emissions using a difference-in-difference analysis with randomization as our assignment. In the first month of treatment, participants who receive the Carbon app significantly reduce carbon emissions from groceries relative to participants who receive the Spending app. We estimate an average decrease of 5.8 kg (kg) in weekly carbon emissions ($p=0.003$). The size of the reduction corresponds to a 27% of the baseline emission. The magnitude is equivalent to reducing driving by 49 km (30 miles) per week. During this period, the Carbon app decreases both overall purchases and emissions per purchase, with an estimated 45% decrease in emissions from beef ($p=0.019$), which has been the focus of prior work 6). However, over the full 19-week treatment period, the impact of the Carbon app is smaller – an estimated 2.4 kg per week decrease – and not statistically significant.

The pattern of treatment effects over time mirrors the pattern of app usage over time. Engagement in the app is concentrated in the first four weeks of treatment with over half of total app usage taking place in the first month. App usage is similar in the first four weeks for the Carbon and Spending treatments with participants in both groups checking the app on average a little over once a week. Over the full treatment period, app usage is lower in the Carbon treatment than in the Spending treatment, though the differences are not statistically significant. We find similar results for the treatment group that received access to both apps. Providing both apps increases total app usage but crowds out the usage of the individual apps, particularly for longer-term usage of the Carbon app. Over the 19-week treatment period, usage of the Spending app is almost 40% higher than the Carbon app ($p=0.074$). These results suggest a weak preference for spending information compared to emissions information over the longer term.

We make several contributions to the existing literature. Our experiment is the first to test the impact of an *individualized* intervention, in a real-life setting, aimed at decreasing the environmental impact of regular grocery purchases. Previous studies have tested

---

[1] Carbon emissions were calculated the app company and based on the 'Food Climate Imprint' publication from the Danish Ministry of the Environment and Food, the Unilever's CO2 calculator, and the tool Food-emissions.com.

[2] Milne-Ives et al. (2020) review a larger literature on the use of mobile apps in the context of physical and mental health, while Ytreberg et al. (2023) find that several retailers-based studies use apps to convey decision information type nudges in their effort to map the digital climate nudges in Nordic online grocery stores. In an environmental context, Brandon et al. (2022) discuss the research on smart thermostats for home energy usage, which include the option to program the thermostat via a smartphone app.

the impact of carbon labels and information on a limited set of products, on a one-time purchase at an experimental store, through self-reported data, on the intention to consume or in the lab (Brownback et al. 2023; Brunner et al. 2018; Camilleri et al. 2019; Elofsson et al. 2016; Kanay et al. 2021; Muller et al. 2017; Osman and Thornton 2019; Panzone et al. 2021a; Panzone et al. 2021b; Perino et al. 2014; Spaargaren et al. 2013; Suchier et al. 2023; Vlaeminck et al. 2014). In related work, for instance, researchers find that informing students about the environmental consequences of meat consumption reduced the demand for meat at educational institutions (Jalil et al. 2020; Lohmann and Gsottbauer 2022).

We are also the first to estimate the effect of providing individuals with *real-time information* about the climate impacts of their groceries. While interventions aimed at changing environmental behaviors are motivated by the externalities of energy usage, prior work has not provided direct information on how individual food behavior directly translates into environmental impact in a real-time feedback fashion. Instead, prior interventions have largely provided information about individuals' direct costs (Allcott and Knittel 2019; Allcott and Taubinsky 2015; Davis and Metcalf 2016; Jessoe and Rapson 2014). Other work has provided individualized feedback in terms of usage rather than environmental impact or carbon emissions (Brandon et al. 2019; Hahn and Metcalfe 2016; List and Price 2016). Tiefenbeck et al. (2018) for instance couple individualized feedback on water usage with a picture of a polar bear on an ice cap that shrinks as water usage increases (Tiefenbeck et al. 2018). A comparable study employing an app with a carbon footprint calculator of all different household activities has been recently published (Enlund et al. 2023); similarly, the gamification of individual CO_2 emissions related to mobility has been explored by means of a smartphone app by a study at the working paper stage (Goetz et al. 2022). Related work tests general messages about the need for conservation but does not provide individualized information (Ferraro and Price 2013; Ito et al. 2018). Our study examines whether people are responsive to individualized feedback about the externalities of their behavior.

Third, our smart technology allows us to directly measure engagement with the informational interventions. Participants can only receive the information if they open the app, which we track throughout the treatment period. Prior studies that provide individualized feedback over time – for example through smart meters, home energy reports, or robocalls – are not able to measure whether people actually hear or read the information (Allcott and Rogers 2014; Brandon et al. 2019; Ferraro and Price 2013). We measure revealed preferences for the emissions information by comparing engagement with the Carbon app to engagement with the Spending app. Our analysis also examines the relationship between app engagement and treatment impacts on behavior.

Taken together, our findings demonstrate that providing people with personalized emissions information can affect their food purchasing behavior. However, our results also suggest that the impact of the informational intervention requires sustained engagement. In periods with regular app usage, we find meaningful treatment effects on carbon emissions, which decline along with app engagement. We also find suggestive evidence that the impact on emissions is sustained over the longer term for users who remain engaged with the app. The results of this study provide a foundation for the potential of using low-cost, highly scalable informational interventions to shift food purchasing behavior in a green direction.

The rest of the paper is organized as follows. The second section describes the experimental design. The third section discusses the results, and the fourth section concludes.

2 Experimental Design, Materials, and Methods

2.1 Sample and Recruitment

We recruited a national Danish sample to participate in the study.[3] To do so, we worked with Statistics Denmark, which is the Danish governmental organization that creates statistics on Danish society. On our behalf, Statistics Denmark selected a representative sample of 100,000 Danish adults. In two waves, mid-January 2020 and mid-June 2020, we sent an invitation to participate in our study through the mandatory public electronic mail system in Denmark (only 96,324 were effectively reached). The invitation letters included a description of the research project and the requirements for participation, which consisted of answering a brief survey, downloading an app, and set-up a profile to use it. We varied the framing of the language describing the purpose of the study (the title and one sentence in the description changed) across letters using an environmental frame, an economic frame, or a neutral frame to study any potential differences in motivations to participate along these dimensions (see Figure 4 in the Appendix for letters). No significant difference in study participation or impact on consumption was found across the invitation framings. The variation of the letter framings did not depend on and was not linked to the randomization into treatment. We do acknowledge a selection into the study. We give an impression of the ongoing selection by comparing the climate attitudes retrieved from the baseline survey of those who completed the survey (n = 2711) and those who fully participated in the study by using the app (our actual sample, n = 258): we report this evidence in the Descriptive Statistics subsection below.

In order to participate in the study, participants clicked on a link at the bottom of the letter. The link took them to a survey about perceptions and attitudes toward food in relation to health, the environment, and money. We also asked participants to rank five food items (potatoes, beef, chicken, cheese, and orange) on three dimensions: pollution, cost, and health. The survey questions and tasks are listed in Table 4.

Upon survey completion, we randomly assigned participants to receive the Carbon app, the Spending app, or both apps. Respondents downloaded the assigned app(s) to their smartphone, activated an app-user profile that included optional demographic questions, and connected the app to an e-receipt system of widespread use in Denmark. The e-receipt system collects data from all supermarkets in the country using individual payment card data. Participants who did not have the e-receipt system set up yet could easily sign up; a quick guide to doing so was provided in the online survey platform. The automatic e-receipt system registers all food purchases at the individual level without the need for any manual entries. Once set up, none of our participants disconnected the data collection, suggesting that the participants did not perceive the data collection as problematic after experiencing the app(s). Moreover, our data collection provides historical data on grocery purchases prior to the intervention. The app platforms therefore both served as the data collection device and the information provider. Because the app was required to collect the outcome data on grocery purchases, we were not able to include a group that received no app and no information but, given the longitudinal structure of the data, we can observe behavior before the information was provided. We fix the time corresponding to when the participants enter the experiment and start using the app as their individual t_0.


2.2 Treatment and Control Apps

The two apps are identical, the only difference is the consumption metric shown. Respondents in the *Carbon* treatment received the Carbon app, which provides an overview of the CO_2 emission associated with their food purchases. In the *Spending* treatment, participants received the Spending app, which provides an overview of the expenditures related to their own food consumption. In the *Both* treatment, respondents received both apps. The two apps were developed by the same company (*Spenderlog*) and share structural visual design. In both apps, the overview is organized by food groups (e.g., dairy, meat and fish, fruit, and vegetables), individual foods (e.g., cheese, fresh milk, beef, chicken, apples), and item-level purchases. Users can see weekly, monthly, or yearly summaries. It was not possible to include a control group that used a "white" app: the app not only served as a data collector but also provided information, making it challenging to have a separate group solely dedicated to data recording without any informational features. The apps also show comparisons of the user with other households active in the app for all of Denmark (default), as well as by region, household income, household type (e.g., apartment, house), and family type (e.g., single, couple, couple with kids). One additional app feature allows users to set any kind of quantitative goal in relation to their groceries (e.g., reduction in candy consumption). Therefore, the effect of the Carbon treatment could be the sum of the emission information provision, social comparison, and goal setting. We would, however, like to note that only 6 participants were setting climate-related goals, suggesting that the treatment is unlikely influenced by this. Furthermore, while the potential for engaging in social comparisons was present across all treatments, variations in the content of these comparisons may have arisen from the differing information provided and subsequent behavioral responses. We do not have access to details regarding these social comparisons, so we cannot observe the impact. However, it is worth noting that within our sample, participants with initial emission levels above the median exhibited no distinct treatment effects compared to those below the median. This observation indicates that social comparisons might not have had a central influence.

Figure 1 shows examples of the app layout from the Carbon app. The analogous screenshots for the Spending app are in Figure 5 in the Appendix.

Within this framework, the Spending app provides cost information about participants' grocery purchases and the Carbon app shows the carbon footprint linked to each item. In the Carbon app, the emissions information is shown both in terms of kilograms of CO_2 and kilometers driven by an average passenger vehicle, which is a common measure to ensure that non-experts can relate to the data.[4] The calculations are based on the methodology developed by the International Organization for Standardization (ISO 2006a), which estimates the greenhouse gas (GHG) emissions of the full food production process and supply chain. This includes the environmental impact of land use change, farming, inputs (e.g., imported feed and fertilizer), outputs (e.g., livestock manure sold to another holding), processing, and transportation. To provide a summary measure, the GHG emissions are converted into kilograms of Carbon Dioxide equivalents (kg of CO_2e). Major greenhouse gases such as methane and nitrous oxide are expressed in terms of their effect relative to one kg of CO_2. For example, since methane is 25 times more efficient at retaining heat in

---

[4] Interestingly, Allcott and Knittel 2019) benchmark cost information about vehicle fuel efficiency against the cost of groceries (e.g., in terms of gallons of milk).
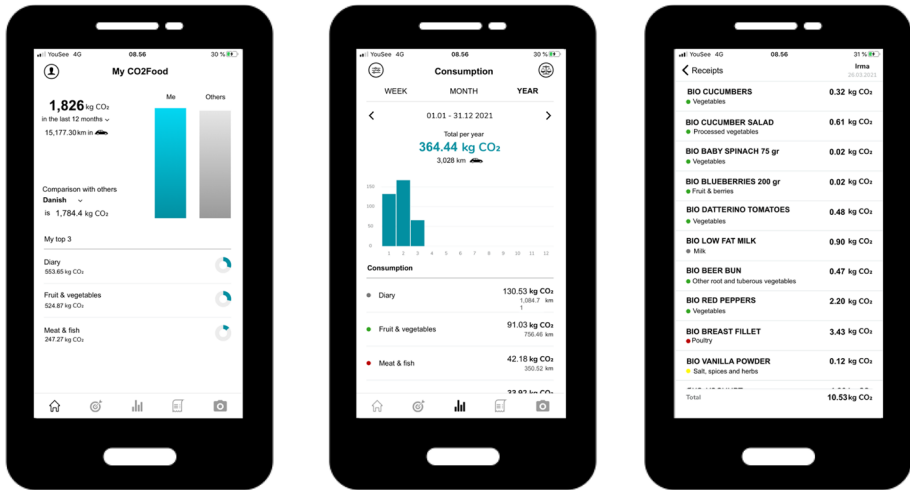
Fig. 1 Carbon app screenshots. *Notes:* The figure includes three screenshots taken from the Carbon app which illustrate the layout and content of the app

the atmosphere than CO_2, one kg of methane corresponds to 25 kg of CO_2-equivalents. Similarly, one kg of nitrous oxide equals 298 kg of CO_2-equivalents (Boardman 2008; Mogensen et al. 2009; ISO 2006). The authors were not responsible for the emission calculations, which were developed by the partner company.

