



HAL
open science

Once a quality-food consumer, always a quality-food consumer? Consumption patterns of organic, label rouge, and geographical indications in French scanner data

Mathieu Lambotte, Stéphane de Cara, Valentin Bellassen

► To cite this version:

Mathieu Lambotte, Stéphane de Cara, Valentin Bellassen. Once a quality-food consumer, always a quality-food consumer? Consumption patterns of organic, label rouge, and geographical indications in French scanner data. *Review of Agricultural, Food and Environmental Studies*, 2020, 10.1007/s41130-020-00121-z . hal-02924042

HAL Id: hal-02924042

<https://hal.inrae.fr/hal-02924042v1>

Submitted on 27 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Once a quality-food consumer, always a quality-food consumer? Consumption patterns of organic, *label rouge* and geographical indications in French scanner data

Accepted, unedited version (postprint)

Journal: Review of Agricultural, Food and Environmental Studies

Authors : Mathieu Lambotte^{a*}, Stéphane De Cara^b and Valentin Bellassen^a.

^a CESAER, INRAE, Université Bourgogne Franche-Comté, AgroSup Dijon, 21000 Dijon, France.

^b Economie Publique, INRAE, AgroParisTech, Université Paris-Saclay, 788850 Thiverval-Grignon, France.

* Corresponding author, email address: mathieu.lambotte@inrae.fr

Abstract

The aim of this study is to analyze the behavior of French consumers with respect to food products under various quality labels (organic, *label rouge*, and geographical indications). In particular, we investigate if consumers who purchase once a product under a given label tend to purchase a large fraction of this product (and other products) under the same label.

Using a large scanner database, the regularity of quality-food consumption is analyzed through the relative frequency of conventional and quality purchases. The respective roles in regular consumption of product attributes, availability and household characteristics are then examined using a random utility model.

Regular organic consumers purchase around 28% of the organic market value, with variations depending on products. We find that product attributes are more related to regular organic behavior than household characteristics. In particular, product availability and product family (vegetables, eggs, milk, etc.) play a key role whereas low-price organic products are not associated with more regular consumption.

Acknowledging the existence of regularity in organic consumption and understanding its variation between product categories should lead public policies to more often target specific products in order to develop quality-food consumption.

Keywords: quality food, consumption behavior, organic, regularity, Lancaster.

Acknowledgments

We are thankful to Stéphan Murette, Céline Bonnet, Christine Boizot-Szantai, Burkhard Schaer, Sophie Legras, Georges Giraud and Louis-Georges Soler for their fruitful discussions and comments on the manuscripts. We are also grateful to ALISS and especially Nicolas Guinet for facilitating the access to the panel data and providing support to its analysis

Declarations

Funding

The research leading to these results was funded by the REGION Region Bourgogne Franche-Comté and project DIETPLUS, grant ANR-17-CE21-0003 financed by the French National Agency for Research (ANR). The authors only are responsible for any omissions or deficiencies. Stéphane De Cara acknowledges additional support from the CLAND grant (ANR-16-CONV-0003) funded by the French National Research Agency under the "Investissements d'avenir" program.

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

1. Introduction

Over the last few years, the consumption of organic food has experienced a rapid and steady growth. In France, its value rose from € 5.9 billion in 2015 to € 9.7 billion—i.e., 5% of food expenditures—in 2018 (Agence Bio, 2019b). In a recent French survey, about one in eight respondents declares eating at least one organic product a day (Agence Bio, 2019a). Other quality labels—such as Protected Designation of Origin (PDO), Protected Geographical Indication (PGI) or *label rouge*—also represent a substantial and increasing share in food consumption. For some products, these labels dominate organic ones as a quality sign. This is notably the case for cheese, for which geographical indications represent 11% of the French market, a much higher share than that of organic cheese (2%).

The rising demand for quality-food products—defined here as products certified as organic, PDO, PGI or *label rouge*—impacts farming practices, land use, and value chains. As an illustration, the area devoted to organic agriculture in France has increased by 48.3% from 2015 to 2018, reaching 7.5% of the total agricultural area. As for other quality signs, between 2015 and 2019, the volume of meat and dairy products, fish, and eggs produced under these labels has increased from 602 to 665 thousand tons, reaching 1.5% to 1.8% of market share in these categories (INAO, 2019).

Providing consumers with healthier and higher-quality food while leveraging the potential health and environmental benefits from less intensive agricultural practices without affecting too negatively farmers' income has attracted renewed interest from policy makers. The development of the production and consumption of quality-food products (organic food, but also other quality signs) has become a central objective in agricultural and food policies. EGalim, a French law on food and agriculture passed in 2018, provides a good illustration of this trend in public policies.

An adequate understanding of the demand-side determinants of quality-food consumption is thus critical to assess the potential impacts of such policies. In particular, it is important to determine whether the demand for these products comes primarily from a large base of occasional consumers or a small base of 'regular' ones. The recent increase in quality-food consumption is consistent with both a larger consumer base and an increasing share of quality-food in individual purchases. Which explanation dominates remains however an open question. This raises further questions regarding consumers' behavior with respect to quality-food products. Do consumers tend to routinely and systematically purchase a product under the same label? If so, is this behavior restricted to some products, or consistent across the food basket? The answers to these questions may have strong implications for the design and targeting of the policies aiming to develop quality-food consumption.

In this article, we define a 'regular' consumer—as opposed to an 'occasional' one—as a consumer who predominantly buys a given product with a consistent quality. Regularity in quality-food consumption is thus related to how consumers perceive and value quality labels. In this regard, previous research has shown that credence attributes—i.e., attributes that the consumer cannot identify through search nor experience—play a major role in the decision to buy quality-food products (Massey et al., 2018; Padilla et al., 2013; Rana and Paul, 2017). If the demand for quality-food products is primarily driven by credence attributes, one may expect that regularity should be the rule rather than the exception.

Yet, this intuition is at odds with recent evidence. In a study of organic food purchasing behaviors in France, Boizot-Szantai et al. (2017) find that organic food represents only 8.1% of total food expenditures of the consumers whose organic budget share is above the top quintile. This suggests that, even for the consumers who value organic food the most, a large share of their purchases goes to non-organic products. That is not to say, however, that regularity is necessarily absent from consumers' behavior. A consumer may well always choose to buy the organic version of some products, and systematically prefer the conventional version (or another quality label) of other products. The aggregation of all food expenditures at the food-basket level therefore may thus mask regularity patterns that occur at the product level.

Our main objective is to identify and document regularity patterns in food purchasing behaviors at various levels of aggregation and for various quality labels, and to study the interplay with other drivers of quality-food consumption such as price, availability and socio-demographic characteristics of quality-food consumers.

Our study thus contributes to the large body of literature—in economics, but also in sociology and marketing science—that has examined consumers' behavior and attitude towards quality-food products. The vast majority

of this literature is focused on organic consumption, which covers a wider range of food products than other quality signs. This literature has investigated the effects of price, income, and socio-demographic variables on organic food consumption, as well as the role of the relative preferences for products objective characteristics (taste, color, etc.) and considerations regarding health, the environment, ethics, or animal welfare.

Quantitative studies in this literature are based on two broad types of approaches. First, a number of studies use surveys or controlled experiments. These studies rely on stated preferences to elicit consumers' preferences or willingness to pay for various goods with different characteristics or attributes. The experimental design makes it possible to disentangle potential confounding factors. However, they often rely on small sample sizes and/or may be subject to sampling biases. Furthermore, stated intentions to buy quality food may overestimate actual purchases (Sun et al., 2010). Studies in the second category are based on observational data. Those relying on scanner data are particularly relevant for the present study. By using detailed records of actual and repeated purchases by a large and representative sample of consumers, such studies allow to overcome some of the limitations faced by survey-based and experimental approaches. Observed purchases do not, however, directly inform on consumers' preferences. In addition, such data mostly pertain to purchases for food-at-home consumption at the household level, which may differ from actual consumption at the individual level.

This literature has shed interesting light on the socio-demographic profile of quality-food consumers, as well as on the influence of product attributes on consumption behavior. For example, findings suggest that organic food consumers tend to be more educated and have higher income than conventional food consumers. They are more likely to be female and to have children, especially newborns (Gracia et al., 2008; Hughner et al., 2007). They also spend more on food and have a healthier diet (Boizot-Szantai et al., 2017). Several studies underline the key importance of credence attributes in attitudes toward organic food. Consumers tend to place more value on credence attributes of organic products than on their search and experience attributes (Massey et al., 2018; Rana et al., 2017). The influence of credence attributes on the probability of purchasing organic products has been found to be larger than that of education, marital status, and income (Padilla Bravo et al., 2013). The review by Rödiger and Hamm (2015) shows inconclusive evidence with respect to the effect of price on organic consumption. Results from observational studies indicate that the effect of economic determinants of organic consumption, most notably price and availability, varies across products and types of retail stores (Buder et al., 2014; Dimitri et al., 2012; Padel et al., 2005).

Regularity in organic food consumption has been much less studied than the influence of the socio-demographics variables, price, income, or product attributes. Some survey-based studies have examined the factors influencing the probability that a consumer defines himself or herself as a regular organic consumer (Barrena et al., 2010; Baudry et al., 2019; Kesse-Guyot et al., 2013; Oates et al., 2012; Onyango et al., 2007; Pearson et al., 2013; Treu et al., 2017). However, as these studies are based on stated consumption averaged at the whole food basket level, it is difficult to draw robust conclusions on actual purchases at finer aggregation levels. Observational studies based on scanner data (Boizot-Szantai et al., 2017; Lacour et al., 2018) provide some interesting findings about purchasing patterns, but, the focus in these studies is on the share of organic food in total food expenditures. Considering all food products together may mask regularity patterns at a more disaggregated level.

Although expectedly important, the availability of quality-food products has also received less attention than other potential determinants (Dimitri et al., 2012). One reason lies in the difficulty to accurately quantify availability. As a consequence, availability has been overlooked in most survey-based and observational studies of organic consumption.

In this paper, we use a large scanner dataset of food purchases in France, provided by Kantar®, which provides detailed records (quantities, price, type of retail store, etc.) of purchases of 237 food products by an unbalanced panel of 12,453 French households between 2011 and 2016, along with socio-demographic information about households (age, income, *département* of residence, household size, etc.). This dataset enables to decompose household food purchases at a fine level of disaggregation in terms of products. Importantly, it informs about the labels (organic, PDO, PGI, *label rouge*), if any, attached to each purchased product.

In order to identify regularity, we use two complementary approaches. For a given product or category of products, we first examine whether the distribution of purchases under a given label is compatible with a large base of occasional consumers, or a small base of quality regulars. In a second step, we then investigate the influence of various determinants of the probability that a consumer is regular for a given product and label. The

variables accounted for in this analysis include various characteristics of products (price, retailer's brand, availability), retail stores, and households (socio-demographic variables, income, city size, etc.).

