Choice averse behavior and sampling risk: a field experiment with actual shoppers

DAVID ONG AND MENGXIA ZHANG[1]

April 14, 2018

How people react to more choices is of fundamental importance to the welfare implications of free markets and many government policies. A large body of research has attempted to demonstrate that people can exhibit choice averse behavior (CAB) when faced with many choices. However, meta-analyses of these studies (each of a small number of product lines) reveal conflicting results. Findings of CAB with some lines are balanced by findings of choice loving behavior (CLB) with others, with no predictor of which result occurs. The mean effect is zero. We posit that both CAB and CLB are moderated by unfamiliar shoppers' ex-ante belief about the risk of getting an ex-post undesired product. We constructed an estimator of such beliefs about sampling risk by surveying 1,440 shoppers for their "likes", "neutrals", "dislikes", and "tried" for 339 varieties across 24 product lines at a large supermarket. We then recorded 35,694 shoppers pass by, stop, and purchase after we randomly reduced the varieties they faced on shelves. As found in the meta-studies, we observed both CLB and CAB. The mean effect of reduced-variety is also zero. However, we show that both the probability and intensity of CAB/CLB is predicted by our measure of sampling risk. Our findings suggest the possibility of reconciling and predicting what had been regarded as conflicting results in the literature on CAB/CLB.

Keywords: field experiment, choice overload, risky decisions, choice averse

JEL Codes: C93, D81, M3

[1] We are grateful to United States National Science Foundation grant SES-08-51315 for financial support. Ong: Economics Department, Peking University HSBC Business School and Chinese University of Hong Kong, (corresponding author email: dvdong@gmail.com); Zhang: email: Mengxia.Zhang.2020@marshall.usc.edu, Marketing Department, University of Southern California. Earlier versions of this paper were circulated under the title "Can there ever be too many flavors of Häagen-Dazs?", which itself is an extension of Ong & Niu (2011) by the same title.

I. Introduction

How people react to more choices is of fundamental importance to the welfare implications of free markets and many government policies, e.g., health care (Ketcham, Lucarelli, & Powers, 2015). Standard economic theory implies that more choices cannot make people worse off. Among other things, more choices offer the possibility of better matches and greater flexibility. All else being equal, rational decision makers should exhibit choice loving behavior (CLB). However, a growing body of research in psychology, marketing, and economics suggest that people can at least act *as if* more varieties would make them worse off. Much of this literature was initiated by a field experiment with jams at a grocery store (Iyengar & Lepper, 2000). In it, shoppers were more likely to visit a special jam tasting display when it had 24 rather than six varieties. However, they were more likely to purchase at the shelves when only six varieties were at the display. Evidence supporting this "choice overload" hypothesis, where consumers probability of purchasing decreases on the number of varieties they face, which we will term "choice averse behavior" (CAB) hypothesis has come from many follow-up studies with different products (Chernev, Böckenholt, & Goodman, 2014; Scheibehenne, Greifeneder, & Todd, 2010).

Many theories have been put forward to explain CAB. These include actual psychological "overload" from the presumed increased psychological cost of ranking a larger number of varieties (Chernev et al., 2014; Scheibehenne et al., 2010). This psychological overload can be modeled as search costs (Kuksov & Villas-Boas, 2010). It can also be modeled as "contextual inference", where consumers are theorized to believe that their odds of success from sampling untried varieties is greater for product lines with fewer varieties, because the firm will prioritize the introduction of more popular varieties[2] (Kamenica, 2008; Kuksov & Villas-Boas, 2010). Given knowledge of the idiosyncraticness of their own preferences, consumers are also predicted to have an optimal finite number of varieties above which they experience CAB and below which they experience CLB (Kuksov & Villas-Boas, 2010). The existence of such an optimal number has been confirmed in a recent study exploiting voter self-selection into more or less refined rankings of candidates in Australian elections (Nagler, 2015).

[2] Such inferences help to explain why additional funds in an individual's 401(k) plan is associated with a greater allocation to money market and bond funds at the cost of equity funds (Iyengar & Kamenica, 2010). Such inferences may further influence the firm's choice of the number of varieties to offer because greater variety could signal higher levels of surplus extraction by the firm, given that a more tailored fit can sustain higher prices (Villas-Boas, 2009).

