Ecological Economics 241 (2026) 108844

Contents lists available at ScienceDirect

Ecological Economics

journal homepage: www.elsevier.com/locate/ecolecon

Unpacking the side effects of social norm nudges☆

Julien Picard [a,b,c,*], Sanchayan Banerjee [d]

[a] Politecnico Di Milano, Via Lambruschini 4/b, Milan, 20156, Italy
[b] CMCC Foundation - Euro-Mediterranean Center on Climate Change, Italy
[c] RFF-CMCC European Institute on Economics and the Environment, Italy
[d] King's College, 22 Kingsway, London, WC2B 6NR, United Kingdom

ARTICLE INFO

Keywords:
Social norm
Behavioural spillovers
Policy spillovers
Pro-environmental behaviour
Environmental policies

ABSTRACT

Growing evidence shows that green nudges can affect secondary (non-targeted) behaviours. However, it is unclear whether such spillovers arise as a result of engaging in the targeted behaviour or if they are a by-product of policies. In a survey experiment with 2775 English respondents, we test whether a social norm nudge promoting vegetarianism alters a non-targeted action: environmental donations. Social norm messaging increases intended vegetarian choices, but not uniformly. A subgroup identified with machine learning drives this effect. However, we find the social norm nudge crowds out donations from respondents in this subgroup. Our experimental design allows us to rule out the possibility that this crowding-out effect is driven by respondents substituting climate-friendly food choices with environmental donations. This suggests that social norm nudges may work, but at the risk of discouraging other environmentally friendly actions.

1. Introduction

Climate change is one of the most critical challenges of the 21st century. Growing evidence indicates we are at the "brink of an irreversible climate disaster" (Ripple et al., 2024). In response, policymakers have increasingly turned to behavioural interventions — such as nudges and informational campaigns — to promote sustainable habits (Sunstein, 2020). Yet, mitigating climate change requires drastic habit changes in many domains. Therefore, it is unclear whether these policy interventions can lead to systematic behaviour shifts (Chater and Loewenstein, 2023).

Initial findings suggest they might. For instance, social norm interventions that promote water conservation have also been found to reduce electricity use (Jessoe et al., 2021; Bonan et al., 2024; Carlsson et al., 2021). But these positive "spillover" effects are typically reported without distinguishing whether they stem from the behaviour itself (e.g., conserving water) or the intervention promoting it (e.g., the social norm message). This distinction matters for policy design.

For instance, consider a social norm intervention that encourages households to use less water. Suppose that after complying, these households become less likely to recycle—a negative spillover. If both the behaviour and the policy contribute to this decline, it might indicate that saving water makes people feel they have "done their part". The policy reinforces this mindset, suggesting that either the behaviour or the message should be reconsidered. On the other hand, if conserving water tends to reduce recycling but the policy helps counteract that tendency, then the target behaviour should be reconsidered or the intervention improved further. Conversely, if saving water leads to more recycling, but the policy undermines this positive pattern, then the policy needs to be redesigned. In all three cases, the overall spillover effect is the same, but the underlying causes and the appropriate policy responses differ.

In this paper, we use an experimental strategy to distinguish these sources of spillovers and a theoretical framework to inform their mechanisms. Then, we use this design in a survey experiment testing the effect of social norm messaging promoting vegetarianism on a sample of 2775 non-vegetarian English respondents. Meat production accounts for approximately 14% of global greenhouse gas emissions and is the largest source of methane emissions, representing a significant challenge for mitigation (Poore and Nemecek, 2018; Green et al., 2015; Riahi et al., 2022). However, internalising these negative externalities through Pigouvian prices is likely to be met with resistance (Douenne and Fabre, 2022). This makes the use of "soft" environmental policies, such as social norm messaging, appealing in this context. Social norm messages are applied widely in the environmental domain. They have

☆ We thank Eugenie Dugoua, Marion Dumas, Elisabeth Gsottbauer, Sarah Gharbi, Gregor Singer, Susana Mourato, Kelsey Jack, Romain Espinosa, Jean-Marc Bourgeon, Mark Andor, Meryem Yankol-Schalck, and many conference participants for their insightful comments on the paper. We also thank the Royal Geographic Society for its financial support.

* Corresponding author.
E-mail addresses: julienroger.picard@polimi.it (J. Picard), sanchayan.1.banerjee@kcl.ac.uk (S. Banerjee).

https://doi.org/10.1016/j.ecolecon.2025.108844
Received 17 April 2025; Received in revised form 5 August 2025; Accepted 28 October 2025
Available online 3 November 2025

been used to foster recycling (Andersson and von Borgstede, 2010; Bratt, 1999; Fornara et al., 2011; Nigbur et al., 2010), promote sustainable diets (Sparkman and Walton, 2017; Sparkman et al., 2020; Salmivaara and Lankoski, 2019; Testa et al., 2018; Stea and Pickering, 2019; Wenzig and Gruchmann, 2018; Richter et al., 2018), improve water and electricity consumption (Allcott, 2011; Costa and Kahn, 2013; Carrico and Riemer, 2011; Nolan et al., 2008; Handgraaf et al., 2013; Ferraro et al., 2011; Lapinski et al., 2007), and even foster towel reuse in hotels (Reese et al., 2014; Goldstein et al., 2008; Schultz et al., 2008). These messages consist of information on what others do, approve or disapprove of (Bicchieri, 2016). Their widespread use makes the study of their policy spillover effects particularly important.

We model the mechanisms underpinning behavioural and policy spillover effects in a utility maximisation framework. Our theory departs from other models that explain behavioural spillovers through limited attention (Nafziger, 2020; Altmann et al., 2024, 2022). We relate more closely to previous work that explains behavioural spillovers by self-image concerns (Goetz et al., 2024; Alt and Gallier, 2022; Ek, 2018). We differ by modelling behavioural policies as playing on intrinsic or extrinsic motivations, which either strengthen or weaken individuals' pro-environmental identity. We show that policies spill over to non-targeted pro-environmental behaviours through two channels. The first is indirect. Policies foster the initial pro-environmental decision, triggering behavioural spillovers. The second, our "policy spillover", is direct and either amplifies or weakens behavioural spillovers based on whether policies reinforced intrinsic or extrinsic motivations. Policy and behavioural spillovers can, therefore, have opposite signs.

We disentangle these two sources of spillovers with an instrumental variable embedded in the design. Namely, beyond allocating participants into control and treatment groups, we vary the salience of vegetarian items on the restaurant menus on which people make food choices. This alters the likelihood of choosing a vegetarian dish without directly affecting donations to pro-environmental charities, the non-targeted decisions on which we measure spillover effects. We use vegetarian salience as an instrumental variable to estimate the causal effect of choosing a vegetarian meal on donations and disentangle it from the policy spillover effect of social norm messaging.

We further investigate whether policy-driven spillovers are heterogeneous depending on respondents' inclination to follow the norm. In another treatment arm (n = 2782), respondents revealed their inclination to conform to the norm. We use this extra survey data to investigate this heterogeneity in an exploratory analysis. Using machine learning, we classify respondents of the main experiment into a spectrum of profiles ranging from *unwilling* to follow the norm, *hesitating* about following the norm, *trying* to follow the norm, and *following* the norm. This allows us to get a conditional average treatment effect of social norm messaging for each profile.

We find that the social norm message increases the likelihood of choosing a vegetarian item on average. Yet, this result hides heterogeneity. Respondents predicted to be *trying* to follow the norm drive this effect. However, they do not significantly decrease the carbon footprint of their food choices. Conversely, respondents that we predict will *hesitate* to follow the norm make less carbon-intensive food choices when nudged. The nudge does not affect the choices of respondents classified at the extremes of the spectrum: those predicted to be *unwilling* to conform and those predicted to be already *conforming*.

We also find evidence of a positive behavioural spillover effect on average. Namely, respondents choosing vegetarian food are more likely to give to pro-environmental charities. However, this effect is too small to induce the social norm message to increase donations through increasing vegetarian choices. Furthermore, providing social norm information crowds out donations of those predicted to be trying to conform. This negative policy spillover effect dominates the positive behavioural spillover effect. Our model suggests that the nudge pushes this group to act out of extrinsic motivation (e.g., through social pressure). This, in turn, reduces their engagement in the donation task. In other words, there is no "free lunch". Whenever the social norm nudge succeeds in increasing vegetarian food choice intentions, it is at the cost of crowding out this willingness to do more.

We contribute to a burgeoning literature assessing the spillover effects of pro-environmental policies. To our knowledge, only Alt and Gallier (2022), Alt et al. (2024), and León et al. (2024) have endeavoured to disentangle different channels of spillovers. Alt and Gallier (2022) do not rely on an instrumental variable strategy. Instead, they randomly offer participants the targeted behaviours to estimate behavioural spillover effects. This design relies on the strong assumption that choosing (not) to do the targeted behaviour is the same as (not) being proposed to do it. We differ from Alt et al. (2024) and León et al. (2024) in that we employ an empirical strategy that directly stems from our theoretical framework. This enables us to infer the mechanisms of policies from the sign of their policy spillovers. We also examine pro-environmental decisions, while Alt et al. (2024) and León et al. (2024) rely on abstract real-effort tasks. We also explore the heterogeneity in the spillover effects triggered by the social norm message. Thus, our study provides richer insights into the trade-off policymakers may face between nudging pro-environmental behaviours and crowding out others. A limitation of our study is that food choices are only declarative. This, however, gives us full control over the environment in which people choose food. We took advantage of this for our identification strategy. We present reassuring evidence showing that the empirical patterns of food choices are coherent with pre- and post-treatment attitudinal information.

Our estimation strategy expands Bonev (2025)'s recommendations to estimate behavioural spillovers. Bonev (2025)'s paper shows the conditions under which the policy itself can be used as an instrumental variable to estimate behavioural spillovers. We differ by introducing an instrumental variable in the design *on top* of the policy. This design allows us to causally identify both behavioural and policy spillovers. It corresponds to the second part of the parallel design proposed by Imai et al. (2013), and as such, we do not implement a version of the experiment without the instrumental variable. This choice increases statistical power, but requires assuming no interaction effects between the policy and the instrumental variable to identify the total spillover effect of the policy on donations, as well as the indirect effect of the policy on donations through food choices.

