

# Market stealing and market expansion: an examination of product introductions in the organic coffee market

Richard Friberg<sup>1</sup> · Mark Sanctuary<sup>2</sup>

Received: 29 August 2016 / Accepted: 18 August 2017 / Published online: 24 August 2017  
© The Author(s) 2017. This article is an open access publication

**Abstract** Market shares for organic products remain modest despite broad professed consumer interest in purchasing organic products. We use counterfactual product introductions in an estimated demand model to systematically explore the links between product assortment and the market share for organic products. The demand estimates on which we base the counterfactuals use 3 years of household panel data on retail coffee purchases in Sweden, combining household's stated and revealed behavior in a discrete choice model. The predicted market shares of new organic products are highly dependent on which brand that the organic label partners with. Introduction of a new organic product is predicted to increase in-sample organic market share from around 5% up to a maximum of around 8%. The market expansion effect dominates the market stealing effect for almost all entrants. The largest market share gains for organic products are not to be had amongst the keenest organic households, but rather amongst the moderately keen organic households.

**Keywords** Co-branding · Organic · Coffee · Market stealing · Market expansion

**JEL Classification** D12 · D40 · Q13

---

✉ Mark Sanctuary  
mark.sanctuary@hhs.se

<sup>1</sup> Stockholm School of Economics, Stockholm, Sweden

<sup>2</sup> CEPR, London, UK

## 1 Introduction

Many people claim to be willing to buy environmentally friendly products. For instance, Eurobarometer (2011, p 76) reported that 72% of respondents are ready to buy environmentally friendly products even if they cost more.<sup>1</sup> Despite stated intentions, the market share for products that carry an organic label remains relatively small. For example, the estimated market share of organic coffee in Germany and Italy is 3 and 0.5%, respectively.<sup>2</sup>

One potential reason for the discrepancy is that consumers find that the organic assortment provides a weak match with their preferences in other dimensions. A consumer may, for instance, value an organic label but also value a particular brand, and if that brand does not offer an organic variety then consumers will choose non-organic varieties even if they are keen to buy organic: the brand trumps the organic characteristic.

In this article, we provide a systematic examination of the role brand-organic partnerships play in determining the market share of organic products. New product introductions will partly crowd out sales of incumbent organic products (“market stealing”) and partly expand the organic market share (“market expansion”). We use a structural model of demand to examine the market stealing and market expanding effects of new product introductions in the organic coffee segment. We are also interested in identifying the consumers from which a market expanding effect would come from. Is it the most avid organic shoppers or consumers with a more moderate organic profile?

We study this issue using the retail coffee market in Sweden. We use a consumer-scan-panel of Swedish households’ coffee purchases and observe the coffee varieties that households buy at the bar-code level as well as a number of demographic variables. We focus on coffee purchases because it is a market segment with a large number of differentiated products where brands are easy to identify, and its importance as a commodity. Once a year participating households fill in a questionnaire and answer whether they, to the extent feasible, try to buy organic products when shopping. We are thus able to relate a measure of the preferences regarding organic products to the same household’s actual shopping choices. The source of data is the market research firm GfK.

To systematically investigate coffee choice, we estimate a discrete choice, conditional logit, and model of demand. We establish that household willingness to pay is in line with their survey responses: households that said they try buying organic products have—as demonstrated through the shopping choices they made in the market—higher choice probabilities for organic coffee products. We use the demand estimates to evaluate counterfactual product introductions: we introduce a synthetic organic product to all shopping trip choice sets, and use this synthetic product to explore all possible brand-organic alliances. For each counterfactual setting, we brand the synthetic product with one of the 35 brands observed in the

---

<sup>1</sup> This result was based on face-to-face interviews with more than 26,000 respondents across the 27 EU member states.

<sup>2</sup> Tropical Commodity Coalition (2012).

data and then predict the resulting counterfactual organic market shares for each brand.

The counterfactual experiments indicate that the effects on organic market shares from introducing new synthetic organic products depends crucially on the brand that new products enter with. The organic market share increases by up to a maximum of around 20–60% among a handful of the brands. However, organic market shares increase by a negligible amount in most brand-organic instances. The actual organic market share observed in the sample is around 5%. Choosing the right brand partner could boost the market share of organic coffee from 5% to a 6–8% at most, illustrating the potential importance of the choice of choosing the right brand to partner with for an organic label that wants to increase organic market shares.

Access to the stated shopping behavior regarding organic products also allows us to explore the role of different types of consumers for expansion of the organic segment. The counterfactuals show that for the keenest organic shoppers, introducing a synthetic organic product will have relatively little effect on organic market shares; the largest predicted increase in market share is around 3 percentage points (which can be related to the in-sample market share of 45% for this group). Among these self-professed organic shoppers the introduction of a new synthetic organic product results in relatively few shoppers switching from non-organic to organic coffee. The successful new organic products largely steal market share from other organic products in this consumer segment.

In contrast, amongst moderately keen organic consumers, the introduction of the synthetic organic product increases organic market shares by up to a maximum of 8 percentage points. Set in relation to the in-sample market share of 13% for these consumers this is a large market expanding effect. These consumers value the organic label, but they also value their brand. From the point of view of the organic label, it is in this niche that there is a large potential for market share gains.

Finally, our demand estimates show that less keen organic shoppers put a negative value on the organic label and here the product introductions are associated with virtually no market stealing and only little market expansion.

Taken together, these results suggest that the best co-branding partner, from the organic marketer's point of view, is determined by the moderately keen organic households, and not the keenest organic households. It is these consumers where the co-brand partnership can have the largest impact on inducing consumers to change their choice from non-organic to organic.