However, not all subsequent findings are confirmatory. A meta-analysis of 63 marketing studies (which are generally of one or two individual product lines) with N=5,036 subjects revealed a "mean effect size of virtually zero" for exogenously assigned number of varieties (Scheibehenne et al., 2010). But, since the treatments across the different experiments were not identical, it is not clear that they should be averaged to measure aggregate treatment effects (Chernev, Böckenholt, & Goodman, 2010). However, the same authors who point out this weakness later confirmed the net zero effect finding with a follow-up meta-study of their own, this time with N=7,202 (Chernev et al., 2014). This last meta-study reveals an additional four moderating factors: task difficulty, set complexity, preference uncertainty, and effort minimization goals, that could explain nearly 70 percent of the variation in past treatment effects (Chernev et al., 2014). To our knowledge, none of these theories have been used to actually predict CAB or CLB in the lab or in the field.

The strongest challenge to the CAB hypothesis comes from a large representative sample of millions of consumers who were found more willing to switch Medicare plans when choosing between more plans for a fixed decrease in the cost of the plans (Ketcham et al., 2015). The authors concluded that there was no such phenomena as CAB. Though, we think this conclusion premature, their finding at the least does highlight the potential limitations of prior tests of CAB/CLB.

Most marketing studies of CAB/CLB use unincentivized surveys of hypothetical choices over verbal descriptions of products in a laboratory setting with marketing or psychology students. However, the shopping choices of ordinary people in a grocery store may tend towards heuristical "system 1" choices (in the sense of Gigerenzer (2007) and Kahneman (2011)) rather than deliberative. The field experiments on CAB are rare and may have other validity issues. The original field experiment which initiated the literature used a special display to attract subjects. The display for the fewer (six) variety treatment was smaller, and thus, may have attracted "motivationally different consumers" than the larger (24) variety treatment (Iyengar & Lepper, 2000). The validity of the other field experimental study we are aware of has also been challenged on crucial methodological grounds (see footnote 43 in Kamenica (2008)). Most importantly, with respect to our hypothesis of differences in heterogeneity in taste across product categories, previous laboratory or empirical studies in both the marketing and the economics literatures tested only one or two product lines or services at a time.

In contrast to the psychological cost literature on CAB, but in line with the contextual inference literature, we focus on the consumer's reaction to the risk of mismatch between the consumer's taste and product characteristics. Given the potential for mismatch, a consumer's choice of whether to purchase unfamiliar varieties is in effect a gamble. As an example of how this risk of mismatch may vary across different product categories, people could be more particular about sour products than creamy products. There could be an evolutionary basis for such preferences; Sourness could be an indicator of toxicity (Breslin, 2013), the validity of which may vary with circumstance, e.g., climate. Should consumers have more idiosyncratic tastes for sour than creamy products, they face more risk of not achieving a required rate of success when they choose randomly over unfamiliar flavors of jam[3] than over unfamiliar flavors of ice cream. Moreover, consumers' reactions to such risks may be non-cognitive, as has been shown for financial professionals (Coates & Herbert, 2008).

We test this sampling risk conception of CAB/CLB with (to our knowledge) the first large scale field experimental study using a large number of shoppers making actual consumption decisions in an ordinary grocery store setting. Our main hypothesis is that there is an interaction between unfamiliar shoppers' belief about their rate of success in sampling, which varies across different product categories, and the number of varieties that they observe. Our prediction is that when shoppers' belief is pessimistic, the interaction is negative: more varieties decrease the probability of purchase. In contrast, when shoppers' belief is optimistic, the interaction is positive: more varieties increase the probability of purchase. We argue that this interaction can be interpreted in terms of the literature on risky assets or gambles. The basic idea is as follows.

We conceive of the varieties (e.g., vanilla Oreo cookies) that shoppers choose to sample as a set of repeated gambles[4]. To do this, we relax the assumption in prior theoretical work (Kamenica, 2008; Kuksov & Villas-Boas, 2010) that shoppers have unit demand. We introduce the assumption that each shopper has a required rate of success per product category or brand[5]. For example, heavy coffee drinkers may be satisfied with a low rate of success for new coffee

[3] Our hypothesis that there is a threshold level of sampling risk was in fact inspired by one coauthor's recollection of his disgust with rhubarb jam when told about the 24 varieties of jam in the original jam study (Iyengar & Lepper, 2000), and his realization that he had no such grounds to fear when choosing over untried flavors of Häagen-Dazs ice-cream.