2.3 Estimations

Of the 100,000 invited participants, 96,324 received the email invitation, 2711 completed the enrollment survey, 332 downloaded their randomly assigned app(s) and 258 created a profile and connected it to their e-receipt system. We randomized participants on a rolling basis using the survey software (Qualtrics). Because randomization occurred before participants downloaded the app, we were not able to block the randomization on demographic characteristics or baseline behavior. We tracked participants for at least 19 weeks after they initially enrolled and installed the app(s). We also include 19 weeks of pre-intervention grocery purchases as the baseline comparison in our analysis.

In a difference-in-difference approach, our primary analysis examines weekly greenhouse gas emissions resulting from food purchases in the 19 weeks prior to and after entering the experiment across the carbon and spending treatments. We estimate changes in emission using the following regression with individual random effects:

$$y_{it} = \alpha_i + \beta_1 \sigma + \beta_2 \gamma + \beta_3 \gamma \sigma + \varepsilon_{it}$$

where the dependent variable y_{it} is the emission for individual i at the weekly level t; α_i captures the random effect for individual i; the dummy σ is an indicator for the intervention phase (19 weeks after the intervention); the dummy γ is the indicator for the Carbon treatment (1 for Carbon treatment, 0 otherwise). We cluster standard errors at the individual level.

3 Results

3.1 Descriptive Statistics

Our experimental sample includes the full distribution of the national adult population of Denmark in terms of geography (drawing from all regions of the country), age (ranging from 20–72), household income and composition. As shown in columns (1) & (2) of Table 1, our sample is not nationally representative. We also note that the demographic data are not complete for all respondents.[5] To account for the non-representative sample and address potential selection into treatment, we reweight our sample to match national averages on a number of dimensions: gender, age, income, employment, children, and region (as shown in the Appendix in Table 6 columns (6) & (12)). The results of the Inverse Probability Weighting (IPW) and treatment effects do not change.

Columns (3)-(5) of Table 1 report baseline characteristics for each treatment group. We report the significance from t-tests of binary differences compared to the Spending app treatment for the Carbon app and the Both app treatment groups. There are no statistically significant demographic differences between the Spending treatment and the Carbon treatment. The proportion of participants aged 50–59 and 70–79 significantly differs between the Both treatment and the Spending treatment at the 10% level.

The middle panel reports baseline grocery purchases for the experimental sample. We focus our experimental analysis at the weekly level in an effort to find a unit that includes at least one grocery shopping trip per individual and is not driven by heterogeneity in how people spread their shopping throughout the week. In the 19 weeks prior to study enrollment, participants averaged about 2.5 grocery trips per week with average weekly spending of \$56 (USD) and weekly carbon equivalent (CO_2-e) emissions of 20.7 kg. The weekly emissions are equivalent to driving 172 km (107 miles), which is about two-thirds of the estimated 252 km that the average Dane drives per week (Christiansen and Baescu 2022). In Figure 6 in the Appendix, a visual representation of the carbon-equivalent emission composition of the average customer's weekly grocery basket is provided. Unsurprisingly, we observe that a large portion of the carbon emission basket is linked to dairy products and meat products, with beef leading the trend on meat, which is the area where we find a significant and persistent reduction in emissions (see Mechanisms section).

There is some baseline imbalance in grocery purchases between the Carbon treatment and the Spending treatment. Participants in the Carbon treatment have higher weekly spending ($p = 0.051$) and higher carbon emissions ($p = 0.073$). As discussed in the Methods section, because we had to randomize participants into the assigned app before we knew their demographic information or baseline purchases, we were not able to block the randomization on baseline characteristics to ensure balance. However, as shown in the Appendix in Table 6, when we include demographic controls in the analysis to address the initial imbalance, the baseline difference in emissions between the Carbon treatment and the Spending treatment is small and not statistically significant (columns (1) & (7)).

---

[5] The following percentages of participants did not provide the demographics in question: gender (14% missing); education (71%); age (9%); income (19.5%); household composition (14%); region (13%); employment (17%).

Table 1 Descriptive statistics of experimental sample, experimental groups and national population

	National Population	Experimental Sample	Spending Treatment	Carbon Treatment	Both Treatment
Sample size (N)	–	258	94	73	91
Demographics					
Female	50.3%	63%	66%	58%	64%
				[0.299]	[0.777]
NA		14%			
Age	42	46.9	45.8	46.8	47.9
				[0.560]	[0.287]
20–29	13%	11%	13%	9%	11%
				[0.460]	[0.692]
30–39	12%	21.4%	22.6%	23.5%	18%
				[0.788]	[0.513]
40–49	13%	25%	28.6%	25%	22%
				[0.738]	[0.350]
50–59	14%	22.6%	16.6%	23.5%	28%
				[0.241]	[0.078]
60–69	11%	16%	12%	17.6%	19.5%
				[0.272]	[0.174]
70–79	10%	3.4%	7%	1.5%	1%
				[0.109]	[0.060]
NA		9%			
Employed	66%	75%	73%	80%	73%
				[0.160]	[0.764]
NA		17%			
Household Income ($)					
<47.4 k	34%	18%	19.4%	19.6%	14.6%
				[0.785]	[0.577]
48–79 k	26%	23%	29%	18%	21%
				[0.236]	[0.419]

Table 1 (continued)

		Experimental Sample	Spending Treatment	Carbon Treatment	Both Treatment
>80 k	39%	59%	51.4%	62%	64%
				[0.441]	[0.276]
NA		19.5%			
Household Type					
Single	24%	22%	28%	17.5%	20%
				[0.180]	[0.327]
Single+children	4%	8%	9%	8%	6%
				[0.882]	[0.590]
Couple	33%	32.6%	29%	32%	37%
				[0.668]	[0.263]
Couple+children	24%	34%	30%	40%	33%
				[0.220]	[0.642]
3 or more adults	15%	4%	4%	3%	4%
				[0.865]	[0.968]
NA		14%			
Capital Region	23%	39%	37.5%	40.6%	39%
				[0.615]	[0.756]
NA	13%	13%			
Baseline grocery purchase (19 weeks)					
Weekly trips		2.32	2.19	2.54	2.25
		(2.46)	(2.76)	(2.37)	(2.21)
Weekly spending ($)		56	53	62*	55
		(44.41)	(50.43)	(42.92)	(38.67)
Weekly CO_2-e emissions (kg)		20.7	19.8	23.1*	19.8
		(16.43)	(18.66)	(16.56)	(13.63)

Table 1 (continued)

Baseline survey responses

Climate Attitude Index (1–5)	4.15	4.14	4.19	4.12
	(0.78)	(0.81)	(0.77)	(0.78)
Food Emissions Awareness Index (1–5)	3.02	2.83	3.17*	3.09
	(0.97)	(0.98)	(0.91)	(0.99)
Environmental ranking mistakes (0–5)	2.19	2.29	2.17	2.09
	(1.34)	(1.36)	(1.28)	(1.38)

The first column of the table reports demographics for the national adult population of Denmark using StatBank.dk. Columns (2–5) report demographics, baseline grocery purchases and baseline survey responses for the full experimental sample and by treatment group. We report demographics based on non-missing responses. Standard deviation is in parentheses. In columns (4) and & (5), we report the results of Ranksum and Chi-squared tests (for percentages) of equality compared to the Spending and to the Carbon treatments with the following symbols: $* p < 0.1$, $** p < 0.05$, $*** p < 0.01$ (Comparison with the Spending treatment); $^\wedge p < 0.1$, $^{\wedge\wedge} p < 0.05$, $^{\wedge\wedge\wedge} p < 0.01$ (Comparison of the Both treatment with the Carbon treatment). We have internally replicated the Ranksum test on the difference between the Both treatment and the Carbon treatment: no test resulted significant. Conversion rate Danish kroner to U.S. dollars (6.33DKK = $1) as of March 30th, 2021

In the bottom panel of Table 1, we report average responses from the baseline survey participants completed prior to receiving the app. We report average responses on a climate attitude index and a food emissions awareness index with responses on a 1–5 Likert scale.[6] Participants score highly on climate attitude with average scores of 4.15, indicating a high willingness to address CO_2 emissions. Scores are lower, an average of 3.02, for the food emissions awareness index (participants in the Carbon treatment have higher self-reported food emissions awareness scores than in the Spending treatment, $p = 0.043$). Consistent with their self-reported lack of food emissions awareness, fewer than 20% of participants correctly rank the emissions impact of five food items (potatoes, beef, chicken, cheese, and oranges). Taken together, these results suggest that participants want to address climate change through their personal behavior but are not fully informed on how to do so through their food purchases.

The baseline survey answers also allow us to study selection in our study. Those who completed the survey but did not participate in the study ($n = 2453$) on average scored lower on the Climate Attitude Index than participants (4.03, compared with 4.15 among the participants, $p = 0.0195$). Non-participants also made more mistakes in the Environmental Ranking, averaging 2.47 mistakes compares to 2.19 mistakes among participants ($p = 0.0014$). Non-participants and participants scored similarly on the Food Emissions Awareness Index, 3.01 vs. 3.02 respectively ($p = 0.4182$). These results suggest that, compared to the broader population, our study participants may be more aware of and motivated to address the climate impact of their food purchases. And more generally that those who are more climate-engaged may be more likely to voluntarily take up informational tools related to climate change. We estimate selection into the experiment using the initial measured attitudes (see Table 7 in Appendix). Only very little selection is observed. To evaluate the influence of the selection on our main conclusions, the selection estimates are used as weights in an Inverse Probability Weighting analysis (we comment more on this at the end of Result Section 3).

3.2 App Engagement

Figure 2 Panel A displays the share of people in each experimental week who checks the app at least once, pooling all treatment groups (in the Appendix Figure 7 shows app usage over time by treatment groups divided across the three panels). App checking is concentrated in the first month of the experiment with more than 90% of participants checking the app at least once in the first week after they set-up their profile in the app, about half of participants checking the app at least once in the second week, almost 40% checking in the third week and a little over a third checking in the fourth week. App usage steadily declines in the second month of the study with an average of 23% checking at least once in a given week and then plateaus at about 10–15% for the

---

[6] The Climate Attitude Index is an average of the responses to: "It is important that we all do our part to reduce CO_2 emissions and take care of the environment", and "If the majority does nothing to reduce CO_2 emissions and take care of the environment, it does not help that I do anything". The responses are scaled so that higher scores indicate more desire to help reduce emissions. The Food Emissions Awareness Index averages: "I think about how much CO_2 has been spent on producing and transporting foods I buy", and "I'm in doubt about how to eat to eat climate friendly". The responses are scaled so that higher scores indicate more awareness of food emissions. See Table 5 in the Appendix for responses to each of the survey questions.
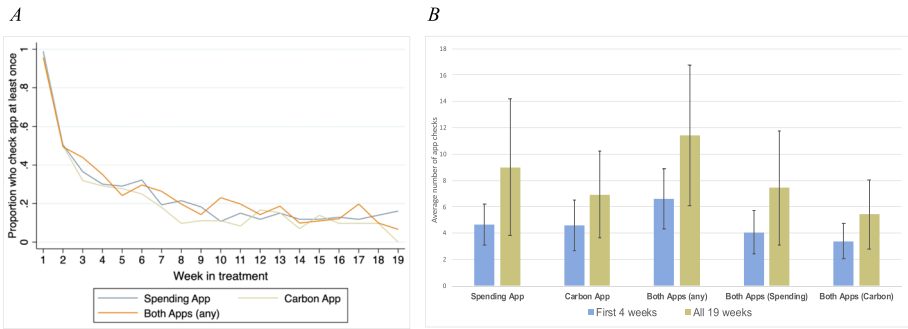
Fig. 2 App engagement. **A** shows the time development of the proportion of participants who check the app at least once in a given week based on each individual participants' date of enrollment. The figure pools all treatments. For the Both treatment group, we measure whether a participant checks either app at least once. **B** shows the average number of app checks by treatment group in the first 4 weeks of treatment and all 19 weeks of treatment. Bars indicate 95% Confidence Intervals

remainder of the study through week 19. Given this pattern of app usage, our analysis focus on both the first month (week 1–4) when there is greater engagement (The short-run analysis is insensitive to a choice of bounding 3, 4, or 5 weeks together.), and the longer run outcomes (week 1–19) with lower average engagement

When participants do check the app, it is generally within a few days of a shopping trip. The average app check is 1.02 days after the previous shopping trip and 4.4 days before the next one. 41.5% of app checks are on the same day as a shopping trip. Of these 41.5% occur before the shopping trip and 58.5% occur after the shopping trip in the overall period. These patterns are similar for the Spending app and the Carbon app, both in the short and longer term.