We relate to several literatures. The demand for organic products has been examined empirically by a number of researchers. Van Doorn and Verhoef (2015) use, as we do, revealed preference data to investigate the supply side factors and consumer characteristics to identify the barriers and drivers of organic purchases. They find that organic products are less popular in vice categories, categories with high promotional intensity but more popular in fresh as opposed to processed categories. On the consumer side, they find that environmental and animal welfare concerns increase organic purchases. In a related study, Ngobo (2011) seeks to identify the determinants of organic purchasing, although his data does not include consumer attitudes. Griffith and Nesheim (2010) combine consumer-scan-panel and survey responses using data to elicit bounds on willingness to pay for organic

produce. The environmental economics literature has also studied demand for organic products. For example, Brooks and Lusk (2010) gauge the social-welfare benefits of organic labeling initiatives; and Bjørner et al. (2004) study the willingness to pay for eco-labeled toilet paper, paper towels and laundry detergents.

We also relate to articles that try to determine the extent of market stealing and market expansion of new products. For example, Berry and Waldfogel (1999) use a structural model to examine the welfare effects of entry in the US radio market. Methodologically, the current article is close to that work, even if the questions posed are distinct. We are not aware of any previous studies that apply, as we do, supply side counterfactual exercises to examine the interaction between branding and organic labeling.<sup>3</sup>

One lens through which to view our results is as an examination of interaction between brands (co-branding or brand alliances), see van der Lans et al. (2014) or Cunha et al. (2015). A central question in this literature is the optimal choice of co-branding partner. For example, Geylani et al. (2008) study a setting in which brands face uncertain consumer beliefs. They suggest that it is not necessarily optimal for a brand to partner with a brand that is strong on the attribute of interest but rather to partner with a moderately strong brand. This shares some of the flavor in our study that the main market share gains are to be had among the consumers with a self-professed moderate tendency to shop organic products.

## 2 The data

We use data collected by GfK, a German-based market-research consultant with an affiliate in Sweden. GfK has assembled a consumer-scan-panel that follows grocery shopping choices of 3000 households across Sweden. The data was collected with an electronic scanner and web-based diary entries. We use observations on each household shopping trip from January 2007 to January 2010. Not all participating households buy coffee. The dataset that we use consists of an unbalanced panel of 2782 households.

The participating households are chosen as a representative sample of the Swedish population, but were sampled using non-probabilistic methods typical for this type of market research data.<sup>4</sup> We observe household characteristics such as the age and level of education of the reference shopper and household annual income. Panel A in Table 1 compares the household characteristics of the sample with national averages in Sweden. There are only small differences compared to the national averages. In sample average annual income is 371,970 SEK, which is higher than the national average of 350,300 SEK during the period. The average reference shopper is slightly older than the average age of the population: 50.6

<sup>3</sup> The application of structural models to predict market shares under various counterfactuals is not new; however, see for instance Hausman (1997) or Petrin (2002), for ex-post examinations of the welfare effects of new product introductions (breakfast cereal and the Chrysler minivan, respectively).

<sup>4</sup> Lusk and Brooks (2010) provide a critical evaluation of the representativeness of such samples, finding that participants in two US household scanning surveys were slightly more price sensitive than a random sample.

**Table 1** Summary statistics—households and their purchasing behavior

| Variable  | Mean    | SD      | National mean     |
|---|---------|---------|-------------------|
| (A) Household characteristics                     |         |         |                   |
| Average annual income (thousand Swedish crowns)   | 371.97  | 182.23  | 350.3             |
| Age in years (above 18)                           | 50.60   | 14.19   | 48.9              |
| Have university education                         | 0.33    | 0.47    | 0.36              |
| Household size                                    | 2.28    | 1.19    | 1.97              |
| (B) Purchasing behavior                           |         |         |                   |
| Shopping trips per household per year             | 7.06    | 6.78    |                   |
| Purchases per household per year (Swedish crowns) | 325.29  | 329.76  |                   |
| Purchases per household per year (g)              | 7104.40 | 7206.81 |                   |
| Purchases per capita per year (g)                 | 3929.02 | 4366.96 | 4500 <sup>a</sup> |

Non-missing observations for income (2678) and for household size (2781). All other variables in panel A and B 2782 observations. Figures on national means from Statistics Sweden, data for 2008

<sup>a</sup> From <http://www.kaffeinformation.se/>

versus 48.9 years, respectively, (the average age of those 18 years and older, since the reference persons in the panel were all 18 or older). The share of households with a university education is lower in the sample than the national average: 33 versus 36%, respectively. The average size of the sampled household is 2.28, compared to the national average of 1.97.

The households in our dataset appear to have diligently reported their retail coffee purchases, as seen in panel B of Table 1. On average, households purchased coffee in retail stores on 7.1 occasions per year, and the average annual household expenditure on retail coffee was 325 Swedish crowns (approximately 36 Euro using July 2008 exchange rates). In 2008, average coffee consumption in Sweden was 9.4 kg/capita/year.<sup>5</sup> Of this, roughly 60% was bought through retail channels for household consumption. The remaining 40% was consumed at work or in restaurants and cafes. Around 12% of the total consumption was instant coffee, which is almost exclusively sold retail. This means that, if our sample was representative and fully diligent in reporting all purchases, we would expect them to consume approximately 4.5 kg/capita/year. Our sample of households purchased an average of 3.9 kg/capita/year, which is close to the expected level of consumption. As with any Homescan data, some degree of under reporting is expected. Einav et al. (2010) compared the recorded purchasing behavior of US households in the Homescan data administered by AC Nielsen, with the purchasing behavior reported

<sup>5</sup> This figure includes children. Source of this and the following statistics: <http://www.kaffeinformation.se/>.

by stores. Overall, the authors found evidence that households are diligent and that Homescan data are a valuable source of information.<sup>6</sup>

GfK questionnaires are completed by households when they join the consumer panel and then again every January, and cover a range of issues related to household shopping preferences. There are 35 questions in the questionnaire, and many questions have multiple alternative responses. One subset of questions relates to household choices of different types of products. We made use of one question regarding organic labeled products. The question was: “When I buy groceries I try, to the extent feasible, to buy organic products”. The respondent can tick one of six boxes; box 1 indicates “Totally Disagree”, box 5 indicates “Totally Agree”, and box 6 indicates “Don’t Know”.

