Available online at www.sciencedirect.com
ScienceDirect
Journal of Retailing 99 (2023) 26–45
www.elsevier.com/locate/jretai

"Sorry, the product you ordered is out of stock": Effects of substitution policy in online grocery retailing

Dong Hoang[a,*], Els Breugelmans[b]

[a] Leeds Business School, City Campus, Leeds Beckett University, Leeds, LS1 3HB, United Kingdom
[b] Department of Marketing, Antwerp Campus, KU Leuven, Hendrik Conscienceplein 8, 2000 Antwerpen, Belgium

Available online 9 July 2022

Abstract

Postpurchase out-of-stock (PP-OOS) often happens in an online grocery context, where products appear to be available at the time a consumer places an order, but become OOS when the order is to be dispatched. This paper investigates two substitution policies that can mitigate negative responses: substitutions can match (i) on the dominant attribute and (ii) with a product from the consumers' past purchase portfolio. According to data collected through two computer-simulated purchase experiments, involving more than 3,000 households and five product categories, matching the substitution on the dominant attribute increases acceptance, but this dominant attribute varies across category differentiation level (flavor for horizontal differentiated categories like cereals or crips vs. brand for vertical differentiated categories like margarine or ketchup). Category differentiation also informs acceptance of national brand or private-label flavor substitutes, such that, same-flavor private label is prefered more in horizontal differentiated categories. Matching on the basis of previous purchases has positive effects for both category differentiation levels, and when combining both policies, the previous purchase matching effect grows stronger for same flavor, rather than same brand, matching. These detailed insights establish several key managerial implications for substitution policies in online grocery contexts.

Keywords: Online grocery; Online retailing; Out-of-stock; Substitution; Product attribute.

The online grocery shopping process consists of two main phases: purchase (ordering) and receipt (pick up or delivery), separated by several hours at least and frequently by a few days, such that many of the growing market of online grocery shoppers place orders in advance of the actual pick-up or delivery date (Kantar 2021). Due to this time gap, it is possible for products to be available at the time orders are placed but then go out-of-stock (OOS) on the dispatch date. When such postpurchase OOS (PP-OOS) occurs, the retailer identifies the inventory issue, typically a few hours before the order pick-up or delivery time but well after the consumer has completed the purchase phase. These PP-OOS events represent one of the top three drawbacks of online shopping experiences, according to surveys of grocery shoppers (Mintel 2021a). On average, more than one-quarter of online grocery shoppers have experienced PP-OOS for at least one item ordered in the preceding six months; for some supermarkets, PP-OOS rates are as high as 48% (Chan 2017).

To deal with the consumer frustration evoked by PP-OOS, online grocery retailers usually replace the item with a substitution, rather than leave the customer empty-handed.[1] When an OOS event occurs prior to purchase, such as when an in-store shopper encounters an empty shelf or when an online shopper is still engaged in the purchase decision process, the

* Corresponding author.
E-mail address: d.hoang@leedsbeckett.ac.uk (D. Hoang).

[1] Online grocery retailers might try to reach consumers to inform them about the OOS and ask for their replacement preferences, but such two-way communication often requires instant responses for it to be effective. Most online grocery orders typically are prepared and dispatched only a few hours in advance (due to the perishable nature of some products), thus, any delayed response from consumers would negatively impact retailers' pick and dispatch operations efficiency.

https://doi.org/10.1016/j.jretai.2022.06.006
0022-4359/

consumer can select a preferred substitution. But for a PP-OOS, the consumer is no longer in the purchase phase and no longer has easy access to remaining choice alternatives. Therefore, the online retailer is responsible for selecting a suitable substitute, offered during the receipt phase (pick up or delivery). At that moment, consumers might accept or reject the offered substitution (Severs 2014). Because retailers, in offering a substitution, seek to prevent or mitigate consumers' dissatisfaction with the OOS (Cocozza 2019), it is crucial for them to have an effective strategy for selecting suitable substitutions on behalf of their consumers. This choice, however, is far from obvious (Mintel 2021a); 30–45% of shoppers express dissatisfaction with the substitution offered following a PP-OOS.

The choice of substitutions in the context of a PP-OOS poses two important challenges for an online grocery retailer that wants to mitigate negative OOS reactions (Cocozza 2019). Firstly, the remaining assortment to choose a substitution from is oftentimes large, and existing literature has not offered a full understanding of which product attributes are determinant in selecting a substitution. Is there a dominant attribute across all categories? Or does the determining attribute for the substitution depend on category characteristics? And, secondly, how (if at all) can online retailers that have the record of consumers' past purchases, use this as a decisive factor in their substitution policy? Are substitutions of past purchased items more effective for some categories or for some attributes?

With these considerations, we seek to investigate consumers' behavioral responses, in the form of their acceptance of a substitution, to two policies that online grocery retailers might adopt. The first policy we test involves the investigation of a dominant attribute in deciding on the substitution, i.e., suggesting a substitution along the attribute which carries more weight in consumers' product evaluation (Batra, Homer, and Kahle 2001; Gutman 1982; Shao 2015). In this research, we manipulate the substitution to match the OOS item on either the brand (i.e., another flavor but same brand) or the flavor (i.e., same flavor from another brand), as brand and flavor are two key attributes present in almost all categories (Campo, Gijsbrechts and Nisol 2003; Bronnenberg, Dubé and Sanders 2020).[2] To gain more insights, we also examine whether consumers' reactions differ for flavor-based substitutions involving a national brand (NB) versus a private label (PL). In the second policy, we explore the effect of offering a substitution that reflects consumers' previous purchases. We also test the interaction effects of the two substitution policies on the acceptance of substitution. We further compare whether the effects of the two substitution policies are the same across product categories that vary in their level of differentiation. We distinguish between vertically differentiated categories (i.e., where brands offer clearly distinct levels of quality and performance) or horizontally differentiated categories (i.e., where no one brand is uniformly better and consumers differ in taste preferences) (Spiller and Belogolova 2017).

Literature review and contributions

In Table 1, we summarize extant literature pertaining to OOS responses published since 2000 (for earlier studies, see the review by Sloot, Verhoef, and Franses 2005). Such studies focus predominantly on physical store environments, though recent studies also address online settings (e.g., Breugelmans, Campo, and Gijsbrechts 2006; Gunness and Oppewal 2020; Jing and Lewis 2011; Ma, Chen and Zheng 2018; Peterson, Kim, and Jeong 2020; Pizzi and Scarpi 2013). Yet, there is little research that has looked at a PP-OOS, which represents a different scenario, because the OOS occurs at the product receipt phase, and consumers can no longer choose a substitution themselves. We know of only one study, by Jing and Lewis (2011), that investigates PP-OOS using online grocery scanner data. It seeks to determine the effects of nonfulfillment of online orders on short- and long-term customer behaviors, such as subsequent purchase volume and incidence. These authors advise reducing stockout rates, especially for particular consumers and categories, which is helpful. But some PP-OOS is likely inevitable (Mintel 2021a), so we seek to advance this field by investigating the effects of the retailer's immediate response in providing a suitable substitution. If retailers can develop effective substitution policies, it would greatly benefit the online grocery sector, such that our unique research promises managerially relevant findings.

Furthermore, Table 1 shows that a few studies have investigated how retailers might mitigate negative OOS responses in online settings. For example, they recommend revealing the OOS prior to consumers' selections or proposing a substitution at the time of purchase, (Breugelmans, Campo, and Gijsbrechts 2006), strategically designing communications (i.e., timing and wording) for OOS disclosures (Kumar, Shamar, and Tapar 2021; Peterson, Kim, and Jeong 2020; Pizzi and Scarpi 2013), and offering financial incentives for future or backorders (Anderson, Fitzsimons, and Simester 2006; Kim and Lennon 2011). A financial compensation strategy reduces immediate dissatisfaction (Kim and Lennon 2011) and encourages backorders, instead of cancellations (Anderson, Fitzsimons, and Simester 2006), but it also is the least profitable solution for retailers, with the potential for negative effects on future demand. Nor can backorders resolve grocery customers' immediate and pressing needs. Recommending a substitution for a PP-OOS is thus an interesting policy and in fact, a common practice amongst online grocery retailers when an ordered item is OOS at the time of pick up or delivery (Cocozza 2019). Logistically, retailers likely need to select the substitution themselves, rather than offer consumers the choice, to maintain the efficiency of their online order handling and limit the time needed per order. We know of no research into retailers' substitution strategies in response

[2] We focus on brand or flavor as focal attributes, rather than attributes such as package size, because substitutions of the same product in another size likely have very strong positive effects, but many items are available in only one size on shelves.

Table 1
Extant OOS literature and contributions.

Authors	Prior-to-purchase OOS		PP-OOS	Study objectives	Strategy to mitigate negative OOS response	Method
	Offline	Online				
Fitzsimons (2000)	X			Consumer's OOS response e.g., satisfaction with the decision process and subsequent store choice behavior		Lab-based experiment
Campo, Gijsbrechts, and Nisol (2000)	X			Consumers' behavioral response to OOS (switching behaviors)		Survey
Zinn and Liu (2001)	X			Effects of consumer, situational, and store characteristics on OOS response		Survey
Campo, Gijsbrechts, and Nisol (2003)	X			Consumers' behavioral response to OOS and impact on purchase quantity		Scanner panel data
Sloot, Verhoef, and Franses (2005)	X			Impact of brand equity and the hedonic level of the product on consumers' OOS response		Survey
Anderson, Fitzsimons, and Simester (2006)		X		Short- and long-run effects of OOS on purchase behavior	Communication strategy: message types; financial incentive for backorders	Panel data
Breugelmans, Campo, and Gijsbrechts (2006)		X		Impact of OOS policy on consumers' category purchase and choice decisions	Non-visible OOS; replacement recommendation	Purchase simulation
Zinn and Liu (2008)	X			Comparison of actual and intended OOS responses (delay, leave, switch store)		Survey
Ge, Messinger and Li (2009)	X			Effects of information about soldout products on consumer choice decision (e.g. deferal or make purchase).		Lab-based experiment
Jing and Lewis (2011)			X	Impact of nonfulfillment of online orders on consumers' subsequent purchase behaviors		Scanner panel data
Kim and Lennon (2011)		X		Effects of managerial response in mitigating the adverse impact of OOS on store image and behavioral intent	Communication strategy: message types; offer 10% discount for next order	Mock website experiment
Che, Chen, and Chen (2012)	X			Effects of OOS on consumer SKU preferences and price sensitivity		Scanner panel data
Puligadda *et al.* (2012)	X			Effects of the interplay of brand and store loyalty on OOS response		Survey + experiment
Helm, Hegenbart, and Endres (2013)	X			Antecedents of OOS response (e.g., store loyalty, presence of suitable alternatives)		Field survey
Diels, Wiebach, and Hildebrandt (2013)	X			Effects of promotion of a substitution on consumers' substitution decision		Online experiment
Pizzi and Scarpi (2013)		X		Effects of managerial response in mitigating the adverse impact of OOS on consumers' decision satisfaction and repatronage intentions	Communication strategy: timeliness and type of OOS announcement	Online experiment
Huang and Zhang (2016)	X			Effects of OOS noticed by consumers without a specific target option in mind on their preference among the in-stock options		Online + Lab-based experiment
Ma, Chen, and Zheng (2018)		X		Consumers' attitude toward OOS products and store		Online experiment

(*continued on next page*)

Table 1 (*continued*)

Authors	Prior-to-purchase OOS		PP-OOS	Study objectives	Strategy to mitigate negative OOS response	Method
	Offline	Online				
Peterson, Kim, and Jeong (2020)		X		Effects of framing product outage (out-of-stock vs. sold out or unavailable) on consumers' satisfaction and behavioral intentions		Online experiment
Gunness and Oppewal (2020)		X		Effects of consumer's mindset and familiarity with a website on OOS response (switching behavior)		Online experiment
Kumar, Shamar, and Tapar (2021)	X			Effects of OOS justifications, product type, and sales level on consumers' perception of products and switching behaviors	Communication strategy: types of OOS justification	Online + field experiment
Tian, Chen and Xu (2022)		X		Effects of the proportion of sold-out options in the choice set on consumer purchase choices		Online experiment
Present study			X	Effects of substitution policies in mitigating the adverse impact of PP-OOS on the probability of acceptance of substitution	Substitution strategy: dominant attribute matching; past purchase matching	Online purchase simulation experiment

Note: Sloot, Verhoef, and Franses (2005) review relevant studies published prior to 2000.

to PP-OOS, despite the regularity of such events in practice (Chan 2017).