We also contribute to a growing body of literature that demonstrates heterogeneity in the effect of a social norm nudge (Helferich et al., 2023; Mundt et al., 2024; Czajkowski et al., 2019; Vesely et al., 2022). Our findings also contribute to improving our understanding of how spillover effects are moderated by one's values and norms (e.g., Thøgersen and Ölander 2003, Nash et al. 2019, Maki et al. 2019). Our approach differs from moderator-based analyses (e.g., Krefeld-Schwalb et al. 2024) in that it does not rely on direct measurements. This sidesteps the challenges of pre-treatment questions that can hint at the study's objectives. Furthermore, unlike other machine learning techniques—e.g., see Athey and Imbens (2015) and Künzel et al. (2019)—the source of heterogeneity is explicitly identified. In our case, heterogeneity stems from people's readiness to conform to the norm. Our results suggest that social norm messaging is only effective for specific social-demographic profiles, albeit at the cost of potentially crowding out further engagement.

This "no free lunch" result resonates with a broader debate questioning nudges and individual-based solutions to address climate change. Most notably, Chater and Loewenstein (2023) warn that focusing on individual actions diverts precious resources which could have been allocated to structural policies with much higher mitigation potential. In this regard, the meta-analysis of Nisa et al. (2019) – which presents alarmingly low effects of green nudges – triggered a prolific debate on the ability of such interventions to produce any improvement (Nisa et al., 2020; van der Linden and Goldberg, 2020; Stern, 2020). Our results suggest that social norm nudges do not generate sustained habitual change and might even morally license people to undertake further pro-environmental efforts (Merritt et al., 2010)

The remainder of this article is articulated as follows. We present our theoretical model in Section 2. In Section 3, we present our empirical framework. In Section 4, we share the experimental results from a social norm nudge promoting vegetarianism. Section 5 discusses key aspects of our design, including the use of the instrumental variable strategy, the implications of the hypothetical nature of food choices, and the application of machine learning to investigate heterogeneity. Section 6 concludes.

2. Conceptual framework

Social psychology provides competing predictions of how one pro-environmental action influences subsequent ones. Some theories, such as moral licensing, suggest that an initial pro-environmental action can license us to do less (Monin and Miller, 2001). Others, including cognitive dissonance theory or social identity theory, predict positive consistency effects (Festinger, 1962; Tajfel et al., 1979).[1] Motivations seem to be a key moderator: the meta-analysis of Maki et al. (2019) suggests that intrinsically motivated actions (done as an "end to themselves") foster consistency. Conversely, extrinsic motivations (e.g., rewards, social pressure) may crowd out intrinsic motives and induce moral licensing (Deci et al., 1999; Cerasoli et al., 2014). In this section, we use a utility-maximisation framework to model these competing dynamics and show how policies shape spillover effects. We can therefore use this model to infer the underlying mechanism of policies from the sign of their spillovers.

We adopt a simple two-period framework. Let us denote by x_1 the pro-environmental effort in period 1 and x_2 the pro-environmental effort in period 2. Individuals' payoff from pro-environmental efforts is:

$$u(x_t, I_t, \eta_t, \epsilon_t) \tag{1}$$

We denote the periods by $t = 1, 2$. Parameters η_t and ϵ_t capture the strength of intrinsic and extrinsic motivations for making pro-environmental efforts. High η_t implies that individuals make pro-environmental effort x_t as an end in itself (e.g., because they care about the environment). High ϵ_t implies that individuals do x_t as a means to an end (e.g., to be praised). As in Akerlof and Kranton (2000), I_t captures individuals' identity. When $I_t > 0$, individuals define themselves as pro-environmental (protecting the environment guides their actions). When $I_t < 0$, individuals define themselves as pro-self (selfish motives guide their actions).

Assumption 1. We assume that $u(x_t, I_t, \eta_t, \epsilon_t)$ is increasing and concave in $x_t \in \mathbb{R}_+$. Furthermore, $u''_{x\eta} > 0$, $u''_{x\epsilon} > 0$, and $u''_{xI} > 0$, where $u''_{x\eta}$ denotes the cross-derivative $\frac{\partial^2 u(\cdot)}{\partial x_t \partial \eta_t}$.

As in Bénabou and Tirole (2011), we model identity I_t as a "stock" building up based on individuals' past pro-environmental efforts. The way we model how pro-environmental efforts feed this stock reflects two key premises:

Premise 1. Acting pro-environmentally out of intrinsic motivation strengthens individuals' pro-environmental identity.

Premise 2. Acting pro-environmentally out of extrinsic motivation weakens individuals' pro-environmental identity.

[1] See Truelove et al. (2014) and Dolan and Galizzi (2015) for complete reviews of the literature on this question.

The first premise reflects the prediction made by cognitive dissonance theory: making pro-environmental efforts out of conviction should reinforce these convictions (Festinger, 1962). The second premise implies that pro-environmental efforts done as a means to an end (e.g., to be praised or gain social status) crowd out the pro-environmental identity as the effort is not self-driven (Deci et al., 1999; Cerasoli et al., 2014). These dynamics are captured as follows:

$$I_t = (\eta_{t-1} - \epsilon_{t-1}) \cdot x_{t-1} + \rho \cdot I_{t-1} \tag{2}$$

The "return" of pro-environmental efforts is positive when one is intrinsically motivated ($\eta_{t-1} > \epsilon_{t-1}$, doing the effort for its own sake), and negative when extrinsically motivated ($\epsilon_{t-1} > \eta_{t-1}$, doing the effort as a means to an end). Parameter ρ captures the depreciation rate of one's pro-environmental identity.

We also introduce the marginal cost of pro-environmental efforts at period t:

$$c(x_{t-1}, \delta_t) = a \cdot x_{t-1} + \delta_t \tag{3}$$

Parameter δ_t captures ancillary features of the choice environment, altering the difficulty of pro-environmental effort x_t (e.g., the salience of climate-friendly options). When $a < 0$, there is learning-by-doing: doing a first pro-environmental effort makes it easier to do more. When $a > 0$, efforts exerted in the previous period make the next effort costlier (e.g., individuals get tired).

Assumption 2. We assume that the marginal cost is always positive: $c(\cdot) > 0$

We focus on period 2: individuals choose their levels of pro-environmental effort based on their actions during the first period.[2] They solve:

$$\frac{\partial(u(\cdot) - c(\cdot) \cdot x_2)}{\partial x_2} = 0 \Leftrightarrow u'_{x_2}(\cdot) - c(\cdot) = 0 \tag{4}$$

To alleviate notations, denote by V the left-hand side of Eq. (4):

$$V \equiv u'_{x_2}(\cdot) - c(\cdot) \tag{5}$$

We want to know how the behavioural policies of period 1 can influence the decisions of period 2. In our model, policies either strengthen motivations (i.e., parameters η_1 or ϵ_1) or alter ancillary features of the choice environment δ_1 to increase pro-environmental efforts x_1. This subsequently affects the level of pro-environmental effort chosen in period 2.

Proposition 1. *Policies targeting $\theta_1 \in \{\eta_1, \epsilon_1\}$ alter efforts of period two through two channels:*

$$\underbrace{\frac{dx_2^*}{d\theta_1}}_{\text{Net spillover}} = \underbrace{\frac{V'_{x_1}}{-V'_{x_2}}}_{\text{Behavioural spillover}} \cdot \underbrace{\frac{\partial x_1^*}{\partial \theta_1}}_{\text{Main effect}} + \underbrace{\frac{V'_{\theta_1}}{-V'_{x_2}}}_{\text{Policy spillover}} \tag{6}$$

Policies targeting δ_1 only affect efforts of period 2 through a behavioural spillover effect.

Proposition 2. *Policies targeting δ_1 only alter efforts of period 2 through one channel:*

$$\underbrace{\frac{dx_2^*}{dc_1}}_{\text{Net spillover}} = \underbrace{\frac{V'_{x_1}}{-V'_{x_2}}}_{\text{Behavioural spillover}} \times \underbrace{\frac{\partial x_1^*}{\partial \delta_1}}_{\text{Main effect}} \tag{7}$$

[2] Individuals' sophistication in period 1 does not affect the results presented in this section. Therefore, we focus solely on period 2 as the rest fall outside the scope of this paper. See Picard (2023) for a complete discussion of the model and its implications.

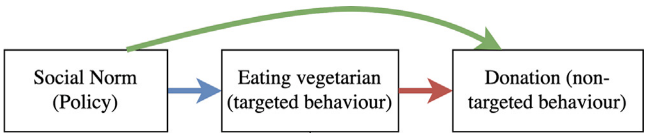

Fig. 1. Side effects a policy.
Note: Causal mechanisms of the effects of a policy on non-targeted efforts. The red arrow represents behavioural spillovers, whilst the green arrow represents policy spillover effects.

See Section 1 of the Online Appendix for proofs of propositions 1 and 2. Policies targeting intrinsic or extrinsic motivations affect the efforts of period 2 through two channels as described by Eq. (6). Fig. 1 illustrates these channels in the context of our experiment. The first channel is indirect: policies alter targeted pro-environmental effort (x_1), which subsequently influences non-targeted pro-environmental efforts (x_2). We refer to this effect as an "indirect spillover": it is the product of the *main effect* of the policy on the targeted decision and a *behavioural spillover* effect, represented by the red arrow in Fig. 1. The second channel is direct, through the effect of the policy on individuals' motivations. The green arrow represents this "policy spillover" in Fig. 1.

In the model, behavioural spillovers are the effect of pro-environmental efforts of period 1 on the marginal utility of doing pro-environmental efforts in period 2. The sign of the behavioural spillover effect depends on whether the initial effort was made out of intrinsic or extrinsic motivation ($\eta_t - \epsilon_t$ higher or lower than zero) and on whether the first effort facilitates ($a < 0$) or hardens further efforts ($a > 0$):

$$V'_{x_1} = V'_{I_2} \cdot (\eta_1 - \epsilon_1) - a \tag{8}$$

The policy spillover captures the effect of the policy on period two choices through its impact on motivations:

$$V'_{\eta_1} = V'_{I_2} \cdot x_1 > 0, \quad V'_{\epsilon_1} = -V'_{I_2} \cdot x_1 < 0 \tag{9}$$

When playing on intrinsic (extrinsic) motivations, policies increase (decrease) the marginal utility of making pro-environmental effort x_2.

Proposition 3. *Policies that play on intrinsic (or extrinsic) motivations trigger positive (or negative) policy-spillover effects.*

Proposition 3 implies that the sign of policy-spillover effects informs us of the mechanism through which policies induce behavioural change. In the next section, we propose an experimental design to estimate policy and behavioural spillover effects.