To examine revealed preferences for receiving spending information compared to receiving emissions information, Fig. 2 Panel B shows the average number of total app checks in the first 4 weeks and the full 19 weeks of the intervention by treatment group. Across all groups, over half of total app checks for the 19-week treatment period take place in the first month. Comparing the Spending treatment and the Carbon treatment, there is little difference in average app engagement in the short term. App usage in both groups averages a little over once a week during the first month of treatment, 1.15 ($p = 0.57$ from a *Ranksum* test of differences across treatments). Over the longer term, however, there is greater engagement with the Spending app than the Carbon app. In the full 19-week treatment period, usage of the Spending app averages about 0.47 times per week compared to 0.36 checks per week for the Carbon app, though the differences are not statistically significant ($p = 0.25$).

A similar pattern emerges for participants who have access to both apps. Usage of the Carbon and Spending apps are similar in the first month of treatment, averaging about once per week ($p = 0.31$) but engagement with the Spending app is almost 40% higher than the Carbon app over the 19-week treatment period, averaging 0.39 and 0.28 checks per week respectively ($p = 0.074$). Providing both apps increases overall app usage compared to providing either of the apps alone ($p < 0.01$ for all comparisons in the both the short and longer-term). However, it crowds out the usage of the individual apps. In particular, the usage of the Carbon app is about 30% lower in the short term
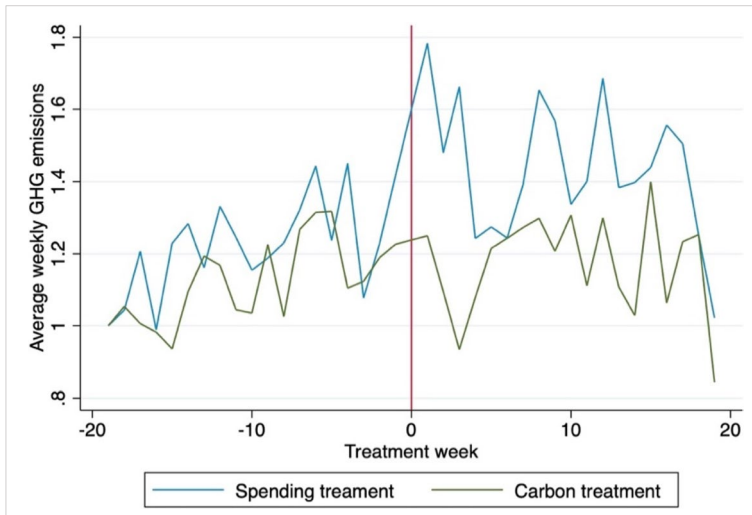
Fig. 3 Weekly greenhouse gas emissions. *Notes:* The lines are indexed in week -19 due to initial emission imbalances. A non-indexed graphical representation of the emissions can be found in the Appendix in Figure 8

and about 25% lower over the full treatment period, when participants receive both apps compared to receiving the Carbon app alone ($p = 0.016$ in the short term and $p = 0.082$ in the longer term).

Taken together, our results suggest a weak preference for spending information compared to emissions information over the longer term. This could reflect that people like the spending information more, or that once people see the emissions information, they learn about climate impacts of different food choices, and do not need repeated engagement with the app. In the Appendix, Table 6, we explore correlates of above-median app usage and do not find strong associations between demographics and app engagement. While finding a positive association of app engagement for the carbon treatment with food awareness, we find suggestive evidence that climate attitude is negatively associated with app engagement in the carbon treatment, potentially because climate-interested people gain less new knowledge from the carbon app compared to people who are not initially climate interested.

3.3 Treatment Effects on Carbon Emissions

Our main analysis estimates the difference-in-difference effect on carbon equivalent (CO_2-e) emissions of providing participants with the Carbon app compared to providing them with the Spending App. Figure 3 displays, across the Spending and Carbon treatments, the average weekly greenhouse gas emissions 19 weeks prior to receiving the app and 19 weeks after receiving the app. Average emissions are increasing over the baseline period, perhaps because our experiment took place during COVID-19, which shifted consumption from restaurants and cafes to groceries. During the treatment period, participants assigned to the Spending app continue to increase their emissions whereas emissions flatten among participants assigned to the Carbon app.

Table 2 Effect of carbon app on weekly carbon emissions

	First four weeks			Overall treatment period		
	Carbon treatment	Lower engagement	Higher engagement	Carbon treatment	Lower engagement	Higher engagement
	(1)	(2)	(3)	(4)	(5)	(6)
After	5.171***	3.403**	6.938***	3.447***	2.686**	4.349*
	(1.517)	(1.671)	(2.533)	(1.294)	(1.340)	(2.356)
Treatment	0.521	2.775	-1.099	0.992	0.220	-0.455
	(2.544)	(2.898)	(4.148)	(2.558)	(2.804)	(3.555)
After X Treatment	-5.837***	-3.192	-8.558***	-2.438	-0.164	-5.635*
	(1.985)	(2.361)	(3.215)	(1.805)	(2.022)	(3.178)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-27.622*	2.639	-8.556	-19.782	-5.880	-32.441**
	(15.418)	(24.850)	(14.678)	(13.843)	(19.466)	(15.657)
Baseline weekly emissions	21.247	20.828	21.681	21.247	20.428	22.327
Average app checks	1.16	0.59	1.77	0.43	0.18	0.84
Observations	3841	1955	1886	6346	3610	2736
Participants	167	85	82	167	95	72

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

Individual random effects regressions with weekly carbon-equivalent emissions as the dependent variable. Observations at the weekly level. Standard errors clustered at the participant level in parentheses. The dependent variable is measured as the weekly individual average of the GHG emissions, converted into kilograms of Carbon Dioxide equivalents (kg of CO_2e), connected to each grocery item. The detailed estimates of the contro in models 1–6 can be seen in Appendix Table 6

An intuitive illustration of the weekly emissions estimates across all three treatments, before and after the experiment was launched, is shown in Appendix Figure A.5. It is clear from the illustration that the estimated emissions in the Spending treatment increases while staying constant in the Carbon treatment (with a tendency to decrease). The Both treatment also shows an increase, although of a smaller magnitude than the increase in the Spending treatment.

In Table 2, we report results from difference-in-differences individual-level random effects regressions estimating the impact on weekly CO_2-equivalent (CO_2-e) emissions of the Carbon app treatment compared to the Spending app treatment. Each participant-week is an observation, and we cluster standard errors at the individual level. The first three columns restrict the sample to the first month of treatment. The last three columns include the full 19-week treatment period. Columns (1) & (4) include all participants. Columns (2–3) and columns (5–6) split the sample based on app engagement, measured as being above or below median app usage during the relevant period. All regressions include 19 weeks of pre-intervention observations.[7]

Table 2 column (1) reports the estimated effects of providing the Carbon app for the first month. Our coefficient of interest is the interaction term of *After X Carbon Treatment* which estimates the pre- vs. post-treatment difference in weekly emissions in the Carbon app treatment compared to the difference in the Spending app treatment: this is our difference-in-difference weekly emission measure. We estimate that providing information about CO_2-e emissions via the app reduces the subsequent CO_2-e emissions of weekly food purchases by about 5.8 kg ($p=0.003$) compared to the weekly CO_2-e emissions in the Spending Treatment. It is noteworthy that in absolute terms, the emissions levels are stable in the Carbon treatment, but increases in the Spending Treatment after the intervention is introduced, leading to a relative reduction of weekly emissions in the Carbon treatment, compared to the Spending treatment. The size of the reduction corresponds to 27% of the pre-treatment baseline emissions of 21.25 kg in the Carbon Treatment. However, the impact of the Carbon app does not persist over time. As shown in column (4), over the 19-week experiment period, we estimate a decline in emissions of about 2.4 kg, a 11.3% decrease that is not statistically significant ($p=0.177$).

We also present estimates of providing both apps compared to the Spending app alone, as well as estimates the Both treatment and Pooled treatment (corresponding to the Carbon and Both app groups together) (Table 6, columns (4–5), and (9–10), in the Appendix). The pattern of results is similar, though the effect sizes are smaller and not statistically significant for the Both treatment. The larger impact of the Carbon treatment compared to providing both apps may be due to the higher engagement with the Carbon app when it is provided alone. There also may be effects of providing the Spending app to participants that interact with the impact of the Carbon app.

---

[7] In the Appendix, we report the complete estimates, including controls coefficients (Table 6 in the Appendix, columns (1) & (7)), from the random effects models (RE) estimated in Table 2 but also the sensitivity checks on fixed effects (FE) regressions: the results do not change (Table 6 in the Appendix, columns (3) & (9)). While in the FE models individually constant covariates are omitted, in the RE models we include controls for the demographic characteristics reported in Table 1, the recruitment wave and the Likert-scale responses to the baseline survey questions reported in Table 4 in the Appendix (we transform all non-continuous variables into dummies, and we use an indicator variable for missing covariates). Excluding these covariates from the RE models does not affect the results (Table 6 in the Appendix, columns (2) & (8)).

Relatedly, we note that across all regressions, there is a positive and significant coefficient for the variable *After*, which suggests that weekly CO_2-e emissions are increasing during our study period. As we discussed above, that may be because we implemented the study during the onset of the COVID-19 pandemic when grocery purchases increased (Chenarides et al. 2021), which mechanically increases CO_2-e emissions from groceries. An alternative interpretation is that providing the Spending app to participants affects grocery purchases. With our data, we cannot explicitly disentangle the effect of the Spending app from seasonality (including potentially the COVID-19 effect) because we do not have a participant group that receives no information and no app. Seasonality is, however, unlikely to influence our findings as seasonality affects the participants in our treatments equally, and comparison, therefore, cancels out this influence. We compare the Carbon app users against the Spending app users for both technical and policy reasons. First, as mentioned in the design section, we could not include a no-app control group since our app served also as a data collector and we did not have the availability of a "white" app that could just record data without providing any information. Moreover, from a policy perspective, most interventions in this context have focused on providing spending and usage information (Allcott and Knittel 2019; Brandon et al. 2019; Davis and Metcalf 2016; Hahn and Metcalfe 2016; Jessoe and Rapson 2014; List and Price 2016). Our results suggest that, if anything, in our context, this may increase carbon emissions and that providing Carbon information is a more effective intervention. However, the effects fade out over time.

The fade-out in treatment effects corresponds with a fade-out in engagement as discussed in the section above. To further explore the role of engagement, we split the sample by above and below median app usage. Given that the engagement with the app is endogenous, our analysis is correlational. In the first month, above median and below median users check the app on average 1.77 times and 0.59 times per week, respectively. Over the full treatment period, above-median users sustain their usage at 0.84 times per week compared to below median users who check on average 0.18 times per week during a 19-week intervention.

We estimate that among highly engaged (above-median) participants, providing information about CO_2-e emissions reduces weekly CO_2-e emissions by about 8.5 kg ($p = 0.008$) in the first month of treatment compared to the spending treatment. A 38% decrease compared to pre-treatment baseline weekly emissions of 22.3 kg (column (3)). That compares to the insignificant 3.2 kg reduction in CO_2-e emissions for those who are below the median app usage, corresponding to a 16% decrease (column (2)). When we examine the longer-term effects of the Carbon app, we find suggestive evidence that the most engaged participants who sustain their engagement also sustain meaningful treatment effects. As shown in column (6), we estimate that above-median users reduce their CO_2-e emissions by an average of 5.6 kg ($p = 0.076$) per week over the 19-week intervention, which is similar to the short-term effect estimated for the full sample. For the least-engaged, the long run reduction in CO_2-e emissions is smaller than for the whole sample and not statistically significant (0.16 kg; $p = 0.935$). These results suggest that the most engaged users are driving the treatment impact of the Carbon app, though we caution that the estimated treatment effects are statistically indistinguishable across subgroups.

To address the concern that above median users of the Carbon app may not be comparable to above-median users of the Spending app, we also split the sample based

on predicted engagement with the Carbon app. In Table 8 in the Appendix, we regress our models' covariates on usage among those who received the Carbon app and then use the coefficients to create a predicted Carbon app usage score for all participants. We then split the sample by those predicted to be above- or below-median users and replicate our analysis in the Appendix in Table 9. Among the predicted engaged compliers, the Carbon app reduces CO_2-e emissions by an estimated 9.7 kg ($p = 0.002$), which corresponds to a 45% reduction in the first month of treatment. In the long run, those predicted to be above-median users show a reduction of 4.7 kg in CO_2 emissions ($p = 0.095$). For the less engaged, the reductions are not significant in either the short or long run.

As an additional robustness check, we have also compared the Carbon and the Spending treatments with the Both treatment. It is important to note that the impact of the information provision in the Both treatment is influenced by the fact that participants endogenously chose what information to attend to. In contrast, the information is exogenously imposed in the two other treatments. With this caution in mind, we compare the three treatments in regressions similar to those in Table 2 (see Table 10 in the Appendix). In a regression with the Both treatment as the base category, the weekly emissions after the experiment started were not significantly different between the Spending Treatment and the Both treatment, but significantly smaller (borderline) in the Carbon treatment, compared to the Both treatment. This finding is unsurprising given that participants in the Both treatment attend significantly more towards the spending app than the carbon app, and as such participants in the Both treatment are exposed more to information similar to that of the Spending treatment.