[4] We assume that dynamic gambles can be modelled as a sequence of static gambles, which is our focus. Dynamic gambles would only introduce complex forms of updating, which would not, in our view, fundamentally change the incentives we model. Thus, we also avoid the intricacies of narrow and broad framing in the behavioral literature (Thaler & Tversky, 1997).

[5] We regard brands as sub-categories of products categories, to which our theory would also apply. For the purpose of ease of exposition, we make the convenient assumption that within-category (.e.g., cookies) across-brand (e.g., Oreo biscuits) heterogeneity in taste is smaller than across-category heterogeneity in taste, aggregating across brands.

varieties tried, because they will purchase many more units of each variety they end up liking. That greater purchasing intensity means that their stream of utility from these successful trials will compensate for the cost of a greater share of unsuccessful trials. Moreover, we posit that the shopper's decision to purchase within a given product brand/category will be based on the relation between their required rate of success and their believed rate of success.

When the shopper's belief about their rate of success is low with respect to their required rate of success, i.e., optimistic, the shopper should purchase more units (if they purchase at all) when they face more varieties in order to exploit the averaging effect of many repetitions of a positive expected value of gamble (e.g., 100 repetitions of a 50% chance of -$10 and 50% chance of $11 is to be preferred to 1 repetition). In contrast, when the shopper's belief about their rate of success is high relative to their required rate of success, i.e., pessimistic, the shopper should purchase fewer units (if they purchase at all) when they face fewer varieties, just as they would avoid the averaging effect of many repetitions of a negative expected value gamble[6]. Though, we do not observe shopper's required rate of success for a product category, those required rates of success may be randomly distributed across shoppers within a given product category. In that case, average shoppers' belief about their rate of success across different product categories alone may be sufficient to predict CAB and CLB across different product categories.

To make predictions based upon this sampling risk conception of CAB/CLB, we constructed a popularity measure of heterogeneity in taste across many different product categories (e.g., cookies) by surveying 1,440 ordinary shoppers at a large supermarket in Shenzhen (a city of 14 million), China, for their "likes", "neutrals", "dislikes", and "untried" for 339 varieties (e.g., vanilla) across 24 product lines (e.g., Oreo biscuits), which we use as proxies for product categories. We use the percentage of likes for a given product line, which is the count of likes divided by the sum of the count of likes, neutrals, and dislikes to estimate the odds that a random shopper among those considering purchasing the product, i.e., stopped in front of the shelf, will like a random variety in the product line. We then secretly observed 35,694 shoppers over a two week period who passed by the shelf, stopped in front of the shelf, or purchased an item, after we reduced the varieties they faced on the shelves.

---

6 The purchase of negative expected value gamble lottery tickets is a recognized phenomena. Though, there is no established explanation for such purchases (Barberis, 2007), it is obvious that people would never buy all or even most of the tickets, even on an unlimited budget, since that would result in a sure loss. Hence, with negative expected utility gambles, we expect that even those who do buy lottery tickets to recognize that less repetitions can be better than more.

Consistent with the conclusion of the meta-studies, we find that the reduction in variety has no net effect on the average probability of purchase *for the fixed set of varieties* available across both the reduced- and full-variety treatments. However, our treatments were identical addressing the issue of comparability of treatments. We identify an average level of sampling risk in taste (60 percent) above which more varieties lead to a higher rate of purchase (CLB), and below which more variety leads to a lower rate of purchase (CAB). We also show that CLB (CAB) is indeed due to decreased (increased) purchasing intensity (i.e., each shopper who purchases purchasing more units), as posited in our theory, rather than to a greater (fewer) number of shoppers purchasing, or to each shopper purchasing with merely greater (lower) probability.

Our findings suggest the possibility of a simple-to-measure predictor of CAB and CLB. Thus, our findings may help reconcile conflicting results of prior studies as special cases of risk averse[7] behavior in the context of different levels of sampling risk across different product categories. In that case, the four factors identified in the meta-study can also be interpreted from the perspective of potential mismatch due to heterogeneity in taste. Task difficulty, set complexity, preference uncertainty, and effort minimization goals, either measure the effort cost of matching or the risk of lost surplus from not matching. Even the apparently strongly disconfirmatory study of Medicare plan choice (Ketcham et al., 2015) is consistent with sampling risk being a factor as to whether more options lead to CAB or CLB. The authors also find that surveys of those consumers yielded a satisfaction rate of 85 percent. In light of our conceptual framework, such a high rate of satisfaction would predict CLB -- their main finding. Hence, our findings suggest that CAB is not inconsistent with the implication of standard economic theory that more choice is always better. Rather, more untried varieties are not just more choices, but more gambles.