Next to the OOS literature, the phantom decoy literature is a related, albeit different stream of relevant literature. This literature has examined the effect on choice decisions when a preferred option (phantom) is unavailable in a choice set of a few alternatives (typically, three), where each alternative can be described on two dimensions with a clear order of value, from below average/low score to excellent/high score (Hedgcock, Rao, and Chen 2009; Pratkanis and Farquhar 1992). Consumers in that case tend to base their choices on the similarity on the most important product attribute of the phantom to the remaining options (Adam, Wessel, and Benlian 2019; Evangelidis, Levav, and Simonson 2018; Friedman, Savary, and Dhar 2018). These predictions offer some insights, but the context is quite different from the post-purchase setting we consider. That is, for a PP-OOS, consumers cannot access or make the decision among alternatives in the choice set (i.e., unavailability happens *after* the purchase phase). Moreover, the set of remaining alternatives, from which the retailer makes the choice, generally is quite large, and the most important attribute in the category is not evident, nor is the order of attribute values (e.g., there is not one flavor outperforming other flavors).

In sum, this study contributes to the OOS and phantom decoy literature in several ways. First, this study advances prior OOS research to investigate PP-OOS, a situation in which retailers must select the substitution and compensate for the inconvenience of the OOS. Secondly, we establish novel theoretical and managerial evidence on substitution policy strategies related to how product attributes might determine the best substitution and which attributes dominate, depending on the type of category differentiation (vertical vs. horizontal). This study also clarifies how retailers can leverage data about consumers' past purchases to inform their substitution decisions and in which context using dominant attributes is more advantagous to past purchase substitution. The managerial outcome is a dashboard for retailers to implement a suitable substitution policy that satisfies consumers' needs, mitigates their dissatisfaction, and maintains sales.

Conceptual framework

Previous research has shown that when considering a substitution, consumers tend to focus on (i) the alternative's similarity to the item they intended to purchase (Campo, Gijsbrechts, and Nisol 2003; Müller and Diels 2016) and (ii) whether they are familiar with the alternative, such as due to prior experience with it (Bronnenberg and Vanhonacker 1996). Drawing on these key factors, we develop our conceptual framework in Figure 1, which underlies our investigation of the effects of dominant attribute similarity (policy 1), of having purchased the substitution in the past (policy 2), and their interaction. We also control for several covariates (see the Model Development section).

Policy 1: similarity of dominant attribute

Similarity is defined by common attributes or characteristics between a substitute and the OOS item (Tversky 1977). Phantom effects research has established that similar alternatives are preferred to dissimilar ones when the preferred item (phantom) is not available (Hedgcock, Rao, and Chen 2009, 2016; Pratkanis and Farquhar 1992). Similarity reduces cognitive decision-making effort (Fitzsimons 2000), by indicating an interchangeable choice option (Arens and Hamilton 2016;

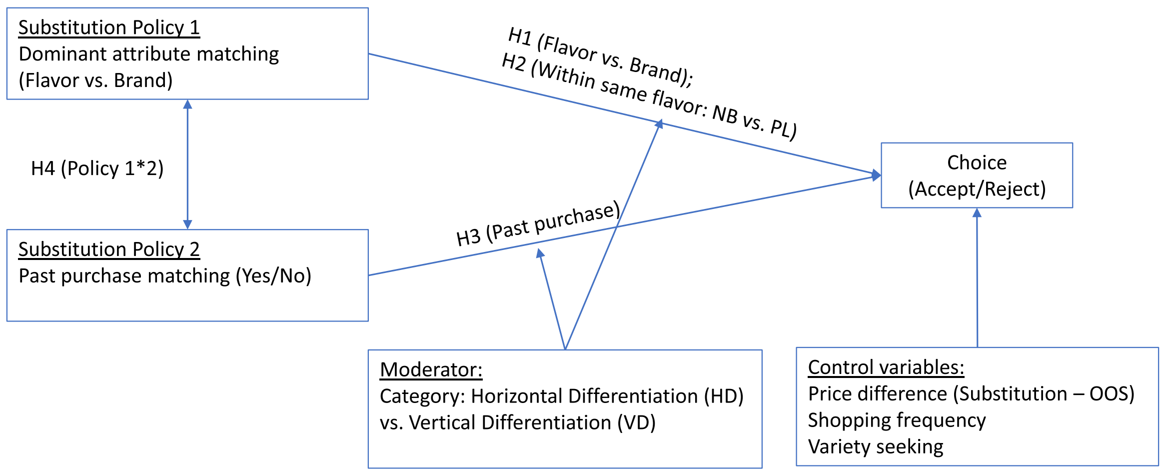

Fig. 1. Conceptual framework.

Huber, Payne, and Pluto 1982; Tversky 1972). If consumers must make a substitution choice quickly, they tend to select options that are easy to justify and reduce their likelihood of error or regret (Dhar and Simonson 2003).

However, similarity is not homogeneous; it might pertain to a wide range of attributes, such as flavor or brand (Arens and Hamilton 2016; Campo, Gijsbrechts, and Nisol 2003). A study by Bronnenberg, Dubé and Sanders (2020) has hinted that flavor and brand are the two attributes that play a focal role in consumers' product choices. Prior research has also found that the predominant reaction from consumers when facing an OOS is switching to another alternative, either a different brand or a different flavor (Campo, Gijsbrechts and Nisol 2000; Sloot, Verhoef and Franses 2005). Still, little is known whether consumers prefer to stick with the same brand but a different flavor or switch to the same flavor but another brand, and whether these preferences also depend on category characteristics. Prior research has suggested that in some categories, a particular attribute might carry more weight (Batra, Homer, and Kahle 2001; Gutman 1982), and in turn, consumers might prioritize these dominant attributes in their decision, such as emphasizing brand before flavor, or vice versa (Shao 2015).

We speculate that the attribute chosen to match the OOS item plays an important role in consumers finding the substitution similar. This determination also might depend on the category and its primary differentiation features. Two main differentiation strategies are vertical and horizontal (Hauser and Shugan 1983; Moorthy 1985; Sayman, Hoch, and Raju 2002; Shao 2015). *Vertical differentiation* (VD) occurs when products can be ordered easily according to the levels of quality or performance they provide (Sayman, Hoch, and Raju 2002; Spiller and Belogolova 2017). Consumers choose products in VD categories mostly on the basis of their perceptions of overall quality and price (Nowlis and Simonson 2000). Examples of VD categories include margarine and ketchup. They span a range of brands, positioned from standard to premium levels, each of which offers a few varieties (e.g., original and light) (Mintel 2019). Consumers may have a preferred brand or a level of acceptable quality and price when choosing certain products in a VD category depending on their value for money perceptions and willingness to trade off overall quality and price (Luce, Payne and Bettman 1999; Orchant 2013).

In contrast, *horizontal differentiation* (HD) exists if products in the category generally can be differentiated by attributes that cannot be objectively ordered, such as flavor. Diverse consumer tastes lead them to buy certain products rather than others (Spiller and Belogolova 2017), as might occur in categories involving cereals, pizza, or crisps (Hauser and Shugan 1983; Hotelling 1929). These categories often are dominated by a few brands, offering a wide range of flavors to target different taste preferences (Mintel 2020).[3]

In turn, we expect that in a VD category, the brand attribute, reflecting differences in value for money (Nowlis and Simonson 2000), dominates perceptions of whether a substitution is similar to the OOS item. In an HD category though, we anticipate that the flavor attribute dominates these perceptions. Then, the dominant attribute in each category should determine which substitution, as provided by the retailer, appeals more to the consumer. If the PP-OOS involves a product in a VD category, the substitute likely is accepted more if it matches on brand rather than on flavor, but in HD categories, the acceptance is likely higher if the products match on flavor rather than brand. Formally,

> *H1:* Substitutions of the same flavor (vs. brand) have a higher probability of being accepted in a predominant HD (vs. VD) category.

When offering a substitution that matches on brand, there is no real brand choice, as the substitution is from the same

[3] A VD category is not determined by the mere presence of PLs in the assortment. Nowadays, PLs are present in almost all grocery categories (Mintel 2021b). Instead, it is the relative dominance of the brand vs. the flavor attributes in a category that play a determining role in the VD vs. HD classification. In VD categories, most brands (including the PL brand) offer just a few flavors; in HD categories, each brand (including the PL) wants to offer a larger set of flavors.

brand as the OOS. This is different when offering a substitution that matches on flavor because the same flavor substitution comes from a different brand and can be from a different brand type, i.e. a private label versus a national brand. In this context, retailers must decide whether to offer a standard PL brand[4] or an NB of the same flavor. Retailers might be motivated to promote their own PL to customers as a substitution (CBinsights 2018) but the strength of consumers' preferences for PLs varies by category (Choi and Coughlan 2006; Sayman and Raju 2004), so we speculate and test for potential differences between VD and HD categories in this respect. We thus develop a hypothesis where we contrast the NB vs. PL substitution (in the same flavor condition) across VD and HD categories.

Retailers' PLs often seek to mimic competing NBs, such as by copying their packaging, typeface, labeling, and range of flavors (Sayman, Hoch, and Raju 2002). In a VD category, consumers tend to evaluate alternatives based on brand (Render and O'Connor 1976), and NBs usually are considered superior in terms of quality to PLs (Richardson, Dick, and Jain 1994). Accordingly, in a VD category, offering the same-flavor NB substitution might be preferred over a PL alternative. In contrast, in an HD category, preferences tend to based on the flavor, and the PL's mimicking strategy signals to consumers that it offers similar, comparable flavors, but with a better value for money proposition than the NB (Choi and Coughlan 2006; Sayman, Hoch, and Raju 2002; Schmalensee 1978). The same-flavor PL then should be more appealing to consumers in terms of value for money compared with same-flavor NB alternatives, in HD (vs. VD) categories. Therefore,

H2. Substitutions of same-flavor national brands (vs. same-flavor private labels) have a lower probability of being accepted in a predominant HD (vs. VD) category.

Policy 2: previously purchased items

Studies of choice heuristics suggest that prior experience with a product influences the types of information that people subsequently process when making comparisons and choices (Wedel *et al.* 1998). When consumers face an OOS situation and consider a familiar substitution that they purchased and consumed in the past, they can use their memory of that previous experience to aid their decision (Hoeffler and Ariely 1999). People tend to prefer familiar over unfamiliar products, because familiarity eases the choice (Fitzsimons 2000), whereas unfamiliar products create uncertainty and increase decision complexity (Dhar 1997). In addition, consumers often have a set of product options in mind that can satisfy their needs (Hamilton *et al.* 2014). Being added to the consideration set indicates some overall utility, which meets a consideration threshold established by consumers (Roberts and Lattin 1991). Therefore, substituting a product that the consumer has purchased previously, regardless of the category differentiation, should lead to greater acceptance than one the consumer has no experience with. In both VD and HD categories, consumers thus should be more likely to accept a substitution of a product they purchased in the past. So,

H3. In both predominant VD and HD categories, substitutions of products purchased in the past have a higher probability of being accepted than substitutions not previously purchased by the consumers.