On the product side, the data was matched to European Article Numbers (EANs), providing a description of each coffee product bought by the household, including the package size, brand name, whether it was labeled organic, Fairtrade, as well as other product characteristics. Table 2 summarizes the characteristics of the coffee products in our sample. Around 7% of the available choices were organic. We use data on purchases of all ground and bean coffee. Instant coffee was excluded.

The data on households and the data on product varieties are linked via a database of market transactions. These market transactions describe the price and quantity purchased for each variety of coffee on a particular shopping trip date at a particular store by a particular household. There are 11 grocery store chains, each with varying store formats, which we group into four different classes: large supermarket, supermarket, discount store, and other. The combined dataset, therefore, includes household statistics, coffee product descriptions and a record of market transactions.

## 2.1 Choice sets

The dependent variable in our estimation of the demand system is the choice made by the consumer. For each shopping trip by each household, we construct a set of coffee products from which the consumer chooses, i.e., the choice set. The choice variable is discrete and binary: it is equal to 1 when a household purchases a particular variety and equal to 0 otherwise.

Homescan data provides observations on choices actually made by the consumer, but does not provide observations on choices that are not made. Hence, we cannot directly observe the coffee varieties amongst which the household can choose from a given shopping trip. However, the data are detailed enough for us to make use of observed coffee purchases by *other* households. We therefore construct the choice set for each shopping trip using the purchasing data of other household purchases from the same chain and store format (11 chains across 4 store format types is 44 combinations in all) for a given type of municipality (4 types) within a three month

<sup>6</sup> One important discrepancy that they found is that around 20% of purchases on a given shopping trip are not recorded. As seen, this does not appear to be a significant concern with respect to our data. Second, they found that consumers who use loyalty cards to receive discounts often fail to report the discounts. In Sweden, the use of such rebates and coupons to achieve discounts is less prevalent than in the US; we do not expect any such misreporting to have had an important impact on our results.

**Table 2** Summary statistics—coffee product characteristics, stores and choice sets

| Variable                                | Mean   | SD     | Min  | Max   |
|---|--------|--------|------|-------|
| Hedonic price (SEK per 100 g)           | 5.166  | 1.596  | 1.57 | 26.60 |
| Organic                                 | 0.070  | 0.2554 | 0    | 1     |
| Fairtrade                               | 0.033  | 0.1789 | 0    | 1     |
| Package size: less than 250 g           | 0.011  | 0.107  | 0    | 1     |
| 250 g                                   | 0.015  | 0.121  | 0    | 1     |
| 251–499 g                               | 0.033  | 0.181  | 0    | 1     |
| 500 g                                   | 0.939  | 0.238  | 0    | 1     |
| Package type: bag in bag                | 0.266  | 0.463  | 0    | 1     |
| Bag in box                              | 0.088  | 0.283  | 0    | 1     |
| Can                                     | 0.002  | 0.043  | 0    | 1     |
| Monobag                                 | 0.302  | 0.459  | 0    | 1     |
| Other                                   | 0.311  | 0.463  | 0    | 1     |
| Roast: medium roast                     | 0.5645 | 0.4958 | 0    | 1     |
| Dark roast                              | 0.3472 | 0.4761 | 0    | 1     |
| Other roast                             | 0.0882 | 0.2836 | 0    | 1     |
| Grinding: fine ground                   | 0.8099 | 0.3923 | 0    | 1     |
| Coarse ground                           | 0.1376 | 0.3444 | 0    | 1     |
| Whole bean                              | 0.0525 | 0.2230 | 0    | 1     |
| Country of origin: unspecified/multiple | 0.924  | 0.265  | 0    | 1     |
| Columbia                                | 0.030  | 0.170  | 0    | 1     |
| Other Latin America                     | 0.002  | 0.046  | 0    | 1     |
| Ethiopia/Kenya                          | 0.025  | 0.156  | 0    | 1     |
| Indonesia                               | 0.019  | 0.136  | 0    | 1     |
| Decaffeinated                           | 0.016  | 0.127  | 0    | 1     |
| Store type: large supermarket           | 0.5159 | 0.4997 | 0    | 1     |
| Supermarket                             | 0.2376 | 0.4256 | 0    | 1     |
| Discount store                          | 0.1861 | 0.3892 | 0    | 1     |
| Other store                             | 0.0604 | 0.2381 | 0    | 1     |
| Urbanization (4 is most urban)          | 2.89   | 0.9402 | 1    | 4     |
| Number of products in choice set        | 35.26  | 11.50  | 1    | 60    |
| Number of organic coffees in choice set | 2.50   | 1.521  | 0    | 7     |

Number of observations in sample: 1,225,533

window. For example, a shopper buying a coffee product in a large store belonging to the “ICA” chain, in Stockholm in early June of 2008 faces a choice set of 30 coffee products. We observe only the choice made by this particular shopper. We identify the other 29 coffee products that are part of this choice set from the choices of other shoppers buying coffee at large stores belonging to the “ICA” chain, in Sweden’s largest cities, between mid-April and mid-July of 2008. A manual comparison with the assortment in some selected stores pointed to our generated choice sets as giving a generally accurate representation of the assortment.