We take advantage of the richness of this information to address four main questions. First, can the consumption of a given (set of) product(s) be characterized as 'regular'? In other words, do consumers purchase mostly one version of the product (conventional, organic, PDO, PGI, or *label rouge*) or do they buy a mix of various versions? Second, is regularity consistent across the food basket? Put differently, are quality regulars for a given product also regular for most of all other products they purchase? Third, how much of the total demand for quality-food products comes from regular consumers? Fourth, how do product attributes and household's characteristics interplay in the probability that a given household is a quality regular for a given food product?

Our contribution is threefold. First, we address these questions for a wide range of products sold under various labels. Contrary to most of the literature, which is predominantly focused on organic products, we extend the scope of the analysis to other quality labels such as PDO, PGI, or *label rouge*. Products under these labels have several important features in common with organic products. In particular, their production is often based on extensive farming practices with a potential to improve the sustainability of food systems (Arfini and Bellassen, 2020) and credence attributes play a major role in consumers' decision to buy these products. They do however differ from organic products in a number of dimensions, such as the range of products covered by the label, availability, price, perception of environmental and health benefits, organoleptic characteristics, aspect, etc. We investigate whether different labels can be associated with different purchasing behaviors in terms of regularity. Besides, previous studies were often focused on only one or a few products. The large number of products (237) examined in this article allows us to further study how consumers' behaviors varies with the type of products (Buder et al., 2014; Chekima et al., 2017; Padel et al., 2005).

Second, we identify regularity patterns in quality-food consumption, but not for all products nor all labels. Such findings would not have been possible with the level of aggregation and coverage used in previous studies (Boizot-Szantai et al., 2017; Lacour et al., 2018).

Third, we propose a new indicator of availability of quality-food products. The main difficulty is that, availability of products that are not bought by consumers in the sample cannot be directly observed. We circumvent this difficulty by constructing a novel indicator based on web-scraping techniques applied to online catalogues of retail stores. This indicator allows us to determine the share of retail stores where any given product can be found within a given region in the total number of stores visited by a given consumer. This indicator is found to be strongly and positively correlated to the probability that a consumer is an organic regular.

In sum, after detailing our methodology, we present the results organized in two axis, the first one on the identification of quality-food regular consumption and the second on its determinants, then discuss them and conclude.

2. Methodology

2.1. Theoretical framework

Consumers often have the choice between several versions of the same food product, e.g. between conventional, organic, PDO, PGI, or *label rouge*. Although different versions of the same product share some common characteristics, they may also differ along various attributes. Some of these attributes are specific to the product itself, whereas others are attached to perception of the label by the consumer. Some can be observable through search and experience. Others remain unobservable and are therefore akin to credence attributes.

Previous research has shown that credence attributes are major drivers of organic and other quality-food purchases (Gracia and Magistris, 2008; Barrena and Sánchez, 2010; Massey et al., 2018; Rana and Paul, 2017). Most of the attributes that define a product—regardless of the presence of a label—can be identified by search and/or experience. In contrast, most of the distinctive attributes of quality-food items are credence attributes. Note that some of the credence attributes associated with a given label may be shared by all products under the same label. It is thus important to disentangle consumers' preferences towards (i) food products regardless of their quality label and (ii) credence attributes associated with quality labels (or absence thereof) across all products. Note also that preferences for a quality label may vary from one product to the other.

It is useful to illustrate the consumer's problem by a simple theoretical framework inspired by Lancaster's approach to consumer demand (Lancaster, 1966) and akin to the individual utility model of the BLP model (Berry et al., 1995). This simple model accounts for search and experience as well as credence attributes. Consider the choice faced by consumer i over a set J of food products (indexed by j). Food products are available in various qualities indexed by l in L , including a conventional version (indexed by $l = c$) and different quality labels. Assume that the consumer's preferences can be represented by a utility function

$U_i(q_{i0}; \{u_{ijl}(q_{ijl})\}_{j \in J, l \in L})$, where q_{i0} is the quantity of a composite non-food item, q_{ijl} is the quantity of j -th product of quality l , and $u_{ijl}(\cdot)$ is the respective subutility function. For simplicity, assume that $u_{ijl}(q_{ijl})$ takes the following simple Cobb-Douglas form:

$$u_{ijl}(q_{ijl}) = a_{ijl} \ln(q_{ijl}) = a_{ij}^{se} \ln(q_{ijl}) + a_{il}^{cr} \ln(q_{ijl}) + a_{ijl}^{se \times cr} \ln(q_{ijl}) \quad (1)$$

The formulation in Eq. (1) characterizes the consumer's utility with respect to product j of quality l . It is parameterized by $a_{ijl} > 0$, which can be decomposed into three terms capturing the utility derived from search and experience attributes (se), credence attributes (cr), and an interaction term ($se \times cr$). The parameter a_{ij}^{se} parametrizes the utility for the j -th product regardless of its quality. It captures how much the consumer values product-specific attributes that can be apprehended by search and experience (e.g., in the case of an apple, characteristics such as the fact that it is a rather small, round, acid and sweet fruit). a_{ik}^{cr} parameterizes the utility for quality l across all products. This parameter is thus label-specific and captures how consumer i values the credence attributes attached to label l (e.g., in the case of organic products, the implications that farming practices do not rely on pesticides nor synthetic fertilizers). The third term $a_{ijl}^{se \times cr}$ captures the interaction between the search and experience attributes of product j with the credence attributes associated with quality l . This reflects that the credence attributes attached to a specific label may vary from one product to the other (e.g., the consumer may value more organically-produced apples than organically-produced eggs). Further assume that the price of the conventional version of the product ($l = c$) is lower than the price of all other quality versions of the product. For simplicity, we normalize the parameters in Eq. (1) so that $a_{ic}^{cr} = a_{ijc}^{se \times cr} = 0$ for all j .

The quantities q_{ijl} that maximize the consumer's utility subject to the budget constraint depend on prices and income, as well as on the elasticities of substitution between products and qualities implied by the form of $U_i(\cdot)$. One expects that the demand in product j of quality $l \neq c$ to be close to 0 if a_{il}^{cr} and $a_{ijl}^{se \times cr}$ are both close to 0—i.e., if the credence attributes attached to the label l are perceived as unimportant by the consumer relative to the search and experience attributes of the product. If this condition holds for all $l \neq c$, the product is predominantly purchased in its conventional version and the consumer is said to be a conventional regular. Conversely, if $a_{il}^{cr} + a_{ijl}^{se \times cr}$ is sufficiently large relative to a_{ij}^{se} , the consumer is a quality l regular for product j .

Moreover, if a_{il}^{cr} is sufficiently large relatively to $a_{ijl}^{se \times cr}$ for all j —that is, if the consumer mostly values the credence attributes associated with the quality label l —and $U_i(\cdot)$ is such that the consumer's preferences exhibit

some taste for product variety, one should expect the quantities of purchased food items of quality l to be positively correlated across products. If, to the contrary, the consumer mostly values the credence attributes when they are associated with a specific subset of search and experience attributes—corresponding to $a_{ijl}^{se \times cr}$ being much larger than a_{il}^{cr} for a few values of j , one should expect little correlation between the purchased quantities of products of quality l .

2.2. Distribution of quality-food consumption

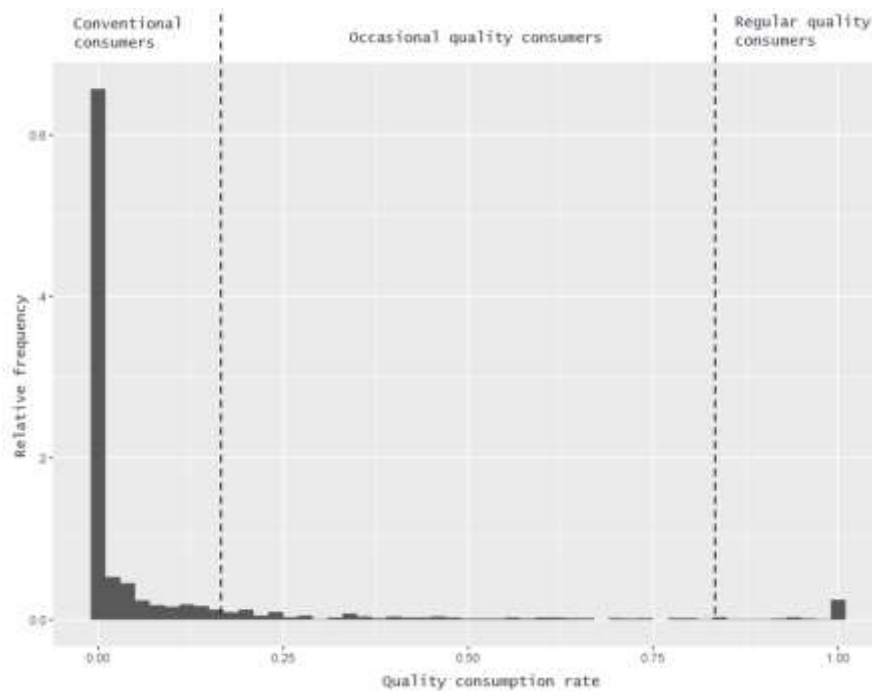
We first investigate regularity by analyzing the distribution of the quality consumption rate of a given product, i.e. the ratio of purchases of the quality version of the product over the total number of purchases during the time span the household remains observed in our dataset. Individuals who almost never buy the quality version of the product are characterized by a quality consumption rate close to 0. They are said to be ‘conventional regulars’; those who mostly purchase the quality version (i.e., quality consumption rate close to 1) are said to be ‘quality regulars’.

The first part of our investigation is organized around the three following hypotheses:

- For a given product, consumers are regulars: they either purchase mostly its conventional version or its quality version, but not a mix of both (H1);
- Quality regular consumption is consistent across the food basket: a consumer who is a *quality regular* for a product is a *quality regular* for an important part of his food basket (H2);
- The total turnover of quality products purchased by regulars is substantial (H3).

If H1 holds true, the distribution of the quality consumption rate among the sample should be bimodal, with a mode around 0 (conventional regulars) and another around 1 (quality regulars). Figure 1 illustrates such a case. The dashed lines represent the two thresholds under and above which a consumer is considered as a conventional or organic regular, respectively.

Figure 1. Bimodal distribution and classification of quality-food consumers



Following Silverman (1981), it is possible to test if the density of a distribution has more than k modes using kernel density estimation. Indeed, the kernel density estimate for window width h based on the observations X_1, \dots, X_n is defined by

$$f(t, h) = n^{-1} h^{-1} \sum_{i=1}^n K\{h^{-1}(t - X_i)\}, \quad (2)$$

where K is the normal density function in our case. The window width h is a parameter controlling to which extent the observations are smoothed in order to obtain the kernel estimate and t is the endogenously estimated mean of the kernel. If the data is multimodal, a large value of h will be needed to obtain a unimodal kernel estimate.