Interaction of policies

In general, product familiarity (policy 2) should exert positive effects, but the positive influence of familiarity also might be strengthened with greater dominant attribute similarity (policy 1). As previous research indicates, familiarity influences how consumers use extrinsic and intrinsic cues in their product evaluations (Borgogno *et al.* 2015; Vizcaíno and Velasco 2019). Extrinsic cues are product-related attributes that are not part of the physical product (e.g., brand name); intrinsic cues are product attributes inherent to the objective nature of the product (e.g., a certain flavor), which consumers might need to experience to establish their preference (Szybillo and Jacoby 1974). In case of brand substitution, consumers make comparisons based on the brand, an extrinsic cue (Dodds, Monroe, and Grewal 1991; Richardson, Dick, and Jain 1994), which enables them to assess the appropriateness of that substitution relatively easily. Having bought the product in the past is therefore less insightful and helping. In contrast, for a flavor substitution, the evaluation relies more on intrinsic cues, gained from past experience (Hoeffler and Ariely 1999). Prior purchases of the substitution product therefore facilitate decisions and assessments of the appropriateness of the substitution. Hence,

H4. The positive effect of past-purchased substitutions on the probability of being accepted is stronger for same-flavor substitutions than for same-brand substitutions.

Data

To collect data for this study, we used an online purchase experiment that mimicked an online grocery shopping environment. We collected data across two waves. Wave 1 included two categories (cereals and margarine); Wave 2 featured three categories (pizza, crisps, and ketchup). Both waves rely on the same experimental procedures and design, so we discuss the method together, then present each wave's estimation results separately.

The computer-simulated shopping experiment offers several advantages over studies based on survey or scanner panel data (Breugelmans, Campo, and Gijsbrechts 2006; Müller and Diels 2016; Pecher and van Dantzig 2016). First, with a simulated experiment, we can implement manipulations (i.e., different substitution policies) affordably (Breugelmans, Campo,

[4] In almost all categories, a standard PL is available; economy and premium PL products are not as common (Mintel 2021b). Further research should address the implications of substituting these PL variations.

and Gijsbrechts 2006; Massara, Melara, and Liu 2014). Second, it prompts realistic buying behavior especially when cues (e.g., brand, price, assortments) closely resemble those of a real shopping environment (Burke *et al.* 1992; Campo, Gijsbrechts, and Guerra 1999). Yet, we still acknowledge openly that, despite our efforts to create a very realistic shopping environment, the simulated experiment remains hypothetical and cannot fully capture real-world scenarios.

In this section, we outline the assortment selection, substitution policy manipulations, experimental design, and sample characteristics for both waves. A full list of the assortments in all categories, a detailed overview of the substitution policy algorithm, and a clear outline of the survey flow, including all instructions and questions in the analyses, are available in the Web Appendix.

Product categories and assortments

In the purchase simulation, we presented shopping trips for five frequently purchased categories: breakfast cereals and margarine in Wave 1 and frozen pizza, crisps, and ketchup in Wave 2. Category characteristics, such as the average number of stockkeeping units (SKUs) per brand and the ratio of the number of flavors to the number of brands, serve as the cues for category differentiation (Hotelling 1929; Spiller and Belogolova 2017). Cereals, pizza, and crisps are typical HD categories, dominated by a couple of leading brands that each offer a wide range of flavors to target different taste preferences. Consumers' preferences tend to be based on their personal perceptions of the superior taste of a particular flavor. The metrics in Web Appendix Table WA.1 verify that these categories tend to offer a large number of SKUs per brand and exhibit a large average flavors-to-brands ratio. In contrast, margarine and ketchup are VD since they include a larger range of brands, clearly positioned from standard to premium, which offer relatively few flavors. Consumers' preferences tend to be based on perceptions of superior quality or value for money where consumers trade off quality and price (Orchant 2013). The metrics verify that these VD categories include fewer SKUs per brand and a much smaller average flavors-to-brands ratio.

The offered assortment in each category represents SKUs of brands with high market shares. In Wave 1, we included 23 SKUs (3 brands, 8 flavors, market share 74%[5]) of cereals and 16 SKUs (8 brands, 2 flavors, market share 87%) of margarine. In Wave 2, we listed 20 frozen pizza SKUs (4 brands, 9 flavors, market share 93%), 24 crisps SKUs (4 brands, 6 flavors, market share 82%), and 10 ketchup SKUs (4 brands, 3 flavors, market share 70%). Each category includes a set of NBs and one (hypothetical) PL. To avoid limiting the store in the experiment to one retailer (and thus increase response rates) or a biasing impact of prior experience with retail chains, we create a hypothetical retailer with its own PL brand, which we asked respondents to imagine was the one where they frequently shopped. Therefore, the PL products in the assortments were referred to as "Your supermarket's own-label [category] product" (e.g., "Your supermarket's own-label vegetable pizza").

Before each wave, we conducted a pretest to identify the category penetration of the selected assortment, involving 138 UK consumers for cereals and margarine in Wave 1 and 223 UK consumers for pizza, crisps, and ketchup in Wave 2. For all five categories, this penetration was very high. For example, 88% of pretest consumers bought at least one of the products in the cereal assortment in the 12 months prior to the experiment (92% for margarine, 84% for frozen pizza, 97% for crisps, 94% for ketchup). Thus, the assortments are relevant for most respondents, which enhances the realism of our purchase simulation.

Substitution policies manipulations

To manipulate the attribute matching substitution policy, we used either flavor or brand attributes. In the same-brand substitution condition, the substitutions represented the same brand but had a different flavor. For VD categories such as margarine and ketchup, in which most brands provide few flavors, this same-brand substitution is the one other flavor; for HD categories such as cereals, pizza, and crisps, most brands have multiple SKUs, so we paired the products with the flavor as closely as possible. When multiple substitutions were available, we randomly picked one. For example, the same-brand substitution for Kellogg's Special K might be Kellogg's Cornflakes or Kellogg's Wheats (same brand, neutral flavor). In such cases, we randomly assigned one. In the flavor matching condition, we provided a product with a similar flavor but a different brand. For example, Kellogg's Cornflakes would be substituted with Nestle Cornflakes or the PL version of cornflakes. If multiple NBs are available, we again randomly picked one of the possible options. By randomly assigning respondents to one substitution, even if multiple alternatives are available, we increase the robustness of our findings, such that they cannot depend on any particular substitute.

For the past purchase matching policy, we checked the substitutes assigned against the consumer's prior purchase portfolio, after the data collection. To capture whether consumers received a substitution that matched their prior purchases, we collected information about which products they bought in the 12 months before the experiment in the self-report portion of the questionnaire, administered in the presimulation stage (see Study design section). In total, we have six experimental conditions: 3 dominant attribute matching (same brand, same flavor NB, same flavor PL) × 2 past purchase matching (yes, no).

Study design

The design, adapted from Breugelmans, Campo, and Gijsbrechts (2006), consists of three parts: presimulation questionnaire, purchase and delivery simulation task, and postsimula-

[5] The market share estimation is based on the sum of the market share of individual brands in each assortment, obtained from Mintel market reports, 2019–2020, for the individual categories.

tion questionnaire. In the first part, we screen the respondents and collect necessary information for the simulation task. The screening questions confirm the respondents are the main grocery shoppers in their household, establish that they are familiar with online grocery shopping in general, and request their prior purchase portfolio in the focal categories. Specifically, respondents indicated which SKUs in the offered assortments they had bought in the previous 12 months. If they had bought products in more than one category, they were randomly allocated to complete the purchase task in one of the categories; if they only bought items from one category, they were assigned to it. Respondents who did not buy any SKUs in any category were thanked and excused, as were those who were not the main grocery shoppers in their household. We also asked about grocery shopping frequency, which we use as a control variable.

In the second part, the purchase and delivery simulation section, respondents were instructed to shop in the assigned product category as they would in real life, prompted by the message "Imagine you are making an online grocery order with a retailer you often shop with. [Category name] is one of the items on your shopping list. Please go to the next page to select the product you want to add to your basket." Because our objective is to test the impact of substitution policies following a PP-OOS, respondents only completed one purchase task in the assigned category. The next screen featured an online shelf, with images of all SKUs in the assortment, product names, and prices,[6] similar to real-world grocery websites. Respondents could select multiple SKUs by clicking on product images. The selected products then appeared on the next screen, as a confirmation of chosen items in the basket. Then respondents selected a delivery date and time for the grocery order to be delivered to their homes and checked out (without any actual payment procedure). An order confirmation appeared, noting that their order would be delivered on their chosen date and time. These steps in the simulation help increase its realism and engage respondents with the online shopping task.

After they completed the shopping task, respondents had to imagine that, prior to the delivery time, they received a notification that the item they ordered (or one of them, if the respondent ordered multiple items) was out of stock, and the retailer had selected a substitution. If the respondent ordered multiple items, the simulation randomly chose one item to be OOS. The notification read: "The item you ordered was: [name of OOS item]. The substitution item given is: [name of substitution]." Both OOS and substitution items were presented in detail, including product name, image, and shelf price, exactly as respondents saw them when they placed the order. The chosen substitution was randomly selected to reflect one of the three attribute matching conditions: same brand, same flavor NB or same flavor PL. We used the prior purchase portfolio, captured in the presimulation phase, to assess posterior the past purchase matching. After seeing the OOS and suggested substitution, respondents indicated if they wanted to accept or reject the substitution. Finally, following this decision, respondents completed a short postsimulation questionnaire, which included measures that capture the respondents' perceptions of the received substitution (distributive fairness, quality and value perception), and their perceptions of the retailer (trust, satisfaction), the share-of-wallet among products purchased in the past and some sociodemographics (see Web Appendix, Table WA.4).

Sample

In the first wave, we gathered responses from 2,113 respondents, related to cereals (n = 1,154) and margarine (n = 959); in the second wave, we questioned 1,292 respondents, pertaining to pizza (n = 371), crisps (n = 341), and ketchup (n = 580). The sample largely reflects the general profile of UK online grocery shoppers in terms of demographics (Statista 2020, 2022), as Table 2 shows.

Model development

To test the effects of the two substitution policies on acceptance probability, we employ a binary logit model, in which the probability that household h accepts (vs. rejects) a substitute item i as a replacement for the OOS item j is:

$$P_i^h(accept) = \pi_i = \frac{\exp\left(U_i^h\right)}{1+\exp\left(U_i^h\right)},$$

where P_i^h is the probability of household h accepting substitute item i, and U_i^h is the deterministic portion of utility that household h obtains from accepting substitute item i.

For parsimony, we estimate one pooled model per wave to test the hypotheses, with the following equation:

$$\begin{aligned} U_i^h = {} & \beta_0 + \beta_{FL} FL_{i-j}^h + \beta_{FLNB} FLNB_{i-j}^h + \beta_{PP} PP_i^h \\ & + \beta_{HD} HD^h + \beta_{FL*HD}(FL_{i-j}^h * HD^h) \\ & + \beta_{FLNB*HD}\left(FLNB_{i-j}^h * HD^h\right) + \beta_{PP*HD}(PP_i^h * HD^h) \\ & + \beta_{FL*PP}\left(FL_{i-j}^h * PP_i^h\right) + \beta_{Control} Controls^h + \varepsilon, \end{aligned}$$

where FL_{i-j}^h is a dummy variable equal to 1 when household h receives a substitute item i that has a similar flavor as the OOS item j and 0 if the substitute item is from the same brand; $FLNB_{i-j}^h$ is a dummy variable equal to 1 when household h receives a substitute item i that is another NB with the same flavor as the OOS item j, and 0 otherwise; PP_i^h is a dummy variable equal to 1 when the substitute item i was purchased in the past by household h and 0 if household h did not buy the substitute item i in the past; and HD^h is a category differentiation dummy variable that equals 1 when household h got assigned to a HD category, whether cereals (Wave 1), pizza, or crisps (Wave 2), but 0 if household h got assigned to a VD category, whether margarine (Wave 1) or ketchup (Wave 2).

[6] The prices of the SKUs represent regular retail shelf prices charged by a large UK grocery retailer at the time of the experiment.

Table 2
Sample representativeness.