3. Empirical strategy

Estimating behavioural and policy spillovers is not trivial. Two complications arise. First, getting a *causal* estimate of behavioural spillovers is difficult. Unobserved variables can affect several pro-environmental actions simultaneously (e.g., values and beliefs). Second, a policy can increase intrinsic motivations for some people and extrinsic motivations for others. Thus, policy spillovers can differ from one person to another. This section presents our empirical framework to address these two issues.

Addressing omitted variable biases. We assume a population of N individuals indexed by i. Individuals are randomly assigned to a policy that fosters a given pro-environmental action. Denote by $\boldsymbol{x}_1$ the $N \times 1$ vector capturing individuals' decision to do the targeted pro-environmental action. Denote by $\boldsymbol{\theta}_1$ the $N \times 1$ vector capturing their treatment status, where $\theta_{1i} \in \{0, 1\}$. We assume that the relationships between the policy, the targeted pro-environmental action, and the non-targeted pro-environmental decision $\boldsymbol{x}_2$ can be described by a system of linear equations:

$$x_{1i} = \alpha^{ME} + \beta^{ME}\theta_{1i} + \varepsilon_i^{ME} \tag{10}$$

$$x_{2i} = \alpha^{NSE} + \beta^{NSE}\theta_{1i} + \varepsilon_i^{NSE} \tag{11}$$

Here, $\hat{\beta}^{ME}$ is the estimate of the effect of the policy on the targeted decision, $\boldsymbol{x}_1$. We refer to it as the *main effect* of the policy. $\hat{\beta}^{NSE}$ is the estimate of the effect of the policy on the non-targeted decision, $\boldsymbol{x}_2$. We refer to it as the *net spillover effect* of the policy. These estimates are unbiased under the following assumptions:

Assumption 3. The stable unit treatment value assumption holds, and the error terms ε_i^{ME} and ε_i^{NSE} are such that $cov(\boldsymbol{\varepsilon}^{ME}, \boldsymbol{\theta}_1) = cov(\boldsymbol{\varepsilon}^{NSE}, \boldsymbol{\theta}_1) = 0$

The equality $cov(\boldsymbol{\varepsilon}^{ME}, \boldsymbol{\theta}_1) = cov(\boldsymbol{\varepsilon}^{NSE}, \boldsymbol{\theta}_1) = 0$ holds when the policy is randomised. As we showed in Section 2, the net spillover of policies is composed of a behavioural spillover effect and a policy spillover effect. A naive approach to dissociate behavioural from policy spillovers consists of fitting the following linear model:

$$x_{2i} = \tilde{\alpha} + \tilde{\beta}^{BS}x_{1i} + \tilde{\beta}^{PS}\theta_{1i} + \tilde{\varepsilon}_i \tag{12}$$

$\hat{\tilde{\beta}}^{BS}$ is a naive estimate of the behavioural spillover. $\hat{\tilde{\beta}}^{PS}$ is the naive estimate of the policy spillover. These estimates are biased when unobserved variables simultaneously affect $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, implying $cov(x_{1i}, \tilde{\varepsilon}_i) \neq 0$. This omitted variable bias can be solved with an instrumental variable. A good instrumental variable alters $\boldsymbol{x}_1$ without changing people's intrinsic or extrinsic motivations to do $\boldsymbol{x}_1$. This is equivalent to randomly allocating people to a pure *choice-architecture* nudge. Denote by $\boldsymbol{\delta}_1$ the $N \times 1$ vector capturing people's allocation to this *choice architecture* nudge, where $\delta_{1i} \in \{0, 1\}$. We can then get unbiased estimates of behavioural and policy spillovers with two-stage least squares:

$$\begin{aligned}&\text{Stage 1: } x_{1i} = \alpha + \beta_1\delta_{1i} + \beta_2\theta_{1i} + \varepsilon_i\\&\text{Stage 2: } x_{2i} = \alpha' + \beta^{BS}\hat{x}_{1i} + \beta^{PS}\theta_{1i} + \varepsilon'_i\end{aligned} \tag{13}$$

Where $\hat{x}_{1i}$ are the predicted values for the first stage. To yield unbiased estimates of policy and behavioural spillovers, we need to make the following assumptions:

Assumption 4. The instrumental variable should be:

1. relevant ($cov(\boldsymbol{\delta}_1, \boldsymbol{x}_1) \neq 0$),
2. exogenous ($cov(\boldsymbol{\delta}_1, \boldsymbol{\varepsilon}') = 0$),
3. homogeneous ($x_{1i}(\overline{\delta}_1) \geq x_{1i}(\underline{\delta}_1)\ \forall i \in [1, \dots, N]$ and $\overline{\delta}_1 > \underline{\delta}_1$)

We also make the following assumption:

Assumption 5. Exposure to the policy does not alter the magnitude and the sign of the behavioural spillover effect for all $i \in [1, N]$.

Assumption 5 is akin to the assumption of "no interaction effect" of Imai et al. (2013). Together, assumptions Assumptions 4 and 5 ensure that we causally identify the policy and the behavioural spillovers.

One can draw a parallel with causal mediation analyses: the treatment is the policy θ_{1i}, the mediator is the targeted decision x_{1i}, and the outcome is the non-targeted decision x_{2i}. Our experiment aligns with the second condition in the parallel encouragement design proposed by Imai et al. (2013), where the encouragement is the variation in the choice architecture δ_{1i}. To estimate the net spillover effect, Imai et al. (2013) rely on another experimental condition where only the treatment is randomly manipulated. With our design, identifying the net spillover effect requires an additional assumption:

Assumption 6. The effect of the choice architecture on the targeted decision does not interact with the effect of the treatment for all $i \in [1, N]$.

Taken together, Assumptions 3, 4, 5, and 6 allow us to characterise the sign of the indirect spillover effect, i.e., the product of the main effect and the behavioural spillover:

Proposition 4. *Under assumptions Assumptions 3, 4, 5, and 6, the estimands of models (10), (11), and (13) are identified, such that:*

$$\underbrace{\hat{\beta}^{NSE}}_{\text{Net Spillover}} = \underbrace{\hat{\beta}^{BS}}_{\text{Behavioural spillover}} \times \underbrace{\hat{\beta}^{ME}}_{\text{Main effect}} + \underbrace{\hat{\beta}^{PS}}_{\text{Policy Spillover}} \tag{14}$$

See Section 1 of the Online Appendix for the proof. Proposition 4 shows that, under several assumptions, we can use Eq. (6) derived in Section 2 to interpret our estimates.

Addressing heterogeneity. People may react differently to a policy. To explore this heterogeneity, we define types based on characteristics that influence people's reactions to a policy. We then collect two data sets: a *main* sample and a *training* sample. In the *main* sample, we randomise the policy $\boldsymbol{\theta}_1$ and a choice architecture nudge $\boldsymbol{\delta}_1$. In the *training* sample, we elicit the types of new respondents. We use the *training* data to train an algorithm to predict these types. We then predict the types of respondents in the *main* sample with the algorithm.

Let us index by $j \in [1, \dots, N']$ the N' observations in the *training* sample where each observation's type y_j is known. Denote by $\boldsymbol{W}$ and $\boldsymbol{W'}$ the $N \times M$ and $N' \times M$ matrices of covariates of the *main* and the *training* samples. In three steps, we estimate the conditional average treatment effects of policy θ_1. First, estimate the function $y_i = f(W_i')$ such that:

$$\hat{f} \in \arg\min_{f} L(y_i, f(W_i')) \tag{15}$$

Where $L(\cdot)$ is a loss function. Then, predict the types of observations in the *main* sample:

$$\hat{y}_i = \hat{f}(W_i) \tag{16}$$

Finally, the treatment effects for each type are estimated.

The following section presents the results of our experiment studying the side effects of social norm messaging.

4. Experiment

In the rest of this article, we study whether a social norm message promoting vegetarian diets triggers any side effects on a non-targeted behaviour, donations to environmental charities, and we disentangle the channels of these side effects. In Section 4.1, we detail the design of this experiment. The results are presented in Section 4.2. Finally, we explore heterogeneity in Section 4.3.

4.1. Experimental design and data collection

We designed the survey experiment on Qualtrics and recruited respondents via Prolific. The experiment lasted approximately 10 min. We paid respondents according to Prolific's standard payment rate, £5 per hour. Upon completing the survey, respondents have a 1 in 100 chance to win a £20 voucher. In total, we recruited a sample of 5557 English respondents. They were divided between a *main* sample (n = 2775) and a *training sample* (n = 2782).

Respondents in the *main* sample took part in the main experiment. Its timeline is presented in Fig. 2.[3] We use the *training* sample to look at the heterogeneity in our treatment effects as part of an exploratory analysis (see Section 4.3).

Table 1
Sample sizes of treatment groups.

		Policy	
		Control	Treatment
Instrumental variable	Plant-intensive	690	693
	Meat-intensive	694	698

Policy. The policy of interest in this experiment is social information provision. More precisely, we consider the following dynamic social norm message[4]:

> *A study published in The Lancet Planetary Health found that the share of British people who stopped eating meat has increased by more than 50% from 2008 to 2019. More and more people are choosing plant-based dishes that are kinder to the planet and in turn, are becoming climate-friendly.*

Its formulation is like the one used by Blondin et al. (2022). The authors find that this message effectively increases intentions to choose vegetarian food. We randomly divided respondents into a treatment group, where they saw the message before making food choices (n = 1391), and a control group (n = 1384).

Instrumental variable. As explained in Section 3, estimating behavioural spillovers requires embedding an instrumental variable in the design. To do this, we vary the salience of vegetarian options when respondents make food choices. Respondents first see a subset of food items presented as the chef's selection (see Figures 3 and 4 in Section 5 of the Online Appendix). Half of the respondents see a selection containing mostly meat-based items (n = 1383). The other half saw a selection containing mostly vegetarian options (n = 1392). Respondents can choose an item from this selection or opt out and access the main menu, which contains all the items. We expect that respondents are more likely to choose a vegetarian item when vegetarian items are salient. Table 1 presents the sample size of each subgroup formed by the interaction between allocation to social norm messaging and the selections.

Targeted pro-environmental behaviour. We reproduce an online food order environment where participants choose a dish from a restaurant menu. The targeted pro-environmental decision is whether participants choose vegetarian food. We designed 24 versions of the main menu, varying the ordering and appearance of the items. In all menus, we label food items with pictures of footprints ranging from green to red. An explanation indicates that green footprints mean "completely climate-friendly" and red footprints mean "not climate-friendly at all" (see Figure 2 in Section 5 of the Online Appendix). As such, all participants have the same information on the environmental consequences of their choices. Table 14 in Section 5 of the Online Appendix presents the characteristics of the dishes in the menus.