The test relies on the comparison between the estimated h for a unimodal distribution and h_{crit} , the k-critical window width, defined as:

$$h_{crit} = \min\{h, f(\cdot, h) \text{ has at most } k \text{ modes}\}.$$

The lowest h_{crit} gives the most likely number of modes for the distribution. Using the Silverman test, we can thus test whether the distribution of quality-food consumption is bimodal for each product and product family.

To test H2, we define three types of behavior toward quality and organic food, illustrated in Figure 1:

- *conventional regulars*, who are almost never buying the quality version of a product (0-20% of quality purchases for a given product);
- *quality-food regulars*, who are almost always buying the quality version (more than 80% of quality purchases for a given product);
- and *occasional quality-food consumers* (20-80% of quality purchases for a given product).

The 20% and 80% thresholds are partly arbitrary. However, when a beta distribution is fitted to the organic consumption rate of products with a bimodal distribution, around 90% of the values are either above or below these thresholds (see supplementary materials (SM) 1 for the detailed procedure). Through this typology, we also assess the duplication of quality consumption behavior across different products (H2): if I am an organic regular for eggs, am I likely to be an organic regular for milk? (see Monier et al., (2009) for an example of this question in a two-products case).

2.3. Determinants of the probability to be a regular organic consumer

In a second step, we assess which product attributes or household characteristics influence whether a quality-food consumption (i th consumer x j th product) is regular. Because no regular behavior is identified for geographical indications and *label rouge* products (see section 4.1.1), this second step is restricted to organic products. We use the random utility discrete choice framework (McFadden, 2001) as an application of the theoretical model. In this model, the utility function is assumed to be well-behaved (preferences are complete, reflexive and transitive) and known by the consumer. Thus, the consumer can compare the organic and conventional alternatives of a product and rank them in order to purchase the product that maximizes his utility. However, some parts of the utility function, essentially the preference for the credence attributes *per se*, cannot be separated from other elements in scanner data. This is why the random utility model is used here to identify which product attributes or household characteristics influence quality-food regularity rather than a structural estimate of the coefficients in equation (1) of our theoretical framework. The random utility model describes the utility (U_{ijo}) of the i th consumer purchasing the j th product in its organic version as the sum of the observed attributes (V_{ij}) and a random component (ε_{io}):

$$U_{ijo} = V_{ijo} + \varepsilon_{ijo} \quad (3)$$

Similarly, the utility (U_{ijc}) of the conventional product j 's choice is described as:

$$U_{ijc} = V_{ijc} + \varepsilon_{ijc} \quad (4)$$

The i th consumer will be an organic regular for the j th product at time t if $U_{ijto} > U_{ijtc}$ and the probability that this consumer will be an organic regular can be written as:

$$P(Y_{ijt} = 1) = P(U_{ijto} > U_{ijtc}) = P(\varepsilon_{ijtc} - \varepsilon_{ijto} < V_{ijto} - V_{ijtc}) = P(\varepsilon < V_{ijto} - V_{ijtc}) \quad (5)$$

where Y_{ijt} is a binary choice variable:

$$Y_{ijt} = \begin{cases} 1 & \text{if the } i\text{th consumer is an organic regular for the } j\text{th product at time } t \\ 0 & \text{otherwise} \end{cases}.$$

Defining $f(\varepsilon_i)$ as the density function of ε_i , (3) becomes:

$$P(Y_{ijt} = 1) = \int_{\varepsilon} Z_{ijt} (\varepsilon_{ijt} < V_{ijto} - V_{ijtc}) f(\varepsilon_{ijt}) d\varepsilon_{ijt}, \quad (6)$$

where Z_{ijt} is a binary variable indicating if the term inside the parenthesis is true ($Z_{ijt} = 1$) or false ($Z_{ijt} = 0$), i.e. if the utility derived from the regular organic choice exceeds the one from the conventional choice.

Furthermore, Z_{ijt} can be empirically described as depending on the i th consumer's characteristics and the j th product attributes at time t . Because most households remain in the panel for several years – on average 5 years – a random household effect is introduced to correct for the within-household correlation in the error terms:

$$Z_{ijt} = \beta X_{ijt} + u_i + \varepsilon_{ijt} \quad (7)$$

where $X_{ijt} = (x_{ijt1}, \dots, x_{ijt k})$ is a matrix of variables explaining the choice of being a regular organic consumer, i.e. they represent the j th product attributes (price ratio between the organic and the conventional version, a binomial variable indicating if the product is processed or not, the type of shop where the product is bought...) or the i th consumer's characteristics (age of children, income, size of the city in which they live, their socio-professional category...), $\beta = (\beta_0, \beta_1, \dots, \beta_k)$ is a vector of parameters to be estimated, ε_{ijt} captures the idiosyncratic residuals and u_i is a random effect related to the i th household, accounting for the unobserved heterogeneity of the households and the correlation among the ε_{ijt} .

To estimate Eq. (7) with scanner data, we assume that ε_{ijt} follow a logistic distribution and thus that ε_{ijtc} and ε_{ijto} are identically and independently distributed as type I extreme value (Onyango et al., 2007). This hypothesis on the distribution of the errors terms may sometimes be violated in empirical analysis, as the unobserved portion of utility, captured by the errors terms, can be correlated among the different consumer x product couples, but the potential induced bias is limited (Train, 2003). Under this assumption, $P(Y_{ij} = 1)$, the probability that the consumer i is an organic regular for the product j is given by the following logit model:

$$P(Y_{ij} = 1) = F(Z_{ij}) = F(\beta X_{ijt} + \varepsilon_i) = \frac{1}{1 + \exp(-\beta X_{ijt} + \varepsilon_i)}. \quad (8)$$

Furthermore, we construct two other logit models, one—Logit Product—using only the information on the product attributes and the other—Logit Household—using only the variables describing the consumers' characteristics. Using a likelihood ratio test, we then compare the goodness of fit of these two models, in order to assess which set of characteristics—products' vs households'—is the most important driver of regular organic behavior.

We repeat the analysis on a restricted sample containing only households who purchased at least 20% the product of interest as organic. This allows to examine whether the same variables are correlated to the probability of being a regular organic consumer for a product, compared to being either a conventional or occasional consumer (called thereafter “full sample logits”) or to the probability of being a regular organic consumer compared to being an occasional consumer only (“restricted sample logits”).

3. Data

This paper uses French data from Kantar WorldPanel which contains food-at-home purchases of French households. For the analysis, we use the number of organic, quality and conventional products purchased, filtering out households who bought less than three times the considered products. To define the product families (level 1, 23 families), the classification from (Boizot-Szantai et al., 2017) is reused, with more detailed groups for fruits and vegetables. We also created a more in-depth classification (level 2, with 237 categories/products), which differentiate each fruit and vegetable (carrots, potatoes...) and types of meat (pork, beef...). In this panel data, we also have socio-economic information on 12,453 yearly active households, surveyed one or several years from 2011 to 2016 (5 years on average) whose descriptive statistics are presented below (Table 1). In addition, a map of the spatial distribution of our sample in France is available in SM 3.

Table 1. Descriptive statistics of the households in the sample (relative frequencies)

Years		Children		Professions and socio-professional categories		City		Family status		Income category	
2011	15.0%	No children	67.7%	Managers, shopkeepers and heads of companies	34.3%	City of less than 2000 inhabitants	27.7%	Couple	61.8%	Superior middle income	39.9%
2012	14.9%	Youngest child <25 month	5.5%	Farmers	0.8%	City of 2000-4999 inhabitants	6.9%	Single	38.2%	High income	48.5%
2013	15.3%	Youngest child >25 months and <5 years	7.9%	Managers and superior professions	9.1%	City of 5000-9999 inhabitants	6.0%			Inferior middle income	9.3%
2014	17.8%	Youngest child >6 years and <10 years	10.3%	Employes	20.9%	City of 10 000-19 999 inhabitants	5.5%			Low income	2.3%
2015	16.5%	Youngest child >11 years and <15 years	8.7%	Manual occupations	16.2%	City of 20 000-49 999 inhabitants	6.7%				
2016	20.5%			Intermediate professions	15.2%	City of 50 000-99 999 inhabitants	6.6%				
				Retired	1.8%	City of 100 000-199 999 inhabitants	4.7%				
				Unemployed, student	1.6%	City of more than 200 000 inhabitants	20.2%				
						Paris agglomeration	15.7%				

As shown in the introduction, the organic products' availability and their price differences compared to conventional versions are presented as key drivers of organic consumption.

However, in the recent literature on organic consumption, availability is most often neglected. In order to assess the role of product availability on consumption behavior, an indicator of the availability of quality-food—organic, red label or geographical indications—is developed for each consumer. This indicator is the share of shops which offer the quality food of interest out of all shops where the household is going and the exhaustive results are available in the supplementary data (SD) 1. It is computed after the following steps:

- For each household, we define the set of shops where it is going as the shops where the household went at least three times.
- For each shop, we estimate that a given quality product is available if it has been purchased by at least one household in a shop of the same retail chain, same size (hypermarket/supermarket) and in the same region. These three variables are indeed the most important in predicting the availability of organic products: together, they explain 68% of the variance in the number of organic items available for online purchase in the Burgundy region (see SM 2 for details on data collection and regression).
- At shops where there are few purchases for a given product, the absence of organic purchases may be an artefact caused by data scarcity rather than actual unavailability. These shop x product combinations are filtered out as “no data” to avoid false negatives. For this purpose, the act of buying an organic product in a given shop is assumed to follow a binomial law using the nationwide average share of organic purchases for this product as its probability of success. We determine the critical (minimal) sample size, for a given significance level of 0.05 and a power level of 0.8, for which the hypothesis that the product is available but has not been purchased can be rejected.
- Then, if the observed number of purchases of a product is above the critical sample size and if the organic version has never been purchased, we can assert that the organic version of the product is not available. However, if the organic version has been purchased at least one time, we directly assess that it is available in the shop considered.
- The indicator of availability for each household is then defined as the proportion of shops offering the organic product among all shops in which the household purchases food. The indicator ranges between 0—the organic version of the product is never available to the household, and 1—the organic version of the product is available in all the shops the household attends to.

Similarly, we develop a price ratio indicator (organic price per kg divided by conventional price per kg) between organic and conventional versions of a same product, which varies depending on the region, the shop and the product considered. In addition, we create an absolute price difference indicator (organic price per calorie – conventional price per calorie, the harmonization per calorie making it comparable across product families).

4. Results

4.1. Bimodality of quality-food consumption at different aggregation levels

4.1.1. Consumption behavior of geographical indications and *label rouge* products is occasional

For the few products for which information on quality purchases other than organic is available (labelled meat, fish, eggs, processed meat and cheese), these quality purchases are not bimodally distributed, with the exception of processed meat (*label rouge* and PDO matured ham) which has a mode above 80% of quality purchases (Table 2). It must be noted that *label rouge* and PDO/PGI information are lacking for some products (i.e. meat, where certification information is available for chicken only). Consumers thus buy these products as exceptional purchases, possibly for special occasions. H1 is therefore invalidated for geographical indications and *label rouge*, and so is H2. Moreover, only 3% of the market value of these certified products is purchased by regular consumers, which invalidate H3 (for an example, see Table 4). As geographical indication and *label rouge* products are not subject to regular consumption, we do not perform the second part of the analysis (estimation of the determinants of regular consumption) for them.