	Wave 1 (n = 2,113)	Wave 2 (n = 1,292)	Total (n = 3,405)	UK online grocery shoppers
Age				(Statista 2020)
Age group 18-34	24.9% (526)	14.0% (181)	20.8% (707)	29% (16-34)
Age group 35-44	18.8% (398)	21.6% (279)	19.9% (677)	19%
Age group 45-64	42.6% (900)	44.7% (576)	43.4% (1,476)	30%
Age group 65+	13.6% (288)	19.6% (253)	15.9% (541)	5%
Missing	(3)	(1)	(4)	
Gender				(Statista 2022)
Female	63.6% (1,342)	55.7% (718)	60.6% (2,060)	56%
Social status[a]				UK census 2011 (ONS 2011)
AB	34.4% (726)	41.8% (523)	37.1% (1,249)	22.17%
C1	24.9% (526)	19.9 (249)	23.1% (775)	30.84%
C2	8% (170)	10.1% (126)	8.8% (296)	20.94%
DE	32.7% (691)	28.2% (353)	31.0% (1,044)	26.05%
Missing		(41)	(41)	

[a] The British National Readership Survey (NRS) defines six categories of social grading: A = higher managerial, administrative, professional occupations; B = intermediate managerial, administrative, professional occupations; C1 = supervisory, clerical & junior managerial, administrative, professional occupations; C2 = skilled manual occupations; D = semi-skilled; E = unskilled manual occupations or unemployed. (Ipsos Mori: https://www.ipsos.com/ipsos-mori/en-uk/social-grade).

To test our hypotheses, we are mainly interested in the following coefficients. The coefficient of the term $(FL_{i-j}^{h} * HD^{h})$ captures the dominance of the flavor (brand) attribute in an HD (VD) category (substitution policy 1); we expect it to be positive (H1). The coefficient of the term $(FLNB_{i-j}^{h} * HD^{h})$ captures the difference of NB versus PL within the same-flavor substitution in the HD category, and it is expected to be negative (H2). The main effect of past purchase subsitution PP_{i}^{h} (substitution policy 2) is expected to be positive while the coefficient of the term $(PP_{i}^{h} * HD^{h})$ is expected to be non-significant, given preferences for a familiar substitution item, regardless of category differentiation (H3). Finally, the coefficient of the interaction effect between the two policies $(FL_{i-j}^{h} * PP_{i}^{h})$ is expected to be positive, such that familiarity with a substitute item bought in the past is especially important for same-flavor, rather than same-brand, matching (H4).

As control variables, we consider price difference, shopping frequency, and variety seeking. Price difference $(PRICEdf_{i-j}^{h})$ is a measure of the difference in regular retail price between the substitute item i and OOS item j for household h and is expected to have a negative effect on acceptance (Breugelmans, Campo, and Gijsbrechts 2006). Shopping frequency $(SHPFQ^{h})$ is an one-item measure of grocery shopping frequency of household h in an average week. It is commonly used as a control variable in OOS studies (e.g., Campo, Gijsbrechts and Nisol 2000; Sloot, Verhoef and Franses 2005), and likely has negative effects on the probability of substitution acceptance, because more frequent shoppers can more easily delay their purchase. Finally, variety-seeking behavior (VAR^{h}) is a behavioral measure that equals the total number of SKUs bought by household h in a category in the past 12 months, collected via the self-reported prepurchase simulation questionnaire. It is expected to have a positive effect on acceptance, in that variety-seeking consumers are more likely to accept substitutions (Van Trijp, Hoyer, and Inman 1996). Table 3 summarizes all the variables used in the model.

Results

Descriptive statistics

Table 4 lists the frequencies with which we observe impacts of each substitution policy, revealing substantial variation across categories and different policies. In line with our expectations, we find opposing trends for HD and VD categories: A substitution in HD categories is more likely to be accepted if it matches the flavor of the OOS (Wave 1 = 40.66%; Wave 2 = 71.84%) rather than the brand (Wave 1 = 20.97%; Wave 2 = 48.80%). A substitution in VD categories instead is more likely to be accepted if it offers the same brand (Wave 1 = 64.97%; Wave 2 = 71.01%) instead of the same flavor (Wave 1 = 55.50%; Wave 2 = 54.50%). Category differentiation also plays a role for same-flavor substitutions, such that NB substitutions are less acceptable than PL alternatives in HD categories (Wave 1 = 39.50% vs. 44.51%; Wave 2 = 69.24% vs. 76.40%), but the outcomes reverse for VD categories (Wave 1 = 58.70% vs. 47.83%; Wave 2 = 59.40% vs. 48.02%). The results involving the previous purchase matching policy show that respondents are more likely to accept a substitution that matches what they bought before (Wave 1 = 73.05%; Wave 2 = 84.04%), compared with a substitution not previously purchased (Wave 1 = 39.80%; Wave 2 = 56.24%). The differences across HD and VD categories are limited though. The descriptive statistics thus match our expectations, so we explicitly test them next.

Table 3
Variable descriptions.

Variable	Description	Adapted from	Wave 1		Wave 2	
			HD (cereal)	VD (margarine)	HD (pizza, crisps)	VD (ketchup)
FL_{i-j}^{h}	Dummy variable that equals 1 if household h receives substitute item i that has a similar flavor as the OOS item j, and 0 if the substitute item is from the same brand[a] as the OOS item. Flavor example: chocolate, honey, frosty or neutral (cereals); original, light (margarine) Brand example: Nestle, Kellogg's, Weetabix (cereals); Flora, Lupark (margarine) (see Web Appendix; Table WA.2).	Attribute asymmetry variable used by Campo, Gijsbrechts, and Nisol (2003)	M = .677 Std.Dev. = .468	M = .815 Std.Dev. = .388	M = .560 Std.Dev. = .497	M = .709 Std.Dev. = .455
$FLNB_{i-j}^{h}$	Dummy variable that equals 1 if household h receives substitute item i that is a national brand with the same flavor as the OOS item j, and 0 otherwise	Consideration of brand type in choice utility by Sayman, Hoch, and Raju (2002)	M = .520 Std.Dev. = .500	M = .576 Std.Dev. = .495	M = .336 Std.Dev. = .473	M = .403 Std.Dev. = .491
PP_{i}^{h}	Dummy variable that equals 1 if the substitute item i was bought by household h in the 12 months prior to the study and 0 otherwise. We collected the purchase history of household h via the self-reported prepurchase simulation questionnaire.	Last purchase variable in Campo, Gijsbrechts, and Nisol (2003)	M = .150 Std.Dev. = .355	M = .200 Std.Dev. = .403	M = .220 Std.Dev. = .416	M = .130 Std.Dev. = .332
HD^{h}	Dummy variable that equals 1 if household h is assigned to a horizontal differentiated category (cereals, frozen pizza, crisps), and 0 if household h is assigned to a vertical differentiated category (margarine, ketchup).	Category differentiation definition by Hotelling (1929); Spiller and Belogolova (2017)				
PRICEdf_{i-j}^{h}	The difference in regular retail price between the substitute item i and OOS item j for household h. $PRICEdf_{i-j}^{h} = Price_{i}^{h} - Price_{j}^{h}$	Price difference variable suggestion in Breugelmans, Campo, and Gijsbrechts (2006) and Campo, Gijsbrechts, and Nisol (2003)	M = .185 Std.Dev. = .819 Min. = -2.39 Max. = 2.50	M = -.200 Std.Dev. = .733 Min. = -2.25 Max. = .80	M = -.094 Std.Dev. = .636 Min. = -2.00 Max. = 2.00	M = -.155 Std.Dev. = 1.129 Min. = -2.95 Max. = 2.95
$SHPFQ^{h}$	One-item measure of grocery shopping frequency of household h in an average week: "How many times do you shop for grocery in an average week?" (1 = not shopping every week, 5 = shopping 4 times and more).	Based on shopping frequency by Sloot, Verhoef, and Franses (2005) and insights from market research by Mintel (2021a)	M = 3.046 Std.Dev. = 1.134 Min. = 1 Max. = 5	M = 3.066 Std.Dev. = 1.105 Min. = 1 Max. = 5	M = 2.785 Std.Dev. = 1.090 Min. = 1 Max. = 5	M = 2.757 Std.Dev. = 1.086 Min. = 1 Max. = 5
VAR^{h}	Behavioral measure equal to the total SKUs bought by household h in a category in the past 12 months, collected via the self-reported prepurchase simulation questionnaire. Participants indicated which products in the assortment they bought in the last 12 months. 70% of participants (n = 2,370) bought multiple SKUs in the assigned category.	Adapted from variety in a portfolio model by Kahn and Lehmann (1991)	M = 3.275 Std. Dev. = 2.523 Min. = 1 Max. = 23	M = 2.941 Std. Dev. = 2.171 Min. = 1 Max. = 16	M = 3.548 Std.Dev. = 2.863 Min. = 1 Max. = 24	M = 1.660 Std.Dev. = 1.095 Min = 1 Max = 10

[a] Flavor and brand substitution in most cases cannot coincide, because it is not possible to have a substitution with the same flavor and from the same brand, unless it changes the package size. For this study, a same-flavor substitution involves a different brand, and a substitution with the same brand provides a different flavor, because we only offer the most common pack size per brand

Table 4
Frequency distribution of substitution acceptance rate.

	Wave 1					Wave 2									
	HD		VD		Overall	HD						VD		Overall
	Cereals		Margarine			Pizza		Crisps		Pizza + Crisps		Ketchup		
	n	%	n	%	%	n	%	n	%	N	%	n	%	%
Overall	**1,154**	**34.32%**	**959**	**57.25%**	**45.79%**	**371**	**70.35%**	**341**	**54.84%**	**712**	**62.60%**	**580**	**59.31%**	**60.95%**
Same Flavor (FL)	**782**	**40.66%**	**782**	**55.50%**	**48.08%**	**168**	**79.17%**	**231**	**64.50%**	**399**	**71.84%**	**411**	**54.50%**	**63.17%**
FL + Past Purchase Yes	105	79.05%	163	84.05%	81.55%	49	93.88%	39	87.18%	88	90.53%	46	80.43%	85.48%
FL + Past Purchase No	677	34.71%	619	47.98%	41.35%	119	73.11%	192	59.90%	311	66.51%	365	51.23%	58.87%
FL + National Brand	600	39.50%	552	58.70%	49.10%	93	79.57%	146	58.90%	239	69.24%	234	59.40%	64.32%
FL + Private Label	182	44.51%	230	47.83%	46.17%	75	78.67%	85	74.12%	160	76.40%	177	48.02%	62.21%
Same Brand (BR)	**372**	**20.97%**	**177**	**64.97%**	**42.97%**	**203**	**63.05%**	**110**	**34.55%**	**313**	**48.80%**	**169**	**71.01%**	**59.91%**
BR + Past Purchase Yes	66	39.39%	32	84.38%	61.89%	51	82.35%	19	68.42%	70	75.39%	27	88.89%	82.14%
BR + Past Purchase No	306	16.99%	145	60.69%	38.84%	152	56.58%	91	27.47%	243	42.03%	142	67.61%	54.82%
Past Purchase (PP)														
Past Purchase Yes	171	61.99%	195	84.10%	73.05%	100	88.00%	58	81.03%	158	84.52%	73	83.56%	84.04%
Past Purchase No	983	29.20%	764	50.39%	39.80%	271	63.84%	283	49.47%	554	56.66%	507	55.82%	56.24%

Hypotheses testing

To build a pooled model per wave, we started with a model that includes policy 1 (i.e., dominant attribute matching; same flavor vs. same brand) and its interaction with the HD/VD dummy variable. Then we add the brand type dummy variable (i.e., same flavor from NB) and its interaction with the HD/VD dummy. Next, we include policy 2 (past purchase matching, yes/no) and its interaction with HD/VD. We then add the interaction of policies 1 and 2. Finally, we include the control variables. We present detailed results in Table 5. We also estimate models for the five product categories individually, using a similar model (yet, without the category differentiation dummy and its interactions), with the results in Table 6.