Non-targeted pro-environmental behaviour. At the end of the survey, we ask participants if they want to donate an amount between £0 and £10 to a pro-environmental charity.[5] This task is our non-targeted pro-environmental behaviour. We use donations to proxy respondents' willingness to do extra pro-environmental behaviours. Donations are deducted from the £20 voucher.

[3] The pre-treatment and post-treatment survey questionnaires can be found in Section 3 of the Online Appendix. We pre-registered the experimental design, power analysis, empirical strategy and instrumental variable strategy on Open Science Framework (here). The pre-analysis plan describes a broader project encompassing three different studies: (1) the effect of familiar food choices on one's inclination to choose vegetarian food; (2) the effect of reflection on the effectiveness of social norm nudges; and (3) the present study. When reporting our results, we correct for the pre-registered hypotheses across the three studies. Deviations from the pre-analysis plan are documented and justified here.

[4] We construct it using the study of Stewart et al. (2021) analysing UK meat consumption trends using data from the National Diet and Nutrition Survey.

[5] Respondents are offered to give to the following charities: World Wide Fund (WWF), Friends of the Earth, Carbon Fund, Campaign against Climate Change, The Vegetarian Society, The Vegan Society, Extinction Rebellion, Woodland Trust. Alternatively, they can select "other" and write the name of their chosen charity.

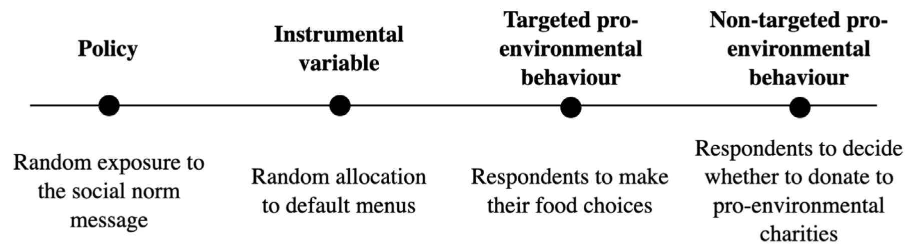

Fig. 2. Timeline of the experiment.

Table 2
Descriptive statistics.

	Control group (n = 1384)	Social norm group (n = 1391)	p-value
Age			0.139
Mean	38.6 years old	37.9 years old	
Median	36 years old	35 years old	
Income			0.920
< £10,000	18.6%	17.9%	
£10,000 - £15,999	11.5%	12.5%	
£16,000 - £19,999	11.3%	10.8%	
£20,000 - £29,999	27.2%	28.1%	
£30,000 - £39,999	16.2%	14.9%	
£40,000 - £49,999	8.5%	8.4%	
£50,000 - £69,999	4.5%	4.6%	
£70,000 - £89,999	1.5%	1.9%	
£90,000 - £119,999	0.6%	0.5%	
£120,000 - £149,999	0.2%	0.2%	
More than £150,000	0.0%	0.2%	
Gender			0.450
Female	48.3%	51.0%	
Male	50.7%	48.2%	
Other	1.0%	0.7%	
Education			0.961
No education	0.1%	0.1%	
Primary education	0.2%	0.1%	
Lower secondary education	2.5%	2.6%	
Upper secondary education	22.6%	21.9%	
Post-secondary non-tertiary education	15.6%	15.0%	
Short-cycle tertiary education	5.5%	6.6%	
Bachelor or equivalent	40.2%	39.4%	
Master or equivalent	11.9%	12.9%	
Doctoral or equivalent	1.5%	1.4%	

Note: descriptive statistics per treatment group. We use a Wilcoxon test to compare the age differences between the treatment and control groups. We use a Chi-square test for gender differences. We use trend tests to test the differences in education and income between the two groups.

Sample characteristics. We collected data from March 1st to April 24th of 2022. We pre-screened participants to select only native English speakers. We also excluded vegetarian and vegan participants. Attrition is low: 0.8% of respondents did not finish the survey. We excluded them. Table 2 shows descriptive statistics per treatment group. The median respondent is 35 years old, earns between £20,000 and £30,000 per year and has a Bachelor's degree. There sample is balanced in terms of gender, with 49.9% of females, 49.2% of males and 0.9% of respondents considering themselves genderfluid or agender. Comparisons using the UK census data and the survey of personal income suggest that our sample is younger, slightly poorer and more educated than the UK population (see Figure 6 in Section 5 of the Online Appendix). Randomisation was successful. No significant differences exist across the treatment groups in terms of age, gender, income, and education. About 98.28% of the *main* sample has passed an attention check placed at the beginning of the survey. From these 98.28%, 99.75% passed a focus check we placed after the pre-treatment questionnaire. Furthermore, 81.69% of the participants passed a manipulation check between the food choice and the donation task.

Table 3
ATE of the social norm message.

Outcome	Chose vegetarian food (binary)	Food choice in kgCO2-eq
Columns	(1)	(2)
Baseline	0.135***	5.850***
	(0.012)	(0.218)
Social norm	0.067***	−0.688**
	(0.016)	(0.232)
Plant-intensive	0.115***	−1.969***
Chef's selection	(0.016)	(0.232)
Num.Obs.	2775	2775
R2	0.025	0.028

Note: This table presents the effect of the social norm message and the effect of making vegetarian items salient on the likelihood of choosing a vegetarian food item (first column) and on the carbon footprint of food choices (second column). Coefficients are estimated using OLS. Robust standard errors are displayed in parentheses. We apply Benjamini and Hochberg (1995) correction to conventional p-values (p).
* p < 0.1, ** p < 0.05, *** p < 0.01.

4.2. Results

Main effect. Table 3 presents the effect of the social norm message on food choices estimated by fitting model (10) with ordinary least-squares (OLS).[6] The nudge increases intentions to choose vegetarian food by 6.7 percentage points and reduces the carbon footprint of food choices by 11.8%. Results are robust to non-linear probit specifications (see Table 3 in Section 4 of the Online Appendix).

Side effects. Table 4 displays the spillover effects of social norm messaging on the binary decision to donate (Panel A) and the amount donated (Panel B). The first column contains the results of the OLS specification (11), where we regress donations on exposure to the social norm message. This coefficient corresponds to the net spillover effect of the nudge. In both panels, these net spillovers are not significantly different from zero.

The second and fourth columns display the results obtained from fitting specification (12) with OLS. They correspond to the naive approach for disentangling behavioural and policy spillovers. The third and fifth columns display the results obtained from the two-stage least squares regression (13), where we instrument food choices. The effect of the social norm message on donations when controlling for food choices is not significantly different from zero, whether or not we instrument food choices. Thus, we do not find evidence of policy spillovers.

We find suggestive evidence of a positive behavioural spillover effect. The correlation between food choices and donations is statistically significant (columns two and four, specification (12)). When instrumenting food choices, we find that choosing a vegetarian dish increases the likelihood of giving by 36 percentage points (column three, specification (13)) and reducing the emissions of one's food choices by on kgCO2-eq increases the likelihood of giving by 2 percentage points

[6] Analyses were conducted on R using the package *estimatr* (Blair et al., 2022)

Table 4
Net spillovers, behavioural spillovers and policy spillovers.

Columns	(1)	(2)	(3)	(4)	(5)
Panel A					
Decision to donate (binary)					
Baseline	0.477***	0.443***	0.408***	0.525***	0.578***
	(0.013)	(0.014)	(0.035)	(0.015)	(0.049)
Social norm	0.008	−0.004	−0.016	0.002	−0.006
	(0.019)	(0.019)	(0.022)	(0.019)	(0.020)
Food choice		0.178***	0.357*	−0.010***	−0.021*
		(0.022)	(0.166)	(0.001)	(0.010)
			[0.005; 0.708]		[−0.040; −0.001]
Indirect spillover			0.024		0.014
(studentized bootstrap 95%-CIs)			[−0.025, 0.102]		[−0.037, 0.103]
R2	0.000	0.022		0.015	
Panel B					
Amount donated (in £)					
Baseline	3.309***	3.023***	2.870***	3.695***	3.956***
	(0.108)	(0.111)	(0.272)	(0.124)	(0.389)
Social norm	−0.009	−0.109	−0.163	−0.063	−0.100
	(0.151)	(0.150)	(0.175)	(0.151)	(0.161)
Food choice		1.490***	2.286	−0.079***	−0.133
		(0.187)	(1.309)	(0.012)	(0.077)
			[−0.494; 5.065]		[−0.289; 0.023]
Indirect spillover			0.154		0.091
(studentized bootstrap 95%-CIs)			[−0.328, 0.672]		[−0.459, 0.586]
R2	0.000	0.024		0.015	
Food choice		Chose vegetarian food		Food choice in kgCO2-eq	
Specification	OLS	OLS	2SLS	OLS	2SLS
Num.Obs.	2775	2775	2775	2775	2775

Note: This table displays the effect of food choices and the effect of the social norm nudge on the decision to donate (Panel A) and on the amount donated (Panel B). The first column shows the net spillovers of the social norm nudge on donations. The other columns show estimates of behavioural and policy spillovers. The second and fourth columns show estimates of the social norm nudge and food choices on donations. The third and the fifth columns display the same estimates, where this time, we instrument food choices. In columns two and three, we use a dummy equal to one when respondents choose vegetarian food and zero otherwise. In columns four and five, we use the carbon emissions of food choices in our regressions. Robust standard errors are displayed in parentheses. We apply Benjamini and Hochberg (1995) correction to conventional p-values (p). The brackets display 95% confidence intervals adjusted with Lee et al. (2022)'s procedure.
* p < 0.1, ** p < 0.05, *** p < 0.01.

(column five, specification (13)). There is no statistically significant effect on the amount donated after p-value correction. We do not observe a statistically significant difference between the instrumented and non-instrumented coefficients. The signs and magnitudes of our estimates are robust when adding controls and when using non-linear specifications (see Tables 4 and 5 in Section 4 of the Online Appendix). Finally, we compute the indirect spillover effect (i.e., behavioural spillover scaled by the main effect) using the relationship given by Proposition 4 and 2000 studentized bootstraps for the confidence intervals.[7] Indirect spillovers are not statistically different from zero.