II. Outline of conceptual framework

We test predictions based on our measure of sampling risk using different product lines, which are also across different product categories, to ensure sufficient variation in this measure. Thus, we generally use the product category and the product line in the category interchangeably, though "product line" more accurately describes the actual elements in our experiment. Also, the literature on choice overload generally compares purchases of all available varieties in each of

[7] Both expected utility theory and the theory of loss aversion (Rabin, 2000) predict that more repetitions should generally be preferred to less repetitions of a positive expected value gamble (Palacios-Huerta & Serrano, 2006; Rabin, 2000).

the full- and reduced-variety treatments. In contrast, we compare the probability of purchase for the *same* constant set of varieties (i.e., the varieties in the reduced-variety treatment) across the reduced-and the full-variety treatments. Moreover, within this fixed set of varieties, we make within-variety comparisons of the treatment effect of reduced-variety (e.g., comparing purchases of chocolate Oreo biscuits when there are 12 varieties available and when there are 6 varieties, including chocolate) because that both maximize comparability and statistical power since we have 339 varieties but only 24 product lines.

We collected two kinds of data: video data of shoppers purchasing behavior and in-store sales data, which the store management gave us access to. We approximate the standard measure of CAB by using our own video footage of shopper purchasing behavior at the shelf to measure the probability of purchase. We acquired in-store sales data because the shopper's body often obscures the view of our ceiling mounted cameras of the actual varieties they purchase. This in-store sales data allows us to rule out the potential substitution of shoppers who are familiar with the varieties they want to purchase towards their less preferred varieties in the reduced-variety treatment from their most preferred varieties in the full-variety treatment as the driver our findings.

Importantly for the testing of our hypothesis that shoppers are attempting to achieve a required rate of success in sampling untried varieties, we do not observe the individual shopper's required rate of success. However, if their required rates of success are exogenously or randomly distributed across shoppers for a given product line, we can use average beliefs about their rate of success p_i for a product line i to predict the average decision to sample in product line i. We define $buys_{i,j}$ as the number of purchases of variety j from product line i (using in-store sales data). We condition $buys_{i,j}$ on $stops_i$, which is the count of shoppers who stop in front of the shelf (for more than three seconds using video footage data) to derive the probability of purchase, which we measure by the buys per stop ratio for variety j of product line i

$$BS_{i,j} = \frac{buys_{i,j}}{stops_i}.$$

To avoid making strong functional form assumptions about CAB/COB behavior across product lines, we dichotomize outcomes and define the variable $CAB_{i,j}$ as an indicator of choice averse behavior for variety j in product line i for F=full-variety and H=reduced-(approximately half of the original number of varieties) treatments as

$$CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,H} > BS_{i,j,F} \\ 0, if\ BS_{i,j,H} \leq BS_{i,j,F} \end{cases}$$

We interpret CLB as the opposite of CAB: CLB = 1-CAB.

By definition, shoppers have outside options in the full-variety treatment choice set that are not included in the reduced-variety treatment set. Thus, the number of purchases of the fixed varieties in the reduced-variety condition can be greater than in the full-variety condition merely because shoppers who would have purchased the missing variety, now purchase the variety available in the reduced-variety treatment, creating the appearance CAB. To normalize for the fact that the odds of an individual variety in the reduced-variety treatment being purchased in the full-variety treatment is lower due to dilution by the other varieties that are only available in the full-variety treatment, we multiply the BS ratio in the full-variety treatment by the ratio of the number of varieties in the full-variety treatment over the number of varieties in the reduced-variety treatment to arrive at the

$$Adjusted\ BS_{i,j,F} = BS_{i,j,F} \cdot \frac{\#\ varieties\ in\ full\ treatment}{\#\ varieties\ in\ reduced\ treatment}.$$

and

$$Adjusted\ CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,H} > Adjusted\ BS_{i,j,F} \\ 0, if\ BS_{i,j,H} < Adjusted\ BS_{i,j,F} \\ null, if\ BS_{i,j,H} = Adjusted\ BS_{i,j,F} \end{cases}.$$

Given this setup, we now proceed to predictions about the average effect of reduced-variety for the average shopper over all of our product lines. We formulate our hypotheses in terms of increased varieties in order to be consistent with the conception of CAB in the literature.

