Environmental Labelling and Consumption Changes: A Food Choice Experiment

Laurent Muller[1] · Anne Lacroix[1] · Bernard Ruffieux[1]

Abstract

This study examines how environment-related food labels affect shopping behaviour. Using an experimental store consisting of nearly 300 available food items, we observe the food purchases of consumers before and after each product is tagged with one of three environmental labels: *single traffic lights*, *multiple traffic lights* and the *kilometric* format. The kilometric format indicates the greenhouse gas emissions for each product by indicating the equivalent number of kilometres driven by an average car. The other two formats are based on the traffic-light rating system used for nutrients by the British *Food Standard Agency*. Multiple traffic lights present three traffic lights simultaneously: one each for greenhouse gas emissions, water eutrophication and air acidity. The single traffic light label displays one unique traffic light for greenhouse gas emissions. All three types of environmental labels lead consumers to purchase more environmentally friendly food baskets (i.e., significant decreases in carbon dioxide, nitrogen and sulfur dioxide emissions). Labelling, however, does not affect the price of selected food baskets (in Euros per 100 g) or their nutritional content. By generating more product substitutions between unlabelled and labelled baskets, multiple traffic lights are more effective in reducing GHG emissions, eutrophication and acidification.

Keywords Labelling · Framed-field experiment · Food · Environment · Greenhouse gas emissions

✉ Laurent Muller
laurent.muller@inra.fr

Anne Lacroix
anne.lacroix@inra.fr

Bernard Ruffieux
bernard.ruffieux@grenoble-inp.fr

[1] Grenoble Applied Economics Laboratory (GAEL), INRA, Univ. Grenoble Alpes, Grenoble INP, CNRS, 38000 Grenoble, France

1 Introduction

Food production and consumption have a significant impact on the environment, with food systems contributing 19–29% of global anthropogenic greenhouse gas (GHG) emissions (Vermeulen et al. 2012). In Europe, food is responsible for 20–30% of the impact of household final consumption (Hertwich et al. 2010), with meat and dairy products together accounting for a significant share of total GHG emissions (Notarnicola et al. 2017). Dietary change is therefore one of the most important ways that sustainable lifestyles can be achieved in developed countries (Carlsson-Kanyama 1998). One way to motivate changes in food consumption is through environmental food labels, although there is much debate over the design and effectiveness of environmental food labels. Most existing labels focus on the carbon footprints, and are developed by private and voluntary initiatives (Ernst and Young 2010). The Grenelle Environmental initiative in France advocates for the simplification and the harmonization of European environmental labelling in order to avoid the proliferation of environmental claims that could undermine the credibility of the process.[1] In Finland, the importance of the environmental dimension in the choice of consumer products is increasing (Rokka and Uusitalo 2008), but there are still significant misunderstandings of some fundamental environmental concepts such as the carbon footprint (Hartikainen et al. 2014). In 2009, a survey showed that 72% of EU citizens believe that a carbon footprint label should be mandatory in the future (Eurobarometer 2009). However, this desire for information does not necessarily imply a willingness to change food consumption habits. In the United Kingdom, while consumer demand for carbon labels is also relatively high, the proliferation and technical content of these labels lead to misinterpretations and misunderstandings that could hinder their impact on purchasing behaviours (Gadema and Oglethorpe 2011). Although UK consumers show very diverse perceptions of environmental labels, they generally tend to be skeptical about unfamiliar labels and general claims (Sirieix et al. 2013). While consumers' desire for simple environmental labelling seems widespread, the question of its format and its real impact on food purchases remains to be answered and requires further research.

To help inform these policy debates, this study assesses the extent to which simple environmental labelling on the front side of packaging can shift food consumption patterns towards greater sustainability and what format is most effective for achieving this shift. We test three labelling formats: the *single traffic lights* format (sTL), the *multiple traffic lights* format (mTL) and the *kilometric* format (KM). KM indicates the carbon dioxide (CO_2) impact for each product by converting that number into the equivalent number of kilometres driven by an average car. The other two formats are based on the traffic-light rating system used for nutrients by the British *Food Standard Agency*. mTL presents three traffic lights simultaneously: one each for GHG emissions, water eutrophication and air acidity. sTL displays a single traffic light for GHG emissions. To assess the relative impact of each format, we have set up an innovative experiment that replicates online shopping platforms. In an experimental store that offers 282 products, we observe the food purchases of 275 consumers before and after one of the three aforementioned environmental labels is affixed to each product. Changes generated by labels on food

---

[1] See the report from the French National Assembly on environmental labelling (http://www.assemblee-nationale.fr/14/rap-info/i1562.asp, 2013).

shopping baskets are therefore examined under strict *ceteris paribus* conditions. In addition to providing high control and avoiding confounding factors, this laboratory experiment enables the truthful elicitation of demand by properly incentivizing the purchase of food items.

Many studies in marketing science have sought to understand consumers' intentions towards eco-labelled products. Such studies have since been widely criticized (Auger et al. 2003; Auger and Devinney 2007) because they are based on unconstrained survey response methods. Respondents are attracted to socially acceptable responses, which lead them to overstate their intentions. In response, many studies have measured consumers' willingness to pay a premium for eco-labelled products. In his review, Gallastegui (2002) notes that these estimates are not reliable because they were mostly derived from consumer statements (such as the contingent valuation method). Such statements are known to be biased (e.g., social desirability bias). Later, studies using the discrete choice method revealed that most consumers prefer eco-labelled products and are willing to pay a premium for them (Janssen and Hamm 2012; Johnston et al. 2001; Van Loo et al. 2011). Although the method limits the risk of overstatements by forcing trade-offs, choices still remain hypothetical and thus do not guarantee that they reflect true attitudes and actual purchases.

Two alternatives to survey methods are the analysis of observational data and the use of experimental methods. When using observational data, one is limited to examining the effects of labels currently used in markets. As a result, existing studies have mainly focused on the widespread organic label (Onyango et al. 2007; Wier and Calverley 2002). Because of the halo effect, consumers may opt for organic food not only for its low environmental impact but also for its healthier reputation. This makes interpretations of environmental attitudes ambiguous. Experimentalists are able to focus exclusively on environmental labelling with little confounding. They build controlled and reproducible conditions that replicate food purchases: participants actually buy real eco-labelled products in the laboratory (Bougherara and Combris 2009; Marette et al. 2012; Moser and Raffaelli 2012; Tagbata and Sirieix 2008). These experiments confirm the existence of price premium on eco-friendly food. However, this premium is always evaluated in the context of purchasing single products. Environmental attitudes are largely ignored in the context of food basket purchases.

To our knowledge, only two experiments involving food baskets have been conducted to date. Vanclay et al. (2011) observed food purchases before and after the posting of environmental labelling for 8 weeks. The study used labels on 37 food products from five product lines (milk, butter spread, canned tomatoes, bottled water and non-perishable pet food) in a grocery store in Australia. The only format tested was a tricolour logo (green, yellow and black), in the same vein as our single traffic light logo. Consumers shifted from "black" eco-unfriendly products (−6%) to "green" eco-friendly products (+4%). In their between-subject framed-field experiment, Vlaeminck et al. (2014) tested two labelling formats: a multi-coloured labelling format similar to that of our multiple traffic lights and another format with several raw pieces of environmental information. Labels were affixed to 9 products (3 fruits, 3 vegetables and 3 protein-based products) in a Belgian supermarket. Participants were required to buy at least one of these nine products. The study stresses the importance of labelling formats: while simple logos that signal overall environmental quality in colour have positive effects on food purchases, complex logos that provide multiple pieces of raw information on resources used (water, soil, pesticides, etc.) have no effect at all.

Our experiment differs in several ways. First, with 282 products divided into 36 product categories, our laboratory simulates a small grocery store, which makes the task more

realistic and allows an assessment of the impact of labelling on a complete food basket rather than on a limited set of products. Second, because labels are applied to all available products, we simulate a comprehensive labelling policy. Third, our intra-subject structure controls individual differences: the same consumers are observed with and without labelling, all other things being equal. Fourth, we compare three formats: two color-coded, one with a single piece of information and one with several pieces of information, and one with a single piece of raw information. The experimental design allows a fair comparison of the formats. Finally, the purchasing task is not subject to any constraints (on quantities or budgets). Participants are free to buy what they want, just as they would be in a real grocery store.