There is limited variety in brand-organic combinations facing households when they shop. For example, at the 10th percentile of observations in the sample there was one organic coffee in the choice set. At the median, a household faced 2.5 organic coffee varieties in their choice set. This suggests that while organic coffee is widely available in Sweden, only a few organic coffees make their way into household choice sets.

A total of 43,252 shopping trips were observed in our data. However, the construction of choice sets expanded the size of the dataset to a total of 1,260,081 observations. The descriptive statistics in Table 2 are therefore based on the full, expanded sample.

We observed the actual price of the coffee product when it was purchased. However, estimating the demand system means we also need to infer the price of products at a store that were not purchased. To do this, we used a hedonic regression to generate prices for all of the products in the choice set (i.e., the price of the products that were not purchased by the household). The hedonic regression was run on the 42,143 observations on price. We regressed price per 100 g of coffee on brand fixed effects (35 in all), store fixed effects (by chain and store format: 44 in all), coffee country of origin fixed effects (5 in all), bean roast (3 types in all), monthly fixed effects (34 months), municipal-type fixed effects (4 in all), package size fixed effects (4 types), package type (5 types), and a fixed effect for decaffeinated coffee. The adjusted  $R^2$  of this regression is 0.51 and the  $F$ -statistic for the joint significance of all variables is 320.15. Table 3 summarizes the results of this hedonic regression.

The organic label has a positive and statistically significant coefficient at 0.791 SEK per 100 g of coffee. There is important variation in the estimated value associated with each brand. The omitted (reference) brand in the hedonic regression is Gevalia. The estimated coefficients for each brand therefore give an indication of the value of the brand relative to Gevalia. Lavazza, which is profiled as a high quality luxury Italian coffee, has the highest brand coefficient at 5.2 SEK per 100 g. In contrast Euro Shopper, which is a discount brand, has the lowest coefficient at  $-2.127$  SEK per 100 g. The brand coefficients are in line with expectations.

The hedonic regression is used to predict the hedonic price. The mean hedonic price for the entire sample (Table 2) is around 52 Swedish crowns per kg, with considerable dispersion between the highest and lowest prices.

### 3 The empirical specification

We use a discrete choice model of demand and assume a logit specification (McFadden (1974), Cameron and Trivedi (2009)). Consider household  $i$  facing the choice of a product  $j$  among a set of  $J$  available products on shopping trip  $s$ . The household derives utility  $U_{ijs}$  from its choice  $j$  and chooses the alternative that provides the greatest utility. The behavioral model is, therefore, household  $i$  chooses alternative  $j$  if  $U_{ijs} > U_{iks} \forall j \neq k$ . We express utility as:

**Table 3** Estimating the hedonic price for coffee, 2007–2009

| Observed characteristic | Coefficient | Standard errors |
|-------------------------|-------------|-----------------|
| Organic                 | 0.791       | 0.024***        |
| Fairtrade               | 1.308       | 0.048***        |
| Brand                   |             |                 |
| Gevalia                 | (Omitted)   |                 |
| Aroma                   | (Omitted)   |                 |
| Bellarom                | −0.692      | 0.072***        |
| BKI                     | −1.312      | 0.143***        |
| Blå Mocca               | 0.275       | 0.040***        |
| Bremer                  | −0.692      | 0.111***        |
| Café Organico           | 1.978       | 0.282***        |
| Caffe Musetti           | 0.273       | 0.150*          |
| Classic                 | 0.365       | 0.021***        |
| CO-OP                   | −1.033      | 0.042***        |
| Eldorado                | −1.237      | 0.163***        |
| ELU                     | −1.173      | 0.128***        |
| Euro shopper            | −2.127      | 0.057***        |
| Excellent               | −1.698      | 0.370***        |
| First class             | −0.061      | 0.255           |
| First price             | −0.969      | 0.103***        |
| Guldkaffe               | −0.503      | 0.344           |
| Hemköp                  | 0.036       | 0.085           |
| ICA                     | −0.361      | 0.043***        |
| Ideekaffe               | 4.378       | 0.309***        |
| Kalas Kaffe             | −0.040      | 0.565           |
| Lavazza                 | 5.200       | 0.244***        |
| Lindvalls               | −0.342      | 0.044***        |
| Lyxkaffe                | −1.380      | 0.370***        |
| Löfbergs Lila           | −0.024      | 0.032           |
| Maxwell House           | 1.330       | 0.073***        |
| Melangerie              | −0.965      | 0.330***        |
| Monte Santos            | −1.071      | 0.995           |
| Perfetto                | −1.664      | 0.112***        |
| Queen                   | −0.705      | 0.300**         |
| Repris                  | (Omitted)   |                 |
| Signum                  | −0.454      | 0.066***        |
| Willys                  | −0.862      | 0.057***        |
| X-TRA                   | −1.527      | 0.053***        |
| Zoegas                  | 0.710       | 0.033***        |
| Observations            | 42,143      |                 |
| Adjusted $R^2$          | 0.51        |                 |

Dependent variable: coffee price in SEK per 100 g. The hedonic regression includes fixed effects for: 44 store chain and formats; 5 regions of origin for single-origin coffees; 3 types of bean roast; 3 grades of bean grinding; 4 levels of urbanization where the purchase took place; 4 package sizes; 5 package types; and decaffeinated coffee. 34 monthly fixed effects are also included. White-corrected standard errors

\*\*\* Significance at 1%, \*\* at 5% and \* at 10%

$$U_{ijs} = X_j\beta + HH_{ijs}\gamma + p_{js}\alpha + \varepsilon_{ijs}. \quad (1)$$

$X_j$  is a vector of coffee product characteristics summarized in Table 2 and includes dummy variables to capture whether good  $j$  carries an organic or Fairtrade label; brand (there are 35 brands); a measure of roast (dark roast and other roast, medium roast); the package type; the package size; an indicator for decaffeination; and dummy variables for different national origins of single-origin coffee (Columbia, other Latin America, Ethiopia/Kenya and Indonesia).