Again, assuming that the required rate of success for individual shoppers are randomly distributed across shoppers and product lines, we can use the average of shoppers' belief to predict the average shoppers CAB/CLB from increased varieties. Thus,

Hypothesis I. CAB decreases with shoppers' belief about the rate of success of finding a product that they like (among the constant set of variety in a product line).

More formally, for a variety *j* in product line *i*, we predict

$$Pr(CAB_{i,j} = 1|p_i') \geq Pr(COB_{i,j} = 1|\ p_i)$$

Eq.(1)

for $p_i' < p_i$, where p_i = probability in success in sampling and is inversely related to the shopper's belief about heterogeneity in taste for product line *i*. Note that variety *j* must be among those that are available across both the full- and reduced-variety treatments (i.e., those varieties in the reduced-variety treatment).

The increase in risk from the increase in dispersion of outcomes due to greater variety should induce greater stratification between shoppers who would have purchased before the increase in variety. Those whose belief was barely above the required rate of success to purchase $(p_i \geq \bar{r}_i)$ before the increase in variety may no longer be confident enough to purchase. Those whose belief cleared their required rate of success by a wide margin $(p_i \gg \bar{r}_i)$ should increase the number of varieties they purchase. Again, we predict this for the same reasons that we predict that CLB is increasing on p_i in Hypothesis I. When shoppers' belief is optimistic $(p_i \gg \bar{r}_i)$, more averaging brings about higher expected returns from purchasing. However, since neither our video footage nor our store purchase data record varieties purchased per shopper, we make the following weaker prediction about the number of units shoppers purchase rather than the number of varieties shoppers purchase.

Hypothesis II. The units purchased per shopper in the full-variety treatment is larger than in the reduced-variety treatment (among the constant set of variety in a product line).

More formally, we hypothesize that for variety *j* in product line *i* for *F*=full- and *H*=reduced-variety treatments

$$Average\left(\frac{sales_{i,F}}{buyers_{i,F}}\right) \geq Average\left(\frac{sales_{i,H}}{buyers_{i,H}}\right) \quad \text{Eq.(2)}$$

where $sales_{i,treatment} = \sum_j buys_{i,j}$ and $buyers_{i,treatment}$ is the number of buyers of each product line *i*. Note that with $sales_i$, we sum $buys_{i,j}$ over all of the varieties *j* in product line *i* in the reduced-variety treatment, which is constant across both the full- and reduced-variety treatments, rather than over all varieties in each treatment (in particular, the full-variety treatment).

The increase in segmentation between shoppers who purchase was induced by the increase in variety should be greater in product categories with which the average shopper is already optimistic, with respect to the average required rate of success, than among product categories with which the average shopper is already relatively pessimistic[8]. Thus, we have the following hypothesis about the impact of the interaction between the variety treatment and shoppers' belief about sampling risk.

Hypothesis III. The increase in units purchased in the full-variety treatment is larger when the shoppers' believed rate of success is higher (among the constant set of variety in a product line).

More formally, we hypothesize that for variety *j* in product line *i* for *H*=reduced-and *F*=full-variety treatments,

$$Average\begin{pmatrix}\frac{sales_{i,F}}{buyers_{i,F}}(p_i\ )\\-\frac{sales_{i,H}}{buyers_{i,H}}(p_i)\end{pmatrix}\geq Average\begin{pmatrix}\frac{sales_{i,F}}{buyers_{i,F}}(p_i{}')\\-\frac{sales_{i,H}}{buyers_{i,H}}(p_i{}')\end{pmatrix}\qquad \text{Eq.(3)}$$

for $p_i{}' < p_i$ where $p_i =$ probability in success in sampling, and is inversely related to the shopper's belief about sampling risk for product line *i*. Importantly, this effect is not consistent with standard substitution effects, if the probability of success in sampling is interpreted as a measure of substitutability. In that case, the excluded varieties will make more of a difference when substitutability is low--the opposite of what is predicted here.