To address potential selection into the experiment, we reweight the main treatment analysis with Inverse Probability Weighting method to match the overall survey sample on the individual attitudes measured in the initial survey (Table 11 in the Appendix). The IPW estimations (columns (2) & (5)) replicates our overall main results (columns (1) & (4)), with only marginal changes in the interaction term, suggesting that our conclusions from the main model are not changed when accounting for the selection effect. When adding the socio-demographics controls (columns (3) & (6)), they seem to pick up some of the treatment and the constant variation, without affecting their significance levels.

Taken together, our results suggest that for time periods and people with high app engagement, providing emissions information can have a meaningful and sustained impact on the carbon footprint of grocery purchases.

3.4 Mechanisms

Finally, we explore the mechanisms leading to our observed reduction in weekly CO_2-e emissions when providing the Carbon app. Table 3 has the same structure as Table 2 except that we examine different outcomes, which are reported for each column. As shown in columns (1) & (2), we find that overall purchase quantities and money spent are significantly lower among participants who receive the Carbon app.[8] Purchases decline by about 28% in the first month of the experiment, while smaller and insignificant

---

[8] Quantity is defined as the number of items purchased per week. The unit is at the item or package level and does not depend on weight.

Table 3 Mechanisms

	First four weeks of treatment					Overall treatment period				
	Quantity	Money	CO2 per item	CO2 per dollar	CO2 by beef	Quantity	Money	CO2 per item	CO2 per dollar	CO2 by beef
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After	5.246***	13.754***	0.091*	0.070***	0.416	3.206***	8.232**	0.059	0.046**	0.561*
	(1.239)	(3.788)	(0.047)	(0.022)	(0.349)	(1.188)	(3.203)	(0.037)	(0.019)	(0.292)
Treatment	0.919	1.116	0.078	0.017	0.285	1.363	2.343	0.082	0.016	0.289
	(2.223)	(6.829)	(0.054)	(0.022)	(0.456)	(2.212)	(6.781)	(0.054)	(0.022)	(0.453)
After X Treatment	-5.579***	-16.515***	-0.105*	-0.055*	-1.197**	-2.269	-5.085	-0.059	-0.030	-0.789*
	(1.665)	(5.171)	(0.063)	(0.033)	(0.508)	(1.623)	(4.425)	(0.053)	(0.024)	(0.419)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-18.833	-60.336	0.024	0.103	-3.622	-16.270	-52.619	-0.026	0.080	-3.392
	(12.123)	(42.538)	(0.284)	(0.135)	(2.599)	(11.117)	(40.378)	(0.247)	(0.113)	(2.135)
Baseline weekly value	19.016	56.954	0.891	0.315	2.628	19.016	56.954	0.891	0.315	2.628
Observations	3841	3841	3841	3841	3841	6346	6346	6346	6346	6346
Participants	167	167	167	167	167	167	167	167	167	167

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

Individual random effects regressions. The dependent variable reported for each column is defined as follows: *quantity* is the number of items purchased per week and does not depend on weight (e.g. $2 = 2$ cans of tomato sauce; $1 = 1$ package of dry beans); *money* is the amount of money spent per week; CO_2 *per item* is the relative kilograms of Carbon Dioxide equivalents connected to each item; CO_2 *per dollar* is the relative kilograms of Carbon Dioxide equivalents connected to each dollar spent; CO_2 *by beef* describe the kilograms of Carbon Dioxide equivalents connected to beef items. Observations at the week level. Standard errors clustered at the participant level in parentheses. The controls included in models 1–6 correspond to those included in Appendix Table 6

reductions are observed over the full 19-week intervention (columns (6) & (7)). We note that lower emissions foods also tend to be less expensive so the reduced spending could reflect both changes in quantity and changes in basket composition. Indeed, when decoupling the carbon emissions from the total quantities, we find suggestive evidence that net of total quantities the Carbon app also reduces carbon emissions per item and per dollar spent (columns (3) & (4)).

In particular, we find a large and significant decrease in emissions from beef consumption in both the short and longer run (columns (5) & (10)). We estimate a 1.2 kg per week reduction in CO_2-e emissions from beef in the first month of treatment ($p = 0.019$), a 45% decrease that is equivalent to over 21% of the treatment impact on overall emissions. The Carbon app does not directly highlight beef as a high emissions food group – as prior informational interventions have done (Camilleri et al. 2019; Jalil et al. 2020) and yet has a large impact on this critical target for reducing the carbon footprint of food production and consumption. Achieving such effects through price changes, or taxes, would require an over 30 percent increase in price, based on estimated price elasticities for beef consumption (Taylor and Tonsor 2013). For future research interventions and app designs, our findings suggest that targeting beef consumption would be particularly powerful.

To facilitate a discussion on the substitution dynamics within key product categories, we refer to Table 12 in the Appendix. The exhibit describes the treatment effect on weekly carbon emissions by food category. Compared to the Spending treatment, the Carbon treatment shows a significant, reducing effect on Beef, Charcuterie, Milk, and Processed food in the short run, while only a significant effect on Beef and Milk in the long run. We have encompassed various alternative dairy products within the Plant-Based category, including oat and almond milk. Despite observing a pronounced and negative impact of the Carbon treatment on the consumption of traditional dairy items, we did not find any statistically significant effect on the adoption of plant-based alternatives despite the positive effect. The effect direction shows a spill-over phenomenon from conventional to plant-based products. Similar to the overall treatment effect, for some of the food categories, such as Charcuterie, Milk and Cheese for instance, providing information about CO_2-e emissions via the app reduces the subsequent CO_2-e emissions of weekly food purchases compared to the weekly CO_2-e emissions in the Spending Treatment. It is however noteworthy that such reduction is to be understood in relative terms: the weekly emissions stemming from some food categories increase in the Spending Treatment after the intervention is introduced while the emissions are constant in the Carbon treatment, leading to a relative reduction of weekly emissions in the Carbon treatment, compared to the Spending treatment.

4 Discussion and Conclusion

Our study demonstrates that providing access to personalized emissions information can have a meaningful impact on grocery purchases. The estimated 5.8 kg decrease in weekly CO_2 equivalent emissions, relative to providing spending information, is equivalent to switching from a beef burger to a plant-based burger (e.g. pea-based burger) or reducing

driving by 49 km (30.44 miles) per week (Poore and Nemecek 2018). The magnitude of the short-run effect is almost twice as large as the effect of adding a social comparison to a monthly home energy report (HER). Allcott (2011) estimates that the average treatment effects of HERs translate into 0.62 kWh per day, or 10.4 h of lightbulb use per day for a standard 60-W incandescent lightbulb. Our estimated 5.8 kg decrease in CO_2-e emission per week for the Carbon treatment corresponds to 1.172 kWh per day, or 19.53 lightbulb hours (Environmental Protection Agency 2019).

We emphasize that our study should be considered a first step towards understanding how real-time, individualized information provision can be used to lower carbon emissions from food consumption. In future research, it would be beneficial to address the external validity of our findings, drawing on four conditions proposed by John List (2020), selection, attrition, naturalness, and scaling. We briefly discuss each condition. Selection: despite having a diverse participant sample, it is plausible that the sample composition favored certain individuals. Using inverse probability weighting robustness checks, we found that our main conclusion holds after modeling the selection. Attrition: while our randomized treatment assignment before the app download ensured internal validity, it certainly contributed to the difference between the initial response rate and the actual app usage phase. Naturalness: one of our study's greatest inherent strengths lies in its real-world experimental setting. Participants were aware of their involvement, yet the focus remained solely on their initial engagement, allowing them to make choices within their day-to-day contexts subsequently. This dynamic makes our evidence remarkably externally valid in this regard. Scaling: to scale our information approach as a strategy for food-related carbon emissions, larger samples are needed to quantify how our approach impacts across different and diverse samples.

Our study also highlights the challenges of sustaining the impact of the emissions intervention. Our results suggest that in periods and among people who remain engaged with the app, the emissions information has meaningful effects. However, the impacts fade quickly along with engagement. This differs somewhat from evidence in the home energy context suggesting that decreases in energy usage may be sustained after people stop receiving home energy reports (Allcott and Rogers 2014; Brandon et al. 2019). This may be in part because they received reports for a longer period and built-up habits during this time. It may also be in part due to the nature of the technology. People can reduce their home energy, for example by one-time installations of energy-efficient light bulbs and appliances that have a persistent impact (Brandon et al. 2019). In contrast, grocery purchase decisions are largely made in real time. Future work could examine integrating personalized feedback on climate footprint on grocery receipts or newly implemented scan-and-go tools, which allow you to scan and purchase items with your smartphone in the grocery store. By applying an intervention based on real-time feedback, one could test the impact of receiving the carbon footprint evidence at the moment of purchase, rather than depending on the app engagement. As our work demonstrates, it is critical to understand how engagement with these interventions affects their impact and the external validity of the treatments. In future research, it would also be interesting to identify the effect of using a feedback app per se, something which our experimental design could not detect.

Appendix

Table 4 Survey content

Question	Possible answers (Likert scale)
I keep a close eye on how much money I spend on food	*Totally disagree Partly disagree Neither nor Partially agree Totally agree*
At the end of the month, I often change my grocery shopping to have enough money	*Totally disagree Partly disagree Neither nor Partially agree Totally agree*
It is important to me that my food is healthy	*Totally disagree Partly disagree Neither nor Partially agree Totally agree*
It is important that we all do our part to reduce the greenhouse effect and care of the environment	*Totally disagree Partly disagree Neither nor Partially agree Totally agree*
I am in doubt about how I should eat to eat climate-friendly	*Totally disagree Partly disagree Neither nor Partially agree Totally agree*
I think about how much CO2 is used to produce and transport the food I buy	*Totally disagree Partly disagree Neither nor Partially agree Totally agree*
If the majority does nothing to reduce the greenhouse effect and take care of the environment, there is no point in me doing anything	*Totally disagree Partly disagree Neither nor Partially agree Totally agree*
I am in doubt about how I should eat to eat healthy	*Totally disagree Partly disagree Neither nor Partially agree Totally agree*
Task	Items to rank order
Please put these items in order of how expensive you think they are by dragging them up or down (order the most expensive item as number 1, the second-most expensive item as number 2, etc.)	*Potatoes* *Oranges* *Cheese* *Chicken* *Beef*
Please put these items in order of how healthy you think they are by dragging them up or down (order the healthiest item as number 1, the second-healthiest item as number 2, etc.)	*Potatoes* *Oranges* *Cheese* *Chicken* *Beef*
Please put these items in order of how much CO_2 you think they emit by dragging them up or down (order the most emitting item as number 1, the second most emitting item as number 2, etc.)	*Potatoes* *Oranges* *Cheese* *Chicken* *Beef*

Table 5 Baseline survey responses

	Experimental Sample	Spending Treatment	Carbon Treatment	Both Treatment
It is important that we all do our part to reduce CO_2 emissions and take care of the environment	4.5	4.41	4.63 [0.054]	4.5 [0.336]
I think about how much CO_2 has been spent to produce and transport the foods I buy	3.07	2.82	3.12 [0.131]	3.28 [0.017]
I'm in doubt about how to eat to eat climate friendly (reverse)	2.96	2.84	3.21 [0.042]	2.9 [0.759]
If the majority does nothing to reduce CO_2 emissions and take care of the environment, it does not help that I do anything (reverse)	3.79	3.87	3.75 [0.593]	3.73 [0.320]
I keep a close eye on how much money I spend on food	3.22	3.21	3.30 [0.631]	3.16 [0.816]
At the end of the month, I often change my food purchases to have enough money	2.0	2.05	2.08 [0.984]	1.89 [0.496]
It is important to me that my food is healthy	4.25	4.25	4.21 [0.778]	4.27 [0.572]
I'm in doubt about how to eat to eat healthy	1.84	1.95	1.68 [0.108]	1.85 [0.742]

Average on 1–5 Likert-scale of each baseline survey question for the experimental sample and by treatment. We report p-values in brackets from *Ranksum* tests compared to the Spending treatment