4.3. Heterogeneity analysis

How people perceive the social norm nudge might depend on how much they are willing to follow the norm. For instance, telling respondents that more and more people are quitting meat can lead indecisive participants to change their behaviours, induce convinced meat-eaters to reaffirm their preferences, and be ignored by vegetarians with no room for improvement. In other words, the same social norm nudge likely affects people through different psychological processes. We investigate this heterogeneity by estimating the effect of the social norm message on different preference profiles as part of an exploratory analysis.

Training procedure. In a separate survey, we showed 2782 additional respondents the social norm message and then asked the following question[8]:

Are you trying to change your diet to become more climate-friendly as well?

(a) *No, I am not trying now, and I do not intend to try in future*
(b) *No, I am not trying now, but I might consider changing my diet to be more-climate-friendly in future*
(c) *Yes, I am trying to change my diet now to become more climate-friendly*
(d) *Yes, I have already changed my diet to be more climate-friendly*

We assume that asking this question after the social norm message reveals respondents' inclination to change their diets to follow the norm. It allows us to identify four types: the *transitioned type* is already conforming with the social norm; the *trying type* is inclined to conform; the *hesitant type* considers doing so in the future, and the *unwilling type* does not want to conform.

We use this additional survey data as a *training set* to train a gradient boosting machine classifier (GBM). The goal is to predict respondents' answers using attitudinal measures and social-demographic

[7] We use the R package *systemfit* from Henningsen and Hamann (2008) to estimate the variance–covariance matrix.

[8] This question is part of another treatment arm designed for another research project testing if inducing people to think about their choices increases the effectiveness of social norm nudges.

characteristics as predictors. Then, we use the algorithm to predict what respondents in the *main sample* would have answered.[9]

Like random forests, GBM builds an ensemble of decision trees. However, instead of fitting all trees independently, GBM fits each new tree to correct the errors of the previous ones (Friedman, 2001).

To assess the robustness of our predictions, we also train five alternative classification models: random forest, multinomial logistic regression, ordered logit, linear discriminant analysis, and quadratic discriminant analysis.[10]

We evaluate GBM's performance using nested 10 × 10 folds cross-validation. Overall, GBM performs 1.8 times better than random guessing (F1 Score = 0.433). Section 2 of the Online Appendix describes the performance evaluation in detail, along with the relative importance of each predictor. The four classes predicted by GBM closely resemble the original classes in the training set (see Figures 7, 8, and 9 in Section 5 of the Online Appendix). Table 19 in Section 5 of the Online Appendix provides the socio-demographic and attitudinal profiles associated with each predicted class.

Identification strategy. First, we estimate the effect of the social norm nudge on food choices for each predicted type. We use the *unwilling* type as our reference group and fit the following nested probability linear model[11]:

$$FoodChoice_i = \alpha + \sum_{k \in \Omega\_} \mathbf{1}_k \delta_k + \sum_{k \in \Omega} \mathbf{1}_k \beta_k Norm_i + u_i \tag{17}$$

$\Omega\_ = \{\text{hesitant, trying, transitioned}\}$

$\Omega = \{\text{unwilling, hesitant, trying, transitioned}\}$

And:

$$\mathbf{1}_k = \begin{cases} 1, & \text{if individual } i \text{ type } k \\ 0, & \text{otherwise} \end{cases}$$

Coefficient β_k is the causal effect of the social norm message conditional on being predicted to be of type k. To estimate the effect of the nudge on donation for each predicted type, we fit the following model:

$$Donation_i = \alpha + \sum_{k \in \Omega\_} \mathbf{1}_k \delta_k + \sum_{k \in \Omega} \mathbf{1}_k \beta_k Norm_i + u_i \tag{18}$$

Here again, β_k is the causal side-effect of the social norm nudge conditional on being predicted to be of type k. To investigate heterogeneity in the policy spillover effect, we fit the following model:

$$Donation_i = \alpha + \sum_{k \in \Omega\_} \mathbf{1}_k \delta_k + \sum_{k \in \Omega} \mathbf{1}_k \beta_k Norm_i + \beta_2 \widehat{FoodChoice}_i + \varepsilon_i \tag{19}$$

Here, the coefficient β_k is the policy spillover of the social norm message conditional on being predicted to be of type k. $\widehat{FoodChoice}_i$ captures food choices. We instrument it by the salience of vegetarian items. To check robustness, we fit these models with the predictions of five other algorithms. Furthermore, we re-estimate our GBM algorithm by oversampling the *unwilling* and *transitioned* categories that contain fewer observations. We also re-estimate GBM by adding income and political beliefs to the set of predictors. We previously excluded these variables as they contain too many missing values.

[9] Despite having excluded vegan and vegetarian participants, 12,6% of respondents chose the last answer. We see three explanations for this apparent contradiction. First, the screening was based on social demographic information gathered by Prolific, our data provider. As such, people may have changed their diets between when they answered the Prolific questionnaire and when they took our survey. Second, answers can also capture intentions rather than behaviours. Third, the phrasing of this answer could have been perceived as vague enough to allow non-vegetarian participants to select it without contradicting their actual behaviour.

[10] See Gareth et al. (2013) for a detailed explanation of these methods.

[11] Such a specification is equivalent to fitting four separate linear models for each predicted profile.

Table 5
Main effect of the social norm message conditional on respondents' types.

Outcome	Chose vegetarian food	Food choice in kgCO2-eq
Columns	(1)	(2)
Unwilling (baseline)	0.076***	6.163***
	(0.020)	(0.519)
Hesitant	0.087***	−0.896
	(0.024)	(0.570)
Trying	0.193***	−2.446***
	(0.031)	(0.598)
Transitioned	0.465***	−4.407***
	(0.075)	(0.727)
Social norm × Unwilling	0.005	0.243
	(0.030)	(0.753)
Social norm × Hesitant	0.046*	−0.902**
	(0.020)	(0.319)
Social norm × Trying	0.107**	−0.360
	(0.036)	(0.421)
Social norm × Transitioned	0.041	−0.138
	(0.092)	(0.650)
Specification	Nested OLS	Nested OLS
Num.Obs.	2730	2730
R2	0.070	0.031

Note: This table displays the effect of the social norm nudge on the likelihood of choosing vegetarian food (first column) and on the carbon footprint of food choices (second column) for each predicted type. For instance, coefficients labelled "Social norm × Trying" capture the average effect of the nudge on the predicted *trying* (the difference between control units and treatment units in this subsample). Coefficients labelled *Trying* capture the difference between the control units in the *trying* sample with the control units in the *unwilling* sample, our baseline. Robust standard errors are displayed in parentheses. We apply Benjamini and Hochberg (1995) correction to conventional p-values (p).
* p < 0.1, ** p < 0.05, *** p < 0.01.

Results. As shown in Figs. 3, being predicted to follow the norm positively correlates with the likelihood of choosing a vegetarian dish. It is also negatively correlated with the carbon footprint of food choices. Table 5 displays the results obtained by fitting Eq. (17). The social norm nudge increases the likelihood of choosing vegetarian food for the predicted *trying* (+10.7 percentage points) and only slightly for the predicted *hesitant* (+4.6 pp). The nudge only reduces the emissions of the predicted *hesitant* (−17%). These results are globally robust (see Tables 6 and 7 in Section 4 of the Online Appendix). Coefficients are of the same sign across all the algorithms and of the same order of magnitude.

Table 6 shows the results of regression (18) in the first two columns and regression (19) in the last four columns. The social norm nudge negatively affects the amount the predicted *trying type* gives. When controlling for instrumented food choices, we find that the nudge crowds out the amount they donate by about one pound. It also crowds out the likelihood of donating by around 10 percentage points. This policy spillover effect is globally robust (see Tables 10, 11, 12, and 13 in Section 4 of the Online Appendix).

5. Discussion

This section focuses on three aspects of our experiment. First, we assess the validity of our instrumental variable approach. Then, we discuss the extent to which having self-reported measures for food choices alters the validity of our results. Finally, we discuss the implications of our heterogeneity analysis.

Internal validity and strength of the IV strategy. The internal validity of our results depends on whether the instrumental variable strategy we have adopted is valid. Our instrumental variable should be relevant, exogenous and homogeneous. Regarding relevance, results in Table 3 show a strong and highly significant effect of making vegetarian items salient on the likelihood of choosing vegetarian food. The F-statistic of the IV is 53.400 in the binary case and 71.998 when looking at the

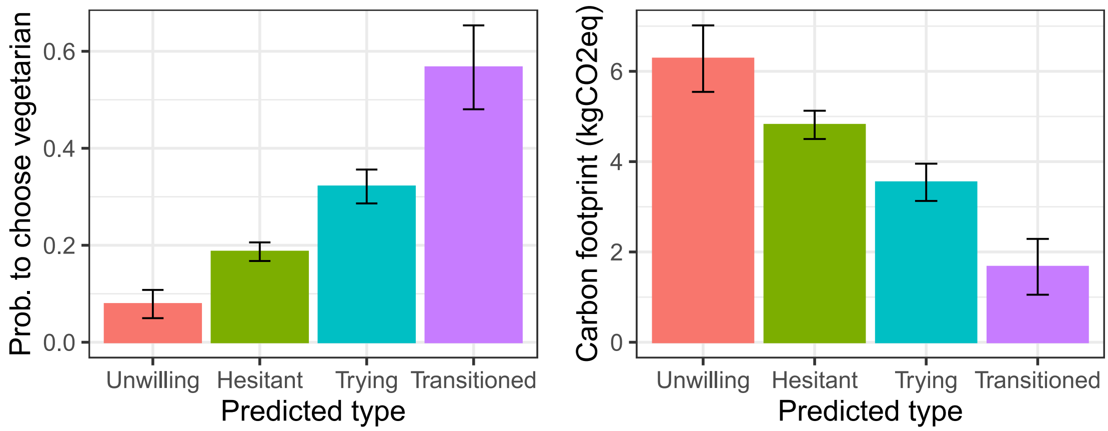

Fig. 3. Food choices of each predicted type.

Table 6
Policy spillover conditional on predicted types.