The gradual build-up of the model reveals that adding both policy effects increases the model fit, as does the inclusion of the control variables. The results remain robust though, without being affected by the inclusion of other variables. Multicollinearity is not a concern; the variance inflation factors are below 10 in both waves. Price difference (*PRICEdf*) and shopping frequency (*SHOPFQ*) do not have significant effects in either wave, perhaps due to the hypothetical situation, such that consumers did not have any financial stakes in the decision or place multiple orders over time. Variety seeking (*VAR*) exerts a significant positive effect on the probability of accepting the substitution in both waves ($W1\beta_{VAR} = .103$, $p = .000$; $W2\beta_{VAR} = .073$, $p = .038$). In what follows, we refer to the coefficients of the models in which we include the control variables.

Effects of dominant attribute substitution policy (H1)

We find similar effects for substitution policy 1 across both waves: significant negative main effects of the same-flavor substitution policy ($W1\beta_{FL} = -.730$, $p = .000$; $W2\beta_{FL} = -.758$, $p = .007$) and a significant positive interaction effect between the same-flavor substitution policy and the HD category dummy ($W1\beta_{FL*HD} = 2.078$, $p = .000$; $W2\beta_{FL*HD} = 1.985$, $p = .000$). That is, substitutions of the same flavor have a higher probability of being accepted, relative to substitutions of the same brand, in an HD category than in a VD category. In both waves, we confirm H1.

Combining the main and interaction effects, we find that in VD categories, consumers are more responsive to a substitution of the same brand rather than the same flavor (increases of 18% in Wave 1 and 19% in Wave 2, all else being equal). In HD categories, consumers are more responsive to a substitution of the same flavor than to one of the same brand (increases of 26% in Wave 1 and 30% in Wave 2).[7] That is, brand is the dominant attribute to be used in suggesting substitutions in VD categories, in which brands signal the overall quality and price acceptable to the consumers, and flavor is the one to be used in HD categories, in which the value of products cannot be objectively determined, and consumer preference heterogeneity for a certain flavor prevails.

[7] We use the model with control variables (Table 5) to calculate the probability of substitution acceptance. To understand how acceptance probability changes for substitution policy 1, we compare acceptance probability per wave for the same-brand substitution with the acceptance probability for the same-flavor substitution, once for a HD category and once for a VD category. We change the parameters of the same flavor dummy and the category differentiation dummy, keeping all other variables at 0. For example, for a VD category in Wave 1 involving a same-flavor substitution, the probability of acceptance is (exp(.44 – .73)/(1 + exp(.44-.73)) = 43%, but if the substitution is from the same brand, the probability of acceptance is (exp(.44)/(1 + exp(.44)) = 61%, where .44 is the constant, and -.73 is the parameter estimate for same flavor. Comparing the two probability estimates, we see an increase of 18% (from 43% to 61%) if the substitution and OOS match on the dominant attribute (brand in a VD category). We repeat these calculations for an HD category in Wave 1 and find the probability of acceptance is 41% (= exp(.44 – .73 + 2.078 – 2.166)/(1 + exp(.44 – .73 + 2.078 – 2.166)) for the same-flavor substitution and 15% (= exp(.44 – 2.166)/(1 + exp(.44 – 2.166)) for the same-brand substitution, or an increase of 26% (from 41% to 15%) when the substitution and OOS match on the dominant attribute (flavor in an HD category). We apply the same procedures to Wave 2 (and subsequent calculations in this section).

Table 5
Pooled model estimation results per wave: Hypotheses testing (H1–H4).

		Wave 1 (n=2,113)					Wave 2 (n=1,292)				
		Policy 1	Policy 1' (Policy 1 + FLNB dummy)	Policy 1' + 2	Policy 1' + 2 + Interaction	Policy 1' + 2 + Interaction + control	Policy 1	Policy 1' (Policy 1 + FLNB dummy)	Policy 1'+ 2	Policy 1' + 2 + Interaction	Policy 1' + 2 + Interaction + control
Same Flavor (vs. Same Brand)		–.397**	–.705**	–.649**	–.720**	–.730**	–.715***	–.975***	–.941***	–.955***	–.758**
		(.173)	(.206)	(.211)	(.213)	(.235)	(.196)	(.227)	(.230)	(.233)	(.279)
Same Flavor × HD	**H1(+)**	1.346***	1.811***	2.021***	1.954***	2.078***	1.473***	2.020***	2.076***	2.070***	1.985***
		(.227)	(.284)	(.296)	(.295)	(.298)	(.252)	(.314)	(.321)	(.322)	(.327)
HD vs. VD		–1.945***	–1.945***	–2.123***	–2.066***	–2.166***	–.774***	–.774***	–.924***	–.923***	–1.033***
		(.203)	(.203)	(.218)	(.215)	(.219)	(.204)	(.204)	(.214)	(.214)	(.221)
Same Flavor National Brand (FLNB)			.438**	.267	.259	.255		.460**	.454**	.454**	.168
			(.158)	(.164)	(.164)	(.195)		(.201)	(.204)	(.205)	(.314)
FLNB × HD	**H2(-)**		–.644**	–.588**	–.600**	–.730**		–.920**	–.968**	–.972**	–.817**
			(.233)	(.241)	(.242)	(.245)		(.306)	(.313)	(.313)	(.331)
Past Purchase				1.635***	1.035**	.887**			1.355***	1.230***	1.087***
				(.211)	(.331)	(.335)			(.331)	(.442)	(.443)
Past Purchase × HD	**H3(n.s.)**			.018	.196	.166			.220	.272	.225
				(.280)	(.293)	(.295)			(.412)	(.431)	(.431)
Same Flavor × Past Purchase	**H4(+)**				.713**	.636**				.172	.164
					(.314)	(.317)				(.413)	(.413)
Price difference						.069					.120
						(.081)					(.109)
Shopping FQ						–.069					.088
						(.043)					(.057)
Variety Seeking						.103***					.073**
						(.024)					(.035)
Constant		.618	.618	.397	.459	.440	.896	.896	.735	.745	.395
		(.158)	(.158)	(.163)	(.165)	(.215)	(.170)	(.170)	(.174)	(.176)	(.239)
R^2		*.099*	*.105*	*.196*	*.199*	*.210*	*.040*	*.050*	*.121*	*.121*	*.130*

*Significant at 10% level, **significant at 5% level, ***significant at 1% level.
Notes: Standard errors are in brackets.

Table 6
Estimation results per category.

	Policy 1	Policy 1' (Policy 1 + FLNB dummy)	Policy 1' + 2	Policy 1' +2 +Interaction	Policy 1' +2 + Interaction + Control
Wave 1					
Cereals (n=1154)					
Same Flavor (vs. Same Brand)	.949*** (.147)	1.106*** (.196)	1.372*** (.208)	1.211*** (.217)	1.440*** (.251)
Same Flavor National Brand (FLNB)		–.206 (.171)	–.321* (.176)	–.345* (178)	–.609** (.227)
Past Purchase			1.652*** (.184)	1.155*** (.294)	1.033** (.304)
Same Flavor × Past Purchase				.844** (.389)	.743* (.393)
Price Difference					.169* (.102)
Shopping FQ					–.016 (.059)
Variety Seeking					.089** (.030)
Constant	–1.327 (.127)	–1.327 (.127)	–1.726 (.145)	–1.586 (.152)	–1.856 (.252)
R2	*.053*	*.055*	*.151*	*.156*	*.169*
Margarine (n=959)					
Same Flavor (vs. Same Brand)	–.397**	–.705** (.206)	–.649** (.211)	–.691** (.218)	–.939** (.272)
FLNB		.438** (.158)	.267 (.164)	.262 (.164)	.514** (.240)
Past Purchase			1.635*** (.211)	1.252** (.516)	1.094** (.521)
Same Flavor × Past Purchase				.450 (.565)	.355 (.571)
Price Difference					–.115 (.135)
Shopping FQ					–.134** (.063)
Variety Seeking					.129** (.041)
Constant	.618 (.158)	.618 (.158)	.397 (.163)	.434 (.170)	.558 (.270)
R2	*.008*	*.018*	*.119*	*.120*	*.138*
Wave 2					
Pizza (n=371)					
Same Flavor (vs. Same Brand)	.800** (239)	.770** (.317)	.763** (.325)	.701** (.340)	.620 (.538)
FLNB		.055 (.382)	.061 (.390)	.062 (.392)	.209 (.718)
Past Purchase			1.420*** (.336)	1.276** (.402)	.919** (.425)
Same Flavor * Past Purchase				.455 (.748)	.603 (.755)
Price Difference					–.078 (.348)
Shopping FQ					.027 (.112)
Variety Seeking					.131** (.059)
Constant	.535 (.145)	.535 (.145)	.242 (.159)	.265 (.164)	–.197 (.371)
R2	*.044*	*.044*	*.124*	*.126*	*.146*
Crisps (n=341)					
Same Flavor (vs. Same Brand)	1.236*** (.243)	1.691*** (.319)	1.870*** (.332)	1.884*** (.345)	2.402*** (.407)
FLNB		–.692** (.299)	–.808** (.307)	–.804** (.307)	–1.450*** (.407)

(*continued on next page*)

Table 6 (*continued*)

	Policy 1	Policy 1' (Policy 1 + FLNB dummy)	Policy 1' + 2	Policy 1' +2 +Interaction	Policy 1' +2 + Interaction + Control
Past Purchase			1.682***	1.744**	1.869**
			(.374)	(.547)	(.557)
Same Flavor × Past Purchase				–.117	–.214
				(.745)	(.758)
Price Difference					.689**
					(.281)
Shopping FQ					.149
					(.115)
Variety Seeking					–.061
					(.051)
Constant	–.639	–.639	–.959	–.971	–1.207
	(.201)	(.201)	(.222)	(.235)	(.411)
R2	*.103*	*.123*	*.208*	*.208*	*.239*
Ketchup (n=580)					
Same Flavor (vs. Same Brand)	–.715***	–.975***	–.941***	–.942***	–.882**
	(.196)	(.227)	(.230)	(.238)	(.306)
FLNB		.460**	.454**	.454**	.380
		(.201)	(.204)	(.204)	(.358)
Past Purchase			1.355***	1.344**	.960
			(.331)	(.638)	(.652)
Same Flavor × Past Purchase				.015	–.178
				(.747)	(.760)
Price Difference					–.005
					(.132)
Shopping FQ					.020
					(.086)
Variety Seeking					.396**
					(.115)
Constant	.896	.896	.735	.736	.098
	(.170)	(.170)	(.174)	(.179)	(.326)
R2	*.032*	*.044*	*.090*	*.090*	*.120*

*Significant at 10% level, **significant at 5% level, ***significant at 1% level.
Notes: Standard errors are in brackets.

Effects of brand type within same-flavor substitution policy (H2)

Within the same flavor substitution policy, retailers can offer a same-flavor substitution from an NB or their PL. To test the brand type effects we predicted in H2, we use a same-flavor national brand (FLNB) dummy variable. We observe non-significant positive main effects (W1β_{FLNB} = .255, p = .191; W2β_{FLNB} = .168, p = .594), but, more importantly, we find significant negative coefficients for the interaction effect of the FLNB dummy with the category differentiation dummy, in both waves (W1$\beta_{\text{FLNB*HD}}$ = -.730, p = .003; W2$\beta_{\text{FLNB*HD}}$ = -.817, p = .013). This suggests that an NB (rather than PL) substitution of the same flavor significantly lowers acceptance probability in HD categories, in support of H2.

Combining the main and the interaction effects, we find that in VD categories, the same-flavor substitution NB achieves a slightly higher chance of being accepted than the same-flavor PL (acceptance rate increases of 6% in Wave 1 and 4% in Wave 2). In HD categories, the acceptance rate increases more if the same-flavor substitution is a PL rather than a NB (increase of 11% in Wave 1 and 16% in Wave 2). This finding indicates that consumers care less about the potentially lower quality of a PL in an HD category for which flavor is the dominant attribute.