$HH_{ijs}$  is a vector of household characteristics summarized in Table 1, interacted with the organic indicator. These interaction include: “Old” households indicating a primary shopper over the age of 55 years; “University” for those with a University degree; and “high income” for households with a combined annual pre-tax income of at least 500,000 Swedish crowns (SEK) (approximately 53,300 euro in July 2008). We interact the organic indicator with the set of five possible household responses to the survey questions. The omitted category is households that answered “neither agree nor disagree”.

$p_{js}$  is the log of the hedonic price of coffee  $j$  on shopping trip  $s$  (predicted from the hedonic regression as explained in Sect. 2).  $\alpha$  is the coefficient capturing sensitivity to price.  $\varepsilon_{ijs}$  is the individual and product specific error term that follows an IID type I extreme value distribution.

We thus estimate the parameters in Eq. (1) with a conditional logit specification, where the estimates are conditional on the shopping trip.

$$\text{Prob}(\text{choice}_{ijs} = 1 | X_j, HH_{ijs}, p_{js}) = \frac{\exp(X_j\beta + HH_{ijs}\gamma + p_{js}\alpha)}{\sum_{k=1}^{J_s} \exp(X_k\beta + HH_{iks}\gamma + p_{ks}\alpha)}. \quad (2)$$

The dependent variable is “choice” which is equal to one if the household chooses variety  $j$  and equal to zero if the variety is not chosen. The details of how choice is computed have been discussed above in Sect. 2.

Our focus is on the interaction between the organic label and household survey responses. For example, we are interested in estimating the impact of a household answering “Totally Agree” in the survey on the probability of choosing an organic coffee. A concern with including survey responses is that the specification in essence regresses outcome on outcome. Note, however, that the question asks about whether the households tries to buy organic when shopping for grocery, as opposed to asking if organic coffee was bought on the shopping trip in question. Our preferred interpretation is therefore that this variable should be seen as capturing preferences.

We know that many households exhibit loyalty—and therefore include a dummy for the particular product purchased on the previous shopping trip.

We make two restrictions on the data: first, for clarity of comparison we drop the observations where households had not responded or answered “don’t know” which leads us to lose 65 households, reducing the number of households in the data from 2785 to 2717; second, we exclude shopping trips where we estimate that consumers are faced with three coffee products or less. This excludes a handful of shopping

trips to pharmacies and petrol stations. As a result, our estimation sample size decreases from 1,260,081 to 1,225,533 observations.

### 4 Estimation results

In Table 4, we report the estimated coefficients of Eq. (1) for two specifications.

In column (1), we report results for the regression where stated preferences for organic are excluded. All the reported coefficients are significant at the 1% level. Standard errors are robust. The price coefficient is negative and high income consumers are less price-sensitive, as expected. The coefficient on organic is negative, suggesting many consumers have an aversion to organic coffee. Moreover, high income consumers have a stronger aversion to organic coffee, but consumers with a university degree have a preference for organic coffee. Older consumers also have an aversion to organic.

In column (2), we report results where stated preferences for organic are included in the regression. First, note that the coefficients on price and price  $\times$  high income are stable. Including the interactions between organic and stated preferences does change the estimated coefficients on organic although the sign and statistical significance of the estimates remains unchanged. Relative to the first specification: the estimate on organic is now slightly more negative; high income consumers have

**Table 4** Conditional logit, estimated demand for coffee, 2717 Swedish households, January 2007–January 2009

| Dependent variable: coffee choice<br>Coefficient | (1)               | (2)               |
|--|-------------------|-------------------|
| Ln(price)  | −0.884 (0.111)*** | −0.885 (0.111)*** |
| Ln(price) $\times$ high income                   | 0.120 (0.011)***  | 0.120 (0.011)***  |
| Organic  | −0.489 (0.042)*** | −0.580 (0.058)*** |
| Organic $\times$ high income                     | −0.237 (0.061)*** | −0.174 (0.066)*** |
| Organic $\times$ university                      | 0.558 (0.051)***  | 0.408 (0.056)***  |
| Organic $\times$ old                             | −0.217 (0.052)*** | −0.230 (0.056)*** |
| Organic $\times$ “totally disagree”              |                   | −0.978 (0.132)*** |
| Organic $\times$ “disagree”                      |                   | −0.711 (0.081)*** |
| Organic $\times$ “agree”                         |                   | 0.634 (0.064)***  |
| Organic $\times$ “totally agree”                 |                   | 1.655 (0.086)***  |
| Observations                                     | 1,225,533         | 1,225,533         |
| Log-likelihood                                   | −89,680           | −89,312           |

Both regressions include fixed effects for: each shopping trip; 35 brands, 5 regions of origin for single-origin coffees, 3 bean roasts, 3 grades of bean grinding, 5 package types, 4 package sizes, a dummy to indicate if the coffee is decaffeinated, and the previous brand purchased. Number of households: 2717. Standard errors derived with Huber/White/sandwich estimator

\*\*\* Significance at 1%

a weaker aversion for organic coffee; and consumers with a university degree have a slightly weaker preference for organic coffee. Interestingly, high-income households are less likely to purchase organic coffee. This stands in contrast to findings presented by other researchers (e.g., Ngobo 2011; Griffith and Nesheim 2010; Kiesel and Villas-Boas 2007). The Swedish income distribution is relatively compressed, which may explain part of the reason for the difference. Another potential explanation is that organic products are simply marketed differently in Sweden. Households with university education are more likely to purchase organic coffee. Older households are less likely to purchase organic coffee.