We find that environmental labelling improves the environmental impact of food baskets without affecting the price per 100 g or the nutritional quality of the food purchased. For instance, GHG emissions decrease between 8% and 12%. Multiple traffic lights are relatively more effective at improving the environmental impact of food baskets. These results converge with the study of Gadema and Oglethorpe (2011), which notes the difficulty consumers experience when processing environmental information. To improve the effectiveness of the labels in the choice process, Gadema and Oglethorpe recommend simplicity and clarity. This recommendation echoes findings on nutritional labels. According to marketing and psychology studies, consumers' attention is best captured by colours and symbols (Muller and Prevost 2016; Argo and Main 2004; Cox III et al. 1997). For example, consumers in supermarkets refer less to nutritional tables than to health logos (van Herpen and van Trijp 2011). Once noticed by consumers, graphic logos such as *traffic lights* or *3-tiered star* icons can even lead to an improvement in purchasing behaviours (Balcombe et al. 2010; Sutherland et al. 2010). However, a critical review of the literature on nutritional labelling (Hieke and Taylor 2012) reaches a mixed conclusion: labels are sometimes useful for some people and only under some circumstances.

2 The Experiment

2.1 Label Formats

We use three indicators to assess the environmental impact of food. One assesses GHG emissions and is expressed in carbon dioxide (CO_2) equivalents (in grams), another assesses marine eutrophication in nitrogen (N) equivalents (in grams), and the last one assesses air acidification in sulfur dioxide (SO_2) equivalents (in grams). These indicators pertain to, respectively, global warming, water pollution and air pollution. The three environmental indicators follow the *lifecycle analysis* approach. This approach has gradually come to be favoured in France[2] and in Europe (Manfredi et al. 2012) because it considers each stage of the product's life cycle (including its packaging) from the manufacturing process to consumption. GHG emissions, eutrophication and acidification are calculated per 100 g of food. In this study, these calculations were carried out by «BIO by Deloitte», specialists in lifecycle analysis.

---

[2] See the report from the French Environment & Energy Management Agency on the "Life cycle assessments of agricultural products" (2008).

We test three labelling formats (sTL, mTL and KM) that differ in their complexity and the degree to which the label provides the consumer with conclusions on the environmental quality of a product (normative information), as opposed to the degree to which the label leaves it to the consumer to draw such conclusions (positive information).

KM reports the GHG emissions expressed as kilometres travelled by a common motor vehicle. This number indicates the distance travelled by a typical car when it emits the CO_2 equivalent needed to produce, package and transport the food product. For example, if a litre of whole milk is reported at 11.0 km, it means that the product's GHG emissions are equivalent to those of a car travelling 11 km. KM is positive and based on a single criterion. sTL displays a single coloured dot. The dot is green when the food belongs to the best one-third of products in terms of GHG emissions compared to products of the same food category. The dot is orange when the food belongs to the middle third and red when it belongs to the bottom third. In other words, because each page of the food catalogue represents one food category, there will be 3 green foods, 3 orange foods and 3 red foods if the page comprises 9 products. sTL is normative and based on single criterion. Finally, mTL displays 3 coloured dots. One dot signals the GHG emissions, another the marine eutrophication and the last one air acidification. The color-coding follows the same rule as for sTL. mTL is normative and based on multiple criteria. The three labelling formats are displayed in Table 1.

2.2 Food Supply

Purchasing behaviours and, therefore, changes in purchasing behaviours induced by labels are inevitably contingent on food supply. Accordingly, choosing foods that will make up the experimental store is a crucial task. First, the food set should be large and varied enough to allow for substitutions between and within food groups. Secondly, the environmental impacts of foods must have sufficient contrast to allow for improvements. Finally, the food set must be realistic and representative of French food consumption. Based on these specifications and with the help of nutritionists and life cycle analysts, we have selected 282 products grouped into 37 categories and 7 meta-categories (see Table 2). Categories refer to self-service store shelves. Within each category, 6 or 9 substitutable products differ in their environmental impacts due to their composition, their preservation (fresh, frozen, and canned), packaging (glass, metal, cardboard) and size (family pack, individual pack).

We use two indicators to assess the nutritional quality of foods: the aggregated LIM index (Darmon et al. 2009; Maillot et al. 2011) and the energy content. The LIM index is a validated nutrient profiling system that calculates the mean percentage for 3 nutrients whose intake should be limited (saturated fatty acids, free sugar and sodium) based on the recommended maximum values. It is also calculated per 100 g. In addition, we also gathered information on weight (in g) and energy content (in kcal). In the absence of information on packaging, this information was collected from the food composition database originally developed for the Su-Vi-Max project.[3] Finally, each product is priced according to its actual retail prices. Prices were taken from the largest French food supermarket (Carrefour) at the time of the experiment. All these indicators are then applied to the baskets by summing the food values according to the total weight of the basket: each indicator is expressed per 100 g.

---

[3] The Su-Vi-Max project was led by the French National Institute of Health and Medical Research in 2006.

Table 1 Label formats

Kilometric format (KM)		GHG (CO_2 equivalent) expressed in kilometres travelled
Single traffic lights (sTL)		GHG (CO_2 equivalent) Colour (green, orange, red) assigned according the item's rank in its food category
Multiple traffic lights (mTL)		GHG (CO_2 equivalent), marine eutrophication (nitrogen equivalent), air acidification (sulfur dioxide equivalent) Colour assigned according to the item's rank in its food category for each of the 3 criteria
Control	(Blank)	–

Table 2 Food categories

		Number of items in the catalogue	Number of items available for purchase
All categories		282	96
Fruits	Fresh fruit (9); fruit desserts (9); citrus juice and nectars (9); fruity beverages (9)	36	10
Vegetables	Fresh vegetables (9); canned vegetables (9); salads (6)	24	6
Meat, fish and eggs	Eggs (3); fish and seafood (9); meat (9); pies and sausages (6); ham and delicatessen (9)	36	12
Mixed dishes	Prepared vegetables (6); soups (6); cheese ready meal (6); meat ready meal (9); fish ready meal (6); pizzas (6); pies and quiches (6); sandwiches (6)	51	6
Cereals, potatoes and legumes	Processed potatoes (9); breads and buns (9); pastry and cakes (6); breakfast cereals (9); pasta and rice (6)	48	25
Dairy products	Milks (9); creams (6); dry cheeses (9); plain dairy (9); dessert creams (6); flavoured dairy desserts (6)	45	17
Snacks and sweets	Sodas (9); cookies (9); spreads (9); salted snacks (9); chocolate and cereal bars (6); chocolate (9)	42	20

Number of products available in the food catalogue and number of products actually available in the experimental store that consumers have to buy if present in their binding basket

Table 3 Environmental, economic and nutritional characteristics of foods in the experimental store

	Number of items	GHG (g per 100 g)			Price (€ per 100 g)			LIM		
		Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
Cereals, potatoes, legumes	39	237.31	140.18	206	0.54	0.35	0.43	15.61	13.54	10.28
Dairy products	45	279.98	235.81	193	0.38	0.26	0.30	14.12	12.19	12.42
Fruits	36	100.67	39.39	94.5	0.27	0.16	0.25	9.48	7.83	10.08
Meat, fish and eggs	36	640.28	347.55	521	1.20	0.63	1.03	14.71	13.99	9.93
Mixed dishes	51	328.18	118.32	340	0.73	0.38	0.63	11.81	6.58	9.90
Snacks and sweets	51	225.86	113.80	225	0.88	0.60	0.80	31.29	21.36	30.17
Vegetables	24	120.92	54.22	117	0.32	0.19	0.32	3.50	2.79	3.84
All	282	282.57	236.11	219	0.64	0.51	0.48	15.59	15.32	10.47

When we look at the 282 foods in the experimental store, we find that there are strong correlations between the three environmental indicators: Spearman's correlation coefficients are 0.91 between eutrophication and GHG, and 0.71 between acidification and GHG. GHG varies widely from a factor of 1–46. GHG heterogeneity is greater for animal foods than for plant products. Meat, fish and eggs are the least eco-friendly foods, followed by ready-made dishes and dairy products. On the other hand, fruits and vegetables are the eco-friendliest. The same observations can be made for eutrophication and acidification. Meat, fish and eggs are also the most expensive, while fruits and vegetables are the least expensive. As for nutrition, snacks and sweets have the worst LIM index, whereas fruits and vegetables are again at the top of the list (see details in Table 3). GHG is positively correlated with price and LIM index: Spearman correlation coefficients are, respectively, 0.68 and 0.42. In other words, foods with a higher environmental impact are more expensive and have lower nutritional value.