Effects of past purchase matching substitution policy (H3)

We observe a significant positive main effect of the past purchase matching policy (W1β_{PP} = .887, p = .008; W2β_{PP} = 1.087, p = .014) while the interaction of this dummy with the category differentiation dummy is not significant (W1$\beta_{\text{PP*HD}}$ = -.166, p = .573; W2$\beta_{\text{PP*HD}}$ = .225, p = .602). These results are in line with our expectation that substitutions matching shoppers' prior purchase portfolio are more likely to be accepted than those that do not match it, regardless of category differentiation. The results thus support H3. Acceptance probability increases by 19% and 18% in Wave 1 if a previously purchased item is suggested in HD and VD categories, respectively. These increases are 32% and 22% for HD and VD categories in Wave 2. Irrespective of category, consumers prefer familiar substitutions.

Interaction of substitution policies (H4)

To test H4, we consider the interaction term of the two substitution policies. The results show positive coefficients in both waves (W1$\beta_{\text{FL*PP}}$ = .636, p = .045; W2$\beta_{\text{FL*PP}}$ = .164,

p = .692). Past purchase matching of same-flavor substitutions have a higher probability of being accepted than past purchase matching of same-brand substitutions. However, this result is only significant in Wave 1, in partial support of H4. To gain more insights into the significant interaction for Wave 1, we calculate the change in acceptance probability for a previously purchased item versus one not purchased before for a substitution with the same flavor, then compare it with the parallel change involving a substitution with the same brand. In Wave 1, acceptance probability increases by 22% if the previously purchased substitute involves the same flavor, compared with a same-flavor substitute not bought previously, whereas this increase is only 18% for the previously purchased same-brand substitution. The result partly confirms our prediction that past purchase matching matters more for the flavor than for the brand attribute.

Individual category effects

We provide the parameter estimates of the main and interaction policy effects per category in Table 6; they are in line with the results of the pooled model. That is, we find positive coefficients for same-flavor substitutions in HD categories such as cereals (Wave 1), pizza, and crisps (Wave 2) and negative coefficients in VD categories like margarine (Wave 1) and ketchup (Wave 2). Regarding the effects of brand type, we observe that a same-flavor substitution of a national brand (FLNB) has positive effects for margarine and ketchup (VD categories) and negative effects for cereals, pizza, and crisps (HD categories). In contrast, the past purchase matching policy has positive effects in all categories. With regard to the interaction effects of the two policies, four of five categories exhibit positive coefficients; however, only the effect in the cereals category is significant.

Robustness checks

To confirm the validity of our findings, we conducted several robustness checks. First, we replaced the HD/VD dummy variable with two alternative metrics (ratio of flavors to brands and average number of SKUs per brand; see Web Appendix, Table WA.1 for details on the measures). The results remained stable for both alternative measures, except that the effect of brand type (H2) became insignificant in Wave 2 when we use the ratio (p = .128), though it stayed in the same direction.

Second, we test the sensitivity of our results to an alternative operationalization of the variety seeking control variable (Kahn and Lehmann 1991). Instead of the number of SKUs in the consumer's prior purchase portfolio, we use separate measures of flavor variety (number of flavors bought in the past) and brand variety (number of brands bought in the past). When we reran the pooled model for each wave, the results did not change.

Third, we replaced our past purchase dummy variable (1/0) with a share-of-wallet measure. In the postpurchase questionnaire, we asked respondents about their preference for previously purchased items, according to a share distribution of 100 points (see Web Appendix, Table WA.4). We reran the pooled models for each wave; the interaction between the two substitution policies (H4) became insignificant in Wave 1, but the other substantive findings remained identical.

Fourth, we used realistic assortments in the purchase experiment that included one PL brand in each category. But if the PL were OOS, there was no option for choosing a same-flavor substitution with another PL, which could affect the results. Therefore, we checked for the stability of our results using a reduced data set that excluded purchases of PL items that became OOS (n = 1,670 in Wave 1, n = 1,066 in Wave 2). The results remained robust.

Fifth, we ran a pooled model with all five categories and added a correction dummy variable for waves and its interactions with the coefficients used to test hypotheses in the pooled model. Except for the interaction of HD and the wave dummy, no other interaction effects were significant. The substantive findings also remained stable, except that we no longer find a significant effect for H4 (parameter was insignificant but in the same direction).

Discussion and conclusion

Using computer-simulated experiments with two representative samples of more than 3,000 UK consumers, we test the effectiveness of two substitution policies in response to PP-OOS that occur after consumers have placed their order, such that retailers must choose a substitute item to replace the OOS item. In contrast to a prior-to-purchase OOS situation where consumers are still in the purchase mode and have access to the remaining choice set for choosing a replacement when being confronted with a stock-out, consumers in an PP-OOS setting must rely on the retailer to select a suitable substitution on their behalf. Their motivation to accept or reject the substitution offered by the retailer is strongly influenced by the evaluation of that single substitution rather than the available assortments as is used in the prior-to-purchase OOS. Previous research claims that most shoppers prefer switching to another item, as a solution to an OOS event at the point of purchase, rather than delaying the purchase or switching stores (Breugelmans, Campo, and Gijsbrechts 2006; Campo, Gijsbrechts, and Nisol 2000, 2003; Sloot, Verhoef, and Franses 2005). Our findings add to this domain, by showing that in a PP-OOS situation, consumers show significant variation in their propensity to accept a substitution offered by the retailer.

We acknowledge upfront that our hypothetical setting could bias consumer responses, yet we do observe important *relative* shifts across scenarios. The descriptive statistics (Table 4) show that the acceptance rate changes depending on the substitution offered. The model results also reveal that implementing a suitable substitution policy, appropriate to the product category, can significantly increase the average probability of acceptance. Although our experiments involve a single purchase incident, without financial implications for the respondents, the variation in acceptance probability indicates that substitution policies can have relevant impacts.

First, with regard to the effect of similarity on a dominant attribute (flavor or brand), we find that similarity effects are not homogenous across categories. Substituting the same flavor is appropriate in an HD category, but in a VD category the substitution should reflect the same brand to increase acceptance. Extant literature has hinted that consumers rely on dominant attributes to evaluate products relative to alternatives in a category (Batra, Homer, and Kahle 2001; Wedel *et al.* 1998). Our study goes further to show that similarity on a dominant attribute positively affects acceptance of a substitute, but the dominant attribute varies between HD and VD categories. For predominantly VD categories, in which brands offer clearly distinct levels of quality performance and price, the brand is the dominant attribute; for predominantly HD categories, where consumers express individual taste preferences, flavor is the more dominant attribute.

Second, the role of this dominant attribute gets reinforced by brand type (NB vs. PL) when substitutions match on the flavor attribute. In HD categories, a PL substitution is preferred to a NB one, but the opposite is true in the VD category although less strong. These findings align with existing evidence in an HD category containing a range of flavors, a PL is considered a close copy of the leading NB (Choi and Coughlan 2006; Sayman, Hoch, and Raju 2002; Schmalensee 1978) and provides a more salient substitution, in terms of flavor similarity, than other NBs in the category, which actively market themselves as more distinctive (Sharp 2016).

Third, offering a substitution that shoppers have bought before increases the acceptance rate significantly and for all categories. This effect likely reflects the locus of familiarity (Block and Johnson 1995; Coupey, Irwin, and Payne 1998; Kumar and Gaeth 1991). That is, familiar products reduce uncertainty and disutility linked to the substitution. In addition, a prior purchase signals a consumer's consideration set, and products in a consideration set are more likely to be accepted than those outside it (Fitzsimons 2000; Fitzsimons and Lehmann 2004; Hauser 2014).

We note that these findings might reflect, as an underlying mechanism, perceptions of fairness, as hinted at in fairness exchange literature (Blodgett, Hill, and Tax 1997; Devlin, Roy, and Sekhon 2014). If PP-OOS represents a service failure that triggers both economic losses (i.e., consumers not getting what they order) and psychological losses (i.e., having to rely on the retailer to suggest a substitution) (Smith, Bolton, and Wagner 1999; Zhu, Sivakumar, and Parasuraman 2004), then a substitution that matches the dominant attribute (policy 1) or that has been bought in the past (policy 2) might appear like a more fair utility exchange and an indication that retailers have devoted adequate attention to selecting suitable substitutions (Namasivayam 2004; Seiders and Berry 1998).

In a posterior analysis, we compare perceived distributive fairness (measured in the postpurchase questionnaire, see Web Appendix, Table WA.4) across substitution scenarios. First, we compare consumers who received a substitution based on the dominant attribute (brand in VD categories, flavor in HD categories) with those who received a substitution based on the other attribute. Results show that distributive fairness is significantly higher in the former case, relative to the latter, though it is only significant in Wave 2 (Wave 1: MDominant = 3.739, MNon-dominant = 3.657, t(2111) = 1.007, p = .314; Wave 2: MDominant = 4.811, MNon-dominant = 4.281, t(1290) = 5.668, p = .000). Second, we compare consumers who received a substitution bought in the past and those who received a substitution not purchased before, and again, the results indicate significantly higher distributive fairness perceptions among the former group (Wave 1: MPastPurchase = 4.777, MNoPastPurchase = 3.467, t(2111) = 12.697, p = .000; Wave 2: MPastPurchase = 5.180, MNoPastPurchase = 4.369; t(1290) = 6.731, p = .000). A more elaborated analyses on mediation effects of distributive fairness and other potential mediators can be found in the Web Appendix, Table W.5. Still, we leave it to future work to fully unravel the underlying mediating mechanism.

Finally, we identify a significant interaction effect between the two substitution policies in Wave 1, such that positive past purchase policy effects grow stronger when the substitution is of the same flavor as the OOS item, rather than the same brand. This difference might exist because an attribute such as flavor requires intrinsic cues based on prior consumption experience to evaluate (Hoeffler and Ariely 1999). Although the results only partially support our prediction, they suggest some insights for further investigations into whether familiarity effects differ across consumers or categories.

Managerial implications

Online retailers can implement effective substitution policies in PP-OOS situations. To facilitate such applications, we prepared a dashboard that translates our results into managerial actions, as depicted in Figure 2.

The recommendations differ depending on whether the retailer needs to replace an OOS item in an HD or VD category. The ratio of the number of brands to the number of flavors or the number of SKUs per brand in a category can indicate category differentiation: HD categories tend to have a fewer brands, each of which offers more variaties (beyond those in our experiments, examples include yogurt, pasta, and tuna), whereas VD categories span many brands, each of which offers a couple of variaties (beyond those in our experiments, examples include shampoo, laundry detergent, orange juice, cooking oil).