Table 6 Sensitivity checks for models robustness

	First four weeks of treatment						Overall treatment period					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Full regression	No covariates	Fixed Effect	Both Treatment	Pooled Treatment	IPW	Full regression	No covariates	Fixed Effect	Both Treatment	Pooled Treatment	IPW
After	5.171***	5.171***	5.171***	5.171***	5.171***	4.967***	3.447***	3.447***	3.447***	3.447***	3.447***	3.488***
	(1.517)	(1.510)	(1.510)	(1.515)	(1.513)	(1.501)	(1.294)	(1.291)	(1.291)	(1.294)	(1.292)	(1.255)
Treatment ^	0.521	3.272	omitted	-2.678	-0.260	0.820	0.992	3.272	omitted	-2.405	-0.074	1.284
	(2.544)	(2.724)		(2.365)	(2.157)	(2.512)	(2.558)	(2.724)		(2.396)	(2.196)	(2.523)
After X Treatment	-5.837***	-5.837***	-5.837***	-1.982	-3.698**	-5.461***	-2.438	-2.438	-2.438	-0.380	-1.296	-2.494
	(1.985)	(1.977)	(1.977)	(1.987)	(1.772)	(2.001)	(1.805)	(1.800)	(1.800)	(1.754)	(1.555)	(1.783)
Climate Attitude	1.098			-1.457	-0.512	0.927	0.954			-1.658	-0.739	0.827
	(1.598)			(1.384)	(1.162)	(1.564)	(1.472)			(1.344)	(1.164)	(1.436)
Food Awareness	1.054			2.165**	1.317	0.992	0.567			1.331	0.839	0.548
	(1.072)			(1.005)	(0.914)	(1.048)	(0.975)			(0.967)	(0.870)	(0.949)
Eye on money	2.008**			0.399	0.975	1.898**	2.178**			0.490	1.023*	2.008**
	(0.868)			(0.692)	(0.616)	(0.848)	(0.849)			(0.690)	(0.602)	(0.828)
End of month	-0.358			-0.389	-0.163	-0.393	-0.543			-0.193	-0.173	-0.602
	(0.961)			(0.909)	(0.807)	(0.926)	(0.897)			(0.862)	(0.764)	(0.873)
Health Attitude	-1.600			-3.625***	-2.136*	-1.432	-1.135			-2.806**	-1.636	-1.026
	(1.273)			(1.254)	(1.124)	(1.285)	(1.196)			(1.250)	(1.087)	(1.191)
Health Doubt	2.064*			0.918	1.205	1.928	1.819*			0.485	0.772	1.686
	(1.196)			(1.091)	(0.986)	(1.194)	(1.074)			(1.029)	(0.925)	(1.073)
Climate Doubt	1.930*			0.025	0.491	1.729	1.694			0.162	0.469	1.477
	(1.146)			(1.001)	(0.827)	(1.111)	(1.095)			(0.943)	(0.794)	(1.066)
Free riding	0.350			-1.192	0.098	0.310	0.397			-0.770	0.356	0.357
	(0.887)			(0.893)	(0.705)	(0.887)	(0.864)			(0.848)	(0.692)	(0.854)
Female	4.456*			1.446	2.247	4.316	5.214**			0.458	1.992	4.947**
	(2.652)			(2.270)	(2.032)	(2.625)	(2.408)			(2.160)	(1.914)	(2.360)

Table 6 (continued)

	First four weeks of treatment						Overall treatment period					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Full regression	No covariates	Fixed Effect	Both Treatment	Pooled Treatment	IPW	Full regression	No covariates	Fixed Effect	Both Treatment	Pooled Treatment	IPW
Female(missing)	6.157			9.974**	9.845*	7.357	9.963			11.458**	11.514*	11.336
	(8.259)			(4.732)	(5.403)	(8.663)	(9.265)			(5.606)	(6.314)	(9.691)
Age	0.261**			0.225**	0.176**	0.247**	0.223**			0.235**	0.160**	0.206**
	(0.110)			(0.098)	(0.083)	(0.106)	(0.105)			(0.095)	(0.079)	(0.100)
Income2	1.988			1.277	1.579	2.085	5.183			2.814	2.802	5.249
	(4.259)			(4.812)	(3.503)	(4.139)	(4.089)			(4.528)	(3.438)	(3.943)
Income3	5.727			-1.797	1.511	6.028	7.983*			0.432	3.372	8.416**
	(4.198)			(4.982)	(3.589)	(4.026)	(4.193)			(4.728)	(3.541)	(3.996)
Income4	9.274*			1.529	4.487	9.903*	10.203**			2.885	5.191	10.826**
	(5.335)			(5.638)	(4.279)	(5.146)	(5.144)			(5.334)	(4.133)	(4.942)
Income5	3.672			4.768	4.749	4.211	5.217			5.774	5.858	5.775
	(5.731)			(6.151)	(4.605)	(5.572)	(5.608)			(5.728)	(4.342)	(5.403)
Income(missing)	5.076			0.109	2.741	5.194	5.981			0.898	3.089	6.195
	(4.960)			(5.093)	(4.244)	(4.811)	(4.653)			(4.558)	(3.821)	(4.488)
Family2	10.010**			9.835***	8.827***	10.104**	7.958**			8.406***	7.969***	8.172**
	(4.060)			(3.630)	(3.362)	(4.161)	(3.567)			(3.000)	(2.881)	(3.654)
Family3	10.115***			7.291**	6.912***	9.664***	9.211***			6.659**	6.334***	8.955***
	(3.052)			(3.098)	(2.456)	(2.973)	(2.795)			(2.886)	(2.295)	(2.700)
Family4	21.187***			12.553***	14.871***	20.383***	20.195***			12.596***	14.278***	19.615***
	(4.099)			(3.648)	(2.942)	(4.010)	(3.794)			(3.251)	(2.750)	(3.692)
Family5	33.303**			32.318**	30.715***	31.618**	33.187**			29.783**	28.661***	31.660**
	(15.262)			(13.371)	(10.749)	(15.208)	(14.550)			(12.472)	(10.124)	(14.522)

Table 6 (continued)

	First four weeks of treatment						Overall treatment period					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Full regression	No covariates	Fixed Effect	Both Treatment	Pooled Treatment	IPW	Full regression	No covariates	Fixed Effect	Both Treatment	Pooled Treatment	IPW
Family(missing)	10.294*			2.022	0.074	8.953	12.993**			0.498	-0.115	11.669**
	(6.031)			(4.339)	(4.034)	(6.004)	(5.319)			(4.498)	(4.240)	(5.131)
Employment2	-2.262			4.168	-1.979	-0.622	-7.079			6.672*	-0.346	-4.918
	(11.014)			(4.697)	(6.772)	(9.855)	(9.545)			(3.788)	(6.722)	(8.751)
Employment3	2.127			15.229	2.207	3.494	-6.171			13.380*	0.714	-4.010
	(12.465)			(9.542)	(8.099)	(11.392)	(10.872)			(7.536)	(7.810)	(10.104)
Employment4	1.198			4.158	-2.475	3.150	-4.106			7.761*	-0.387	-1.746
	(11.473)			(5.284)	(6.901)	(10.382)	(10.035)			(4.570)	(6.951)	(9.269)
Employment5	-2.698			13.636**	5.845	-0.561	-7.797			15.193***	6.613	-5.158
	(12.491)			(5.481)	(7.462)	(11.556)	(10.888)			(5.017)	(7.446)	(10.254)
Employment(missing)	-2.291			11.383**	1.817	-0.441	-10.142			12.838***	1.453	-7.670
	(12.268)			(4.723)	(6.732)	(11.142)	(10.706)			(4.254)	(6.788)	(9.920)
Region2	1.941			-0.741	0.977	2.300	1.379			-0.709	0.915	1.794
	(4.144)			(3.535)	(3.071)	(4.027)	(3.813)			(3.372)	(2.913)	(3.721)
Region3	-4.193			0.868	-1.058	-4.254	-4.739			1.940	-0.744	-4.607
	(3.260)			(3.021)	(2.628)	(3.217)	(3.120)			(3.005)	(2.566)	(3.071)
Region4	-0.434			-1.821	-0.679	-0.004	-0.258			-1.694	-0.617	0.228
	(3.482)			(2.994)	(2.835)	(3.460)	(3.241)			(2.661)	(2.610)	(3.192)
Region5	3.040			5.193	0.439	1.325	1.592			2.523	-1.750	0.058
	(6.220)			(6.800)	(5.108)	(5.769)	(5.461)			(6.381)	(4.654)	(5.057)
Region(missing)	-8.506			-14.332***	-7.674	-8.824	-9.768			-13.299**	-6.414	-10.313
	(8.025)			(4.845)	(6.513)	(8.247)	(9.019)			(6.281)	(8.076)	(9.253)

Table 6 (continued)

	First four weeks of treatment						Overall treatment period					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Full regression	No covariates	Fixed Effect	Both Treatment	Pooled Treatment	IPW	Full regression	No covariates	Fixed Effect	Both Treatment	Pooled Treatment	IPW
Recruitement wave	2.298			1.760	2.280	1.831	-1.217			-1.324	-0.964	-1.518
	(2.325)			(2.305)	(1.767)	(2.260)	(2.145)			(2.160)	(1.643)	(2.089)
Constant	-27.622*	19.817***	21.247***	11.833	0.200	-26.784*	-19.782	19.817***	21.247***	8.036	0.657	-19.299
	(15.418)	(1.921)	(0.177)	(10.233)	(11.084)	(14.485)	(13.843)	(1.920)	(0.455)	(9.111)	(10.192)	(13.134)
Observations	3841	3841	3841	4255	5934	3841	6346	6346	6346	7030	9804	6346
^ treatment:	*Carbon*	*Carbon*	*Carbon*	*Both*	*Pooled*	*Carbon*	*Carbon*	*Carbon*	*Carbon*	*Both*	*Pooled*	*Carbon*

Regression table for the short run and the overall treatment period, which includes the following models: (1,7) full version of random effects regression with control variables and weekly carbon-equivalent emissions as the dependent variable as in Table 2; (2,8) random effects regression results excluding control variables; (3,9) fixed effects regression; (4,10) for the Both treatment; (5,11) for the Pooled treatment (Carbon and Both treatments together); (6,12) Inverse Probability Weighting (IPW) estimations. The dependent variable is measured as the weekly individual average of the GHG emissions, converted into kilograms of Carbon Dioxide equivalents (kg of CO_2e), connected to each grocery item. Standard errors clustered at the participant level in parentheses. The random effects and IPW regressions include controls for: demographics (gender, age, income, household type, employment, region) and recruitment wave as dummy variables (with one category omitted), plus baseline survey answers as categorical variables. See notes on Table 8 for more details on demographics and survey answers

Table 7 Logistic regression model on experimental and treatment selection

	Participation (1)	Spending treatment (2)	Carbon Treatment (3)	Both treatment (4)
Climate Attitude	0.170*	0.033	0.501**	0.064
	(0.094)	(0.151)	(0.202)	(0.153)
Food Awareness	0.011	-0.090	-0.049	0.169
	(0.058)	(0.096)	(0.110)	(0.103)
Eye on money	-0.055	-0.033	-0.053	-0.092
	(0.055)	(0.091)	(0.103)	(0.095)
End of month	-0.080	-0.037	-0.053	-0.147
	(0.055)	(0.090)	(0.098)	(0.099)
Health Attitude	0.024	0.072	-0.055	0.079
	(0.090)	(0.145)	(0.172)	(0.160)
Health Doubt	-0.121*	-0.032	-0.213	-0.170
	(0.066)	(0.106)	(0.134)	(0.111)
Climate Doubt	-0.096	-0.169*	0.078	-0.180*
	(0.061)	(0.102)	(0.114)	(0.103)
Free riding	0.028	0.119	-0.041	-0.009
	(0.056)	(0.097)	(0.100)	(0.095)
Constant	-2.385***	-2.092**	-3.727***	-1.825*
	(0.567)	(0.912)	(1.187)	(0.945)
Observations	2727	939	908	880

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

Regression table with logistic models. The dependent variable is a binary variable ($=0$ if the subject only responded to our online survey; $=1$ if the subject answered to the survey and participated to the whole experiment)