Although these claims are only verified for the 282 foods in the experimental store, we trust they can be generalized to larger sets of foods. However, they are highly dependent on the functional unit. Those correlations are based on 100 g of product. By considering the energy intake or servings as denominators, correlations, for instance, between GHG and price and between GHG and LIM become negative (Spearman, -0.10 and -0.44). This makes sense: meat emits a lot of CO_2 and is expensive, but it also provides many calories. This finding suggests that consumers will find it difficult to improve their environmental impact while sustaining their energy consumption. This result converges with Drewnowski et al. (2015) and Masset et al. (2014): «the sustainability dimensions seemed to be compatible when considering price per kilogram of food. However, this conclusion is too simplistic when considering price per kilocalorie».

2.3 Experimental Design

Participants' task is to do their grocery shopping in our experimental store (see instructions in Appendix 1). The general instruction is to buy food for the household for 2 days. However, as in a real store, the participants are reminded that they are free to buy what and how much they want. If they are not satisfied with the products available in the experimental store or if they are sufficiently supplied at home, they are free to leave the experimental store without purchasing anything. Also, no budgetary constraints are set. To do their shopping, each participant is provided with a food catalogue, a barcode reader and a computer screen. The catalogue includes all 282 food products for sale during the experiment. Each page contains 6 or 9 items from the same category. Each item is displayed with name, brand, front image, price and price per kg or litre, weight or volume and a unique barcode (see Fig. 1). When scanning the barcode, the corresponding item appears on the screen. The on-screen environment is similar to e-shopping sites. Nutritional information (calories, salt, free sugars, saturated fatty acids) and environmental information (GHG emission, marine eutrophication, air acidification) are available on request (by clicking the "see detail" button). If a participant wishes to buy an item, she scans it in the catalogue to make it appear on the screen. She then clicks on the "Add to cart" button. A virtual shopping basket on the left side of the screen reminds users of the items selected (see Fig. 2).

There are several steps in the experiment's procedures, as shown in Table 4. Instructions are displayed both on a large screen and on each participant's computer screen and are read aloud by the facilitator. Participants are told they will have to fill two baskets of food successively, but nothing is said about environmental labelling until part 2 begins. In part 1,
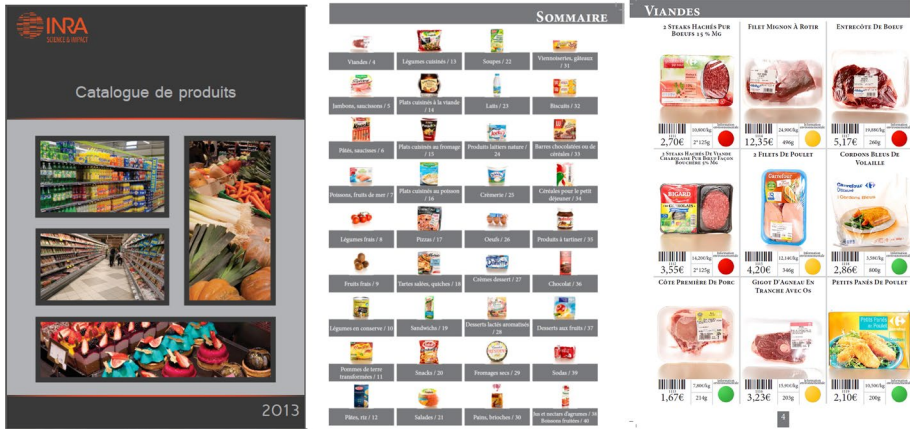
Fig. 1 Catalogue. Left: front page; middle: table of content; right: meat category with the single Traffic Light labelling
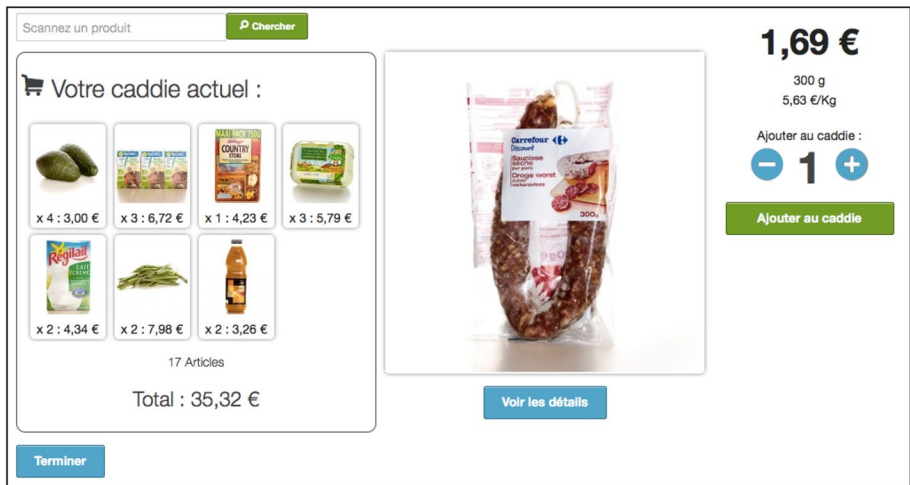

Fig. 2 Computer screen. Example when salami is scanned from the catalogue

the first basket, the *reference basket*, is elaborated without labels. Once all participants have validated their reference basket, a new catalogue is distributed. This catalogue is strictly identical to the previous one, except that all products are now labelled with one label format (sTL, mTL or KM). The corresponding format is briefly described to the participants before each participant is then asked to fill a new shopping basket, the *label basket*. This experimental procedure allows us to observe the impact of environmental labelling on each individual in perfectly identical situations. While this allows a great deal of control over the explanatory variables and thus ensures the nature of the causal relationships, repetition of the task generates a great deal of salience towards labelling and may therefore produce an overestimation of its absolute impact. After the two parts of the shopping experience,

Table 4 Session overview

Step 1	**Welcome speech**—Facilitators give general instructions regarding the upcoming session. Participants receive 35€ to compensate for their attendance
Step 2	**Task instructions**—Shopping tasks and the incentive mechanism are read aloud and projected both on a large screen and on each personal computer screen. Participants receive a food catalogue containing 282 food products with no environmental labels attached
Step 3	**Experiment, part 1**—Participants compose their reference basket
Step 4	**Label presentation**—Participants in the treatment groups receive another food catalogue that is strictly identical to the previous one, except that an environmental label is now applied to each food. The facilitator presents the corresponding label. Participants in the control groups keep their part 1 catalogue
Step 5	**Experiment, part 2**—Participants compose their label basket
Step 6	**Survey**—Participants fill out a survey on socio-demographic characteristics
Step 7	**Draw**—One of the two baskets is randomly drawn for actual purchases
Step 8	**Purchase**—Participants purchase all products from their selected basket that match the products available in the laboratory

participants are invited to complete a socio-demographic questionnaire that provides information on occupation, income, and household size.

The experiment is made incentive-compatible. The instructions emphasize that participants will actually purchase approximately one quarter of the content of one of their two baskets. A quarter of the food is actually available and hidden from participants in an adjacent room. At the end of the experiment, a shopping basket (either the reference basket or the label basket) is randomly drawn. Any drawn basket product that is available in the adjacent room is actually purchased at the end of the experiment. In total, 96 food items can possibly be sold at the end of the experiment. We have taken great care to distribute the available products in all product categories, including fresh and perishable products (see Table 2). Before each session, a new set of 71 products (i.e., one quarter of 282) was chosen among these 96 items. Participants were also compensated 35€ before the start of the experiment. It was explained to them that this money is only used to compensate them for their attendance and has nothing to do with the rest of the experiment. The amount they will spend in the stores is at their own discretion. In fact, after the products actually available were revealed at the end of the experiment, participants never spent more than 10€ on real purchases.

We run three label treatments, one for each labelling format (sTL, mTL and KM) and one control treatment. In the control treatment, participants keep their part 1 catalogue without environmental labelling. The objective is to control for the impact of repetition on choices and thus to adjust any possible artefactual effect. Following the random control trial standard, participants were randomly assigned to one of the four groups. In total, 19 sessions (5 per label treatment and 4 for the control treatment) of approximately 1 h 30 were conducted.

2.4 Participants

A total of 275 participants were recruited through newspapers and flyers in the Grenoble urban area. To be eligible, individuals had to be between the ages of 20 and 65 and responsible for groceries for their household. Following the standards of randomly controlled

trials, participants were allocated to one treatment (sTL, mTL, KM or Control). Each participant received €35 to compensate her participation.