As our results show, retailers should gather consumers' prior purchases to guide their selection of substitutions in a given product category. This information likely is readily accessible to online retailers, so they can achieve high consumer acceptance rates, especially if the substitution item also matches on the most dominant attribute for its category. For VD categories, retailers should always suggest a previously purchased item, even if from another brand but same flavor, because its presence in a consumer's consideration set implies that she or he perceives the brand as acceptable (Erdem *et al.* 2004). For HD categories, previously purchased items should be suggested if they offer the same flavor as the OOS item

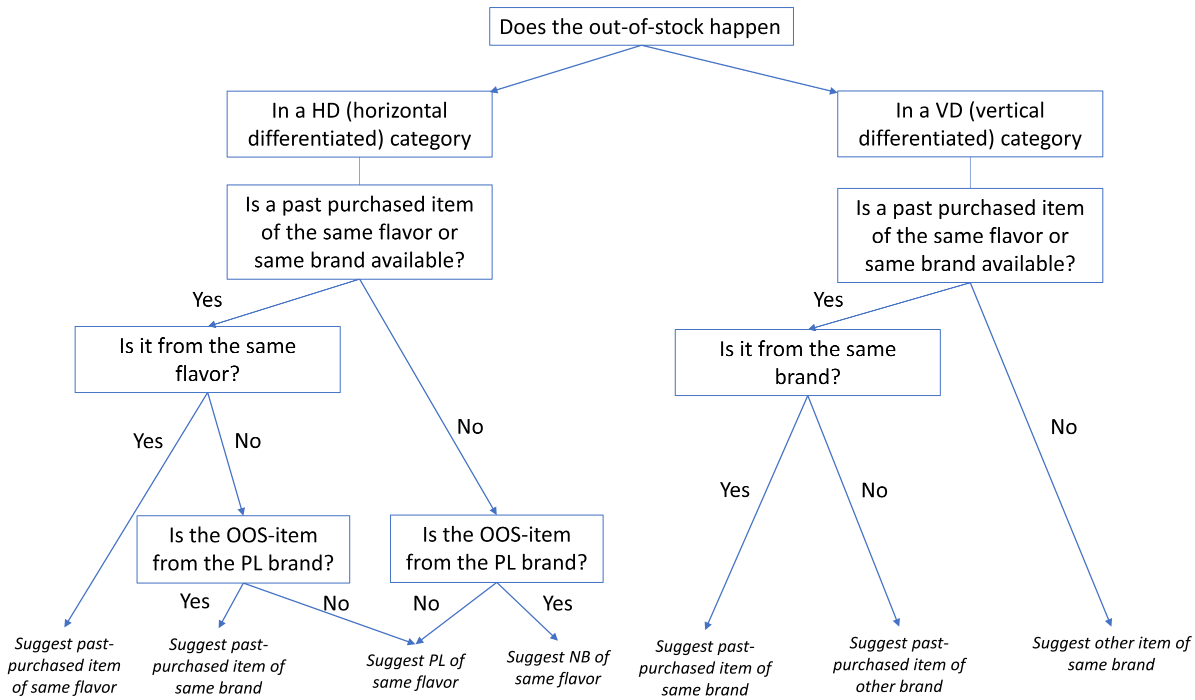

Fig. 2. Substitution policy pathway.

or if both the OOS and the previously purchased items come from the PL. If this is not the case or if no past purchased items that have the same brand or the same flavor as the OOS item are available, the dominant attribute should be the main decision factor.

The past purchase matching policy obviously cannot apply to new customers without any substantial prior purchase record or to very loyal customers whose purchase portfolios are limited to one SKU only. In the growing online grocery shopping sector, especially reflecting the effects of the COVID-19 pandemic, many consumers have only recently started to engage in online grocery shopping (Kantar 2021). Accordingly, in cases where no past purchased item is available of the same brand or the same flavor as the OOS item, retailers may need to rely on the dominant attribute matching policy. If the PP-OOS involves an HD category, they should suggest the same flavor as a substitute, but in a VD category, they should offer the same brand. In the former case, they also can benefit and increase acceptance if they offer a PL brand of the same flavor, if the OOS item is a NB. If the OOS item is the PL though, we advise offering another NB of the same flavor, because retailers do not sell another PL in their stores so another PL of the same flavor is not a viable option.

To test the effectiveness of the substitution policy pathway described above (and presented in Figure 2), we compare – for those consumers who have a known purchase in the category that retailers can use as a substitute – the change in average acceptance rate for a scenario where the retailer randomly picks a past purchased item as a substitute vs. a scenario where the substitute is picked using the most effective policy as suggested by our study. These results show that the average acceptance rate for this group of consumers significantly improves from 66% for the random pick to 75 % when following the substitution policy as suggested by our study. While the absolute acceptance rates must be treated with caution due to the hypothetical nature of the experiment, the relative difference between the two scenarios indicates our policy effectiveness. The improvement in acceptance rate will be even larger between following our substitution pathway and a random pick among the entire assortment when retailers cannot use consumers' past purchased records.

Limitations and further research

This research provides new insights into OOS response strategies in an online grocery retail context, but it also features some limitations. First, reflecting our effort to include realistic assortments in the experiments, rather than hypothetical ones, we could not test a very large set of categories or observe multiple purchase incidences. While our study offers a pathway for substitution decisions that most retailers can apply, large retailers may advance in the future to a more flexible predictive model using machine learning and artificial intelligence to accommodate a large number of categories and transactions. Subsequently, automation may also enable retailers to communicate with consumers instantly when PP-OOS occurs. In this case, multiple substitution options can be presented rather than relying on a single choice investigated in our study. We would also welcome research that examines a wider range of categories or a more realistic purchase setting, in an attempt to validate these substitution policy effects on acceptance. Second, the increased popularity of PL products (Hooker 2018) has prompted the development of different PL tiers, such as premium, standard, and economy (Geyskens *et al*. 2018). We only refer to standard PL and cannot fully generalize the findings to other PL tiers. Third, the experiments do not account for variations observed in reality, such as whether an OOS item was on sale or purchased for the consumer's own use or for others. Furthermore, in the experiments, consumers buy only one product, which keeps us from assessing whether and how the total basket size and multiple fulfillment failures (i.e., ratio of items OOS to the total basket) might exert effects. Continued studies could broaden the scope of the investigation to include these heterogeneous effects. Fourth, we looked into some potential mediators in posterior analyses (and report about these in the Web Appendix, Table W.5). We find indications that the perceived fairness of substitution might be most relevant to mediate substitution acceptance, still we leave detailed investigations of the underlying mechanisms to further research. Fifth and finally, we investigate substitution policy effects in a PP-OOS situation, where on the one hand we assume consumers are more likely to accept a retailer-suggested substitution if the OOS happens after they have completed the purchase phase, rather than if it occurs at the moment of purchase and where on the other hand, reactance may occur because consumers do not receive what they ordered and cannot choose a substitution themselves (Smith, Bolton, and Wagner 1999; Fitzsimons, 2000). It would be interesting to investigate whether the effects of the two substitution policies that we identify still hold in prior-to-purchase OOS situations too, which we leave to further research.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jretai.2022.06.006.

References

Adam, M., Wessel M. and Benlian A. (2019), "Of Early Birds and Phantoms: How Sold-out Discounts Impact Entrepreneurial Success in Reward-based Crowdfunding," *Review of Managerial Science*, 13 (3), 545–60.

Anderson, E.T., Fitzsimons G.J. and Simester D. (2006), "Measuring and Mitigating the Costs of Stockouts," *Management Science*, 52 (11), 1751–63.

Arens, Z.G. and Hamilton R.W. (2016), "Why Focusing on the Similarity of Substitutes Leaves a Lot to Be Desired," *Journal of Consumer Research*, 43 (3), 448–59.

Batra, R., Homer P.M. and Kahle L.R. (2001), "Values, Susceptibility to Normative Influence, and Attribute Importance Weights: A Nomological Analysis," *Journal of Consumer Psychology*, 11 (2), 115–28.

Block, L.G. and Johnson M.D. (1995), "The Locus of Context Effects on Product Proximity Judgments," *International Journal of Research in Marketing*, 12 (2), 121–35.

Blodgett, J.G., Hill D.J. and Tax S.S. (1997), "The Effects of Distributive, Procedural, and Interactional Justice on Postcomplaint Behavior," *Journal of Retailing*, 73 (2), 185–210.

Borgogno, M., Favotto S., Corazzin M., Cardello A.V. and Piasentier E. (2015), "The Role of Product Familiarity and Consumer Involvement on Liking and Perceptions of Fresh Meat," *Food Quality and Preference*, 44 (2015), 139–47.

Breugelmans, E., Campo K. and Gijsbrechts E. (2006), "Opportunities for Active Stock-out Management in Online Stores: The Impact of the Stock-out Policy on Online Stock-out Reactions," *Journal of Retailing*, 82 (3), 215–28.

Bronnenberg, B.J. and Vanhonacker W.R. (1996), "Limited Choice Sets, Local Price Response and Implied Measures of Price Competition," *Journal of Marketing Research*, 33 (2), 163–73.

Bronnenberg, B.J., Dubé J.-P. and Sanders R.E. (2020), "Consumer Misinformation and the Brand Premium: A Private Label Blind Taste Test," *Marketing Science*, 39 (2), 382–406.

Burke, R.R., Harlam B.A., Kahn B.E. and Lodish L.M. (1992), "Comparing Dynamic Consumer Choice in Real and Computer-simulated Environments," *Journal of Consumer Research*, 19 (1), 71–82.

Campo, K., Gijsbrechts E. and Guerra F. (1999), "Computer Simulated Shopping Experiments for Analysing Dynamic Purchasing Patterns: Validation and Guidelines," *Journal of Empirical Generalisations in Marketing Science*, 4 (2), 22–61.

Campo, K., Gijsbrechts E. and Nisol P. (2000), "Towards Understanding Consumer Response to Stock-outs," *Journal of Retailing*, 76 (2), 219–42.

Campo, K., Gijsbrechts E. and Nisol P. (2003), "The Impact of Retailer Stockouts on Whether, How Much, and What to Buy," *International Journal of Research in Marketing*, 20 (3), 273–86.

CBinsights. (2018). *Private Labels Rising: How Retailer's Own Products Are Taking Off And Transforming The CPG Industry*. CBinsights Available from https://www.cbinsights.com/research/private-labels-disrupt-cpg-retail/ [Accessed 8 July 2021].

Chan, S. (2017). 13 of the Worst Supermarket Substitution Fails – Revealed." *Which?* Available from https://www.which.co.uk/news/2017/03/13-of-the-worst-supermarket-substitution-fails-revealed/ [Accessed 22 January 2020].

Che, H., Chen X. and Chen Y. (2012), "Investigating Effects of Out-of-Stock on Consumer Stockkeeping Unit Choice," *Journal of Marketing Research*, 49 (4), 502–13.

Choi, S.C. and Coughlan A.T. (2006), "Private Label Positioning: Quality Versus Feature Differentiation From the National Brand," *Journal of Retailing*, 82 (2), 79–93.

Cocozza, P. (2019). *Are These the Worst Supermarket Substitutions Ever?*. The Guardian Available from: https://www.theguardian.com/lifeandstyle/shortcuts/2019/may/08/are-these-the-worst-supermarket-substitutions-ever [Accessed 05 Aug 2019].

Coupey, E., Irwin J.R. and Payne J.W. (1998), "Product Category Familiarity and Preference Construction," *Journal of Consumer Research*, 24 (4), 459–68.

Devlin, J.F., Roy S.K. and Sekhon H. (2014), "Perceptions of Fair Treatment in Financial Services Development, Validation and Application of a Fairness Measurement Scale," *European Journal of Marketing*, 48 (7-8), 1315–32.

Dhar, R. (1997), "Consumer Preference for a No-Choice Option," *Journal of Consumer Research*, 24 (2), 215–31.

Dhar, R. and Simonson I. (2003), "The Effect of Forced Choice on Choice," *Journal of Marketing Research*, 40 (2), 146–60.

Diels, J.L., Wiebach N. and Hildebrandt L. (2013), "The Impact of Promotions on Consumer Choices and Preferences in Out-of-Stock Situations," *Journal of Retailing and Consumer Services*, 20 (6), 587–98.

Dodds, W.B., Monroe K.B. and Grewal D. (1991), "Effects of Price, Brand, and Store Information on Buyers' Product Evaluations," *Journal of Marketing Research*, 28 (3), 307–19.

Erdem, T., Swait J., Iacobucci D., Mick D.G. and Huber J. (2004), "Brand Credibility, Brand Consideration, and Choice," *Journal of Consumer Research*, 31 (1), 191–8.

Evangelidis, I., Levav J. and Simonson I. (2018), "The Asymmetric Impact of Context on Advantaged versus Disadvantaged Options," *Journal of Marketing Research*, 55 (2), 239–53.

Fitzsimons, G.J. (2000), "Consumer Response to Stockouts," *Journal of Consumer Research*, 27 (2), 249–66.

Fitzsimons, G.J. and Lehmann D.R. (2004), "Reactance to Recommendations: When Unsolicited Advice Yields Contrary Responses," *Marketing Science*, 23 (1), 82–94.