Table 2. Distribution of the quality labels (PDO, PGI, *label rouge*) consumption per product family

	Nb of bimodal categories	Nb of unimodal categories	Nb of multimodal regular categories	Nb of multimodal not regular categories
Meat (Chicken only)		2		
Processed Meat	2	3	2	2
Seafood		2		
Eggs		1		
Cheese		1		

Unimodal = single mode (no regular behavior or only conventional regulars) / Bimodal = two modes with one lower than 20% and the second higher than 80% (regularity, either conventional or quality-food) / Multimodal regular = at least three modes, including one higher than 80% (existence of all consumer types: conventional regulars, occasional and quality regulars) / Multimodal not regular = at least three modes, none higher than 80% (no quality regulars).

4.1.2. Organic regulars are consistent for a given product, but not among product families

56% of the 23 families of products are subject to regular behavior from a significant share of consumers (Table 3): 35% of distributions are bimodal and 21% are multimodal with one mode in the 80-100% range). However, the rationale for agglomerating products into families is not necessarily suited to assess the consistency of consumer behavior. Indeed, consumers are more consistent at the product level (for exhaustive results see the SD 2 and 3): 80% of products are subject to regular behavior from a significant share of consumers (bimodal (44%) or multimodal with a mode in the 80-100% range (36%)). For example, the distribution of the organic fat product family is not bimodal while a deeper look at the distribution of the 5 products categories composing this family shows that 3 fat products categories have a bimodal organic consumption distribution (Figure 2). H1 is therefore validated for organic products.

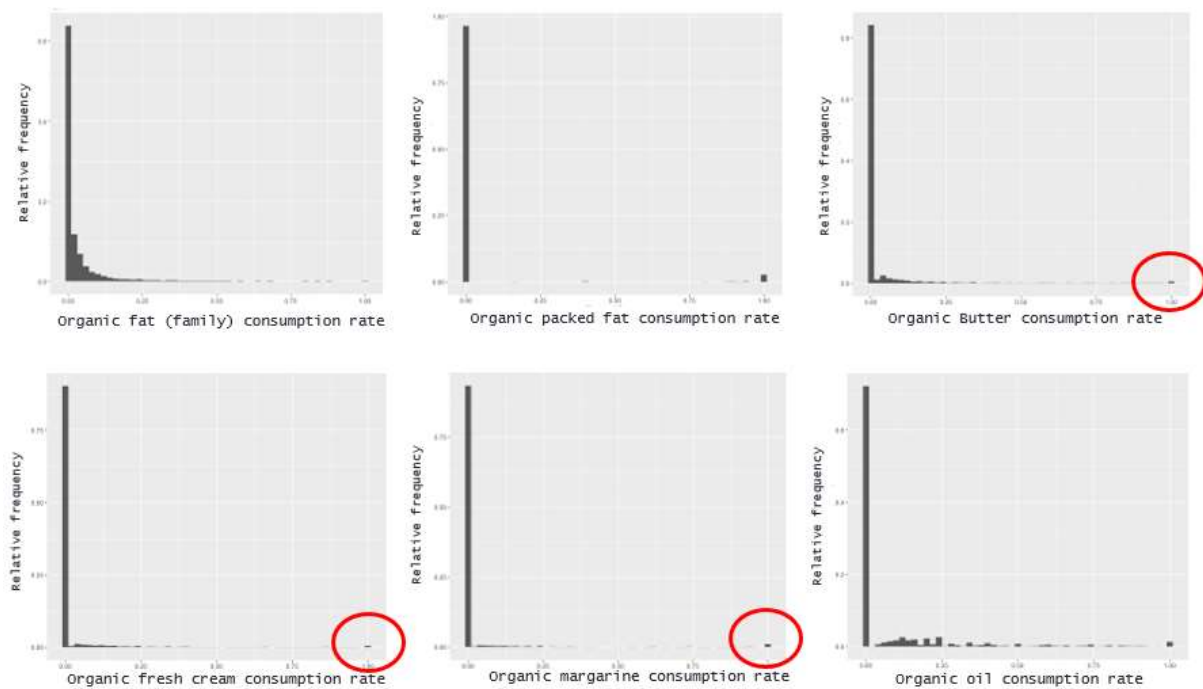
Table 3. Distribution of the organic consumption per product family

	Nb of bimodal distributions at product level	Nb of unimodal distributions at product level	Nb of multimodal distributions with one mode higher than 80%	Nb of multimodal distributions without mode higher than 80%	Family distribution
Alcoholic beverages	1	1	3	3	Multimodal regular
Appetizers	0	0	3	0	Multimodal regular
Baby foods	4	0	2	0	Bimodal

Biscuits, cakes and pastry	1	0	3	0	Unimodal
Bread, flour	1	0	1	1	Bimodal
Cheese	1	1	1	0	Unimodal
Confectionary products	2	1	2	2	Unimodal
Culinary ingredients	11	2	11	2	Unimodal
Desserts	0	0	1	0	Bimodal
Eggs	1	0	0	0	Bimodal
Fat	3	0	1	1	Unimodal
Fresh F&V	45	3	20	6	Unimodal
Hot drinks	1	0	2	1	Unimodal
Prepared Meal	2	1	3	0	Multimodal regular
Meal substitutes	2	0	0	0	Bimodal
Meat	0	0	3	2	Unimodal
Milk	1	0	1	0	Bimodal
Non-Alcoholic beverages	0	1	2	0	Unimodal
Processed F&V	24	1	7	8	Unimodal
Processed Meat	3	1	10	5	Bimodal
Seafood	0	0	4	2	Multimodal regular
Starchy foods	0	1	3	1	Unimodal
Sweeteners	1	0	3	0	Multimodal regular

Unimodal = single mode (no regular behavior or only conventional regulars) / Bimodal = two modes with one lower than 20% and the second higher than 80% (regularity, either conventional or quality-food) / Multimodal regular = at least three modes, including one higher than 80% (existence of all consumer types: conventional regulars, occasional and quality regulars) / Multimodal not regular = at least three modes, none higher than 80% (no quality regulars).

Figure 2. Distribution of the organic consumption of fat products



4.1.3. The basket of organic regulars remains dominated by conventional products

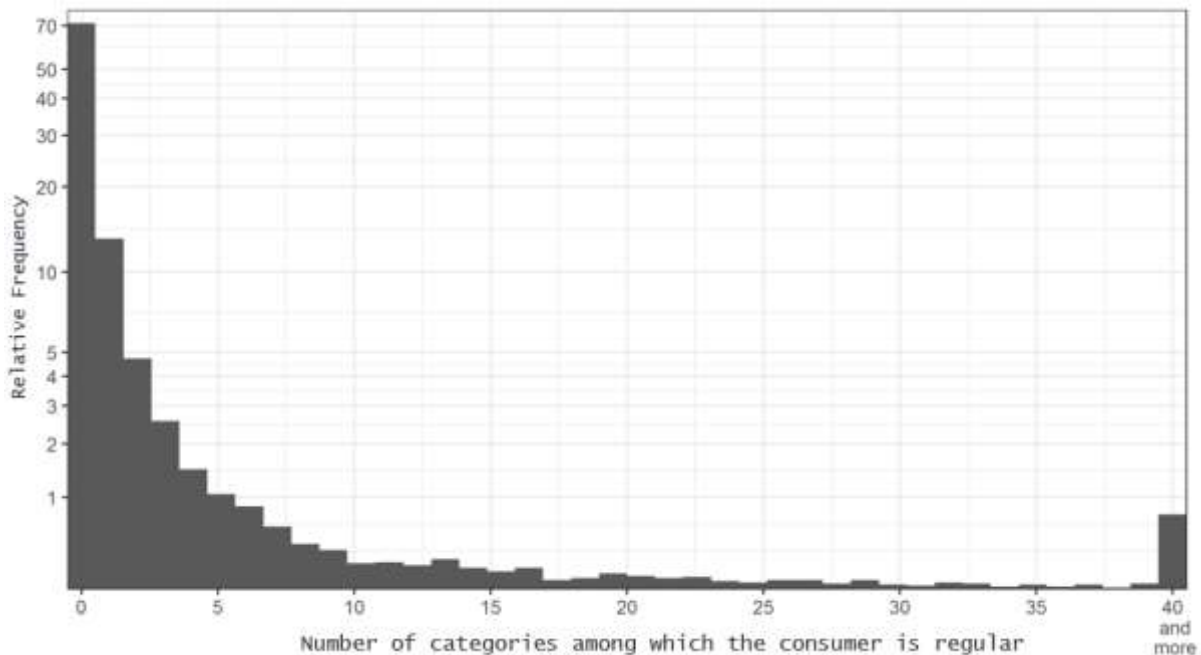
Even if the existence of organic regulars can be observed for 80% of food products, the consumers that purchase several products as organic regulars are scarce. More than 71% of our sample does not buy any product as

organic regulars and only 6% purchase more than 5 products as organic regulars (see SD 4 for exhaustive results). The consumers in our sample purchase on average 80 different products. This shows that most consumers have a regular organic behavior for only a few products (Figure 3). Thus, regular organic behavior is not very consistent across products and H2 is invalidated (this is consistent with the findings of Marian et al. (2014) in the case of organic red meat, chicken, milk and hard cheese).

In our theoretical framework, this indicates strong interaction effects in the consumer utility function (Eq. (1)) between the organic (credence) attribute of a product and the search and experience attributes (roots vegetables, fruits growing in a tree ...). In product families compounded of many categories such as fresh Fruits and Vegetables (F&V), the regular organic behavior seems coherent among very similar products (quasi-substitutes, e.g. lemons and oranges), but not among modestly close products (e.g. carrots and onions). The interaction effect could therefore reveal itself for search and experience attributes slightly more generic than the product level, but certainly more specific than the product family level.

There is however a non-negligible category of consistent regulars: 0.8% of consumers are organic regulars for more than 40 products, that is half of the average basket, noting that many minor products are often not available at all under the organic label. For this consumer category, the appetite for the credence attribute per se is strong ($a_{it}^{cr} \gg a_{ijl}^{se \times cr}$) in our theoretical framework). This consumer category represents 3% of all regulars for at least one product.

Figure 3. Distribution of the number of categories for which consumers are organic regulars



The number of categories is log-transformed and bound at 40 to increase the visibility of the plot.