Table 5 presents the socio-economic and demographic characteristics of participants. Most of the participants were women (nearly 2/3) and were educated (almost 2/3 with higher education), and fewer than half had children (under 14 years of age). Their income level (more precisely the disposable income per consumption unit) is close to the French median income (1610 euros per month in 2010). The results from the Wilcoxon rank sum test or the Pearson χ^2 test show no significant differences (α level of 0.05) between treatment groups and the control group except for education in the sTL treatment. Compared to the French averages, the size of the households in our sample is slightly larger and the income slightly lower (see sociodemographic table for France and Grenoble in Appendix 2).

3 Results

3.1 Data Analysis

Due to the experimental design, we are able to measure labelling impact within subjects and to compare formats between subjects. In strictly identical conditions, participants shop for food before and after one labelling format is implemented on each item. Accordingly, individual differences between the reference basket and the label basket produce the changes in food choices that are induced by the introduction of the label and by repetition of the task. The effect of repetition is controlled by accounting for changes from the control group (who repeats the shopping task twice without the label implementation). The effects are then compared across groups to assess the relative performance between labelling formats. Data analysis is organized in three parts: descriptive analysis, econometric analysis and behavioural analysis.

For each individual, we measure the quantity of the CO_2 equivalent, sulfur dioxide equivalent and nitrogen equivalent (all in grams per 100 g of products), the cost (in Euros per 100 g of products), the energy intake (in kCal per 100 g of products) and the LIM index in both the reference basket and the label basket. We then use the differences between baskets to measure the impact of the labels within subjects. The differentiated impacts of KM, sTL and mTL are then assessed by comparing the individual differences between subjects. We use as statistical tests the Wilcoxon signed-rank test when testing label impact and the Wilcoxon-Mann–Whitney test when comparing treatments.

Second, in order to control for the confounding effects of price, energy and LIM, we run an econometric analysis to examine the impacts of the treatments on environmental footprint. As we have data on the same individuals in both pre- and post-periods, we apply the following model:

$$\Delta y_i = \alpha + \beta D_i + \theta C_i + \varepsilon_i$$

where Δy_i is the difference in the outcome of interest (GHG per 100 g basket, for instance) from the first basket to the second basket for individual i. The term D_i is a treatment dummy variable that takes the value of one if the individual is in the treatment (KM, sTL or mTL) and zero if they are in the control. C_i is a vector of control variables for each individual i.

Table 5 Socio-demographic characteristics of participants (mean ± standard deviation or %)

	ALL	CONTROL	KM	sTL	mTL
Number of subjects	275	59	67	66	83
Female (%)	64.4	64.4	67.2	62.1	63.9
Male (%)	35.6	35.6	32.8	37.9	36.1
Age (years)	40.7±11.8	41.8±12.3	41.4±11.2	41.7±11.5	38.5±12.2
Under 31 years old (%)	28	28.8	23.9	21.2	36.1
Between 31 and 50 years old (%)	47.3	45.8	50.7	54.6	39.8
Over 50 years old (%)	24.7	25.4	25.4	24.2	24.1
Household size	2.7±1.2	2.6±1.2	2.9±1.1	2.8±1.3	2.6±1.2
Household with children (%)	56.3	57.6	47.8	51.5	61.4
Household with no children (%)	43.7	42.4	52.2	48.5	38.6
Income, € per month and per consumption unit	1495.8±640.0	1413.5±692.1	1490.7±565.3	1523.5±617.4	1533.1±680.9
Educational attainment					
No high school diploma (%)	16	20.3	19.4	4.6	19.3
High school diploma (%)	20.7	20.3	22.4	24.2	16.9
Bachelor's degree (%)	36.4	37.3	29.8	42.4	36.1
Graduate degree (%)	27	22.0	28.4	28.8	27.7
Obese or overweight (%)	32.7	30.5	31.3	27.3	39.8

	Reference basket	Label basket	*p* value
Table 6 Means (standard deviations) for environmental, economic and nutritional characteristics (indicators per 100 g) for reference and label baskets; *p* value for the difference between the reference and the label basket according to the Wilcoxon signed rank test			
GHG (g per 100 g)			
Control (*n*=59)	183.61 (45.82)	181.35 (42.50)	0.325
sTL (*n*=66)	178.64 (39.03)	163.21 (35.81)	<.001
mTL (*n*=83)	182.88 (45.61)	160.44 (38.04)	<.001
KM (*n*=67)	185.56 (39.34)	169.18 (39.69)	<.001
Eutrophication (g per 100 g)			
Control (*n*=59)	0.81 (0.28)	0.81 (0.29)	0.949
sTL (*n*=66)	0.80 (0.24)	0.76 (0.24)	0.008
mTL (*n*=83)	0.83 (0.28)	0.74 (0.26)	<.001
KM (*n*=67)	0.81 (0.24)	0.77 (0.19)	0.071
Acidification (g per 100 g)			
Control (*n*=59)	2.01 (0.82)	1.96 (0.74)	0.310
sTL (*n*=66)	2.06 (0.69)	1.81 (0.63)	<.001
mTL (*n*=83)	2.11 (0.86)	1.74 (0.64)	<.001
KM (*n*=67)	2.09 (0.74)	1.89 (0.72)	<.001
Price (€ per 100 g)			
Control (*n*=59)	0.31 (0.07)	0.31 (0.07)	0.298
sTL (*n*=66)	0.31 (0.06)	0.31 (0.05)	0.317
mTL (*n*=83)	0.31 (0.07)	0.30 (0.06)	0.068
KM (*n*=67)	0.33 (0.07)	0.34 (0.07)	0.886
Energy (kCal per 100 g)			
Control (*n*=59)	120.09 (32.23)	117.81 (32.33)	0.237
sTL (*n*=66)	117.14 (25.31)	116.82 (27.67)	0.426
mTL (*n*=83)	120.35 (33.37)	118.84 (28.03)	0.619
KM (*n*=67)	117.66 (27.55)	114.80 (23.60)	0.308
LIM			
Control (*n*=59)	8.07 (3.17)	7.94 (3.03)	0.737
sTL (*n*=66)	8.04 (3.19)	7.73 (2.84)	0.251
mTL (*n*=83)	8.01 (3.20)	7.84 (2.93)	0.332
KM (*n*=67)	8.07 (3.11)	7.68 (3.06)	0.125

Finally, we examine the number and nature of substitutions made by participants between the two baskets. In doing so, we can determine which label triggers the most changes and calculate the environmental impact per change for each label.

3.2 Descriptive Analysis

We first compare the content of the reference baskets to the average content of the experimental store (see Tables 3, 6). Participants choose foods that are more environmentally friendly than the average foods available in the experimental store. While the average GHG emission per 100 g of the 282 products is 283 g, the average GHG is significantly lower in the reference baskets (183 g). Participants also choose cheaper and healthier foods. The average price and LIM index per 100 g are, respectively, 0.31€ and 8.04 in the reference baskets compared to 0.64€ and 15.59 in the experimental store.

Table 7 Individual means of the differences between the label basket and reference basket of the environmental, economic and nutritional indicators (per 100 g) per treatment and per product family