Friedman, E.M.S., Savary J. and Dhar R. (2018), "Apples, Oranges, and Erasers: The Effect of Considering Similar versus Dissimilar Alternatives on Purchase Decisions," *Journal of Consumer Research*, 45 (4), 725–42.

Ge, X., Messinger P.R. and Li J. (2009), "Influence of Soldout Products on Consumer Choice," *Journal of Retailing*, 85 (3), 274–87.

Geyskens, I., Keller K.O., Dekimpe M.G. and de Jong K. (2018), "How to Brand Your Private Labels," *Business Horizons*, 61 (3), 487–96.

Gunness, A. and Oppewal H. (2020), "How Mindset and Store Familiarity Impact Online Stockout Responses," *International Journal of Retail & Distribution Management*, 48 (4), 326–47.

Gutman, J. (1982), "A Means-End Chain Model Based on Consumer Categorization Processes," *Journal of Marketing*, 46 (2), 60–72.

Hamilton, R.W., Thompson D.V., Arens Z.G., Blanchard S.J., Häubl G., Kannan P.K., Khan U., Lehmann D.R., Meloy M.G., Roese N.J. and Thomas M. (2014), "Consumer Substitution Decisions: An Integrative Framework," *Marketing Letters*, 25 (3), 305–17.

Hauser, J.R. (2014), "Consideration-set Heuristics," *Journal of Business Research*, 67 (8), 1688–99.

Hauser, J.R. and Shugan S.M. (1983), "Defensive Marketing Strategies," *Marketing Science*, 2 (4), 319–60.

Hedgcock, W., Rao A.R. and Chen H. (2009), "Could Ralph Nader's Entrance and Exit Have Helped Al Gore? The Impact of Decoy Dynamics on Consumer Choice," *Journal of Marketing Research*, 46 (3), 330–43.

Hedgcock, W.M., Rao R.S. and Chen H. (2016), "Choosing to Choose: The Effects of Decoys and Prior Choice on Deferral," *Management Science*, 62 (10), 2952–76.

Helm, R., Hegenbart T. and Endres H. (2013), "Explaining Costumer Reactions to Real Stockouts," *Review of Managerial Science*, 7 (3), 223–46.

Hoeffler, S. and Ariely D. (1999), "Constructing Stable Preferences: A Look Into Dimensions of Experience and Their Impact on Preference Stability," *Journal of Consumer Psychology*, 8 (2), 113–39.

Hooker, L. (2018). *How UK Shoppers Fell in Love with Own-label Groceries*. BBC Available from https://www.bbc.co.uk/news/business-44684306 [Accessed 8 July 2021].

Hotelling, H. (1929), "Stability in Competition," *Economic Journal*, 39, 41–57.

Huang, Y. and Zhang Y.C. (2016), "The Out-of-Stock (OOS) Effect on Choice Shares of Available Options," *Journal of Retailing*, 92 (1), 13–24.

Huber, J., Payne J.W. and Pluto C. (1982), "Adding Asymmetrically Dominated Alternatives: Violations of Regularity and the Similarity Hypothesis," *Journal of Consumer Research*, 9 (1), 90–8.

Jing, X. and Lewis D. (2011), "Stockouts in Online Retailing," *Journal of Marketing Research*, 48 (2), 342–54.

Kahn, B.E. and Lehmann D.R. (1991), "Modeling Choice Among Assortments," *Journal of Retailing*, 67 (3), 274–99.

Kantar. (2021). *Tentative First Steps Back to The Shops as UK Grocery Market Marks Lockdown Anniversary*. Kantar Available from https://www.kantar.com/uki/inspiration/fmcg/2021-wp-tentative-first-steps-back-to-the-shops-on-lockdown-anniversary [Accessed 8 July 2021].

Kim, M. and Lennon S.J. (2011), "Consumer Response to Online Apparel Stockouts," *Psychology and Marketing*, 28 (2), 115–44.

Kumar, V. and Gaeth G.J. (1991), "Attribute Order and Product Familiarity Effects in Decision Tasks Using Conjoint Analysis," *International Journal of Research in Marketing*, 8 (2), 113–24.

Kumar, M.E., Sharma D.P. and Tapar A.V. (2021), "Out-of-stock Justifications and Consumers' Behavioral Outcomes – Exploring the Role of Product Type and Sales Level Information in Out-of-stock Situations," *Journal of Retailing and Consumer Services*, 60 (2021), 102458.

Luce, M.F., Payne J.W. and Bettman J.R. (1999), "Emotional Trade-Off Difficulty and Choice," *Journal of Marketing Research*, 36 (2), 143–159.

Ma, K., Chen T. and Zheng C. (2018), "Influence of Thinking Style and Attribution on Consumer Response to Online Stockouts," *Journal of Retailing and Consumer Services*, 43 (July), 218–25.

Massara, F., Melara R.D. and Liu S.S. (2014), "Impulse Versus Opportunistic Purchasing During a Grocery Shopping Experience," *Marketing letters*, 25 (4), 361–72.

Mintel. (2019). *Yellow Fats and Oils UK - 2019*. London: Mintel Group Ltd.

Mintel. (2020). *Breakfast Cereals - Inc Impact of COVID-19 UK - August 2020*. London: Mintel Group Ltd.

Mintel. (2021a). *Online Grocery Retailing UK - March 2021*. London: Mintel Group Ltd.

Mintel. (2021b). *Attitudes Towards Private Label Food UK - May 2021*. London: Mintel Group Ltd.

Moorthy, K.S. (1985), "Using Game Theory to Model Competition," *Journal of Marketing Research*, 22 (3), 262–82.

Müller, H. and Diels J. (2016), "Reversing the Similarity Effect in Stock-Outs: A New Look at a Renowned Phenomenon in Consumers' Brand Switching Behavior," *Psychology & Marketing*, 33 (1), 48–59.

Namasivayam, K. (2004), "Action Control, Proxy Control, and Consumers' Evaluations of The Service Exchange," *Psychology & Marketing*, 21 (6), 463–80.

Nowlis, S.M. and Simonson I. (2000), "Sales Promotions and the Choice Context as Competing Influences on Consumer Decision Making," *Journal of Consumer Psychology*, 9 (1), 1–16.

ONS. (2011). *UK Census 2011 - Social Grade*. Office for National Statistics.

Orchant, R. (2013). *Best Butter Taste Test: Can You Tell The Difference Between Cheap And Expensive?*. Huffington Post Available from http://www.huffingtonpost.com/2013/05/31/best-butter-taste-test-cheap-expensive_n_3361785.html [Accessed 02 September 2017].

Pecher, D. and van Dantzig S. (2016), "Replication: The Role of Action Simulation on Intentions to Purchase Products," *International Journal of Research in Marketing*, 33 (4), 971–4.

Peterson, R.A., Kim Y. and Jeong J. (2020), "Out-of-stock, Sold Out, or Unavailable? Framing a Product Outage in Online Retailing," *Psychology & Marketing*, 37 (3), 428–40.

Pizzi, G. and Scarpi D. (2013), "When Out-of-Stock Products DO Backfire: Managing Disclosure Time and Justification Wording," *Journal of Retailing*, 89 (3), 352–9.

Pratkanis, A.R. and Farquhar P.H. (1992), "A Brief History of Research on Phantom Alternatives: Evidence for Seven Empirical Generalizations About Phantoms," *Basic and Applied Social Psychology*, 13 (1), 103–122.

Puligadda, S., Ross W.T., Chen J. and Howlett E. (2012), "When Loyalties Clash Purchase Behavior When a Preferred Brand is Stocked Out: The Tradeoff Between Brand and Store Loyalty," *Journal of Retailing and Consumer Services*, 19 (6), 570–7.

Render, B. and O'Connor T.S. (1976), "The Influence of Price, Store Name, and Brand Name on Perception of Product Quality," *Journal of the Academy of Marketing Science*, 4 (4), 722–30.

Richardson, P.S., Dick A.S. and Jain A.K. (1994), "Extrinsic and Intrinsic Cue Effects on Perceptions of Store Brand Quality," *Journal of Marketing*, 58 (4), 28–36.

Roberts, J.H. and Lattin J.M. (1991), "Development and Testing of a Model of Consideration Set Composition," *Journal of Marketing Research*, 28 (4), 429–40.

Sayman, S., Hoch S.J. and Raju J.S. (2002), "Positioning of Store Brands," *Marketing Science*, 21 (4), 378–97.

Sayman, S. and Raju J.S. (2004), "How Category Characteristics Affect The Number of Store Brands Offered by The Retailer: A Model and Empirical Analysis," *Journal of Retailing*, 80 (4), 279–87.

Schmalensee, R. (1978), "Entry Deterrence in the Ready-to-Eat Breakfast Cereal Industry," *The Bell Journal of Economics*, 9 (2), 305–327.

Seiders, K. and Berry L.L. (1998), "Service Fairness: What It Is and Why It Matters," *The Academy of Management Executive*, 12 (2), 8–20.

Severs, J. (2014). *Analysis: Breaking Down The Barriers of Online Retail*. RetailWeek September 30 2014.

Shao, X.-F. (2015), "Product Differentiation Design Under Sequential Consumer Choice Process," *International Journal of Production Research*, 53 (8), 2342–64.

Sharp, B. (2016). *How brands grow*. Oxford: Oxford University Press.

Sloot, L.M., Verhoef P.C. and Franses P.H. (2005), "The Impact of Brand Equity and The Hedonic Level of Products on Consumer Stock-out Reactions," *Journal of Retailing*, 81 (1), 15–34.

Smith, A.K., Bolton R.N. and Wagner J. (1999), "A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery," *Journal of Marketing Research*, 36 (3), 356–72.

Spiller, S.A. and Belogolova L. (2017), "On Consumer Beliefs about Quality and Taste," *Journal of Consumer Research*, 43 (6), 970–91.

Statista. (2020). *Share of Individuals Who Purchased Food or Beverages From Stores Online in Great Britain in 2019, By Age and Gender*. The Statistic Portal.

Statista. (2022). *Share of UK Consumers Who Usually Shop For Food 2020, By Gender*. The Statistic Portal.

Szybillo, G.J. and Jacoby J. (1974), "Intrinsic versus Extrinsic Cues as Determinants of Perceived Product Quality," *Journal of Applied Psychology*, 59 (1), 74–8.

Tian, J., Chen R. and Xu X. (2022), "A good way to boost sales? Effects of the proportion of sold-out options on purchase behavior," *International Journal of Research in Marketing*, 39 (1), 156–69.

Tversky, A. (1972), "Elimination by Aspects: A Theory of Choice," *Psychological Review*, 79 (4), 281–99.

Tversky, A. (1977), "Features of Similarity," *Psychological Review*, 84 (4), 327–52.

Van Trijp, H.C.M., Hoyer W.D. and Inman J.J. (1996), "Why Switch? Product Category–Level Explanations For True Variety-seeking Behavior," *Journal of Marketing Research*, 33 (3), 281–92.

Vizcaíno, F.V. and Velasco A. (2019), "The Battle Between Brands and Nutritional Labels: How Brand Familiarity Decreases Consumers' Alertness Toward Traffic Light Nutritional Labels," *Journal of Business Research*, 101 (2019), 637–50.

Wedel, M., Vriens M., Bijmolt T.H.A., Krijnen W. and Leeflang P.S.H. (1998), "Assessing The Effects of Abstract Attributes and Brand Familiarity in Conjoint Choice Experiments," *International Journal of Research in Marketing*, 15 (1), 71–8.

Zhu, Z., Sivakumar K. and Parasuraman A. (2004), "A Mathematical Model of Service Failure and Recovery Strategies," *Decision Sciences*, 35 (3), 493–525.

Zinn, W. and Liu P.C. (2001), "Consumer Response to Retail Stockouts," *Journal of Business Logistics*, 22 (1), 49–71.

Zinn, W. and Liu P.C. (2008), "A Comparison of Actual and Intended Consumer Behaviour in Response to Retail Stockouts," *Journal of Business Logistics*, 29 (2), 141–59.