Outcome	Amount (in £)	Decision (binary)	Amount (in £)	Decision (binary)	Amount (in £)	Decision (binary)
Columns	(1)	(2)	(3)	(4)	(5)	(6)
Unwilling (baseline)	1.182***	0.188***	1.038***	0.167***	1.862***	0.290***
	(0.217)	(0.030)	(0.239)	(0.033)	(0.500)	(0.065)
Food choice			1.893	0.283*	−0.110	−0.016*
			(1.251)	(0.158)	(0.073)	(0.009)
Hesitant	1.709***	0.255***	1.544***	0.231***	1.610***	0.241***
	(0.255)	(0.035)	(0.275)	(0.038)	(0.265)	(0.037)
Trying	3.939***	0.483***	3.575***	0.429***	3.670***	0.443***
	(0.315)	(0.039)	(0.394)	(0.050)	(0.361)	(0.046)
Transitioned	3.630***	0.437***	2.750***	0.305***	3.144***	0.364***
	(0.664)	(0.077)	(0.910)	(0.110)	(0.746)	(0.087)
Social norm × Unwilling	0.499	0.068	0.490	0.067	0.526	0.072
	(0.334)	(0.046)	(0.339)	(0.047)	(0.342)	(0.047)
Social norm × Hesitant	0.171	0.025	0.083	0.012	0.071	0.010
	(0.190)	(0.025)	(0.200)	(0.026)	(0.203)	(0.026)
Social norm × Trying	−0.865**	−0.083*	−1.068**	−0.114**	−0.905**	−0.089*
	(0.323)	(0.037)	(0.350)	(0.040)	(0.324)	(0.037)
Social norm × Transitioned	−0.078	0.033	−0.155	0.022	−0.093	0.031
	(0.792)	(0.089)	(0.810)	(0.093)	(0.794)	(0.089)
Food choice			Chose vegetarian food		Food choice in kgCO2-eq	
Specification	Nested OLS	Nested OLS	Nested 2SLS	Nested 2SLS	Nested 2SLS	Nested 2SLS
Num.Obs.	2730	2730	2730	2730	2730	2730
R2	0.072	0.065				

Note: This table displays the effect of the social norm message on the decision to donate (columns 2, 4 and 6) and on the amount donated (columns 1, 3, and 5) for each predicted type. The first two columns show the overall effect of the social norm nudge on donations. The policy spillover of the social norm message is then estimated in the other columns, controlling for instrumented food choices. For instance, coefficients labelled "Social norm × Trying" capture the effect of the social norm on the predicted *trying*. Coefficients labelled *Trying* capture the difference between the control units in the *trying* sample and the control units in the *unwilling* sample, our baseline. Robust standard errors are displayed in parentheses. We apply Benjamini and Hochberg (1995) correction to conventional p-values (p).
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

carbon footprint of food choices.[12] This suggests that our instrument is strong (Bound et al., 1995; Staiger and Stock, 1997).

An instrumental variable is exogenous when it only affects the outcome variable (here, donations) through the endogenous variable (here, food choices). Our instrumental variable is like a default nudge. Respondents must opt out of the selection we show them "by default" to access the full menu. In our context, and using insights from the model of Section 2, exogeneity would be violated if the default altered vegetarian choices by playing on respondents' intrinsic or extrinsic motivations rather than by simply making them more or less accessible. Whilst we cannot formally rule out this possibility, recent empirical evidence suggests that default nudges affect people's decisions unconsciously (Gärtner, 2018; Van Gestel et al., 2020; Ortmann et al., 2023). This confirms priors in the literature (e.g., see Hansen and Jespersen 2013, Thaler and Sunstein 2009). This supports the assumption that the instrumental variable is unlikely to affect donations other than through food choices.

Angrist et al. (1996) show that the homogeneity assumption is threatened in the presence of defiers. In our experiment, defiers systematically choose meat-based options when vegetarian items are salient and vice versa. Angrist et al. (1996) show that biases introduced by defiers are small if the effect of choosing vegetarian food on donations for defiers differs from that of choosing vegetarian food on donations for compliers. The behaviour of defiers in our context would describe that of nonconformists going against suggestions. Such a behaviour is likely orthogonal to the psychological processes driving behavioural spillovers. Angrist et al. (1996) also show that the bias is small when the number of defiers is small. We cannot measure the number of defiers. Nevertheless, we observe that 44% of respondents chose a meat-based item when vegetarian items are salient and vice versa. This subsample

[12] This F-statistic is also robust to adding controls (59.432 in the binary case, 64.140 with carbon footprint).

also contains never-takers (always choosing meat) and always-takers (always choosing vegetarian), reducing the likelihood of having a large share of defiers in our sample. For these reasons, it seems unlikely that the number of defiers is large.

Finally, our empirical strategy relied on assuming no interaction effects, i.e., Assumptions 5 and 6 (see Section 3). Whilst we cannot directly test those assumptions, we would reject them if the interaction between our two treatments significantly differs from zero, which is not the case (see Table 18 in Section 5 of the Online Appendix). Exposure to the social norm nudge increases the likelihood of sticking with the "Chef's selection" when it is predominantly vegetarian and reduces this likelihood when it is meat-intensive (see Table 18 in Section 5 of the Online Appendix). This pattern is logical and does not indicate that varying the salience of vegetarian items interacts with the social norm message.

External validity, hypothetical decisions and experimenter demand. Hypothetical food choices gave us full control over respondents' choice environments, allowing us to apply our identification strategy. Yet, this can limit the external validity of our study by introducing an experimenter demand effect. Answers to a battery of post-treatment questions seem, however, reassuring. First, we ask two questions inspired by the literature on willingness-to-pay estimation (Andor et al., 2017; Ready et al., 2010; Champ et al., 2009; Mohammed, 2012). Namely, participants can revise their choices before continuing the survey. Then, we asked them if they would go to a restaurant offering similar food. We expect that revising one's choice indicates low confidence in one's preferences, increasing the risk of an intention-behaviour gap. Similarly, we expect that not wanting to go to a restaurant that offers similar food would reduce the external validity of our study. Only 1.62% of respondents revised their choices, and only 15.56% of them somewhat disagreed or strongly disagreed with going to a restaurant serving the same menu.

Several correlations indicate that our measure of food choices exhibits empirical patterns that are coherent with the expected behaviour of someone with a preference for meat. We observe a positive correlation between choosing meat and a feeling of sacrifice associated with choosing vegetarian food. Choosing meat is also associated with a habitual order among the predicted *trying* and *transitioned*. Meat choosers also ask for higher monetary compensations when, at the end of the experiment, we prompted them to report how much they would need to accept switching to a vegetarian dish. Importantly, even vegetarian choosers required compensation to switch to another vegetarian dish. If choices were driven solely by an experimenter demand effect, we would expect them to be indifferent. This suggests that their selections reflect genuine preferences rather than compliance with perceived experimental expectations. Furthermore, choosing meat is associated with the belief that British food should be meat-based, preference for British food, and lower climate concerns. Respondents choosing vegetarian food are also more likely to be females, have dietary constraints, and be more politically liberal (see Table 16 in Section 5 of the Online Appendix).

Despite these reassuring patterns, we cannot exclude that an experimenter's demand effect altered the treatment effects reported in this paper. However, we can exclude the possibility that the effect of the social norm nudge on food choices is *fully* driven by experimenter demand, as otherwise, we would expect the different profiles identified with the machine learning procedure to behave similarly. Furthermore, our results are in the same direction as the effects of dynamic social norm messages promoting vegetarian options tested in the field (Sparkman and Walton, 2017; Sparkman et al., 2021).

The fact that food choices are intentional could induce participants who chose vegetarian food to donate because they could not realise their intentions. Nevertheless, choosing a vegetarian item correlates positively with the feeling of having exerted an effort for the environment, which seems to contradict this interpretation (see Table 17 in Section 5 of the Online Appendix). Our results are also consistent with other studies that use an instrumental variable to estimate behavioural spillovers between pro-environmental decisions. For instance, Comin and Rode (2023) find that installing solar panels increase support for pro-environmental policies. Similarly, Alacevich et al. (2021) find that sorting waste leads households to decrease the amount of waste they generate.

Finally, the novelty of our study lies in the fact that we estimate a policy spillover effect, i.e., a spillover effect of our social norm message on donations that is distinct from behavioural spillover effects. To what extent is this result affected by food decisions being hypothetical? We do not have similar studies to compare our results with. Nonetheless, our estimates of the effects of the social norm nudge on food decisions and our estimates of behavioural spillovers are in the same direction as estimates reported in field experiments. It seems, therefore, unlikely that our estimates of policy-driven spillovers do not share the same qualitative characteristics as those that could be estimated in the field. Furthermore, unlike food choices, donations are incentive-compatible, thereby mitigating experimenter demand effects.

Whilst donations may not fully capture the breadth of environmental behaviour, they are commonly employed in the experimental literature as a proxy for pro-environmental engagement (e.g., Carrico et al. 2018, Fanghella et al. 2019, Bergquist et al. 2020, Amaris et al. 2024). Prior research finds that donation behaviour correlates significantly with pro-environmental and biospheric values (Lundberg et al., 2019; Woltin et al., 2022; Kesenheimer and Greitemeyer, 2021). In our own data, we also observe a strong positive association between concern about climate change and the likelihood of donating, suggesting that the donation measure reflects broader environmental orientations (see Table 16 in Section 5 of the Online Appendix).

Interpreting heterogeneity. We find that the nudge is effective for the *hesitant* type and the *trying* type. For the *hesitant* type, the nudge decreases food choices' carbon footprint but only slightly increases the uptake of vegetarian options. This apparent paradox may be caused by the predicted *hesitants* switching from carbon-intensive meat options to less intensive ones. Conversely, the nudge increases the uptake of vegetarian food but does not significantly decrease carbon emissions for the predicted *trying*. This might be because participants classed as *trying* switch from less intensive meat options to vegetarian options. This implies no statistically significant decrease in carbon emissions. Furthermore, it seems that the nudge does not affect respondents predicted to be *unwilling*. Although the absence of evidence is not evidence of the absence, this null result supports a common assumption in the literature that nudges are ineffective for those unwilling to change (Thaler and Sunstein, 2009). Similarly, the nudge does not significantly alter the choices of the predicted *transitioned* type. The *transitioned* respondents have the highest share of controlled units choosing vegetarian food. As such, it may be that *transitioned* respondents have no room for improvement.

We find robust evidence that the social norm message crowds out donations of the predicted *trying* type. Proposition 4 and our model in Section 2 suggest the *trying* respondents may have treated the social norm message as an extrinsic pressure to choose vegetarian food. This would have induced them to slacken once the extrinsic pressure vanishes. The theoretical framework of Truelove et al. (2014) provides a similar interpretation. For the authors, policies can induce people to act to repair a morally threatened identity. This induces moral licensing once the identity is repaired. Interestingly, the social norm nudge does not produce a similar negative policy spillover for the predicted *hesitants*. Respondents classed as *trying* are more aware of the environmental impact of diets. This can make them more prone to guilt when exposed to our message.