The estimates for the interaction between survey responses and organic are all statistically significant at the 1% level. The magnitude and sign of these coefficients fit what we expect: respondents value organic products when they say they do, and do not value them when they say they do not. The omitted category is households that answered “Neither Agree nor Disagree” to the question on their purchasing habits.

## 5 Using the estimates to predict counterfactual choice

We now turn to our analysis of counterfactual supply side scenarios. We want to explore the effect of new product introductions on the market share of organic products. To do so, we introduce a synthetic product to all choice sets. On each shopping trip consumers now meet an additional product. We let this synthetic product be organic and let it have characteristics that are typical in this market: a mid-roasted caffeinated coffee from Colombia, in a 250 g “monobag” package with the most common type of grind.

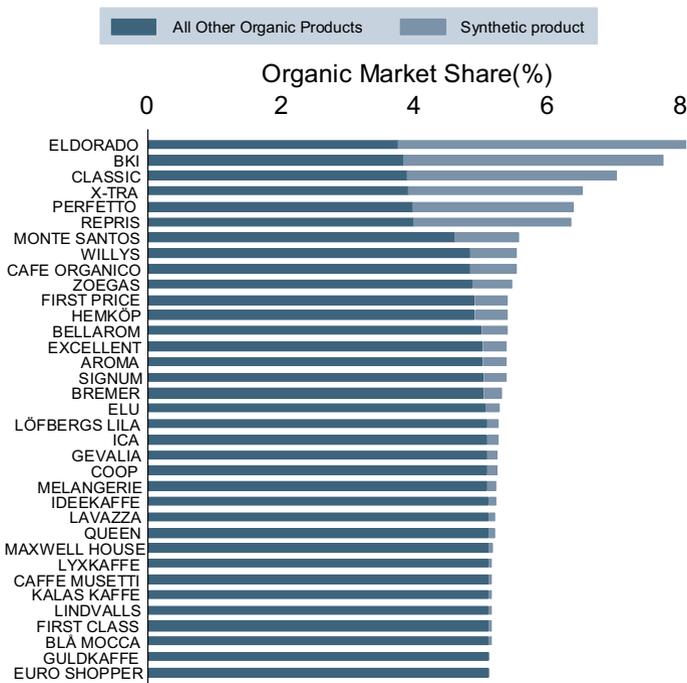
To systematically explore the brand aspect we conduct 35 counterfactual, out of sample, choice predictions: one counterfactual for each brand. The price of the counterfactual coffee brand is predicted using the same hedonic regression discussed in Sect. 2. We use the estimated demand system, summarized in Table 4, column (2) to compute the out of sample predicted choice probabilities.

In a discrete choice setting the product with the highest predicted probability will be chosen. Introducing this new synthetic product changes the ranking of the choice probabilities across households and these new choice probabilities are then used to compute the counterfactual market shares for organic coffee across all brands, presented in Fig. 1. The total length of each bar indicates the organic market share for each counterfactual, and consists of both the non-synthetic and synthetic organic coffees sold. We distinguish between the market share of the synthetic product and the non-synthetic (pre-existing) products. The “brand” next to each bar indicates the brand assigned to the synthetic variety.

We are interested in the change in organic coffee’s market share resulting from the introduction of the synthetic organic variety and can compare to the in-sample market share of organic coffee which is 5.2%. We see that the effect of product introductions is quite sensitive to the brand assigned to the synthetic product. The chosen brand will affect market shares both via the brand fixed effect and via prices. These effects will feed through the demand system and interactions with income,

education, age and answers to the question of stated shopping behavior with respect to organic groceries. Introducing an organic product under the brand of Eldorado or BKI will, as seen, expand the organic market share by some 3 percentage points, whereas at the other end of the scale some brand introductions only achieve a miniscule market share.