(Label basket–reference basket)	GHG (g per 100 g)	Eutrophication (g per 100 g)	Acidification (g per 100 g)	Price (€ per 100 g)	Energy (kCal per 100 g)	LIM
All products						
282 items						
Control	− 2.25[b]	+ 0.00[b]	− 0.06[c]	− 0.00[a]	− 2.28[a]	− 0.13[a]
mTL	**− 22.44[a]**	**− 0.09[a]**	**− 0.36[a]**	− 0.01[a]	− 1.52[a]	− 0.16[a]
sTL	**− 15.43[a]**	**− 0.04[a,b]**	**− 0.25[a,b]**	− 0.00[a]	− 0.32[a]	− 0.31[a]
KM	**− 16.37[a]**	− 0.04[b]	**− 0.21[b,c]**	+ 0.00[a]	− 2.86[a]	− 0.39[a]
Meat, fish and egg						
36 items						
Control	− 10.19[c]	+ 0.07[b]	− 0.49[b]	+ 0.01[a]	− 0.75[a]	+ 0.62[a]
mTL	**− 119.9[a]**	**− 0.32[a]**	**− 2.62[a]**	− 0.05[a]	**− 7.86[a]**	**− 0.96[a]**
sTL	**− 41.09[a,b]**	**− 0.40[a]**	**− 0.90[a,b]**	− 0.02[a]	+ 5.23[a]	+ 0.58[a]
KM	− 47.91[b, c]	+ 0.08[b]	− 0.85[b]	+ 0.01[a]	− 3.09[a]	+ 0.37[a]
Dairy						
45 items						
Control	− 13.33[b]	− 0.02[a]	− 0.20[b]	− 0.02[a]	**− 14.51[a]**	− 1.00[a]
mTL	**− 21.70[a]**	**− 0.04[a]**	**− 0.39[a]**	− 0.00[b, c]	+ 1.42[a]	− 1.04[a]
sTL	**− 23.00[a]**	− 0.03[a]	**− 0.44[a]**	− 0.01[a,b]	**− 11.78[a]**	− 1.34[a]
KM	**− 7.87[b]**	− 0.01[a]	− 0.15[b]	**+ 0.01[c]**	− 5.16[a]	+ 0.04[a]
Fruits and vegetables (fresh and processed)						
60 items						
Control	+ 0.10[c]	+ 0.02[b]	− 0.00[c]	− 0.01[b]	+ 1.77[b]	+ 0.02[a,b]
mTL	**− 12.04[a]**	**− 0.10[a]**	**− 0.07[a]**	− 0.01[a]	− 3.36[a,b]	+ 0.07[a,b]
sTL	**− 5.40[b]**	− 0.02[b]	− 0.03[b]	− 0.00[a,b]	**− 4.05[a]**	+ 0.15[a]
KM	**− 6.48[b]**	− 0.05[b]	**− 0.03[b,c]**	− 0.01[a,b]	− 1.78[a,b]	− 0.03[b]
All others (snacks and sweets; mixed dishes; cereals, potatoes, legumes)						
141 items						
Control	+ 12.82[c]	+ 0.07[b]	+ 0.24[b]	+ 0.01[b]	− 3.13[a,b]	+ 0.15[a,b]
mTL	**− 31.56[a]**	**− 0.15[a]**	**− 0.30[a]**	− 0.03[a]	**− 10.74[a]**	− 0.22[b]
sTL	**− 34.36[a]**	**− 0.09[a]**	**− 0.39[a]**	**− 0.01[a]**	**− 7.75[a]**	**+ 1.20[b]**
KM	**− 20.28[b]**	+ 0.01[b]	− 0.15[b]	+ 0.03[b]	+ 3.28[b]	− 1.08[a]

Differences are statistically significant at the 5% level when figures are bold (Wilcoxon signed-rank test)

[a,b,c]Means per column with the same letter are not significantly different at the 5% level (Wilcoxon Mann–Whitney test)

Environmental labelling, whatever its format, has a positive effect on the environmental quality of baskets (see Table 6). GHG decreases from 179 g in the reference basket to 163 g per 100 g of products in the label basket with the sTL format, from 183 to 160 g with mTL and from 186 to 169 g with KM. All differences are significant at the 1% level according to Wilcoxon signed rank tests. The environmental improvement does not affect cost and nutritional content. Price, energy and LIM index do not significantly vary between the two baskets below the 5% level. While the impact seems bigger with mTL (− 22 g vs. − 15 g and − 16 g for sTL and KM for each individual, on average), relative differences are not statistically significant according to Wilcoxon Mann–Whitney tests (Table 7). Eutrophication and

acidification also significantly decrease with labels, except for KM on eutrophication. Relative decreases are significantly higher with mTL than with KM.

The product family grouping meat, fish and egg categories is clearly the one that allows the biggest environmental improvement (Table 7). For that set of products, mTL generates a larger GHG decrease than KM does. KM does not even perform statistically better than the control.[4] The superiority of mTL over KM can be generalized for all product families: mTL is always among the two labels that generate the largest decrease in GHG emissions, eutrophication and acidification, and KM is among the two smallest decreases (Table 7). sTL is intermediate. Eutrophication and acidifications are always improved with mTL and sTL and never with KM except for fruits and vegetables. As for economic and nutritional impact, there is no explicit trend.

3.3 Econometric Analysis

In order to control for price, energy, LIM and various sociodemographic characteristics, we run a regression on the differences in, respectively, GHG, eutrophication and acidification between the reference basket and the label basket (results of the estimation are shown in Table 8). Labels, whatever the format, induce a GHG decrease: between 14 and 19 g per 100 g. mTL decreases GHG to a larger extent than do sTL and KM. mTL also provides statistically significant reduction of eutrophication and acidification. This finding is not surprising because only mTL informs on eutrophication and air acidification. Although sTL does not inform on acidification, the improvement is nonetheless significant due to a strong correlation between acidification and GHG (0.91 with both Pearson and Spearman coefficients).

The consumption of more expensive and higher calorie products significantly increases GHG emissions, eutrophication and acidification. Sociodemographic moderators do not affect the environmental impacts of baskets. Variables related to income, weight, sex, age and education are not statistically significant, with one exception: baskets of participants with a master's degree appear to generate less eutrophication but more acidification.

3.4 Behavioural Analysis

The rate of changes between baskets varies according to the treatment (see Table 9). The participants in the control treatment nevertheless make many changes: They replace almost 7 products and remove 0.5 items from their original total of 18 (60%). No label generates more changes. On average, participants keep 63.7%, 67.2% and 69.2% of their reference basket items in their baskets labelled with mTL, sTL and KM, respectively. However, participants in the control groups do not change for the better: each change increases GHG emissions by 84 g on average.

With labels, each change considerably reduces GHG emissions: -323 g with mTL, -290 g with sTL and -463 g with KM. While previous analyses show that KM is the label with the least impact, here we show that each item replacement reduces GHG emissions significantly more than when mTL and sTL are used. However, at the same time,

---

[4] Although sTL presents a slighlty lower mean than does KM (-41 g vs. -48 g), sTL is significantly different from the control group. This counter-intuitive result is explained by a much larger spread in KM groups: The standard deviations of individual meat GHG differences between the reference and labelling baskets are 133 and 180 for sTL and mTL respectively.

Table 8 Regression estimation of the model $\Delta y_i = \alpha + \beta D_i + \theta C_i + \varepsilon_i$

	Δy_i		
	Greenhouse gas[1] CO_2/100 g	Eutrophication[1] N/100 g	Acidification[1] SO_2/100 g
α (intercept)	−2.231 (5.823)	0.046 (0.043)	−0.107 (0.122)
β			
KM[2]	−14.525*** (4.408)	−0.047 (0.033)	−0.141 (0.092)
sTL[2]	−14.747*** (4.483)	−0.039 (0.033)	−0.219** (0.094)
mTL[2]	−19.251*** (4.198)	−0.094*** (0.031)	−0.290*** (0.088)
θ			
Price[1]	170.262*** (35.429)	0.881*** (0.263)	1.245* (0.742)
Energy[1]	0.299*** (0.101)	0.002*** (0.001)	0.006*** (0.002)
LIM[1]	1.404 (1.048)	0.003 (0.008)	0.049** (0.022)
Income below the 2nd decile[2]	−5.548 (3.969)	0.012 (0.029)	−0.072 (0.083)
Income above the median[2]	−6.246 (4.037)	−0.009 (0.030)	−0.082 (0.084)
Obese or overweight[2]	1.508 (3.396)	0.024 (0.025)	−0.025 (0.071)
Male[2]	3.370 (3.258)	−0.037 (0.024)	0.092 (0.068)
Under 31 years of age[2]	−4.745 (3.905)	−0.005 (0.029)	−0.099 (0.082)
Over 50 years of age[2]	1.168 (3.911)	−0.040 (0.029)	0.006 (0.082)
High school diploma[2]	3.090 (5.029)	0.012 (0.037)	0.081 (0.105)
Bachelor's degree[2]	6.985 (4.538)	0.002 (0.034)	0.144 (0.095)
Graduate degree[2]	6.993 (4.268)	−0.057* (0.032)	0.179** (0.089)
Household with children[2]	0.447 (3.414)	−0.008 (0.025)	0.042 (0.071)
	$n = 275$	$n = 275$	$n = 275$
	$R^2 = 0.2901$	$R^2 = 0.2240$	$R^2 = 0.1926$

Values in brackets are standard errors. *$p < 0.10$; **$p < 0.05$; ***$p < 0.01$

[1]Values expressed as differences between the reference basket and the label basket

[2]Dummy variable

participants in the KM treatment replace significantly fewer items (4.8 vs. 5.7 in mTL and 5.2 in sTL) and are the only ones who have more items in the label basket (+0.15 vs. −0.61 in mTL and −0.12 in sTL). Finally, only 60% of the items replaced in KM are within the same food category, compared to 70% and 72% in mTL and sTL. These observations can more or less be generalized to all food meta-categories.