A concern that may arise when investigating heterogeneity is the reduction in statistical power that results from examining the effect of our treatments on subsets of our samples. A post-hoc power analysis conducted for the effects estimated on these subsamples indicate an achieved power of 85% at 5% significant level for the effect of the nudge on vegetarian choices found for the predicted *trying* (91% at α =10%) and 81% for the effect of the norm nudge on the carbon intensity of food found for the predicted *hesitant* (88% at α =10%). The achieved power is of 62% for the net spillover effect on the decision to donate (73% at α =10%) and of 76% for the net spillover effect on the amount donated (85% at α =10%). Results on the effect of the social norm message on donations appear to be underpowered. Nonetheless, they are broadly robust across several specifications and machine learning techniques.

6. Conclusion

Do policies trigger side effects on non-targeted behaviours that are distinct from behavioural spillovers? Section 2 shows that policy-driven spillovers arise through their effect on intrinsic or extrinsic motivations. We present an experimental design to disentangle policy spillovers from behavioural spillovers and estimate these two channels in an online experiment involving 2775 English respondents. We tested whether social norm messages promoting vegetarianism spilt over to another non-targeted decision: green donations.

We find that social norm messaging is effective in increasing vegetarian choices. Yet, the effect of this nudge is heterogeneous. The nudge works for individuals who are predicted to be trying to change their diets to follow the norm, as well as those who are hesitant about doing so. We only observe a decrease in the carbon footprint of food choices for the predicted *hesitants*. We also find that choosing a vegetarian item increases the likelihood of giving to a pro-environmental charity. However, this effect is too small to induce the social norm message to increase donations through food choices indirectly. Furthermore, the message crowds out the predicted *tryings*' donations. This negative policy spillover outweighs the positive behavioural spillover. We do not observe a similar crowding-out effect on the respondents who were predicted to be hesitant. This suggests that policymakers seeking to use social norm messaging to reduce the environmental impact of food choices should target this population segment. When it comes to increasing the uptake of vegetarian choices, our experimental findings indicate no "free lunch". When the social norm message effectively fosters vegetarian food choices, it is at the cost of crowding out further engagement.

This result calls for more empirical evidence on whether other policies in other contexts yield similar effects. The main concern is that the effectiveness of such green nudges might decrease over time. We have laid out a theoretical and empirical framework to assess this question. We hope this paper will provide a useful methodological framework for further research.

CRediT authorship contribution statement

Julien Picard: Writing – review & editing, Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Sanchayan Banerjee:** Writing – review & editing, Funding acquisition, Conceptualization.

Informed consent

All participants gave informed consent to complete this study.

Ethics approval

This study was approved by LSE's Institutional Review Board vide reference 38224.

Code availability

The codes and data sets for the analysis are available here.

Funding

This study was funded by the Royal Geographic Society (RGS–IBG) vide reference FSPA 05.21.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Julien Picard reports financial support was provided by Royal Geographical Society. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ecolecon.2025.108844.

Data availability

Link to online data folder and replication package shared in the manuscript.

References

Akerlof, G.A., Kranton, R.E., 2000. Economics and identity. Q. J. Econ. 115 (3), 715–753.

Alacevich, C., Bonev, P., Söderberg, M., 2021. Pro-environmental interventions and behavioral spillovers: Evidence from organic waste sorting in Sweden. J. Environ. Econ. Manag. 108, 102470.

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

Alt, M., Bruns, H., Della Valle, N., 2024. The more the better? synergies of prosocial interventions and effects on behavioural spillovers. J. Environ. Econ. Manag. 128, 103061.

Alt, M., Gallier, C., 2022. Incentives and intertemporal behavioral spillovers: A two-period experiment on charitable giving. J. Econ. Behav. Organ. 200, 959–972.

Altmann, S., Grunewald, A., Radbruch, J., 2022. Interventions and cognitive spillovers. Rev. Econ. Stud. 89 (5), 2293–2328.

Altmann, S., Grunewald, A., Radbruch, J., 2024. The double dividend of nudges. Technical report, CRC TRR 190 Rationality and Competition.

Amaris, G., Vesely, S., Hess, S., Klöckner, C.A., 2024. Can competing demands affect pro-environmental behaviour: a study of the impact of exposure to partly related sequential experiments. Ecol. Econom. 216, 108023.

Andersson, M., von Borgstede, C., 2010. Differentiation of determinants of low-cost and high-cost recycling. J. Environ. Psychol. 30 (4), 402–408.

Andor, M.A., Frondel, M., Vance, C., 2017. Mitigating hypothetical bias: Evidence on the effects of correctives from a large field study. Environ. Resour. Econ. 68 (3), 777–796.

Angrist, J.D., Imbens, G.W., Rubin, D.B., 1996. Identification of causal effects using instrumental variables. J. Amer. Statist. Assoc. 91 (434), 444–455.

Athey, S., Imbens, G.W., 2015. Machine learning methods for estimating heterogeneous causal effects. Stat 1050 (5), 1–26.

Bénabou, R., Tirole, J., 2011. Identity, morals, and taboos: Beliefs as assets. Q. J. Econ. 126 (2), 805–855.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Stat. Methodol. 57 (1), 289–300.

Bergquist, M., Nyström, L., Nilsson, A., 2020. Feeling or following? A field-experiment comparing social norms-based and emotions-based motives encouraging pro-environmental donations. J. Consum. Behav. 19 (4), 351–358.

Bicchieri, C., 2016. Norms in The Wild: How to Diagnose, Measure, and Change Social Norms. Oxford University Press.

Blair, G., Cooper, J., Coppock, A., Humphreys, M., Sonnet, L., 2022. Estimatr: Fast estimators for design-based inference. https://declaredesign.org/r/estimatr/, https://github.com/DeclareDesign/estimatr.