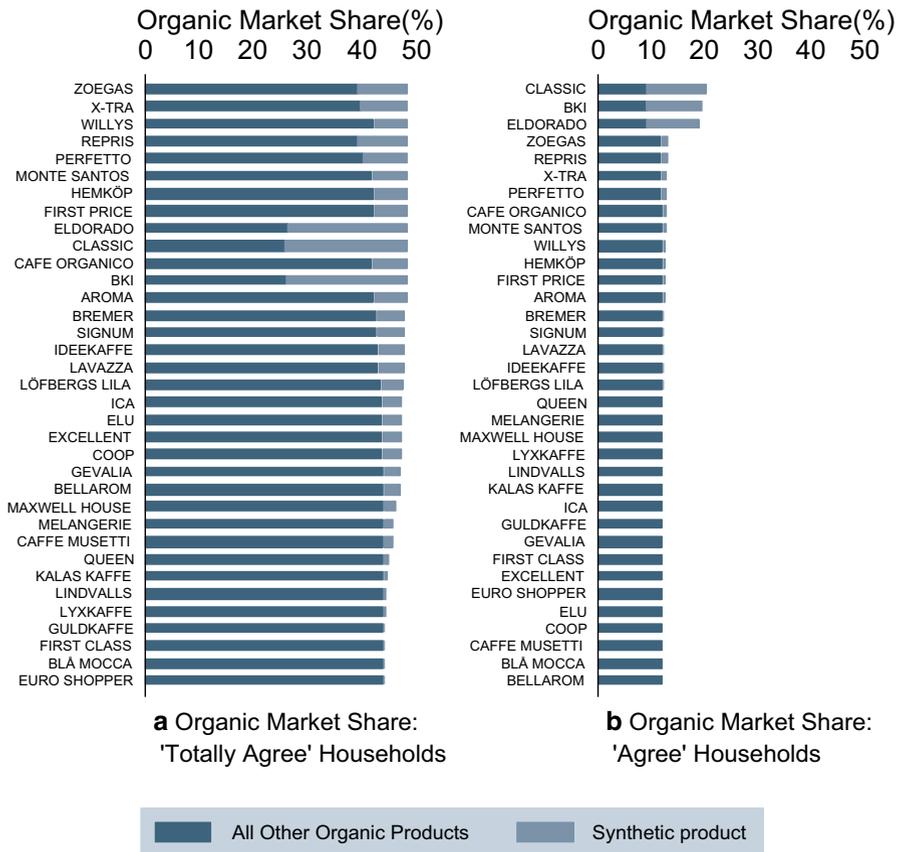
We define two effects of a product’s introduction: market expansion and market stealing. In this case, *market expansion* refers to an expansion of the organic segment of the market, the increase in sales that results from taking market shares from non-organic products. *Market stealing* refers to market shares that come at the expense of lowering the market shares of other organic products. This includes both market stealing from other brands as well as cannibalization within the brand for product introductions for brands that already have a pre-existing organic product. An organic label that wants to maximize market expansion would clearly prefer new products to expand the organic segment rather than reshuffle market shares within the segment. For almost all new product introductions, the market expanding effect is stronger than the market stealing effect. We see that a handful of products have a sizeable effect, where organic market shares increase by 20–60%, from around a 5.2% to a 6–8% market share. Most product introductions have very modest effects. It is also interesting to note that market expansion dominates market stealing in all



**Fig. 1** Counterfactual organic coffee market shares. The counterfactuals are computed using the estimates from Table 2, column (2)

counterfactual settings, except for one: a new organic Monte Santos coffee result in more market stealing than market expansion.

Which consumer groups should the organic label target? Intuitively, one would want to design a new organic product that targets the keenest organic consumers. In Fig. 2 we break down the overall impact that we explored above by different consumer groups. The in-sample predicted market shares for the most keen organic households is 45% and as seen the introduction of the synthetic organic coffee product has a relatively weak impact on overall organic market shares. The share of organic coffee purchased by these households increase from the in-sample predicted 45% to around 48% for most counterfactuals. In several of the counterfactuals, these households switch towards the synthetic product at the expense of the market share of the pre-existing organic coffee products. Consider the impact of the introduction of an “Eldorado” synthetic organic coffee for example. Overall market shares of



**Fig. 2** Counterfactual organic coffee market shares by household survey response across brands for households that respond “Totally Agree” (a), “Agree” (b), “Neutral” (c), and “Disagree” (d). The brand next to each bar indicates the brand assigned to the synthetic product for each counterfactual setting. The counterfactuals are computed using the estimates from Table 2, column (2)

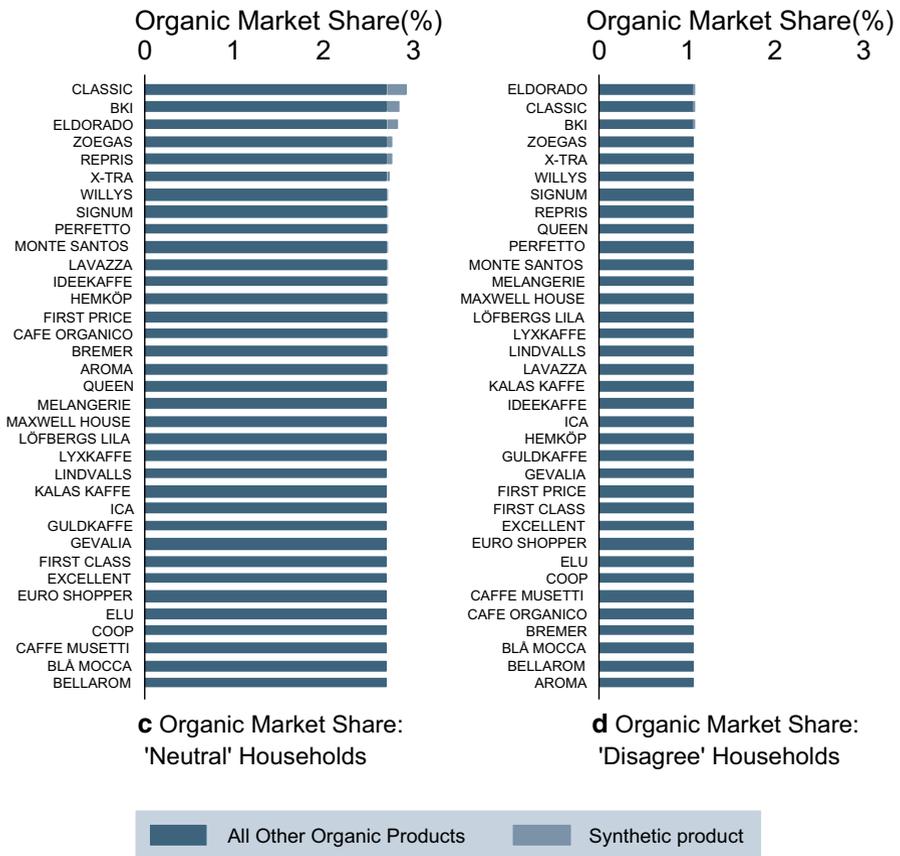


Fig. 2 continued

organic coffee increase from 45 to 48% but the market share of the non-synthetic organic coffee falls from 45% to less than 30%. While there is some market expansion the introductions that garner the largest market share in this segment do so mainly by market stealing (Eldorado, Classic, BKI).