4.1.4. 28% of the organic market is purchased by regular consumers

The share of the organic market in France purchased by organic regulars averages at 28%, which tends to validate H3. However, this average is heavily influenced by the market values of the products, which strongly differs between product families. For example, the market for organic wine is five times larger than the one of organic broccolis, but the share of the organic market purchased by regulars is only 8% for wine compared to 75% for broccolis. Thus, the size of the market and the unit values of products strongly drive down the average share of the organic market purchased by regulars. Among the detailed results, 40%, 33% and 43% of the organic eggs, organic F&V and the organic milk markets are purchased by regulars, respectively (Table 4).

Table 4. Examples of the organic market share represented by regular consumers (exhaustive results are available in SD 2)

		Share of the consumers	Share of organic market	Average frequency of organic purchases
Eggs	Conventional regulars	85%	10%	2%
	Occasional consumers	11%	50%	44%
	Organic regulars	4%	40%	92%
Milk	Conventional regulars	90%	11%	2%
	Occasional consumers	7%	46%	43%
	Organic regulars	3%	43%	92%
<i>Label rouge</i> raw ham	Conventional regulars	63%	20%	0.6%
	Occasional consumers	35%	72%	38%
	Quality regulars	2%	8%	90%

4.2. Determinants of regular organic consumer behavior

4.2.1. Product availability and product family are key determinants

Table 5. Selected results of the Logit models

	Full Logit	Full Logit Household	Full Logit Products	Restricted Logit	Restricted Logit Household	Restricted Logit Products
	(1)	(2)	(3)	(1)	(2)	(3)
Hard-discount	0.453*** (0.063)		0.427*** (0.063)	-0.097 (0.064)		-0.104 (0.064)
Hypermarket	0.557*** (0.044)		0.548*** (0.044)	0.197*** (0.046)		0.203*** (0.046)
Organic Shop	3.360*** (0.052)		3.357*** (0.052)	1.809*** (0.051)		1.813*** (0.051)
Retailer's brand	0.817*** (0.024)		0.815*** (0.024)	0.704*** (0.027)		0.705*** (0.027)
Price ratio between organic & conventional	0.018 (0.015)		0.017 (0.015)	0.181*** (0.033)		0.182*** (0.033)
Absolute price difference between organic & conventional	0.031* (0.016)		0.031* (0.016)	-0.101 (0.079)		-0.101 (0.079)
Youngest child <25 month	0.065 (0.078)	0.232*** (0.076)		0.030 (0.074)	0.094 (0.072)	
Youngest child >11 years and <15 years	-0.072 (0.071)	-0.128* (0.069)		0.015 (0.070)	-0.022 (0.068)	
Managers and superior professions	0.325*** (0.086)	0.315*** (0.085)		0.153** (0.074)	0.205*** (0.073)	
Manual professions	-0.326*** (0.077)	-0.267*** (0.077)		0.035 (0.069)	-0.052 (0.068)	
Retired	-0.241 (0.192)	-0.223 (0.195)		-0.246 (0.166)	-0.203 (0.168)	
City of 5000-9999 inhabitants	0.081 (0.119)	0.113 (0.121)		-0.012 (0.099)	0.012 (0.099)	
City of more than 200 000 inhabitants	0.330*** (0.078)	0.335*** (0.079)		0.066 (0.063)	0.158** (0.064)	
Paris agglomeration	0.191** (0.088)	0.180** (0.089)		-0.066 (0.071)	0.065 (0.070)	
Single	-0.170*** (0.053)	-0.109** (0.051)		-0.100** (0.050)	-0.130*** (0.049)	
High income household	0.109**	0.090*		-0.052	-0.027	
Low income household	-0.183 (0.172)	-0.170 (0.174)		-0.155 (0.148)	-0.111 (0.146)	
BMI	-0.198*** (0.028)	-0.216*** (0.028)		-0.020*** (0.006)	-0.031*** (0.006)	
Availability	0.335*** (0.016)		0.340*** (0.016)	0.574*** (0.069)		0.576*** (0.069)
Baby foods	0.884*** (0.266)		0.911*** (0.266)	0.610** (0.287)		0.634** (0.286)
Bread, flour	0.826***		0.826***	0.547**		0.547**

	(0.255)		(0.255)	(0.274)		(0.274)
Eggs	2.283***		2.284***	1.992***		1.998***
	(0.052)		(0.052)	(0.058)		(0.058)
Meat	0.101		0.103	0.198		0.202
	(0.128)		(0.128)	(0.129)		(0.129)
Milk	2.496***		2.496***	2.321***		2.327***
	(0.053)		(0.053)	(0.060)		(0.060)
Processed F&V	0.159		0.161	0.381		0.380
	(0.256)		(0.256)	(0.277)		(0.277)
Starchy foods	0.283		0.283	0.186		0.186
	(0.256)		(0.256)	(0.275)		(0.276)
Processed Product	0.318		0.317	0.257		0.260
	(0.250)		(0.250)	(0.269)		(0.269)
Constant	-8.753***	-6.666***	-8.276***	-3.569***	-2.223***	-3.397***
	(0.174)	(0.077)	(0.062)	(0.161)	(0.065)	(0.060)
Number of households	12453	12453	12453	8854	8854	8854
Household random effect's standard deviation	2.162	2.222	2.228	1.162	1.174	1.207
Observations	1,221,430	1,221,430	1,221,430	70,549	70,549	70,549
Log Likelihood	-47,362	-53,323	-47,493	-31,058	-33,801	-31,084
Akaike Inf. Crit.	94,858	106,714	95,071	62,251	67,670	62,255

Notes:

***, **, * Significant at the 1, 5, 10 percent levels respectively. The values between brackets refer to the standard deviation of the coefficients. This table presents a selection of the variables used in the models based on their interests for the discussion. The exhaustive results are available in SM 4.

Retailer's brand is a dummy that takes the value 1 if more than 50% of the purchases of a product were from the retailer's brand.

A household of 4 members is considered of low income if its monthly income is under 2094€ and of high income if its monthly income is above 5808€.

Dummy variables	Reference level
Shop type	Open-air market
Youngest child	No children
Working occupation	Artisan and craftsman
Size of City	<2000 inhabitants
Economic status	Superior middle income
Product family	Fresh F&V

The logit model on the whole sample (Table 5, model 1) shows that regular organic behavior is more likely to occur when a consumer makes most of its purchases for a given product at a specialized organic store and if he purchases the organic product from the retailer's brand. The price ratio between the organic and conventional versions of product is not significantly correlated with the probability of being a regular organic consumer while the absolute price difference between the two versions is positively correlated with it: a higher the price difference is associated with a higher share of organic regulars.

The typical organic regulars are single, without children and have a lower Body Mass Index (BMI), indicating that they have a healthier diet or practice more physical activities. They usually also have a high income, a managerial or superior professions or are own-account workers and live in large cities. However, consumers living in rural areas (reference level) are more likely to be organic regulars than the ones living in small cities. The logit regressions reveal that some product families are more likely to be purchased by organic regulars than others: baby foods, biscuits, cakes and pastry, bread and flour, confectionary products, desserts, eggs, fat, beverages, meal substitutes, milk, sweeteners, processed F&V and starchy food. The availability of the organic version of a product in the shops the consumers usually visit appears as a significant driver of regular behavior. When using a smaller sample, composed of occasional and regular organic consumers only, the results are similar (Table 5, model 4). Nevertheless, in these restricted logit models, the role of prices differs: when comparing occasional and regular organic consumers, a higher price ratio between organic and conventional

products is associated with a higher probability of being a regular consumer whereas the correlation with the absolute price difference becomes non-significant. In addition, in all logit models, the standard deviation of the households' random effect is high, of comparable size with the largest treatment effect (eggs family), which show an important household-related variability of the probability of being an organic regular.

4.2.2. Products' attributes matter more than households' characteristics

Two other logit models are also computed, one using as predictor variables the information related to the products only (price ratio, frequency, processed or not...) and the other using the information on the household only (number of children, wealth, professional status...). Performing a Likelihood Ratio test, we conclude that the model using product information predicts better the regular consumer behavior (Table 6). These results still stand when using the restricted sample of occasional and regular organic consumers.

Table 6. Likelihood Ratio tests between the models of product attributes and household characteristics, whole sample

Model	Numbers of Variables	LogLikelihood Value	Logit prediction's accuracy	Sensitivity	Specificity
Logit Full	60	-47,362***	98.9%	69%	99%
Logit Products	36	-47,493***	98.9%	69%	99%
Logit Household	24	-53,323***	98.7%	60%	99%
Restricted Logit Full	60	-31,058***	82.3%	72%	84%
Restricted Logit Products	36	-31,084***	82.1%	72%	84%
Restricted Logit Household	24	-33,801***	79.6%	67%	82%

***Likelihood ratio test significant at the 1 percent level.

5. Discussion

5.1. Strong interrelation between credence and search & experience attributes in determining regular organic consumption.

The first part of this paper, focusing on the bimodality of organic consumption, outlines the product families for which regular consumer behaviors can be found: hot drinks, milk, eggs, baby foods, meal substitutes, desserts (which include dairy products) and bread and flour (Table 3). The results are confirmed by the logits on regular organic behavior, in which these products families have positive regression coefficients (Table 5). Except F&V and starchy food, these product families are the most purchased ones in the organic market (Agence Bio, 2019a; Hill et al., 2002). Clearly, product availability plays a role: product families that are difficult to find organically in conventional supermarkets (meat or seafood for example) are not dominantly subject to regular consumption behavior: their share of organic purchases is unimodal. Also, the fact that some product families are compound of more products than others (for example, fresh F&V family has 71 products while milk has only two) influence the distribution of organic consumption, as shown in Figure 2. Moreover, the characteristics of organic products (price, quality, availability) likely vary between outlets. Consumers are likely to be influenced by the choice offered at the outlet where they shop most often, which may explain why consumers who mostly shop at specialized organic shops are more likely to be regulars than those shopping mostly in hard discounts. Moreover, all consumers do not have access to all outlet brands, depending on where they live. The availability indicator that we develop captures this disparity and appears as a key variable in the logit regressions on regular organic behavior (Table 5). Indeed, a lower availability of organic products compared to conventional ones is correlated with a lower organic consumption (Buder et al., 2014; Dimitri and Dettmann, 2012; Massey et al., 2018).

In the same sense, very few consumers are organic regulars for a large number of product categories, as only 6% of them purchase as regular more than 5 product categories and only 0.7% of them have a regular behavior on more than 50% of the products they purchase (Figure 3).

Following our theoretical framework, the most likely interpretation of this result is a strong interaction effect between the credence attribute (the product is organic) and the search and experience attributes (the product is an apple) in the consumer's utility. Most consumers do not value much the organic sign *per se*, but they value it strongly for a few specific products. Relating that to the dominant health driver of organic consumption (Buder et al., 2014; Padel et al., 2005), a possible explanation is that consumers are mostly concerned about the healthiness of a few specific products, e.g. because they have seen a documentary on the amount of pesticides in lemons or read a newspaper article on the amount of antibiotics in milk. Accordingly, they may become organic regulars for these specific products, for which they received health information, but this change may not spillover on their broader feeding routines.