Table 8 Predicting engagement

	All treatments	Carbon & Spending	Spending	Carbon	Both (Spending)	Both (Carbon)
Baseline Emissions	0.009	0.006	0.003	-0.002	0.032	0.017
	(0.007)	(0.008)	(0.016)	(0.016)	(0.025)	(0.022)
Climate Attitude	0.031	-0.124	0.030	-1.272**	0.006	-0.031
	(0.145)	(0.197)	(0.282)	(0.578)	(0.342)	(0.351)
Food Awareness	-0.010	0.059	-0.114	0.812**	-0.428	-0.532*
	(0.087)	(0.111)	(0.178)	(0.319)	(0.273)	(0.292)
Eye on money	0.018	0.059	-0.100	0.346	0.247	0.527**
	(0.075)	(0.100)	(0.156)	(0.242)	(0.226)	(0.247)
End of month	0.089	-0.075	-0.228	-0.291	0.311	-0.521*
	(0.084)	(0.103)	(0.177)	(0.243)	(0.248)	(0.286)
Health Attitude	0.277**	0.198	0.116	0.822*	1.274**	0.952**
	(0.121)	(0.158)	(0.260)	(0.434)	(0.537)	(0.465)
Health Doubt	-0.107	-0.033	-0.237	0.432	0.285	-0.193
	(0.107)	(0.132)	(0.207)	(0.352)	(0.296)	(0.294)
Climate Doubt	0.010	0.059	0.124	0.298	-0.492**	0.026
	(0.091)	(0.118)	(0.196)	(0.300)	(0.231)	(0.218)
Free riding	0.052	-0.028	-0.068	-0.214	0.090	-0.083
	(0.078)	(0.096)	(0.179)	(0.244)	(0.237)	(0.230)
Female	-0.291	-0.004	0.488	-0.844	0.707	1.382**
	(0.225)	(0.278)	(0.461)	(0.592)	(0.609)	(0.663)
Female(missing)	9.590	9.479	9.641	5.498	-9.131	11.626
	(690.217)	(486.220)	(582.950)	(356.933)	(1164.566)	(2706.462)
Age	0.008	0.003	0.045*	-0.002	0.010	-0.043
	(0.011)	(0.014)	(0.024)	(0.027)	(0.030)	(0.035)
Income2	0.009	0.505	2.251	-1.635	-2.437	-1.601
	(0.598)	(0.681)	(1.432)	(2.080)	(1.976)	(1.830)
Income3	0.883	0.522	0.515	1.106	-1.054	2.084
	(0.603)	(0.647)	(1.252)	(1.692)	(2.068)	(2.118)
Income4	0.959	1.043	0.715	1.390	-2.317	2.151
	(0.645)	(0.702)	(1.397)	(1.675)	(2.255)	(2.284)
Income5	0.671	0.903	0.788	0.656	-2.867	1.876
	(0.675)	(0.744)	(1.396)	(1.829)	(2.324)	(2.338)
Income(missing)	0.818	1.044	1.154	1.595	-2.551	-8.982
	(0.636)	(0.705)	(1.267)	(2.118)	(2.779)	(646.604)
Family2	-0.117	0.207	1.034	0.136	-1.995*	-0.898
	(0.426)	(0.533)	(0.910)	(1.373)	(1.100)	(1.098)
Family3	-0.200	-0.143	0.547	-1.425	-0.338	-0.327
	(0.312)	(0.379)	(0.655)	(0.868)	(0.942)	(0.943)
Family4	-0.958**	-0.784*	-0.407	-1.529	-0.500	-0.914
	(0.372)	(0.459)	(0.773)	(0.996)	(1.143)	(1.148)
Family5	-1.307*	-0.627	0.000	-0.244	0.000	0.000
	(0.706)	(0.886)	(.)	(1.958)	(.)	(.)
Family(missing)	-5.796	-5.415	-6.240	-6.339	-7.526	-7.934
	(564.018)	(297.861)	(385.778)	(356.931)	(719.077)	(2286.738)
Employment2	8.799	0.275	-0.234	4.899	7.601	6.670
	(576.915)	(0.819)	(1.067)	(356.926)	(546.556)	(955.216)

Table 8 (continued)

	All treatments	Carbon & Spending	Spending	Carbon	Both (Spending)	Both (Carbon)
Employment3	8.478	1.011	1.135	6.537	5.868	10.674
	(576.915)	(1.029)	(2.054)	(356.928)	(546.555)	(955.219)
Employment4	8.844	-0.237	-0.644	0.000	4.675	7.485
	(576.915)	(1.033)	(1.495)	(.)	(546.557)	(955.218)
Employment5	8.631	0.452	-1.024	3.645	5.621	5.745
	(576.915)	(0.962)	(1.362)	(356.928)	(546.555)	(955.217)
Employment(missing)	8.316	0.306	0.633	4.075	11.252	12.714
	(576.915)	(1.066)	(1.749)	(356.929)	(988.881)	(2830.304)
Region2	-0.102	-0.457	0.066	-1.065	0.197	1.040
	(0.318)	(0.417)	(0.708)	(0.920)	(0.940)	(1.125)
Region3	0.834***	0.877**	1.706**	1.467*	-1.217	-0.774
	(0.290)	(0.357)	(0.683)	(0.819)	(0.900)	(0.772)
Region4	-0.481	-0.285	-0.371	0.112	-1.601*	-0.602
	(0.293)	(0.362)	(0.586)	(0.850)	(0.861)	(0.797)
Region5	0.362	1.336**	0.000	1.656*	0.000	0.000
	(0.375)	(0.548)	(.)	(0.962)	(.)	(.)
Region(missing)	-3.880	-4.266	-3.590	0.000	11.267	0.406
	(397.848)	(384.303)	(437.044)	(.)	(1092.756)	(2531.640)
Recruitement wave	-0.094	0.300	0.435	1.209**	-0.873	-0.950*
	(0.195)	(0.250)	(0.423)	(0.594)	(0.538)	(0.532)
Constant	-10.742	-1.882	-3.096	-7.030	-9.252	-8.409
	(576.916)	(1.541)	(2.370)	(356.946)	(546.561)	(955.221)
Observations	258	167	87	71	79	78

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

The table presents correlates of above-median app usage. The dependent variable of each column is a dummy indicating the above median of weekly app usage calculated based according to the treatment of interest for each model: e.g., the dependent variable of the second column is calculated on the usage of the Carbon and Spending app together. All Probit model regressions include controls for: demographics (gender, age, income, household type, employment, region) and recruitment wave as dummy variables (with one category omitted for each demographic), and baseline survey answers as categorical variables. Standard errors clustered at the participant level in parentheses

The controls include the following information:

Income1 (omitted) = "< 150,000 DKK a year"; Income2 = "150,000 – 299,000 DKK"; Income3 = "300,000 – 499,000 DKK"; Income4 = "500,000 – 799,000 DKK"; Income5 = "> 800,000 DKK". Family1 (omitted) = "Single"; Family2 = "Single with kids"; Family3 = "Couple"; Family4 = "Couple with kids"; Family5 = "3 or more adults". Employment1 (omitted) = "Self-employed"; Employment2 = "Employed"; Employment3 = "Unemployed"; Employment4 = "Student"; Employment5 = "Senior Citizen". Region1 (omitted) = "Capital Region"; Region2 = "Zealand": Region3 = "Southern Denmark"; Region4 = "Mid Jutland": Region5 = "North Jutland". Climate Attitude = likert score to "It is important that we all do our part to reduce CO_2 emissions and take care of the environment"; Free Rider = likert score to "If the majority does nothing to reduce CO_2 emissions and take care of the environment, it does not help that I do anything"; Food awareness = likert score to "I think about how much CO_2 has been spent on producing and transporting foods I buy"; Climate doubt = likert score to "I'm in doubt about how to eat to eat climate friendly"; Eye on money = likert score to "I keep a close eye on how much money I spend on food"; End of month = likert score to "At the end of the month, I often change my food purchases to have enough money"; Health Attitude = likert score to "It is important to me that my food is healthy"; Health Doubt = likert score to "I'm in doubt about how to eat to eat healthy"

Table 9 Treatment effect on split sample based on engagement prediction

	First four weeks		Overall treatment period	
	Low predicted engagement	High predicted engagement	Low predicted engagement	High predicted engagement
After	3.309*	7.113***	1.859	5.104***
	(0.061)	(0.004)	(0.266)	(0.010)
Treatment	-0.023	6.725*	1.022	8.689**
	(0.995)	(0.055)	(0.788)	(0.022)
After X Treatment	-2.482	-9.694***	-0.358	-4.725*
	(0.314)	(0.002)	(0.877)	(0.095)
Controls	*Yes*	*Yes*	*Yes*	*Yes*
Constant	7.865	-59.951***	9.795	-53.998***
	(0.794)	(0.000)	(0.704)	(0.000)
Observations	2047	1794	3382	2964
Participants	89	78	89	78

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

Regression table with random effects models for the first four weeks and the overall period, with a sample split based on the predicted Carbon app usage score (p-values reported in parenthesis). The dependent variable is measured as the weekly individual average of the GHG emissions (kg of CO_2e) linked to grocery choices

Table 10 Random effect regression models with all treatment combinations together

	First four weeks			Overall treatment period		
	Base treatment: Spending	Base treatment: Carbon	Base treatment: Both	Base treatment: Spending	Base treatment: Carbon	Base treatment: Both
	(1)	(2)	(3)	(4)	(5)	(6)
After	5.171***	-0.667	3.189*	3.447**	1.009	3.066**
	(1.513)	(1.278)	(1.283)	(1.292)	(1.255)	(1.183)
Carbon Treatment	1.918		3.847	2.206		4.036
	(2.565)		(2.239)	(2.604)		(2.232)
After X Carbon Treatment	-5.837**		-3.856*	-2.438		-2.057
	(1.980)		(1.811)	(1.801)		(1.725)
Both Treatment	-1.929	-3.847		-1.830	-4.036	
	(2.313)	(2.239)		(2.341)	(2.232)	
After X Both Treatment	-1.982	3.856*		-0.380	2.057	
	(1.983)	(1.811)		(1.752)	(1.725)	
Spending Treatment		-1.918	1.929		-2.206	1.830
		(2.565)	(2.313)		(2.604)	(2.341)
After X Spending Treatment		5.837**	1.982		2.438	0.380
		(1.980)	(1.983)		(1.801)	(1.752)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Constant	1.331	3.249	-0.598	1.728	3.934	-0.102
	(11.024)	(11.494)	(11.281)	(10.067)	(10.519)	(10.367)
Observations	5934	5934	5934	9804	9804	9804
Participants	164	185	167	164	185	167

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

Regression table with random effects models for the first four weeks and the overall period, with all combinations of the Spending, Carbon and Both treatment alternating as base group. The dependent variable is measured as the weekly individual average of the GHG emissions (kg of CO_2e) linked to grocery choices. Standard errors clustered at the participant level in parentheses. The controls included in models 1–6 correspond to those included in Appendix Table 6

Table 11 Inverse probability weighting for experimental selection

	First four weeks			Overall treatment period		
	Main Analysis	IPW	IPW w/ controls	Main Analysis	IPW	IPW w/controls
	(1)	(2)	(3)	(4)	(5)	(6)
After	5.171***	5.428***	5.428***	3.447***	3.538***	3.538***
	(1.517)	(1.508)	(1.512)	(1.294)	(1.287)	(1.289)
Treatment	0.521	2.726	0.940	0.992	2.726	1.334
	(2.544)	(2.681)	(2.573)	(2.558)	(2.681)	(2.553)
After X Treatment	-5.837***	-5.964***	-5.964***	-2.438	-2.532	-2.532
	(1.985)	(1.985)	(1.991)	(1.805)	(1.794)	(1.797)
Controls	*Yes*	*No*	*Yes*	*Yes*	*No*	*Yes*
Constant	-27.622*	19.980***	-8.105	-19.782	19.980***	-4.053
	(15.418)	(1.907)	(10.516)	(13.843)	(1.907)	(10.389)
Observations	3841	3841	3841	6346	6346	6346
Participants	167	167	167	167	167	167

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

Observations at the weekly level. Standard errors clustered at the participant level in parentheses. The dependent variable is measured as the weekly individual average of the GHG emissions, converted into kilograms of Carbon Dioxide equivalents (kg of CO_2 e), connected to each grocery item
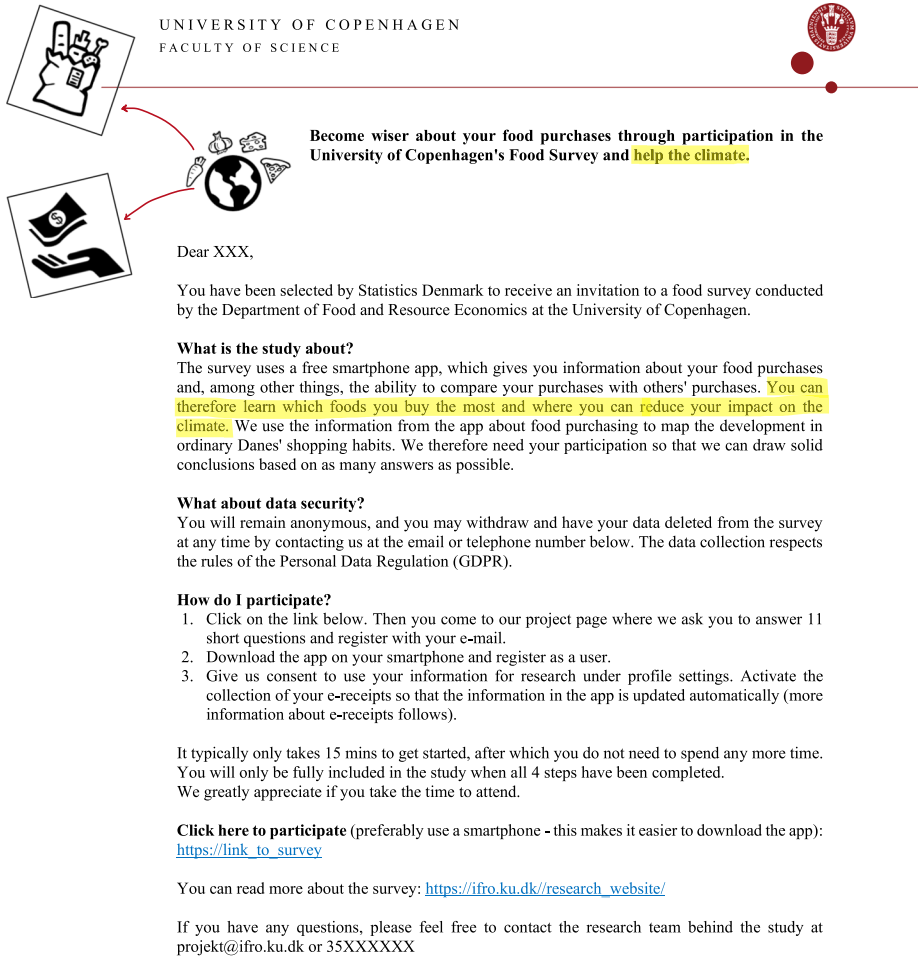
Table 12 Effect of carbon app on weekly carbon emissions by food category