Amongst the more moderately keen organic households (Fig. 2b), the outcome is different. The share of organic coffee purchased by these households increases from the in-sample predicted 13% to close to 20%. The successful brands that generated substantial market stealing in the keenest segment generate market expansion in this less keen segment. As these households are more numerous than the most keen organic consumers this is also the source of the demand response that dominates the aggregate pattern that we documented in Fig. 1. This suggests that in this niche, there are relatively more households that switch away from conventional coffee to organic coffee, when they find an organic coffee that carries brand that they value. In this market, niche consumers substitute towards organic coffee products and it is in this niche where the organic label has the potential to increase its market share.

Turning next to even less organically oriented households the effects are small; a few brand introductions generate a market share expansion that is non-trivial but in none of the cases is the introduction able to push the market share for organic above 3% as illustrated in Fig. 2c. For these consumers there is little market stealing, they do not value an organic label per se, and thus there is no particular market stealing effect of a new organic product from other organic products but rather the market stealing stems from all products. Finally, consumers that state that they disagree with the statement that they try to purchase organic products nevertheless purchase around 1%. For this group new product introductions are not able to overcome the aversion to organic.

## 6 Concluding remarks

We estimate a structural model of demand that we use to explore the performance of brand-organic alliances across consumers with heterogeneous preferences. Survey responses on whether households strive to purchase organic grocery products have important predictive power for the retail coffee purchases made by these households. Our counterfactual simulations show that key to new organic products leading to market expansion of the organic segment is success among the households that have moderate self-reported intent purchase organic groceries. The findings highlight that it is not the absolute level of fit between a product and consumer preferences that are crucial for market success, but rather the ability to induce a switch among consumers that are close to indifferent between different choices.

A contribution is that we show how counterfactual simulations can be used to first identify constraints to organic market share in a market with differentiated retail products, and second identify the best brand partner for the organic label. This relates to a more general issue of brand partnerships (van der Lans et al. (2014), Cunha et al. (2015) and Geylani et al. (2008)). Ours is the first study to examine brand partnerships from the point of view of the ecolabel. We further hope to have illustrated the usefulness of counterfactual product introductions in a discrete choice setting with access to self-reported measures of preferences/shopping habits.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Berry ST, Waldfogel J (1999) Free entry and social inefficiency in radio broadcasting. *Rand J Econ* 30(3):397–420
- Bjørner TB, Hansen LG, Russell CS (2004) Environmental labeling and consumers choice—an empirical analysis of the effect of the Nordic swan. *J Environ Econ Manag* 47(3):411–434

- Brooks K, Lusk J (2010) Stated and revealed preferences for organic and cloned milk: combining choice experiment and scanner data. *Am J Agric Econ* 92(4):1229–1241
- Cameron A, Trivedi P (2009) *Microeconometrics using stata*, vol 5. Stata Press, College Station
- Cunha M, Forehand MR, Angle JW (2015) Riding coattails: when co-branding helps versus hurts less-known brands. *J Consum Res* 41(5):1284–1300
- Einav L, Leibtag E, Nevo A (2010) Recording discrepancies in Nielsen homescan data: are they present and do they matter? *Quant Mark Econ* 8(2):207–239
- Eurobarometer (2011) Attitudes of European citizens towards the environment. Special Eurobarometer 365. European Commission, Brussels
- Geylani T, Inman JJ, Hofstede FT (2008) Image reinforcement or impairment: the effects of co-branding on attribute uncertainty. *Mark Sci* 27(4):730–744
- Griffith R, Nesheim L (2010) Estimating households' willingness to pay. CeMMAP working papers
- Hausman J (1997) Valuation of new goods under perfect and imperfect competition. In: Bresnahan T, Gordon R (eds) *The economics of new goods*. University of Chicago Press, Chicago
- Kiesel K, Villas-Boas S (2007) Got organic milk? Consumer valuations of milk labels after the implementation of the USDA organic seal. *J Agric Food Ind Organ* 5(1):1152
- Lusk J, Brooks K (2010) Who participates in household scanning panels? *Am J Agric Econ* 93(1):226–240
- McFadden D (1974) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed) *Frontiers in econometrics*. Academic, New York, p 1974
- Ngobo PV (2011) What drives household choice of organic products in grocery stores? *J Retail* 87(1):90–100
- Petrin A (2002) Quantifying the benefits of new products: the case of the minivan. *J Polit Econ* 110(4):705–729
- Tropical Commodity Coalition (2012) Coffee barometer (online). [http://www.newforesight.com/sites/default/files/newforesight/TCC\\_CoffeeBarometer2012.pdf](http://www.newforesight.com/sites/default/files/newforesight/TCC_CoffeeBarometer2012.pdf). Accessed 27 Mar 2012
- van der Lans R, Van den Bergh B, Dieleman E (2014) Partner selection in brand alliances: an empirical investigation of the drivers of brand fit. *Mark Sci* 33(4):551–566
- Van Doorn J, Verhoef PC (2015) Drivers of and barriers to organic purchase behavior. *J Retail* 91(3):436–450