However, at least two other reasons can explain this weak consistency of regular behavior across products. The first is again availability, as some products can only be found in specialized shops, but most of the consumers purchase food in several types of shops (conventional supermarket, open air market, specialized shops...). As the logit results show, if a product is mainly purchased at a specialized shop, the probability that the consumers will follow a regular behavior increases drastically. So, as consumers shop in different outlets, even if they choose to be regulars on some products that they usually purchase at their main shopping source, they may have to buy these products in other shops (supplemental purchases, oblivion, unexpected meal to prepare...) but the organic versions of these products may be unavailable, more expensive or of a lesser quality and they will not purchase it. This interpretation is supported by the lower numbers of regulars found when the analysis is based on frequency (number of purchases) rather than volumes (liters or kilograms). Indeed, supplemental purchases weigh more on the shares of organic purchases when computed with frequencies. For the F&V family for example, we find that more households following a regular organic behavior for at least one product when the shares are computed with quantities (1548 households) than in frequencies (1303 households). Similarly, the number of F&V categories purchased by regular organic households is larger when they are computed in quantities (3317) than frequencies (2614).

The second reason is that the data we use is collected at the household level, and even if the main shopper is identified, there can be another shopper (husband/wife, teenagers...) that also sometimes purchases food. This occasional shopper may not have the same purchasing behavior or may not go to the same shopping source, and thus the household may not be regular even though each of its members are regulars of different styles. This effect seems weak however: couples are more likely to be organic regulars than singles (Table 5).

Finally, consumers are more likely to be organic regulars if they purchase products from retailer brands, which are generally cheaper than other organic brands (Ngobo, 2011). However, the price ratio between the organic and conventional version of a product is not significantly correlated with regular behavior, only the absolute difference in prices is significantly, and positively, associated with regular organic consumption. This surprising result—i.e., the more expensive the organic version of a product is, compared to the conventional version, the more likely the consumer will be an organic regular—can be understood as a weak price-elasticity of regular consumers to organic products prices. Indeed, if a consumer is an organic regular, most likely for health consideration, he will not be as reactive to the relative price of organic products than occasional or conventional consumers. This result is strengthened by the restricted logit models, in which the regression coefficient of the relative price of organic is positive and significant.

5.2. The characteristics of actual organic regulars are similar to those of “declared” pro-organic consumers

The description of regular organic consumers which can be drawn from the logit models is comparable with the main findings of the literature on the socio-economic characteristics of organic consumers (Kesse-Guyot et al., 2013). Indeed, we corroborate that regular organic consumers have a higher income, are more urban and have a higher profession position than non-regular organic consumers (Table 5). Comparing regular consumers with occasional consumers only, we uncover that regular consumers are more represented in the upper-middle class and high professional status (Table 5).

Nevertheless, the comparison of the logit models with product attributes to the logit models with household characteristics only demonstrates that product attributes explain better regular organic behavior than the household characteristics (Table 6).

5.3. Public Policy implications: product-specific targets

Of course, what matters most to policy makers is the total amount of organic production produced and sold, relatively to conventional production and the negative health and environmental externalities associated with it. However, an increase of organic production can be reached by sustaining organic consumption, which is currently done by French policies (“Ambition Bio 2022”, “Egalim” law in 2018). Our results from the analysis of organic consumption and their regular behavior advocate for more product-specific policies. Indeed, as much as 28% of organic value is purchased by households with a regular consumption pattern for a few specific products. Accordingly, public initiatives aiming at increasing organic consumption should focus both on promoting the organic label *per se* and on promoting specific organic products. They may also promote a “regular” attitude towards some organic products—e.g. “remove pesticides from your morning orange juice: buy organic oranges”. Targeting product families for which a regular consumption behavior is already frequent may be particularly promising, as regular organic consumption exists (H1 is validated) but is not yet very consistent across the food basket (H2 is invalidated). This is supported by our finding that product characteristics explain better the regular organic behavior than the household characteristics (Table 6). If there is a choice to be made between targeting specific products or specific consumers, one should go for specific products although some consumer segments—higher income, higher professional status and more urban—are likely to be more receptive (Apostolidis et al., 2016).

Geographical indications and *label rouge* products on the other hand are purchased occasionally, making them somewhat comparable to luxury products. Marketing and promotion actions for these products may therefore want to get inspiration from non-food luxury products.

6. Conclusion

This paper demonstrates that quality-food with similar market shares can be subject to widely different consumption behaviors. Consumption of geographical indications and *label rouge* is always occasional while consumer attitude towards organic food is often regular: for a given product, some consumers tend to either purchase it always organic or always conventional. Indeed, conversely to previous studies using scanner data or surveys, we use observable information, the distribution of quality purchases, to categorize regular and occasional quality consumers. Doing so leads to a better comprehension of quality consumers' behavior, i.e. that an important part (29% of our sample) is regular for at least one organic product, but that regular organic consumers for their whole food basket are scarce. Moreover, regular organic consumers purchase 28% of the total value of the organic market and up to 50% for some F&V, eggs or milk. In this sense, public policies should target product categories instead of the whole food basket and develop organic regular behavior, using as levers the product attributes and household characteristics our analysis revealed as strongly related with regular organic consumption. However, other quality products, such as Geographical Indications, are purchased as extraordinary goods and so policies aiming at increasing their consumption should relate to those on luxury goods.

More precisely, we illustrate that the main product families for which organic consumption distributions are bimodal, i.e. for which regular consumer behaviors can be found, are eggs, milk, baby foods, meal substitutes, desserts, bread and flour. When organic consumption is analyzed at product level, one sees that its distribution is bimodal also for most fruits and vegetables (processed or raw).

Besides, households which exhibit regular organic consuming behavior are richer, more urban, have a higher profession statute than the others. These organic regulars also have a higher propensity to be in couples and to have fewer children.

The products categories that are consumed by organic regulars are more available and organic regulars do not seem to be influenced by the price of organic products. To the contrary, a higher price difference between an organic product and its conventional counterpart is associated with a higher share of organic regulars, possibly because organic regulars are willing to buy the organic alternative at all costs. Moreover, product attributes explain better organic regular behavior than household characteristics. Further research may explore the interaction effects in the consumer utility function between the organic attribute and other product attributes (thickness of skin for fruits, roots vegetables...). Computing price and income elasticities of quality-food products would also shed light on consumer behavior, especially in defining which products are luxury goods. Lastly, bridging the gap between organic production and consumption may help designing efficient policies: in light of our work, policies could focus on securing the regular consumption of products whose production in organic systems has high environmental or health value-added (production nearby water catchments for example).

References

- Agence Bio. (2019a). Baromètre 2018 de consommation et de perception des produits biologiques en France. Retrieved from <https://www.agencebio.org/wp-content/uploads/2019/02/AgenceBio-DossierdePresse-Barometre2019.pdf>
- Agence Bio. (2019b). Les chiffres 2018 du secteur bio. Retrieved from https://www.agencebio.org/wp-content/uploads/2019/06/DP-AGENCE_BIO-4JUIN2019.pdf
- Apostolidis, C., & McLeay, F. (2016). Should we stop meating like this? Reducing meat consumption through substitution. *Food Policy*, *65*, 74–89. <https://doi.org/10.1016/j.foodpol.2016.11.002>
- Barrena, R., & Sánchez, M. (2010). Frequency of consumption and changing determinants of purchase decision: from attributes to values in the organic food market. *Spanish Journal of Agricultural Research*, *8*(2), 251. <https://doi.org/10.5424/sjar/2010082-1178>
- Baudry, J., Pointereau, P., Seconda, L., Vidal, R., Taupier-Letage, B., Langevin, B., ... Kesse-Guyot, E. (2019). Improvement of diet sustainability with increased level of organic food in the diet: Findings from the BioNutriNet cohort. *American Journal of Clinical Nutrition*, *109*(4), 1173–1188. <https://doi.org/10.1093/ajcn/nqy361>
- Berry, B. Y. S., Levinsohn, J., & Pakes, A. (1995). Automobile Prices in Market Equilibrium. *Econometrica*, *63*(4), 841–890.
- Boizot-Szantai, C., Hamza, O., & Soler, L. G. (2017). Organic consumption and diet choice: An analysis based on food purchase data in France. *Appetite*, *117*, 17–28. <https://doi.org/10.1016/j.appet.2017.06.003>
- Buder, F., Feldmann, C., & Hamm, U. (2014). Why regular buyers of organic food still buy many conventional products: Product-specific purchase barriers for organic food consumers. *British Food Journal*, *116*(3), 390–404. <https://doi.org/10.1108/BFJ-04-2012-0087>
- Chekima, B., Oswald, A. I., Wafa, S. A. W. S. K., & Chekima, K. (2017). Narrowing the gap: Factors driving organic food consumption. *Journal of Cleaner Production*, *166*, 1438–1447. <https://doi.org/10.1016/j.jclepro.2017.08.086>
- Dimitri, C., & Dettmann, R. L. (2012). Organic food consumers: What do we really know about them? *British Food Journal*, *114*(8), 1157–1183. <https://doi.org/10.1108/00070701211252101>
- Gracia, A., & de Magistris, T. (2008). The demand for organic foods in the South of Italy: A discrete choice model. *Food Policy*, *33*(5), 386–396. <https://doi.org/10.1016/j.foodpol.2007.12.002>
- Hill, H., & Lynchehaun, F. (2002). Organic milk: Attitudes and consumption patterns. *British Food Journal*, *104*(7), 526–542. <https://doi.org/10.1108/00070700210434570>
- Hughner, R. S., McDonagh, P., Prothero, A., Shultz II, C. J., & Stanton, J. (2007). Who are organic food consumers? A compilation and review of why people purchase organic food. *Journal of Consumer Behaviour*.
- INAO. (2019). Les produits sous signe d'identification de la qualité et de l'origine. Chiffres-Clés 2018. Retrieved from https://www.inao.gouv.fr/content/download/3091/28236/version/1/file/INAO_CH_CLE_2019_BD.pdf
- Kesse-Guyot, E., Péneau, S., Méjean, C., Szabo de Edelenyi, F., Galan, P., Hercberg, S., & Lairon, D. (2013). Profiles of Organic Food Consumers in a Large Sample of French Adults: Results from the Nutrinet-Santé Cohort Study. *PLoS ONE*, *8*(10). <https://doi.org/10.1371/journal.pone.0076998>
- Lacour, C., Seconda, L., Allès, B., Hercberg, S., Langevin, B., Pointereau, P., ... Kesse-Guyot, E. (2018). Environmental Impacts of Plant-Based Diets: How Does Organic Food Consumption Contribute to Environmental Sustainability? *Frontiers in Nutrition*, *5*(February), 1–13. <https://doi.org/10.3389/fnut.2018.00008>
- Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*(2), 132–157.
- Marian, L., Chrysochou, P., Krystallis, A., & Thøgersen, J. (2014). The role of price as a product attribute in the organic food context: An exploration based on actual purchase data. *Food Quality and Preference*, *37*, 52–60. <https://doi.org/10.1016/j.foodqual.2014.05.001>