Apart from the control treatment, items in the reference basket are replaced on average by less expensive ones. Each substitution made with KM saves significantly more money (−0.33€ vs. −0.08€ and −0.09€ with mTL and sTL).

Both detailed nutritional and environmental information was available to participants by clicking a button on their computer screen. Participants did not use the buttons. They checked the information, on average, 1.6 and 1.4 times for nutritional and environmental information, respectively, on their reference baskets and 0.4 and 0.3 times, respectively, for their labelled baskets. Less than 10% of the participants checked information at least 5 times. With so few information-viewers, we are unable to statistically test possible behavioural differences between viewers and non-viewers and between treatments.

Table 9 Description of the changes between the reference and the label basket and the impact per substitution on GHG and cost —individual means

	Number of items…					Impact per substitution	
	…in the reference basket (1)	…remaining in the label basket (2)	…substituted (3)	…substituted within food category (4)	…added (+) or suppressed (−) (5)	GHG (g) (6)	Price (€) (7)
All products 282 items							
Control	18.05[a]	10.10[a]	6.97[b]	3.93[a,b]	−0.54[a]	+84.4[c]	+0.09[b]
mTL	18.08[a]	11.52[a,b]	5.66[b]	3.95[b]	−0.61[a]	−322.7[b]	−0.08[b]
sTL	18.51[a]	12.44[b]	5.21[a,b]	3.76[b]	−0.12[a,b]	−289.8[b]	−0.09[b]
KM	17.48[a]	12.10[b]	4.81[a]	2.93[a]	+0.15[b]	−463.4[a]	−0.33[a]
Meat, fish and egg 36 items							
Control	3.22[a]	1.66[a]	1.14[b]	0.78[a,b]	−0.16[a]	+151.4[b]	+0.26[b]
mTL	3.04[a]	1.87[a,b]	0.92[b]	0.76[b]	−0.10[a]	−719.4[a]	−0.19[a,b]
sTL	3.26[a]	2.13[a,b]	0.94[a,b]	0.71[a,b]	−0.03[a]	−365.4[a,b]	−0.53[a]
KM	3.15[a]	2.20[b]	0.69[a]	0.57[a]	+0.02[a]	−840.9[a]	−0.70[a]
Dairy 45 items							
Control	3.25[a]	1.84[a]	0.95[a,b]	0.60[a]	−0.25[a]	+40.4[c]	−0.12[a]
mTL	3.57[a]	2.15[a,b]	1.19[b]	0.93[b]	−0.11[a]	−186.8[b]	+0.15[b]
sTL	3.60[a]	2.2[a,b]	1.14[b]	0.92[b]	−0.11[a]	−311.6[a,b]	−0.12[a]
KM	3.46[a]	2.41[b]	0.72[a]	0.62[a]	−0.13[a]	−433.2[a]	−0.13[a,b]
Fruits and vegetables (fresh and processed) 60 items							
Control	6.15[a]	3.49[a]	2.11[b]	1.54[b]	−0.08[a,b]	−30.7[c]	−0.02[a]
mTL	5.89[a]	3.95[a,b]	1.39[a,b]	1.07[a,b]	−0.41[b]	−305.5[a]	−0.24[a]
sTL	6.50[a]	4.70[b]	1.20[a]	1.03[b]	−0.36[a,b]	−203.2[b]	−0.10[a]
KM	5.66[a]	3.93[a,b]	1.25[b]	0.94[a]	−0.03[a]	−154.3[a,b,c]	−0.08[a]

Table 9 (continued)

	Number of items…					Impact per substitution	
	…in the reference basket (1)	… remaining in the label basket (2)	… substituted (3)	… substituted within food category (4)	… added (+) or suppressed (−) (5)	GHG (g) (6)	Price (€) (7)
All others (snacks and sweet; mixed dishes; cereals, potatoes, legumes) 141 items							
Control	5.59[a]	3.20[a]	1.81[a]	1.05[a,b]	− 0.07[a]	+ 122.8[c]	− 0.02[b]
mTL	5.67[a]	3.60[a]	1.76[a]	1.22[b]	+ 0.00[b]	− 244.9[b]	+ 0.01[b]
sTL	5.49[a]	3.63[a]	1.61[a]	1.17[b]	+ 0.38[a,b]	− 453.0[a,b]	− 0.07[a,b]
KM	5.40[a]	3.70[a]	1.51[a]	0.84[a]	+ 0.31[b]	− 439.8[a]	− 0.30[a]

[a,b,c]Means (per column and per food category) with the same letter are not significantly different at the 5% level (Wilcoxon Mann–Whitney test)

Table 10 Description of the changes between the reference and the label basket and the impact per substitution on GHG for the three labelling treatments—individual means

	Number of items…				Impact per substitution GHG (g)
	…in the reference basket	… remaining in the label basket	… substituted	… added (+) or suppressed (−)	
Fresh products[a] (n=101)	10.51	7.12	2.70	−0.38	−295.67
Ready meal[b] (n=36)	1.14	0.68	0.32	+0.23	−436.08
Long-life products[c] (n=121)	5.77	3.84	1.55	+0.07	−402.17
Frozen products (n=24)	1.49	0.88	0.33	−0.04	−312.42

[a]All fresh fruits and vegetables; all salads; all meat, fish and eggs except 3 frozen products; all dairy products except milk powder

[b]All ready mixed dishes except 15 frozen products

[c]All processed fruits; all canned vegetables; all cereals, potatoes and legumes except 6 frozen products; all snacks and sweets; all beverages; milk powder

3.5 Demand Elicitation

Incentive compatibility is key in experimental economics. The incentive mechanism has to be consequential but also manageable for both participants and experimenters in order to be applicable. We thus opted to incentivize only a subset of the food supply. As previously explained, a quarter of the food supply was actually made available for actual purchase at the end of the experiment. While participants did not know which items were actually available, they knew that they had to pay for items from their experimental shopping basket that belonged to the available subset; i.e., they knew they had to actually buy approximately one quarter of their basket. Theoretically, this mechanism should not distort their preferences. Indeed, if $x \succ y$ then $p.x \succ p.y$ (Savage 1954) with here $p = 1/4$. However, participants may have beliefs about the probability distribution and may think that $p_x \neq p_y$. If subjective probability is nil for some items, then their purchase becomes hypothetical, and experimental results could thus be distorted by demand effects (Zizzo 2010). Although much effort has been made to distribute the products available for purchase across all food categories (see Table 2, last column) and over shelf-life length (fresh, frozen, canned, etc.), participants may still have beliefs about the distribution of the available products. For instance, participants may think that fresh and frozen foods are less likely to be stored in the experimental store for practical reasons.[5] In that case, participants may react more to the labelling of short-lived products. To verify this, we perform a robustness check by carrying out the same behavioural analysis on the three labelling treatments as in subsection 3.4, distinguishing fresh products, ready meals, long-life products and frozen products (Table 10). We find no such evidence. On average, participants keep 68% of their fresh products from their reference basket in their label basket. This proportion reaches 60%, 66% and 59%, respectively, for ready meals, long-life products and frozen products. No

---

[5] Although during the instructions, participants were informed orally that the laboratory was equipped with a refrigerator and freezer.

two-by-two comparisons are significantly different (Mann–Whitney test). Similarly, the GHG impact per substitution is also not in favour of fresh products (−296 g vs. 436 g, 402 g and 312 g). Again, there is no significant difference. Purchases therefore do not appear to be distorted by the demand effect.

4 Conclusion and Discussion

In this study, we identify the causal effects of providing food labels that inform consumers of environmental impacts. To do this, we construct an experimental store that offers 282 food items, and we conduct an experiment that allows us to observe actual purchasing behaviour with and without the presence of environmental labels. Most closely related work relies instead on hypothetical choice surveys or observational data. In contrast to surveys, our experimental approach avoids issues associated with hypothetical bias. Relative to observational studies, our methods allow us to study labels that do not currently exist in markets, and our methods control for other factors that may vary between labelled and unlabelled products. In contrast to other experimental studies based on real purchases, we employ a large set of products, and thus our setting better represents a natural shopping experience.