First four weeks of treatment

	Beef	Pork	Chicken	Charcuterie	Goat	Fish	Milk	Cheese	Plant based	Processed
After	0.416	0.165	0.027	0.226**	-0.127***	0.012	0.495***	0.730**	0.004	0.092*
	(0.349)	(0.166)	(0.091)	(0.105)	(0.039)	(0.032)	(0.154)	(0.295)	(0.006)	(0.083)
Treatment	0.285	-0.005	0.291*	-0.013	-0.038	-0.028	-0.128	0.365	0.005	0.023
	(0.456)	(0.224)	(0.156)	(0.151)	(0.079)	(0.031)	(0.337)	(0.346)	(0.008)	(0.729)
After X Treatment	-1.197**	-0.125	-0.087	-0.300**	0.089	-0.025	-0.461**	-0.595	0.012	-0.161**
	(0.508)	(0.246)	(0.130)	(0.136)	(0.091)	(0.041)	(0.195)	(0.383)	(0.019)	(0.022)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.622	-2.579**	0.531	-0.679	-0.181	-0.295*	-2.264	-1.282	-0.013	-0.797*
	(2.599)	(1.177)	(1.722)	(0.863)	(0.339)	(0.151)	(1.824)	(2.196)	(0.042)	(0.084)
N	3841	3841	3841	3841	3841	3841	3841	3841	3841	3841

Overall treatment period

	Beef	Pork	Chicken	Charcuterie	Goat	Fish	Milk	Cheese	Plant based	Processed
After	0.561*	0.047	0.030	0.252***	-0.043	0.019	0.343**	0.381*	0.003	0.048
	(0.292)	(0.129)	(0.066)	(0.093)	(0.046)	(0.026)	(0.134)	(0.199)	(0.003)	(0.033)
Treatment	0.289	0.012	0.319**	0.013	-0.050	-0.019	-0.074	0.434	0.005	0.021
	(0.453)	(0.224)	(0.156)	(0.149)	(0.077)	(0.031)	(0.339)	(0.342)	(0.008)	(0.066)
After X Treatment	-0.789*	0.036	-0.050	-0.067	0.133	-0.040	-0.387**	-0.265	0.006	-0.032
	(0.419)	(0.163)	(0.091)	(0.141)	(0.111)	(0.032)	(0.193)	(0.279)	(0.008)	(0.047)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-3.392	-1.996*	0.353	-0.534	-0.368	-0.190*	-2.172	0.246	-0.029	-0.691
	(2.135)	(1.030)	(1.576)	(0.842)	(0.388)	(0.105)	(1.732)	(2.473)	(0.044)	(0.427)
N	6346	6346	6346	6346	6346	6346	6346	6346	6346	6346

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

Random effects regression table for the short run (Panel A) and the overall treatment period (Panel B) with control variables for the following food categories: *beef, pork, chicken, charcuterie, goat meat, fish, milk, cheese, plant based* and *processed food*. The dependent variable is measured as the weekly individual average of the GHG emissions, converted into kilograms of Carbon Dioxide equivalents (kg of CO_2e), connected to each grocery item

Invitation Letter.

UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE

Become wiser about your food purchases through participation in the University of Copenhagen's Food Survey and help the climate.

Dear XXX,

You have been selected by Statistics Denmark to receive an invitation to a food survey conducted by the Department of Food and Resource Economics at the University of Copenhagen.

What is the study about?
The survey uses a free smartphone app, which gives you information about your food purchases and, among other things, the ability to compare your purchases with others' purchases. You can therefore learn which foods you buy the most and where you can reduce your impact on the climate. We use the information from the app about food purchasing to map the development in ordinary Danes' shopping habits. We therefore need your participation so that we can draw solid conclusions based on as many answers as possible.

What about data security?
You will remain anonymous, and you may withdraw and have your data deleted from the survey at any time by contacting us at the email or telephone number below. The data collection respects the rules of the Personal Data Regulation (GDPR).

How do I participate?
1. Click on the link below. Then you come to our project page where we ask you to answer 11 short questions and register with your e‑mail.
2. Download the app on your smartphone and register as a user.
3. Give us consent to use your information for research under profile settings. Activate the collection of your e‑receipts so that the information in the app is updated automatically (more information about e‑receipts follows).

It typically only takes 15 mins to get started, after which you do not need to spend any more time. You will only be fully included in the study when all 4 steps have been completed.
We greatly appreciate if you take the time to attend.

Click here to participate (preferably use a smartphone ‑ this makes it easier to download the app): https://link_to_survey

You can read more about the survey: https://ifro.ku.dk//research_website/

If you have any questions, please feel free to contact the research team behind the study at projekt@ifro.ku.dk or 35XXXXXX

Fig. 4 Invitation letter. The invitation letter was translated to English from the original language. The treatment variation includes the logo, the last words of the title and the second period of the second paragraph ("You can therefore learn which foods you buy the most and where you can save on your purchases" for the economic version; absent in the neutral version of the letter)
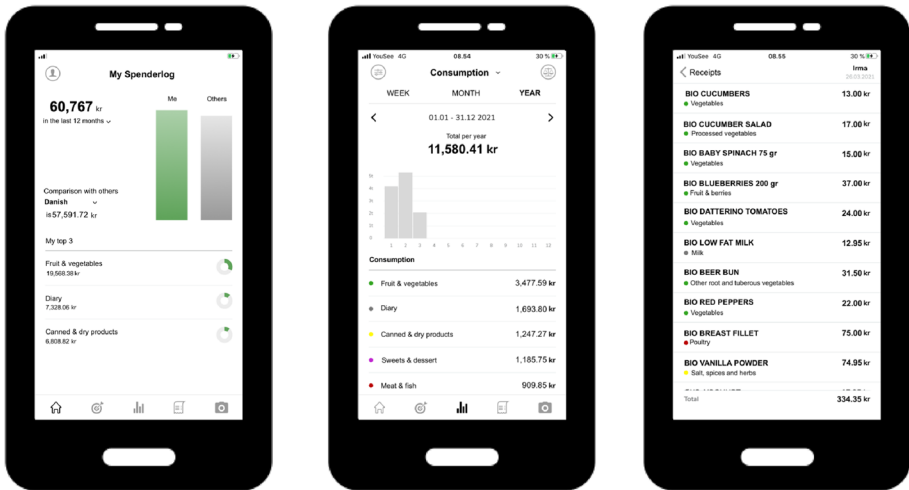
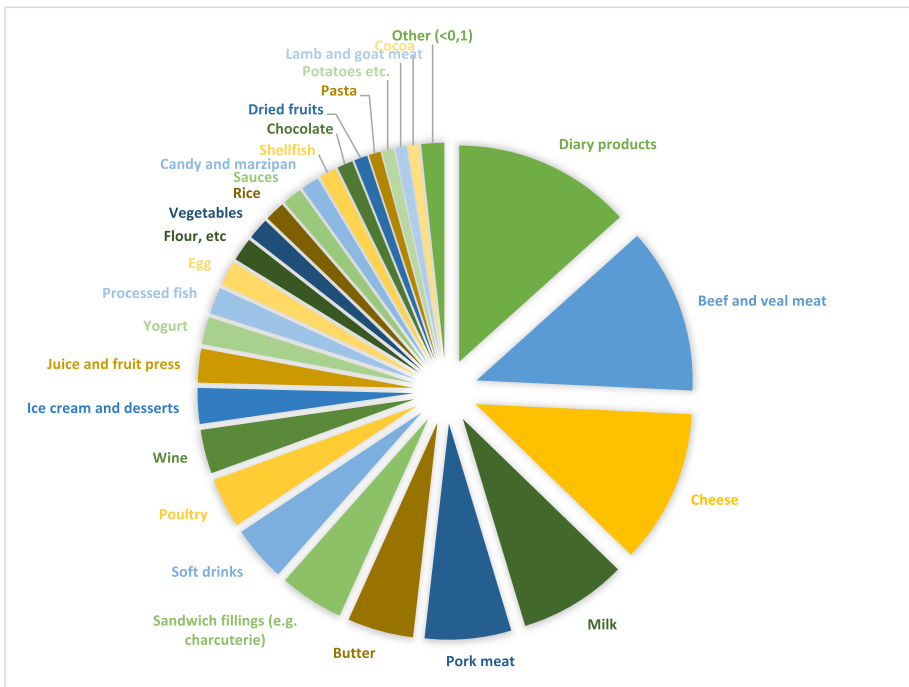**Fig. 5** Spending app screenshots

Fig. 6 Average basket composition of carbon emissions composition before the experiment. The smallest category "others" groups the categories below the 0.8 kg of CO_2-e and includes the following categories: Liquor, Sugar etc., Fish, Chips etc., Fruit, Breakfast products, Marmalade and Honey, Root vegetables, Oil, Alcoholic soft drinks, Margarine, Other meat products, Coffee, Other cereal products, Water and Danish water, Salt, Animal fats, Tea, Artificial sweeteners
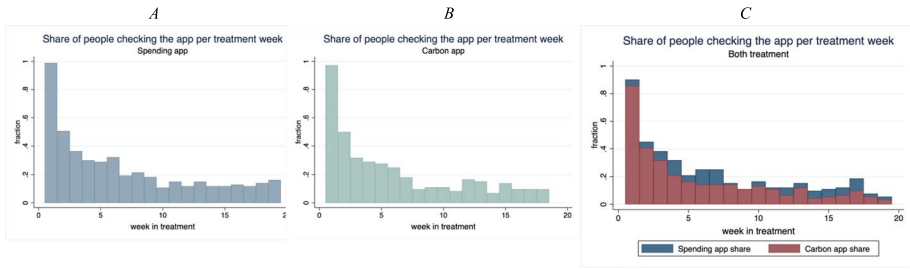
Fig. 7 Share of people checking the app per treatment week by treatment
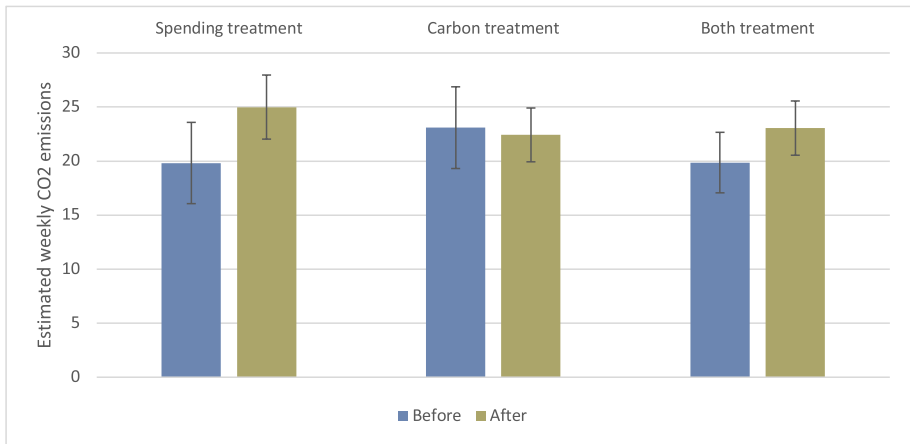
Fig. 8 Weekly greenhouse gas emissions by treatment for the first four weeks of treatment. The bar graph includes the confidence intervals

Declarations

Competing Interest The authors declare that they have no competing interests.