III. Experimental Design

The experiment was performed for the whole day on Wednesday, Thursday, and Friday in two weeks: September 11-13 (week 1) and September 25-27 (week 2), 2013 in a large supermarket in Shenzhen, a city of 14 million. As mentioned, previous studies of CAB or CLB tested *whether* exogenous changes in the number of varieties increase or decrease the probability of purchase for a small number of product lines. In contrast, we use shoppers' belief about rate of success to predict to *what degree* CAB/CLB occur. To estimate shoppers' belief about the probability of

[8] To elaborate, reiterating the argument for the optimal number of gambles, when beliefs about heterogeneity is low relative to the required rate of success, it is beneficial to sample more when choosing over more varieties, because sampling more decreases the variability of outcomes around the expected outcome and increases the odds of getting the required rate of success. In contrast, when heterogeneity is higher, it is less beneficial to sample more when choosing over more varieties, because sampling more decreases the variance but also decreases the odds of getting at least the required rate of success.

having a successful purchase (i.e., the population rate of success p_i for product category *i*), we asked them as they were exiting the supermarket to take a survey for a small gift (e.g., colored pens). In this survey, we asked about their "likes", "neutrals", "dislikes", and "untried" for each variety (e.g., chocolate, vanilla,…) in each of the 24 product lines[9] (e.g., Oreo biscuits). These were recorded with a 0 or a 1 in our survey. We used a binary response rather than one with more values to measure the intensity of preferences within individuals because our focus is on heterogeneity in taste, which is about differences across individuals.

We surveyed 30 shoppers for each treatment and each product line for two treatment days each. Hence, each product line was rated by 60 shoppers, for a total 24*30*2=1,440 independent responses. See Table I for a sample of the survey for the Oreo biscuits product line. Table II lists all product lines surveyed.

[Insert Table I and Table II here.]

We selected the products in Table II using two criteria: a high number of variety and easy re-shelving. We did not attempt to replicate Iyengar and Lepper's (2000) jam result, because jam, like bread and toast, is still a novelty in China. There were only a few varieties of any brand.

We randomly assigned each of our 24 product lines to the same day of the week (e.g. Wednesdays for Oreo biscuits) randomly to full- or reduced-variety in the two weeks of the experiment. We tested approximately eight product lines per day. In the reduced-variety treatments, the supermarket staff, who were uninformed of the hypotheses, moved the least popular varieties off the shelves, filling the rest of the shelf with the remaining varieties from additional stocks. We removed between one-half to one-third of the varieties, depending upon whether and where in the product line there was a steep drop-off in popularity (as measured by store sales data from one day before the experiment) to minimize the imposition on shoppers familiar with those products and who sought them specifically. See the Table A-XIII in Appendix III for the detailed list of the removed items, their share of the product line and their percent of sales per product line, and Table A-XIV, which shows that none of the correlations between these variables are significant. This lack of correlation suggests that the removed items would not affect our results. However, we include the share of the product line removed and its

9 We had to drop six product lines from the original design with 30 product lines due to unexpected problems with our mini- video cameras.

percent of sales in some specifications of the regressions of our main results in any case to further confirm that they do not affect results. We further discuss the regression specification and potential substitution effects of these removed products on our regressions below.

The store is open Monday to Thursday 8:00am-10:30pm and Friday to Sunday 8:00am-11:00pm. All experimental setup for the next day was done at night after the store closed. We installed cameras on the ceilings to record the $pass\ bys$, $stops$, and $buys$ (presumed if taken off the shelf). The floor manager may not have turned on the cameras exactly at store opening. Therefore, to avoid unobserved factors influencing our data, we counted shoppers during the highest traffic time, 5:30pm-9:30pm, and assume that this interval furnishes a representative sample of shoppers for the whole day. The store provided sales data for the whole day, which unlike the $buys$ from the video footage, allows us to reliably identify the varieties purchased within a product line (but not per product line per shopper). In total we observed 35,694 shoppers pass by, of whom 3,291 stopped, and 1,530 purchased.