The experimental food set has well-known characteristics. First, food of animal origin has a much greater impact on the environment than food of plant origin (Carlsson-Kanyama and Gonzalez 2009; Masset et al. 2014; Wickramasinghe et al. 2013). Second, the heterogeneity of the environmental impact is also much greater for animal products (Wickramasinghe et al. 2013). This implies that a transition to a vegetarian diet is not necessary to reduce the environmental impact of diets, despite what some authors have advocated (Berners-Lee et al. 2012; Reijnders and Soret 2003; Stehfest et al. 2009; van Dooren et al. 2014). We also find very strong correlations not only between our three environmental indicators (GHG, eutrophication and acidifications) but also between GHG and price and energy.

The results show that environmental front-of-pack labelling improves the environmental impact of food baskets. All formats lead to a reduction in GHG emissions. This finding corroborates the results of Vanclay et al. (2011) and Vlaeminck et al. (2014) who also found, for a more limited set of products, positive effects of simplified labelling.

mTL format appears to be the one that generates the largest decrease in GHG emissions. This finding was not expected. Simplicity usually improves label effectiveness (Temple and Fraser 2014). Nevertheless, mTL is multi-criteria and thus more complex than sTL. Muller and Ruffieux (2018) used the same experimental store to test several nutritional labels. They found that sTL induced more policy-oriented responses than did mTL. The main difference that could explain these apparently contradictory results is that the nutritional criteria (sugar, fat and salt) are uncorrelated. Inversely, we have seen that GHG, eutrophication and acidification are strongly correlated. As a result, nutritional mTL displays different colours, whereas environmental mTL most often displays the same three colours. As a result, nutritional mTL forces consumers to make trade-offs, while environmental mTL generates more saliency.

KM is mono-criterion and numerical; it enables participants to compare any two of all the products available. In comparison, mTL and sTL only distinguish products belonging to the same food category and whose environmental impact is sufficiently different so as not to have the same colour. Still, KM does not do a better job than traffic lights. Although

each item replacement results in a greater GHG reduction with KM than with mTL and sTL, participants make fewer changes to their initial shopping baskets. This highlights once more the superiority of a simple color-coded logo over a more informative digital one: Consumers comply with normative signals. On the other hand, they are more reluctant to continuously make the necessary efforts to make good use of extensive and positive information. We cannot reject the suggestion that KM's relative underperformance can be explained in part by consumers' difficulty in appreciating the equivalence between car kilometres travelled and CO_2 emissions. It is already the case that consumers are not familiar with carbon footprint labels and may misinterpret them (Hartikainen et al. 2014; Sirieix et al. 2013); further studies are needed to determine whether the use of equivalence for CO_2 emissions adds to consumer confusion.

We have argued that there are many advantages of using a laboratory grocery store to identify the effects of food labelling (more control, presence of counterfactuals, reproducible context, etc.). Nonetheless, there are also limitations. Most prominent, the introduction of labels in the experiment is clear to participants, which, relative to an actual supermarket, likely increases the attention paid to them. Further, there is the potential for experimenter demand effect; i.e., participants form an interpretation of the experiment's purpose and modify their behaviour in response. It is therefore likely that an experiment such as ours provides an upper bound on the effects of introducing environment-related food labels.

Appendix 1: Instructions

Instructions were projected both on a big screen and on each individual computer screen and read aloud by a facilitator. The following text corresponds to the KM treatment.

Welcome to this experiment conducted by the Grenoble Applied Economics Laboratory of the University of Grenoble-Alpes and the French National Institute of Agricultural Research. This experiment focuses on food consumption behaviours. Instructions will be given to you as the experiment unfolds.

During the experiment, you will have to make decisions. If you have any difficulties in understanding, please do not hesitate to contact us with any questions you may have.

In front of you is an envelope containing 35€ corresponding to the compensation for your participation.

You will have the opportunity to buy some products. Regarding these products you will have to buy, it is important that you understand clearly, before we begin, why you will have to make some purchases and how you will do it.

You will actually buy some products at the end of this experience. The purpose of these purchases is to ensure that we are observing your true buying behaviour and not just "buying intentions". As behavioural scientists, we know that a purchase intention may be different from an actual purchase. Our principle is simple: motivate you to make real purchases (almost) as if you were shopping.

In the coming experience, each of you will have to make many purchasing choices. Some of these choices will lead to real purchases. You will leave later with some products, all from the choices you have made, that you will have bought with your own money. Under no circumstances does our laboratory make a profit on sales. The (few) products sold will be sold at supermarket selling prices. You will discover these prices during the experience.

Organization of the session

This experience is composed of several parts. Instructions will be given before each part. In two parts, your task will be to compose food baskets.

The data collected are confidential and will only be used for scientific purposes. No individual data will be accessible. The results will be disseminated and published in a statistical mode. The anonymity of each individual will thus be respected.

PART 1
Your task

During this first part, your task is to buy food for your household for 2 days. You will have to compose a basket as you would in a grocery store. In order to observe real buying behaviour, your decisions will be incentivized (i.e., you will buy certain products from your baskets at the end of the experiment).

As in a real store, you are free to buy what and how much you like. If you are not satisfied with the products in this store or if you are sufficiently supplied at home, you are not obliged to buy! The following products are assumed to be available to you at home and should therefore not be purchased here: Butter—flour—oil—spices—coffee—tea—condiments—water—alcohol—wine—sugar—vinegar—sauces.

Your basket

To compose your basket, you can choose from 282 food products. These products are presented in the catalogue that is placed in front of you. Each page of the catalogue is related to a product category. (A summary is on the first page). Each category contains 6 or 9 products.

A barcode reader is at your disposal in front of you. To select a product, scan the product's barcode and press the button. The product will then appear on your screen.

The products
Pictures represent the front side of the packaging as presented on the shelves of a supermarket. By clicking on it on your computer screen, you can enlarge it.

The posted price corresponds to the prices observed in a local supermarket. Weight and price per kilo are provided.

When a product is selected on the screen, you can click on the 'Details tab'. You will then access an information page that provides access to three sets of information: Ingredients,

Nutritional information, Environmental information (NOTE: these three sets of information are described in detail with some examples).

Your purchases

To ensure that your behaviours today best match your actual behaviours, your decisions will be incentivized. This part or the next, depending on a draw at the end of the experiment, will be considered for actual purchases.

We will use the following terms and conditions:

In the room behind you are 71 products chosen from the 282 in the catalogue. 1/4 of the products in the catalogue are available. You will not know what these products are.

At the end of the experiment, we will identify the products in your shopping cart that are available here: probably 1/4 of your products.

Only one part will be taken into consideration for real purchase. If this part is drawn (1 chance over 2), you will have to buy these few products. For example, if your basket includes 20 products, you will have to buy 4 or 5 of them, which is an expense of a few Euros. You will leave with these products. Please note that the products are sold at the same selling prices as in a store.

In order to familiarize participants with the computer interface, screenshots are projected and described.

PART 2

Your Task

We now ask you to compose another shopping cart, an alternative to your first phase shopping cart. A modified catalogue is distributed to you. It contains the same products as the previous one, in the same provision. A label is now associated with each product.

Incentives

This part is an alternative to the first phase. Only one part will be considered for the sale of the products at the end of the experiment. This one will be randomly drawn.

(NOTE: Part 2 instructions end here for the control groups. For the three treatment groups, the corresponding labelling is described)

Description of the labelling

(KM TREATMENT)
Each product now has a label. Such labelling may appear on the front of the product's packaging, if the public authorities so decide. This labelling is an indicator of environmental quality. It concerns the impact on climate change, measured by greenhouse gas

emissions, measured in carbon dioxide (CO_2) equivalent. To build the indicator, greenhouse gas emissions were assessed at all stages of the product's life cycle (Life Cycle Assessment or LCA): agricultural production, transport, processing, packaging, etc.

In the catalogue, to get a concrete idea of this impact, we use an equivalence. When we drive a car, the engine emits greenhouse gases. The greenhouse gas impact of a food is indicated in kilometer equivalents. For example, if a liter of whole milk is indicated as 11.0 km, this means that, to produce, package and transport this litre of milk, the emissions are equivalent to what a motor car emits when it travels 11 kilometres.

(sTL TREATMENT)

Each product now has a label. Such labelling may appear on the front of the product's packaging, if the public authorities so decide. This labelling is an environmental impact indicator. It concerns the impact on climate change, measured by greenhouse gas emissions, measured in carbon dioxide (CO_2) equivalent. The products in each category are classified into three groups: One-third of the best products in terms of the environment have a green dot; One third of the worst environmental performers have a red dot; The intermediate third has a yellow dot.