References

Allcott H (2011) Social norms and energy conservation. J Public Econ 95(9–10):1082–1095

Allcott H, Knittel C (2019) Are consumers poorly informed about fuel economy? Evidence from two experiments. Am Econ J Econ Pol 11(1):1–37. https://doi.org/10.1257/pol.20170019

Allcott H, Rogers T (2014) The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. Am Econ Rev 104(10):3003–3037. https://doi.org/10.1257/aer.104.10.3003

Allcott H, Taubinsky D (2015) Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. Am Econ Rev 105(8):2501–2538. https://doi.org/10.1257/aer.20131564

Arora NK, Mishra I (2021) COP26: more challenges than achievements. Environ Sustain 4(4):585–588. https://doi.org/10.1007/s42398-021-00212-7

Boardman B (2008) Carbon labelling: Too complex or will it transform our buying? Significance 5(4):168–171. https://doi.org/10.1111/j.1740-9713.2008.00322.x

Brandon A, List JA, Metcalfe RD, Price MK, Rundhammer F (2019) Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity. Proc Natl Acad Sci USA 116(12):5293–5298. https://doi.org/10.1073/pnas.1802874115

Brandon A, Clapp CM, List JA, Metcalfe RD, Price M (2022) The human perils of scaling smart technologies: evidence from field experiments, vol No. w30482. National Bureau of Economic Research

Brownback A, Imas A, Kuhn MA (2023) Behavioral food subsidies. Rev Econ Stat:1–47https://doi.org/10.2139/ssrn.3422272

Brunner F, Kurz V, Bryngelsson D, Hedenus F (2018) Carbon Label at a University Restaurant – Label Implementation and Evaluation. Ecol Econ 146:658–667. https://doi.org/10.1016/j.ecolecon.2017.12.012

Federal Highway Administration (2017) 2017 National Household Travel Survey, U.S. Department of Transportation, Washington, DC. Available online: https://nhts.ornl.gov

Cafaro P (2011) Beyond business as usual: alternative wedges to avoid catastrophic climate change and create sustainable societies. In: The ethics of global climate change, pp 192–215https://doi.org/10.1017/cbo9780511732294.010

Camilleri AR, Larrick RP, Hossain S, Patino-Echeverri D (2019) Consumers underestimate the emissions associated with food but are aided by labels. Nat Climate Change 9(1):53–58. https://doi.org/10.1038/s41558-018-0354-z

Chenarides L, Grebitus C, Lusk JL, Printezis I (2021) Food consumption behavior during the COVID-19 pandemic. Agribusiness 37(1):44–81. https://doi.org/10.1002/agr.21679

Christiansen H, Baescu O (2022) The Danish national travel survey: Annual statistical report for Denmark for 2021

Commission E (2006) Analysis of the life cycle environmental impacts related to the total final consumption of the EU25

Dannenberg A, Weingärtner E (2023) The effects of observability and an information nudge on food choice. J Environ Econ Manag 120:102829. https://doi.org/10.1016/j.jeem.2023.102829

Davis LW, Metcalf GE (2016) Does better information lead to better choices? Evidence from energy-efficiency labels. J Assoc Environ Resour Econ 3(3):589–625. https://doi.org/10.1086/686252

Elofsson K, Bengtsson N, Matsdotter E, Arntyr J (2016) The impact of climate information on milk demand: Evidence from a field experiment. Food Policy 58:14–23. https://doi.org/10.1016/j.foodpol.2015.11.002

Enlund J, Andersson D, Carlsson F (2023) Individual carbon footprint reduction: evidence from pro-environmental users of a carbon calculator. Environ Resour Econ 86(3):433–467https://doi.org/10.1007/s10640-023-00800-7

Environmental Protection Agency (2019) Inventory of U.S. greenhouse gas emissions and sinks: 1990–2017

Ferraro PJ, Price MK (2013) Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment. Rev Econ Stat 95(1):64–73. https://doi.org/10.1162/REST_a_00344

Mogensen L, Kidmose U, Hermansen JE (2009) Fødevarernes klimaaftryk, sammenhæng mellem kostpyramiden og klimapyramiden, samt omfang og effekt af fødevarespild, Nr. 29016, Nr. 541099, 30 s., jun. 15, 2009

Garnett T (2011) Where are the best opportunities for reducing greenhouse gas emissions in the food system (including the food chain)? Food Policy 36:S23–S32https://doi.org/10.1016/j.foodpol.2010.10.010

Goetz A, Marinica I, Mayr H, Mosetti L, Schubert R (2022) Do mobile applications foster sustainable mobility? Evidence from a field experiment. Available at SSRN: https://ssrn.com/abstract=4310386https://doi.org/10.2139/ssrn.4310386

Goldstein B, Gounaridis D, Newell JP (2020) The carbon footprint of household energy use in the United States. Proc Natl Acad Sci USA 117(32):19122–19130. https://doi.org/10.1073/pnas.1922205117

Guthrie J, Mancino L, Lin CTJ (2015) Nudging consumers toward better food choices: Policy approaches to changing food consumption behaviors. Psychol Mark 32(5):501–511. https://doi.org/10.1002/mar.20795

Hahn R, Metcalfe R (2016) The impact of behavioral science experiments on energy policy. Econ Energy Environ Policy 5(2):27–44. https://doi.org/10.5547/2160-5890.5.2.rhah

Hertwich EG, Peters GP (2009) Carbon footprint of nations: A global, trade-linked analysis. Environ Sci Technol 43(16):6414–6420. https://doi.org/10.1021/es803496a

Hummel D, Maedche A (2019) How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. J Behav Exp Econ 80:47–58. https://doi.org/10.1016/j.socec.2019.03.005

Intergovernmental Panel on Climate Change (2018) Global warming of 1.5°C. In: Masson-Delmotte V, Zhai P, Pörtner H-O, Roberts D, Skea J, Shukla PR, Pirani A, Moufouma-Okia W, Péan C, Pidcock R, Connors S, Matthews JBR, Chen Y, Zhou X, Gomis MI, Lonnoy E, Maycock T, Tignor M, Waterfield T (eds) An IPCC special report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty. Cambridge University Press, Cambridge and New York, pp 3–24. https://doi.org/10.1017/9781009157940.001

Intergovernmental Panel on Climate Change (2022) Climate change 2022: impacts, adaptation and vulnerability. In: Pörtner H-O, Roberts DC, Tignor M, Poloczanska ES, Mintenbeck K, Alegría A, Craig M, Langsdorf S, Löschke S, Möller V, Okem A, Rama B (eds) Contribution of working group II to the sixth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge and New York, pp 37–118. https://doi.org/10.1017/9781009325844.002

International Organization for Standardization (2006) Environmental management. Life cycle assessment. Principles and framework. ISO report 14040:2006. Technical Committee : ISO/TC 207/SC 5. ICS : 13.020.10 13.020.60

Ito K, Ida T, Tanaka M (2018) Moral Suasion and economic incentives: Field experimental evidence from energy demand. Am Econ J Econ Pol 10(1):240–267. https://doi.org/10.1257/pol.20160093

Jalil AJ, Tasoff J, Bustamante AV (2020) Eating to save the planet: Evidence from a randomized controlled trial using individual-level food purchase data. Food Policy. https://doi.org/10.1016/j.foodpol.2020.101950

Jessoe K, Rapson D (2014) Knowledge is (Less) power: Experimental evidence from residential energy use. Am Econ Rev 104(4):1417–1438. https://doi.org/10.1257/aer.104.4.1417

Kanay A, Hilton D, Charalambides L, Corrégé JB, Inaudi E, Waroquier L, Cézéra S (2021) Making the carbon basket count: Goal setting promotes sustainable consumption in a simulated online supermarket. J Econ Psychol 83:47–58. https://doi.org/10.1016/j.joep.2020.102348

Lacroix K (2018) Comparing the relative mitigation potential of individual pro-environmental behaviors. J Clean Prod 195:1398–1407. https://doi.org/10.1016/j.jclepro.2018.05.068

List JA, Price MK (2016) The use of field experiments in environmental and resource economics. Rev Environ Econ Policy 10(2):206–225. https://doi.org/10.1093/reep/rew008

Liu PJ, Wisdom J, Roberto CA, Liu LJ, Ubel PA (2014) Using behavioral economics to design more effective food policies to address obesity. Appl Econ Perspect Policy 36(1):6–24. https://doi.org/10.1093/aepp/ppt027

Lohmann PM, Gsottbauer E, Doherty A, Kontoleon A (2022) Do carbon footprint labels promote climatarian diets? Evidence from a large-scale field experiment. J Environ Econ Manag 114:102693. https://doi.org/10.1016/j.jeem.2022.102693

Macdiarmid JI, Douglas F, Campbell J (2016) Eating like there's no tomorrow: Public awareness of the environmental impact of food and reluctance to eat less meat as part of a sustainable diet. Appetite 96:487–493. https://doi.org/10.1016/j.appet.2015.10.011

Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E (2020) Mobile Apps for Health Behavior Change in Physical Activity, Diet, Drug and Alcohol Use, and Mental Health: Systematic Review. JMIR Mhealth Uhealth 8(3):e17046. https://doi.org/10.2196/17046

Muller L, Lacroix A, Lusk JL, Ruffieux B (2017) Distributional Impacts of Fat Taxes and Thin Subsidies. Econ J 127(604):2066–2092. https://doi.org/10.1111/ecoj.12357

Osman M, Thornton K (2019) Traffic light labelling of meals to promote sustainable consumption and healthy eating. Appetite 138:60–71. https://doi.org/10.1016/j.appet.2019.03.015

Panzone LA, Ulph A, Hilton D, Gortemaker I, Tajudeen IA (2021) Sustainable by Design: Choice Architecture and the Carbon Footprint of Grocery Shopping. J Pub Policy Mark 40(4):463–486. https://doi.org/10.1177/07439156211008898

Panzone LA, Ulph A, Zizzo DJ, Hilton D, Clear A (2021) The impact of environmental recall and carbon taxation on the carbon footprint of supermarket shopping. J Environ Econ Manag 109:1. https://doi.org/10.1016/j.jeem.2018.06.002

Lohmann P, Gsottbauer E (2022) Do Carbon Footprint Labels promote Climatarian Diets? Evidence from a large-scale Field Experiment. J Environ Econ Manag 114:102693

Perino G, Panzone LA, Swanson T (2014) Motivation crowding in real consumption decisions: Who is messing with my groceries? Econ Inq 52(2):592–607. https://doi.org/10.1111/ecin.12024

Poore J, Nemecek T (2018) Reducing food's environmental impacts through producers and consumers. Science 360(6392):987–992. https://doi.org/10.1126/science.aaq0216

Spaargaren G, van Koppen CSA, Janssen AM, Hendriksen A, Kolfschoten CJ (2013) Consumer Responses to the Carbon Labelling of Food: A Real Life Experiment in a Canteen Practice. Sociol Rural 53(4):432–453. https://doi.org/10.1111/soru.12009

Suchier J, Demarque C, Waroquier L, Girandola F, Hilton D, Muller L (2023) Effects of intra- and inter-category traffic-light carbon labels and the presence of a social norm cue on food purchases. J Consum Behav 22(3):597–617. https://doi.org/10.1002/cb.2139

Taylor MR, Tonsor GT (2013) Revealed demand for country-of-origin labeling of meat in the united states. J Agric Resour Econ 38(2):235–247. https://doi.org/10.22004/ag.econ.158287

Tiefenbeck V, Goette L, Degen K, Tasic V, Fleisch E, Lalive R, Staake T (2018) Overcoming salience bias: How real-time feedback fosters resource conservation. Manage Sci 64(3):1458–1476. https://doi.org/10.1287/mnsc.2016.2646

UNEP (2017) The emissions gap report 2017. United Nations environment Programme (UNEP), Nairobi. Job number: DEW/2136/PA

Vermeulen SJ, Campbell BM, Ingram JSI (2012) Climate change and food systems. Ann Rev Environ Resour 37:195–222. https://doi.org/10.1146/annurev-environ-020411-130608

Visschers VHM, Siegrist M (2015) Does better for the environment mean less tasty? Offering more climate-friendly meals is good for the environment and customer satisfaction. Appetite 95:475–483. https://doi.org/10.1016/j.appet.2015.08.013

Vlaeminck P, Jiang T, Vranken L (2014) Food labeling and eco-friendly consumption: Experimental evidence from a Belgian supermarket. Ecol Econ 108:180–190. https://doi.org/10.1016/j.ecolecon.2014.10.019

Willett W, Rockström J, Loken B, Springmann M, Lang T, Vermeulen S, Garnett T, Tilman D, DeClerck F, Wood A, Jonell M (2019) Food in the Anthropocene: the EAT–lancet commission on healthy diets from sustainable food systems. Lancet 393(10170):447–492https://doi.org/10.1016/S0140-6736(18)31788-4

Wynes S, Nicholas KA, Zhao J, Donner SD (2018) Measuring what works: Quantifying greenhouse gas emission reductions of behavioural interventions to reduce driving, meat consumption, and household energy use. Environ Res Lett 13(11):113002. https://doi.org/10.1088/1748-9326/aae5d7

Ytreberg NS, Alfnes F, van Oort B (2023) Mapping of the digital climate nudges in Nordic online grocery stores. Sustain Prod Consum 37. https://doi.org/10.1016/j.spc.2023.02.018

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.