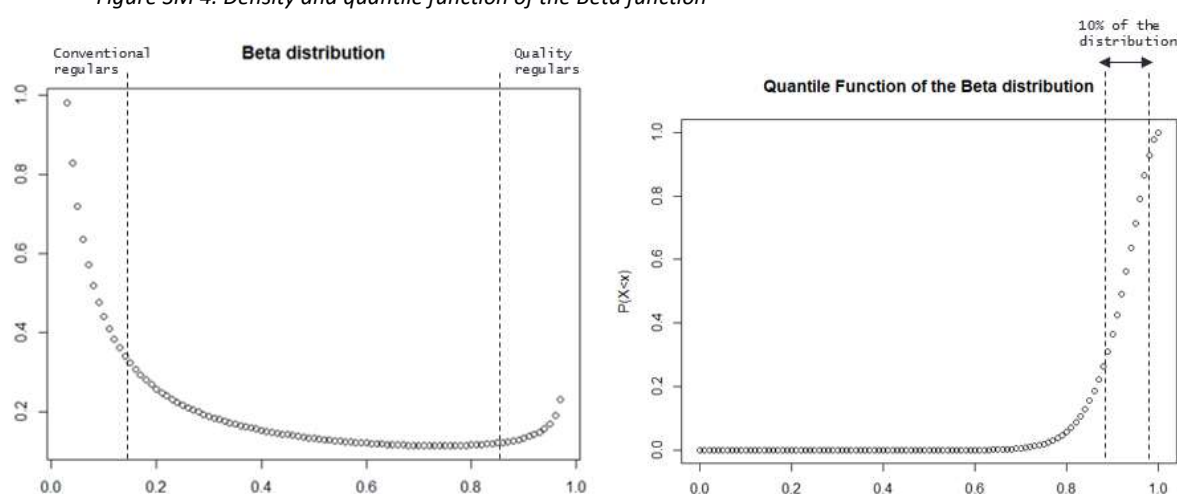
- Massey, M., O’Cass, A., & Otahal, P. (2018). A meta-analytic study of the factors driving the purchase of organic food. *Appetite*, *125*, 418–427. <https://doi.org/10.1016/j.appet.2018.02.029>
- McFadden, D. (2001). Economic Choices. *American Economic Review*, *91*(3), 351–378.
- Monier, S., Hassan, D., Nichle, V., & Simioni, M. (2009). Organic food consumption patterns. *Journal of Agricultural and Food Industrial Organization*, *7*(2). <https://doi.org/10.2202/1542-0485.1269>
- Ngobo, P. V. (2011). What Drives Household Choice of Organic Products in Grocery Stores? *Journal of Retailing*, *87*(1), 90–100. <https://doi.org/10.1016/j.jretai.2010.08.001>
- Oates, L., Cohen, M., & Braun, L. (2012). Characteristics and consumption patterns of Australian organic consumers. *Journal of the Science of Food and Agriculture*, *92*(14), 2782–2787. <https://doi.org/10.1002/jsfa.5664>
- Onyango, B. M., Hallman, W. K., & Bellows, A. C. (2007). Purchasing organic food in US food systems: A study of attitudes and practice. *British Food Journal*, *109*(5), 399–411. <https://doi.org/10.1108/00070700710746803>
- Padel, S., & Foster, C. (2005). Exploring the gap between attitudes and behaviour: Understanding why consumers buy or do not buy organic food. *British Food Journal*, *107*(8), 606–625. <https://doi.org/10.1108/00070700510611002>
- Padilla Bravo, C., Cordts, A., Schulze, B., & Spiller, A. (2013). Assessing determinants of organic food consumption using data from the German National Nutrition Survey II. *Food Quality and Preference*, *28*(1), 60–70. <https://doi.org/10.1016/j.foodqual.2012.08.010>
- Pearson, D., Henryks, J., Sultan, P., & Anisimova, T. (2013). Organic food : Exploring purchase frequency to explain consumer behaviour. *Journal of Organic Systems*, *8*(2), 50–63. Retrieved from <http://www.organic-systems.org/journal/82/8206.pdf>
- Rana, J., & Paul, J. (2017). Consumer behavior and purchase intention for organic food: A review and research agenda. *Journal of Retailing and Consumer Services*, *38*(February), 157–165. <https://doi.org/10.1016/j.jretconser.2017.06.004>
- Rödiger, M., & Hamm, U. (2015). How are organic food prices affecting consumer behaviour? A review. *Food Quality and Preference*, *43*, 10–20. <https://doi.org/10.1016/j.foodqual.2015.02.002>
- Silverman, B. W. (1981). Using Kernel Density Estimates to Investigate Multimodality. *Journal of the Royal Statistical Society. Series B: Methodological*, *43*(1), 97–99.
- Sun, B., & Morwitz, V. G. (2010). Stated intentions and purchase behavior: A unified model. *International Journal of Research in Marketing*, *27*(4), 356–366. <https://doi.org/10.1016/j.ijresmar.2010.06.001>
- Train, K. E. (2003). Discrete choice methods with simulation. *Cambridge University Press*, *9780521816*, 1–334. <https://doi.org/10.1017/CBO9780511753930>
- Treu, H., Nordborg, M., Cederberg, C., Heuer, T., Claupein, E., Hoffmann, H., & Berndes, G. (2017). Carbon footprints and land use of conventional and organic diets in Germany. *Journal of Cleaner Production*, *161*, 127–142. <https://doi.org/10.1016/j.jclepro.2017.05.041>

Supplementary Materials

SM 1. Statistical determination of the regular behavior thresholds

In our empirical analysis of regular consumption behavior, a key step is to determine the thresholds (dash lines in Figure SM 4) under and above which a consumer can be categorized as conventional or quality regular, respectively. To do so, for each product's organic distribution, we fit a beta distribution, the only statistically function with a known density and cumulative distribution function that can easily account for bimodality. Using the quantile function of the fitted beta distribution, we assess the values of each pair of centiles which a 0.1 (10%) difference. In this sense, we evaluate two values, between which only 10% of the theoretical beta distribution lies. To assure that these two values are symmetrically delimiting the two regular behaviors (conventional and quality regularity): noting $threshold_1$ the threshold for the conventional regularity, and $threshold_2$ for the organic regularity, we set the constraint as $threshold_1 = 1 - threshold_2$.

Figure SM 4. Density and quantile function of the Beta function



We compute these thresholds values for 5 key products and use 20% and 80% of quality consumption as the thresholds of conventional regularity and quality regularity respectively (Table SM 7).

Table SM 7. Estimated thresholds of conventional and quality regularities

Product	Threshold conventional regularity	Threshold quality regularity
Milk	0.14	0.86
Egg	0.21	0.79
Oil	0.24	0.76
Lemon	0.20	0.80
Baby Food	0.17	0.83

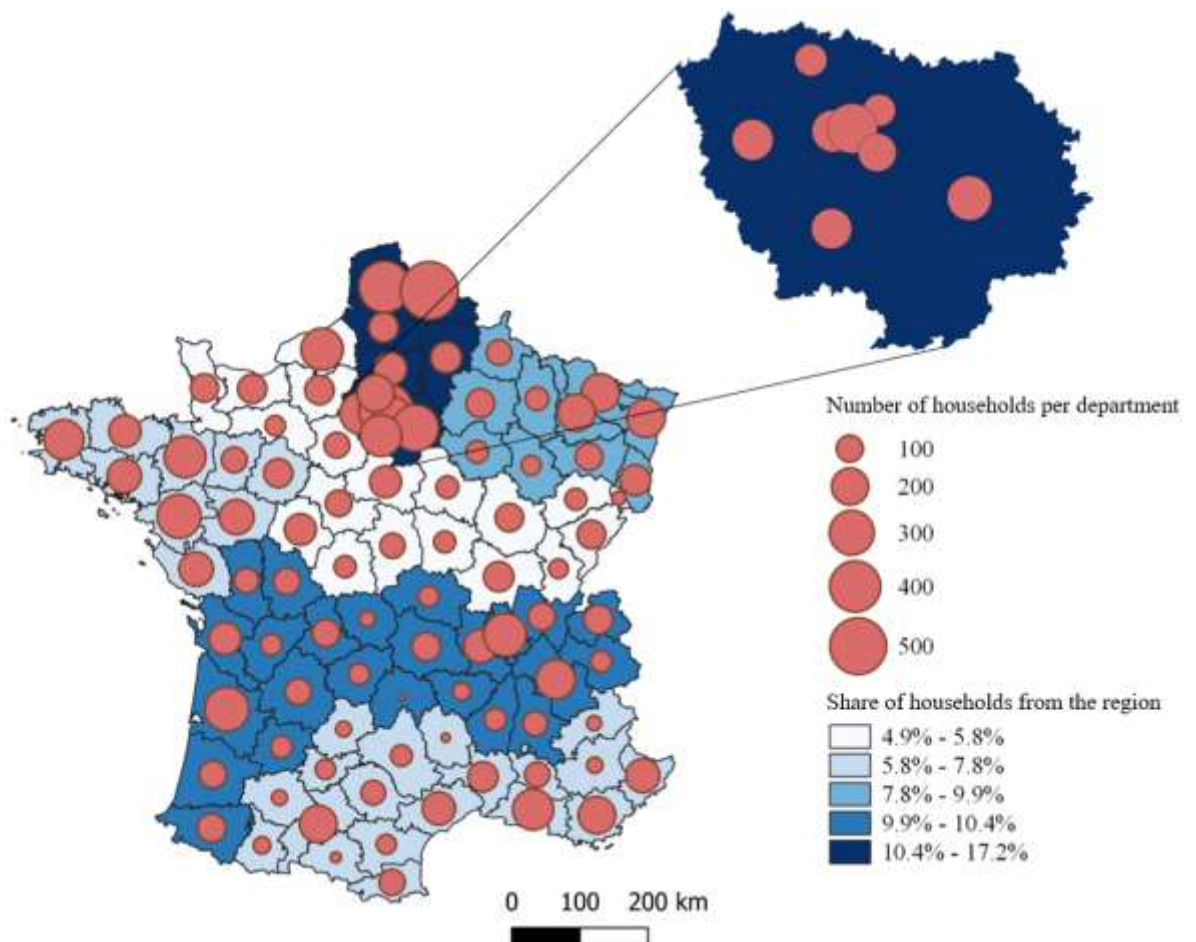
SM 2. Identification of the variables influencing availability

To construct the indicator of organic products' availability, we collected data on which organic products were available in 10 shops of each of the 5 biggest supermarket brands in France in a region (Burgundy) based on their websites. The shops are of different sizes (supermarkets, hypermarkets), are located in 4 different "départements" and in different city sizes. However, they all offer the possibility to purchase food online and pick it up at the shop ("Drive in"). We then run LASSO regression using the number of organic products available in each shop as the dependent variables. We find that supermarkets' brands and size are the only variables that influence the availability of organic products once penalized. When all the variables are included, an ordinary least squares regression explains 86% of the variance of the number of organic products available and when only the supermarkets' brands and size are selected, the R^2 still amounts to 68%.