(mTL TREATMENT)

Each product now has a label. Such labelling may appear on the front of the product's packaging, if the public authorities so decide. This labelling is an environmental impact indicator. It includes 3 colored dots. Each dot informs about an impact. Respectively: On climate change (which is measured in carbon dioxide emission equivalents); On water pollution (phosphate); On air pollution (sulphur dioxide).

For each environmental impact (climate change, water pollution and air pollution) the dot is: Green if the product is in the top third of the category (page); Red if the product is in the lowest third of the category (page); Yellow if the product is in the intermediate third of the category (page).

Appendix 2: France and Grenoble sociodemographic table

	France	Grenoble
Household size, number of persons	2.3	2.3
Income, € per months and per consumption unit	1606€	1558€
Education, %		
No high school diploma	21.6%	26.4%
High school diploma	17.7%	15.3%
BA	39.4%	36.3%
Master and above	21.6%	22.2%

References

Argo JJ, Main KJ (2004) Meta-analyses of the effectiveness of warning labels. J Public Policy Mark 23(2):193–208

Auger P, Devinney T (2007) Do what consumers say matter? The misalignment of preferences with unconstrained ethical intentions. J Bus Ethics 76(4):361–383

Auger P, Burke P, Devinney T, Louviere J (2003) What will consumers pay for social product features? J Bus Ethics 42(3):281–304

Balcombe K, Fraser I, Di Falco S (2010) Traffic lights and food choice: a choice experiment examining the relationship between nutritional food labels and price. Food Policy 35(3):211–220

Berners-Lee M, Hoolohan C, Cammack H, Hewitt C (2012) The relative greenhouse gas impacts of realistic dietary choices. Energy policy 43:184–190

Bougherara D, Combris P (2009) Eco-labelled food products: what are consumers paying for? Eur Revue Agric Econ 36(3):321–341

Carlsson-Kanyama Annika (1998) Climate change and dietary choices—how can emissions of greenhouse gases from food consumption be reduced? Food Policy 23(3):277–293

Carlsson-Kanyama A, Gonzalez AD (2009) Potential contributions of food consumption patterns to climate change. Am J Clin Nutr 89(suppl):1704S–1709S

Cox EP III, Wogalter MS, Stokes SL, Tipton Murff EJ (1997) Do product warnings increase safe behavior? A meta-analysis. J Public Policy Mark 16(2):195–204

Darmon N, Vieux F, Maillot M, Volatier JL, Martin A (2009) Nutrient profiles discriminate foods according to their contribution to nutritionally adequate diets: a validation study using linear programming and the SAIN, LIM system. Am J Clin Nutr 89(4):1227–1236

Drewnowski A, Rehm CD, Martin A, Verger EO, Voinnesson M, Imbert P (2015) Energy and nutrient density of foods in relation to their carbon footprint. Am J Clin Nutr 101(1):184–191

Ernst &Young (2010) Product carbon Footprinting—a study on methodologies and initiatives. European Commission DG Environment

Eurobarometer S (2009) Europeans' attitudes towards climate change. European Comm 29:30

Gadema Z, Oglethorpe D (2011) The use and usefulness of carbon labelling food: a policy perspective from a survey of UK supermarket shoppers. Food Policy 36(6):815–822

Gallastegui I (2002) The use of eco-labels: a review of the literature. Eur Environ 12:316–331

Hartikainen H, Roininen T, Katajajuuri J, Pulkkinen H (2014) Finnish consumer perceptions of carbon footprints and carbon labelling of food products. J Clean Prod 73:285–293

Hertwich E, van der Voet E, Suh S, Tukker A, Huijbregts M, Kazmierczyk P, Lenzen M, McNeely J, Moriguchi Y (2010) Environmental impacts of consumption and production: priority products and materials. International Panel on the Sustainable Use of Natural Resources, UNEP, Paris

Hieke S, Taylor CR (2012) A critical review of the literature on nutritional labeling. J Consum Aff 46(1):120–156

Janssen M, Hamm U (2012) Product labelling in the market for organic food: consumer preferences and willingness-to-pay for different organic certification logos. Food Qual Prefer 25(1):9–22

Johnston RJ, Wessells CR, Donath H, Asche F (2001) Measuring consumer preferences for ecolabeled seafood: an international comparison. J Agric Resour Econ 26(1):20–39

Maillot M, Drewnowski A, Vieux F, Darmon N (2011) Quantifying the contribution of foods with unfavourable nutrient profiles to nutritionally adequate diets. Br J Nutr 105:1133–1137

Manfredi S, Allacker K, Chomkhamsri K, Pelletier N, Maia de Souza D (2012) Product environmental footprint (PEF) guide. European Commission, Ispra

Marette S, Messean A, Millet G (2012) Consumers' willingness to pay for eco-friendly apples under different labels: evidences from a lab experiment. Food Policy 37(2):151–161

Masset G, Soler L-G, Vieux F, Darmon N (2014) Identifying sustainable foods: the relationship between environmental impact, nutritional quality, and prices of foods representative of the french diet. J Acad Nutr Diet 114(6):862–869

Moser R, Raffaelli R (2012) Consumer preferences for sustainable production methods in apple purchasing behaviour: a non-hypothetical choice experiment. Int J Consum Stud 36(2):141–148

Muller L, Prevost M (2016) What cognitive sciences have to say about the impacts of nutritional labelling formats. J Econ Psychol 55:17–29

Muller L, Ruffieux B (2018) Shopper's behavioural responses to 'front-of-pack' nutrition logo formats: GDA Diet-Logo vs. 6 alternative Choice-Logos. WP GAEL

Notarnicola B, Tassielli G, Renzulli P, Castellani V, Sala S (2017) Environmental impacts of food consumption in Europe. J Clean Prod 140:753–765

Onyango B, William H, Bellows A (2007) Purchasing organic food in US food systems: a study of attitudes and practice. Br Food J 109(5):399–411

Reijnders L, Soret S (2003) Quantification of the environmental impact of different dietary protein choices. Am J Clin Nutr 78:664S–668S

Rokka J, Uusitalo L (2008) Preference for green packaging in consumer product choices–do consumers care? Int J Consum Stud 32(5):516–525

Savage L (1954) The foundations of statistics. Wiley, New York

Sirieix L, Delanchy M, Remaud H, Zepeda L, Gurviez P (2013) Consumers' perceptions of individual and combined sustainable food labels: a UK pilot investigation. Int J Consum Stud 37(2):143–151

Stehfest E, Bouwman L, van Vuuren D, den Elzen M, Eickhout B, Kabat P, Vuuren D, Elzen M (2009) Climate benefits of changing diet. Clim Change 95:83–102

Sutherland L, Kaley L, Fischer L (2010) Guiding Stars: the effect of a nutrition navigation program on consumer purchases at the supermarket. Am J Clin Nutr 91(4):1090S–1094S

Tagbata D, Sirieix L (2008) Measuring consumer's willingness to pay for organic and Fair Trade products. Int J Consum Stud 32(5):479–490

Temple Norman J, Fraser Joy (2014) Food labels: a critical assessment. Nutrition 30(3):257–260

van Dooren C, Marinussen M, Blonk H, Aiking H, Vellinga P (2014) Exploring dietary guidelines based on ecological and nutritional values: a comparison of six dietary patterns. Food Policy 44:36–46

van Herpen E, van Trijp HCM (2011) Front-of-pack nutrition labels. Their effect on attention and choices when consumers have varying goals and time constraints. Appetite 57(1):148–160

Van Loo EJ, Caputo V, Nayga R, Meullenet J, Ricke S (2011) Consumers' willingness to pay for organic chicken breast: evidence from choice experiment. Food Qual Prefer 22(7):603–613

Vanclay J, Shortiss J, Aulsebrook S, Gillespie A, Howell B, Johanni R, Maher M, Mitchell K, Stewart M, Yates J (2011) Customer response to carbon labelling of groceries. J Consum Policy 34(1):153–160

Vermeulen S, Campbell B, Ingram J (2012) climate change and food systems. Annu Rev Environ Resour 37(1):195–222

Vlaeminck P, Jiang T, Vranken L (2014) Food labelling and eco-friendly consumption: experimental evidence from a Belgian supermarket. Ecol Econ 108:180–190

Wickramasinghe K, Scarborough P, Goldacre M, Rayner M (2013) Defining sustainable diets by comparing greenhouse gas emissions from different food groups: a systematic review. Lancet 382(Supplement 3):S104

Wier M, Calverley C (2002) market potential for organic foods in Europe. Br Food J 104(1):45–62

Zizzo DJ (2010) Experimenter demand effects in economic experiments. Exp Econ 13(1):75–98