Blondin, S., Attwood, S., Vennard, D., Mayneris, V., 2022. Environmental messages promote plant-based food choices: An online restaurant menu study. World Resour. Inst..
Bonan, J., Cattaneo, C., d'Adda, G., Galliera, A., Tavoni, M., 2024. Widening the scope: the direct and spillover effects of nudging water efficiency in the presence of other behavioral interventions. J. Environ. Econ. Manag. 127, 103037.
Bonev, P., 2025. Behavioral spillovers. J. Econ. Behav. Organ. 229, 106857.
Bound, J., Jaeger, D.A., Baker, R.M., 1995. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. J. Amer. Statist. Assoc. 90 (430), 443–450.
Bratt, C., 1999. The impact of norms and assumed consequences on recycling behavior. Environ. Behav. 31 (5), 630–656.
Carlsson, F., Jaime, M., Villegas, C., 2021. Behavioral spillover effects from a social information campaign. J. Environ. Econ. Manag. 109, 102325.
Carrico, A.R., Raimi, K.T., Truelove, H.B., Eby, B., 2018. Putting your money where your mouth is: an experimental test of pro-environmental spillover from reducing meat consumption to monetary donations. Environ. Behav. 50 (7), 723–748.
Carrico, A.R., Riemer, M., 2011. Motivating energy conservation in the workplace: An evaluation of the use of group-level feedback and peer education. J. Environ. Psychol. 31 (1), 1–13.
Cerasoli, C.P., Nicklin, J.M., Ford, M.T., 2014. Intrinsic motivation and extrinsic incentives jointly predict performance: a 40-year meta-analysis. Psychol. Bull. 140 (4), 980.
Champ, P.A., Moore, R., Bishop, R.C., 2009. A comparison of approaches to mitigate hypothetical bias. Agric. Resour. Econ. Rev. 38 (2), 166–180.
Chater, N., Loewenstein, G., 2023. The i-frame and the s-frame: How focusing on individual-level solutions has led behavioral public policy astray. Behav. Brain Sci. 46, e147.
Comin, D.A., Rode, J., 2023. Do green users become green voters? Technical report, National Bureau of Economic Research.
Costa, D.L., Kahn, M.E., 2013. Energy conservation ''nudges'' and environmentalist ideology: Evidence from a randomized residential electricity field experiment. J. Eur. Econ. Assoc. 11 (3), 680–702.
Czajkowski, M., Zagórska, K., Hanley, N., 2019. Social norm nudging and preferences for household recycling. Resour. Energy Econ. 58, 101110.
Deci, E.L., Koestner, R., Ryan, R.M., 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol. Bull. 125 (6), 627.
Dolan, P., Galizzi, M.M., 2015. Like ripples on a pond: behavioral spillovers and their implications for research and policy. J. Econ. Psychol. 47, 1–16.
Douenne, T., Fabre, A., 2022. Yellow vests, pessimistic beliefs, and carbon tax aversion. Am. Econ. J.: Econ. Policy 14 (1), 81–110.
Ek, C., 2018. Prosocial behavior and policy spillovers: A multi-activity approach. J. Econ. Behav. Organ. 149, 356–371.
Fanghella, V., d'Adda, G., Tavoni, M., 2019. On the use of nudges to affect spillovers in environmental behaviors. Front. Psychol. 10, 61.
Ferraro, P.J., Miranda, J.J., Price, M.K., 2011. The persistence of treatment effects with norm-based policy instruments: evidence from a randomized environmental policy experiment. Am. Econ. Rev. 101 (3), 318–322.
Festinger, L., 1962. Cognitive dissonance. Sci. Am. 207 (4), 93–106.
Fornara, F., Carrus, G., Passafaro, P., Bonnes, M., 2011. Distinguishing the sources of normative influence on proenvironmental behaviors: The role of local norms in household waste recycling. Group Process. & Intergroup Relations 14 (5), 623–635.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 1189–1232.
Gareth, J., Daniela, W., Trevor, H., Robert, T., 2013. An Introduction to Statistical Learning: With Applications in R. Spinger.
Gärtner, M., 2018. The prosociality of intuitive decisions depends on the status quo. J. Behav. Exp. Econ. 74, 127–138.
Goetz, A., Mayr, H., Schubert, R., 2024. One thing leads to another: Evidence on the scope and persistence of behavioral spillovers. J. Public Econ. 236, 105166.
Goldstein, N.J., Cialdini, R.B., Griskevicius, V., 2008. A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. J. Consum. Res. 35 (3), 472–482.
Green, R., Milner, J., Dangour, A.D., Haines, A., Chalabi, Z., Markandya, A., Spadaro, J., Wilkinson, P., 2015. The potential to reduce greenhouse gas emissions in the UK through healthy and realistic dietary change. Clim. Change 129 (1), 253–265.
Handgraaf, M.J., De Jeude, M.A.V.L., Appelt, K.C., 2013. Public praise vs. private pay: Effects of rewards on energy conservation in the workplace. Ecol. Econom. 86, 86–92.
Hansen, P.G., Jespersen, A.M., 2013. Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy. Eur. J. Risk Regul. 4 (1), 3–28.
Helferich, M., Thøgersen, J., Bergquist, M., 2023. Direct and mediated impacts of social norms on pro-environmental behavior. Glob. Environ. Chang. 80, 102680.
Henningsen, A., Hamann, J.D., 2008. Systemfit: A package for estimating systems of simultaneous equations in R. J. Stat. Softw. 23, 1–40.
Imai, K., Tingley, D., Yamamoto, T., 2013. Experimental designs for identifying causal mechanisms. J. R. Stat. Soc. Ser. A: Stat. in Soc. 176 (1), 5–51.
Jessoe, K., Lade, G.E., Loge, F., Spang, E., 2021. Spillovers from behavioral interventions: Experimental evidence from water and energy use. J. Assoc. Environ. Resour. Econ. 8 (2), 315–346.
Kesenheimer, J.S., Greitemeyer, T., 2021. Going green (and not being just more pro-social): do attitude and personality specifically influence pro-environmental behavior? Sustainability 13 (6), 3560.
Krefeld-Schwalb, A., Sugerman, E.R., Johnson, E.J., 2024. Exposing omitted moderators: Explaining why effect sizes differ in the social sciences. Proc. Natl. Acad. Sci. 121 (12), e2306281121.
Künzel, S.R., Sekhon, J.S., Bickel, P.J., Yu, B., 2019. Metalearners for estimating heterogeneous treatment effects using machine learning. Proc. Natl. Acad. Sci. 116 (10), 4156–4165.
Lapinski, M.K., Rimal, R.N., DeVries, R., Lee, E.L., 2007. The role of group orientation and descriptive norms on water conservation attitudes and behaviors. Heal. Commun. 22 (2), 133–142.
Lee, D.S., McCrary, J., Moreira, M.J., Porter, J., 2022. Valid t-ratio inference for IV. Am. Econ. Rev. 112 (10), 3260–3290.
León, A.K., Picard, J., Schobin, J., 2024. Experimental evidence on the role of framing, difficulty and domain-similarity in shaping behavioral spillovers. Sci. Rep. 14 (1), 25124.
Lundberg, P., Vainio, A., Ojala, A., Arponen, A., 2019. Materialism, awareness of environmental consequences and environmental philanthropic behaviour among potential donors. Environ. Values 28 (6), 741–762.
Maki, A., Carrico, A.R., Raimi, K.T., Truelove, H.B., Araujo, B., Yeung, K.L., 2019. Meta-analysis of pro-environmental behaviour spillover. Nat. Sustain. 2 (4), 307–315.
Merritt, A.C., Effron, D.A., Monin, B., 2010. Moral self-licensing: When being good frees us to be bad. Soc. Pers. Psychol. Compass 4 (5), 344–357.
Mohammed, E.Y., 2012. Contingent valuation responses and hypothetical bias: mitigation effects of certainty question, cheap talk, and pledging. Environ. Econ. 3 (3), 62–71.
Monin, B., Miller, D.T., 2001. Moral credentials and the expression of prejudice. J. Pers. Soc. Psychol. 81 (1), 33.
Mundt, D., Batzke, M.C., Bläsing, T.M., Gomera Deaño, S., Helfers, A., 2024. Effectiveness and context dependency of social norm interventions: five field experiments on nudging pro-environmental and pro-social behavior. Front. Psychol. 15, 1392296.
Nafziger, J., 2020. Spillover effects of nudges. Econom. Lett. 190, 109086.
Nash, N., Whitmarsh, L., Capstick, S., Thøgersen, J., Gouveia, V., de Carvalho Rodrigues Araújo, R., Harder, M.K., Wang, X., Liu, Y., 2019. Reflecting on behavioral spillover in context: how do behavioral motivations and awareness catalyze other environmentally responsible actions in Brazil, China, and Denmark? Front. Psychol. 10, 788.
Nigbur, D., Lyons, E., Uzzell, D., 2010. Attitudes, norms, identity and environmental behaviour: Using an expanded theory of planned behaviour to predict participation in a kerbside recycling programme. Br. J. Soc. Psychol. 49 (2), 259–284.
Nisa, C.F., Bélanger, J.J., Schumpe, B.M., Faller, D.G., 2019. Meta-analysis of randomised controlled trials testing behavioural interventions to promote household action on climate change. Nat. Commun. 10 (1), 1–13.
Nisa, C.F., Sasin, E.M., Faller, D.G., Schumpe, B.M., Belanger, J.J., 2020. Reply to: Alternative meta-analysis of behavioural interventions to promote action on climate change yields different conclusions. Nat. Commun. 11 (1), 1–3.
Nolan, J.M., Schultz, P.W., Cialdini, R.B., Goldstein, N.J., Griskevicius, V., 2008. Normative social influence is underdetected. Pers. Soc. Psychol. Bull. 34 (7), 913–923.
Ortmann, A., Ryvkin, D., Wilkening, T., Zhang, J., 2023. Defaults and cognitive effort. J. Econ. Behav. Organ. 212, 1–19.
Picard, J., 2023. Double-edged nudges? Micro-foundations to behavioural interventions and their spillover effects. Available at SSRN 3784009.
Poore, J., Nemecek, T., 2018. Reducing food's environmental impacts through producers and consumers. Science 360 (6392), 987–992.
Ready, R.C., Champ, P.A., Lawton, J.L., 2010. Using respondent uncertainty to mitigate hypothetical bias in a stated choice experiment. Land Econ. 86 (2), 363–381.
Reese, G., Loew, K., Steffgen, G., 2014. A towel less: Social norms enhance pro-environmental behavior in hotels. J. Soc. Psychol. 154 (2), 97–100.
Riahi, K., Schaeffer, R., Arango, J., Calvin, K., Guivarch, C., Hasegawa, T., Jiang, K., Kriegler, E., Matthews, R., Peters, G., et al., 2022. Mitigation pathways compatible with long-term goals. In: IPCC, 2022: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press Cambridge, UK.
Richter, I., Thøgersen, J., Klöckner, C.A., 2018. A social norms intervention going wrong: Boomerang effects from descriptive norms information. Sustainability 10 (8), 2848.
Ripple, W.J., Wolf, C., Gregg, J.W., Rockström, J., Mann, M.E., Oreskes, N., Lenton, T.M., Rahmstorf, S., Newsome, T.M., Xu, C., et al., 2024. The 2024 state of the climate report: Perilous times on planet earth. BioScience biae087.
Salmivaara, L., Lankoski, L., 2019. Promoting sustainable consumer behaviour through the activation of injunctive social norms: A field experiment in 19 workplace restaurants. Organ. Environ. 1086026619831651.
Schultz, W.P., Khazian, A.M., Zaleski, A.C., 2008. Using normative social influence to promote conservation among hotel guests. Soc. Influ. 3 (1), 4–23.

Sparkman, G., Macdonald, B.N., Caldwell, K.D., Kateman, B., Boese, G.D., 2021. Cut back or give it up? The effectiveness of reduce and eliminate appeals and dynamic norm messaging to curb meat consumption. J. Environ. Psychol. 75, 101592.

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

Sparkman, G., Weitz, E., Robinson, T.N., Malhotra, N., Walton, G.M., 2020. Developing a scalable dynamic norm menu-based intervention to reduce meat consumption. Sustainability 12 (6), 2453.

Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65 (3), 557–586.

Stea, S., Pickering, G.J., 2019. Optimizing messaging to reduce red meat consumption. Environ. Commun. 13 (5), 633–648.

Stern, P.C., 2020. A reexamination on how behavioral interventions can promote household action to limit climate change. Nat. Commun. 11 (1), 1–3.

Stewart, C., Piernas, C., Cook, B., Jebb, S.A., 2021. Trends in UK meat consumption: analysis of data from years 1–11 (2008–09 to 2018–19) of the National Diet and Nutrition Survey rolling programme. Lancet Planet. Heal. 5 (10), e699–e708.

Sunstein, C.R., 2020. Behavioral Science and Public Policy. Cambridge University Press.

Tajfel, H., Turner, J.C., Austin, W.G., Worchel, S., 1979. An integrative theory of intergroup conflict. Organ. Identity: A Read. 56 (65), 9780203505984–16.

Testa, F., Russo, M.V., Cornwell, T.B., McDonald, A., Reich, B., 2018. Social sustainability as buying local: effects of soft policy, meso-level actors, and social influences on purchase intentions. J. Public Policy & Mark. 37 (1), 152–166.

Thaler, R.H., Sunstein, C.R., 2009. Nudge: Improving Decisions About Health, Wealth, and Happiness. Penguin.

Thøgersen, J., Ölander, F., 2003. Spillover of environment-friendly consumer behaviour. J. Environ. Psychol. 23 (3), 225–236.

Truelove, H.B., Carrico, A.R., Weber, E.U., Raimi, K.T., Vandenbergh, M.P., 2014. Positive and negative spillover of pro-environmental behavior: An integrative review and theoretical framework. Glob. Environ. Chang. 29, 127–138.

van der Linden, S., Goldberg, M.H., 2020. Alternative meta-analysis of behavioral interventions to promote action on climate change yields different conclusions. Nat. Commun. 11 (1), 1–2.

Van Gestel, L., Adriaanse, M., De Ridder, D., 2020. Do nudges make use of automatic processing? Unraveling the effects of a default nudge under type 1 and type 2 processing. Compr. Results Soc. Psychol. 1–21.

Vesely, S., Klöckner, C.A., Carrus, G., Tiberio, L., Caffaro, F., Biresselioglu, M.E., Kollmann, A.C., Sinea, A.C., 2022. Norms, prices, and commitment: A comprehensive overview of field experiments in the energy domain and treatment effect moderators. Front. Psychol. 13, 967318.

Wenzig, J., Gruchmann, T., 2018. Consumer preferences for local food: Testing an extended norm taxonomy. Sustainability 10 (5), 1313.

Woltin, K.-A., Sneddon, J., Bardi, A., 2022. Effects of nature values and regulatory fit of message framing on message evaluation and actual pro-environmental donations. Environ. Behav. 54 (3), 597–628.