Table SM 8. LASSO regression results on the availability of organic products

Variables	Penalized coefficients
Département1	0.00
Département2	0.00
Département3	0.00
Brand1	449.9
Brand2	-201.2
Brand3	0.00
Brand4	0.00
City Inhabitants : 20k-50k	0.00
City Inhabitants : 50k-100k	0.00
City Inhabitants : 5k-10k	0.00
City Inhabitants : 10k-20k	0.00
City Inhabitants : Smaller than 2k	0.00
City Inhabitants : 2k-5k	0.00
Shop Size : Small	-66.6

SM 3. Map of the spatial distribution of the sample



SM 4. Detailed results of the logit

	Full Logit	Full Logit Household	Full Logit Products	Restricted Logit	Restricted Logit Household	Restricted Logit Products
	(1)	(2)	(3)	(1)	(2)	(3)
Year 2012	0.119*** (0.042)	0.167*** (0.040)	0.116*** (0.042)	0.085* (0.044)	0.097** (0.042)	0.080* (0.044)
Year 2013	0.233*** (0.041)	0.299*** (0.039)	0.225*** (0.041)	0.117*** (0.044)	0.146*** (0.042)	0.107** (0.044)
Year 2014	0.281*** (0.041)	0.350*** (0.038)	0.276*** (0.040)	0.188*** (0.043)	0.220*** (0.041)	0.179*** (0.043)
Year 2015	0.360*** (0.041)	0.425*** (0.039)	0.354*** (0.041)	0.193*** (0.043)	0.235*** (0.041)	0.184*** (0.043)
Year 2016	0.450*** (0.040)	0.585*** (0.038)	0.440*** (0.040)	0.197*** (0.043)	0.274*** (0.040)	0.186*** (0.042)
Number of purchases	-0.248*** (0.019)	-0.250*** (0.016)	-0.247*** (0.019)	-0.008*** (0.001)	-0.009*** (0.001)	-0.008*** (0.001)
Hard-discount	0.453*** (0.063)		0.427*** (0.063)	-0.097 (0.064)		-0.104 (0.064)
Home delivery	0.323*** (0.119)		0.336*** (0.118)	0.012 (0.123)		0.017 (0.123)
HyperMarket	0.557*** (0.044)		0.548*** (0.044)	0.197*** (0.046)		0.203*** (0.046)
Organic Shop	3.360*** (0.052)		3.357*** (0.052)	1.809*** (0.051)		1.813*** (0.051)
Small Supermarket	0.507*** (0.060)		0.501*** (0.060)	0.201*** (0.063)		0.198*** (0.063)
SuperMarket	0.527*** (0.040)		0.522*** (0.040)	0.218*** (0.042)		0.216*** (0.042)
Retailer brand	0.817*** (0.024)		0.815*** (0.024)	0.704*** (0.027)		0.705*** (0.027)
Price ratio between organic & conventional	0.018 (0.015)		0.017 (0.015)	0.181*** (0.033)		0.182*** (0.033)
Average share of organic purchases	0.540*** (0.008)	0.525*** (0.005)	0.539*** (0.008)	0.054*** (0.003)	0.033*** (0.002)	0.053*** (0.003)
Number of times purchased	-0.647*** (0.015)	-0.215*** (0.009)	-0.648*** (0.015)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Youngest child <25 month	0.065 (0.078)	0.232*** (0.076)		0.030 (0.074)	0.094 (0.072)	
Youngest child >25 months and <5 years	-0.062 (0.075)	0.035 (0.074)		-0.028 (0.070)	0.014 (0.069)	
Youngest child >6 years and <10 years	-0.077 (0.074)	-0.053 (0.072)		0.000 (0.069)	0.012 (0.068)	
Youngest child >11 years and <15 years	-0.072 (0.071)	-0.128* (0.069)		0.015 (0.070)	-0.022 (0.068)	
Farmers	-0.648* (0.346)	-0.604* (0.355)		-0.343 (0.302)	-0.386 (0.299)	

Managers and superior professions	0.325*** (0.086)	0.315*** (0.085)	0.153** (0.074)	0.205*** (0.073)
Employees	-0.063 (0.069)	-0.012 (0.068)	0.034 (0.060)	0.027 (0.060)
Manual occupations	-0.326*** (0.077)	-0.267*** (0.077)	0.035 (0.069)	-0.052 (0.068)
Intermediate professions	0.039 (0.074)	0.105 (0.075)	0.107* (0.064)	0.145** (0.064)
Retired	-0.241 (0.192)	-0.223 (0.195)	-0.246 (0.166)	-0.203 (0.168)
Unemployed, student	-0.215 (0.192)	-0.159 (0.186)	-0.083 (0.170)	-0.064 (0.167)
City of 2000-4999 inhabitants	0.151 (0.112)	0.192* (0.113)	-0.048 (0.093)	0.012 (0.094)
City of 5000-9999 inhabitants	0.081 (0.119)	0.113 (0.121)	-0.012 (0.099)	0.012 (0.099)
City of 10 000-19 999 inhabitants	0.284** (0.121)	0.334*** (0.121)	0.110 (0.098)	0.181* (0.098)
City of 20 000-49 999 inhabitants	0.230** (0.116)	0.238** (0.117)	0.130 (0.093)	0.193** (0.093)
City of 50 000-99 999 inhabitants	0.036 (0.115)	0.060 (0.116)	0.013 (0.094)	0.053 (0.094)
City of 100 000-199 999 inhabitants	0.339** (0.134)	0.356*** (0.137)	0.068 (0.106)	0.180* (0.107)
City of more than 200 000 inhabitants	0.330** (0.078)	0.335*** (0.079)	0.066 (0.063)	0.158** (0.064)
Paris agglomeration	0.191** (0.088)	0.180** (0.089)	-0.066 (0.071)	0.065 (0.070)
Single	-0.170*** (0.053)	-0.109** (0.051)	-0.100** (0.050)	-0.130*** (0.049)
High income household	0.109** (0.050)	0.090* (0.049)	-0.052 (0.044)	-0.027 (0.044)
Inferior middle income household	-0.140* (0.085)	-0.115 (0.085)	-0.031 (0.075)	-0.021 (0.074)
Low income household	-0.206 (0.170)	-0.192 (0.173)	-0.163 (0.148)	-0.136 (0.149)
BMI	-0.198*** (0.028)	-0.216*** (0.028)	-0.020*** (0.006)	-0.031*** (0.006)
Absolute price difference	0.031* (0.016)		0.031* (0.016)	-0.101 (0.079)
Availability	0.335*** (0.016)		0.340*** (0.016)	0.574*** (0.069)
Alcoholic beverages	-0.628** (0.273)		-0.629** (0.273)	-0.233 (0.293)
Appetizers	-0.615** (0.240)		-0.613** (0.240)	-0.281 (0.258)
Baby foods	0.884*** (0.266)		0.911*** (0.266)	0.610** (0.287)
Biscuits, cakes and pastry	0.592** (0.255)		0.591** (0.255)	0.612** (0.274)
Bread, flour	0.826*** (0.255)		0.826*** (0.255)	0.547** (0.275)

	(0.255)		(0.255)	(0.274)		(0.274)
Cheese	-0.138		-0.140	-0.069		-0.076
	(0.268)		(0.268)	(0.289)		(0.289)
Confectionary products	0.618**		0.617**	0.677**		0.677**
	(0.257)		(0.257)	(0.277)		(0.277)
Culinary ingredients	0.184		0.184	0.486*		0.486*
	(0.253)		(0.253)	(0.273)		(0.273)
Desserts	0.674***		0.672***	0.615**		0.615**
	(0.257)		(0.258)	(0.277)		(0.277)
Eggs	2.283***		2.284***	1.992***		1.998***
	(0.052)		(0.052)	(0.058)		(0.058)
Fat	1.078***		1.077***	0.941***		0.941***
	(0.253)		(0.253)	(0.272)		(0.272)
Hot drinks	2.345***		2.325***	15.389		15.302
	(0.545)		(0.545)	(105.970)		(106.791)
Meal	0.670**		0.669***	0.931***		0.933***
	(0.256)		(0.256)	(0.276)		(0.277)
Meal substitutes	0.165		0.163	1.712***		1.705***
	(0.411)		(0.411)	(0.589)		(0.588)
Meat	0.101		0.103	0.198		0.202
	(0.128)		(0.128)	(0.129)		(0.129)
Milk	2.496***		2.496***	2.321***		2.327***
	(0.053)		(0.053)	(0.060)		(0.060)
Non Alcoholic beverages	0.648**		0.645**	0.573**		0.571**
	(0.259)		(0.259)	(0.279)		(0.279)
Processed F&V	0.159		0.161	0.381		0.380
	(0.256)		(0.256)	(0.277)		(0.277)
Processed Meat	0.164		0.167	0.372		0.366
	(0.259)		(0.259)	(0.280)		(0.280)
Seafood	-0.837***		-0.834***	-0.358		-0.376
	(0.282)		(0.282)	(0.328)		(0.328)
Starchy foods	0.283		0.283	0.186		0.186
	(0.256)		(0.256)	(0.275)		(0.276)
Sweeteners	0.491		0.162	0.474		0.228
	(0.621)		(0.259)	(0.578)		(0.303)
Processed Product	0.318		0.317	0.257		0.260
	(0.250)		(0.250)	(0.269)		(0.269)
Constant	-8.406***	-6.662***	-8.319***	-2.567***	-0.972***	-3.059***
	(0.090)	(0.077)	(0.061)	(0.178)	(0.161)	(0.089)
Number of households	12453	12453	12453	8854	8854	8854
Household random effect's standard deviation	2.162	2.222	2.228	1.162	1.174	1.207
Observations	1,221,430	1,221,430	1,221,430	70,549	70,549	70,549
Log Likelihood	-47,361.98	-53,322.82	-47,492.58	-31,058.4	-33,801.1	-31,084.5
Akaike Inf. Crit.	94,857.95	106,713.6	95,071.16	62,250.99	67,670.07	62,254.9

Notes:

***, **, * Significant at the 1, 5, 10 percent levels respectively. The values between brackets refer to the standard deviation of the coefficients.

Retailer's brand is a dummy that takes the value 1 if more than 50% of the purchases of a product were from the retailer's brand.

A household of 4 members is considered of low income if its monthly income is under 2094€ and of high income if its monthly income is above 5808€.

Dummy variables	Reference level
Shop type	Open-air market
Youngest child	No children
Working occupation	Artisan and craftsman
Size of City	<500 inhabitants
Economic status	Superior middle income
Product family	Fresh F&V