We estimate the belief of shoppers as to their probability of success in sampling (p_i) using survey data of the rate of success of shoppers who have tried these varieties:

$$\%like_{i,week} = \frac{\#likes}{\#likes+\#neutrals+\#dislikes} = \frac{\#1s}{\#1s+\#0s+\#-1s} \qquad \text{Eq.(4)}$$

for product line i where $i = 1, 2, \dots, 24$ and *week*= $1,\ 2$. The probability is the average over the two treatment weeks:

$$p_i = \%like_i = \frac{\%like_{i,1}+\%like_{i,2}}{2}. \qquad \text{Eq.(5)}$$

$\%like_i$ can be interpreted as our estimate of the probability p_i that a person who would consider purchasing from product line i likes a random variety of product line i. We assume like Kamenica (2008) that consumers who are unfamiliar with product line i choose randomly from it, should they purchase any units at all. Such consumers would face the same risks as someone who was familiar but had a random variety selected for them[10].

We estimated the likely rate of success of those who were considering purchasing using the reported preferences of shoppers who had already purchased. Conceivably, this estimator could

---

[10] $\%like$, being a popularity measure of heterogeneity in taste, does not distinguish between the case where most people like just one variety (type 1 person likes variety 1, type 2 person likes variety 2….) and the case where one type of person likes all varieties, but most other types of persons do not like any variety. To our knowledge, the prior literature in marketing, which assumes and tests for unit demand, has not distinguished between these types of popularity. However, if tastes and incomes are distributed randomly among individuals, then we expect these two types of popularity will converge in outcomes.

be biased because those who consider purchasing (as measured by stopping) need not actually purchase. Those who purchase by definition have revealed a stronger preference. However, we estimate across-product lines variation in purchases between full- and reduced-variety treatments for a fixed set of varieties within each product line. Our estimator would be valid if the difference between those who merely consider and those who actually do purchase is constant across treatments and across product lines. Clearly, the cost of purchasing after stopping before the shelf is constant within a product line across both the full-and half-variety treatments since in both cases, it is the cost of taking something off the same shelf after stopping. The cost of purchasing should also be constant across product lines, because the isles are nearly identically structured for all product lines.

Figure I illustrates how $\%like_{i,week}$ is the area of 1s in the area of 1s, -1s and 0s. The table in Figure I summarizing the data from our actual surveys contain 30 columns, one for each shopper surveyed, and as many rows as there are varieties.

[Insert Figure I here.]

As an example, for Want-want QQ gummies, $\%like = 0.70$, while for Oreo biscuit, $\%like = 0.59$. According to our conceptual framework, we would then predict that shoppers are more likely to experience CAB (less likely to experience CLB) when choosing Oreo biscuits than when choosing Want-want QQ gummies. Table II displays $\%like$s for each product line we used. The correlation between the number of varieties within a product line and it's $\%like$ is insignificant (adjusted R^2=0.01, F-test p-value=0.62).

IV. Results

First, we test whether the reduction in the numbers of varieties has any effect on $stops_i$. We estimate the probability of $stops_i$ for the full- and reduced-variety treatments with Eq. (7).

$$\frac{stops_{i,treatment}}{passbys_{i,treatment}} = \alpha_0 + \beta_0 treatment + \beta_1 price_{i,treatment} + \beta_2 week + \varepsilon_{i,treatment} \quad \text{Eq.(6)}$$

Table III reveals that the coefficient for the treatment is not significant for any model. The implication is summarized in Observation I.

Observation I. The probability of $stops_i$ does not increase with variety.

One possible reason for the contrast between our finding in Observation I and Iyengar and Lepper's (2000) is that, as we mentioned, they used a separate display, the size of which varied with the number of varieties, as well as the regular shelf space, while we used only the regular shelf space, which was fixed in size for both the full- and reduced-variety treatments. Thus, their findings in light of ours confirms their conjecture that physically larger displays with more varieties may draw more shoppers who are less motivated to purchase.

[Insert Table III here.]

We next test the effect of the reduction in the number of varieties on purchasing behavior. Using the video footage we count the number of $buyers_{i,treatmeant}$, which is the number of buyers of each product line in each treatment from 5:30pm-9:30pm. We use this to calculate the probability of purchase per individual shopper per product line *i*, estimated in Eq. (8).

$$\frac{buyers_{i,treatment}}{stops_{i,treatment}} = \alpha_0 + \beta_0 treatment + \beta_1 price_{i,treatment} + \beta_2 week + \varepsilon_{i,treatment} \quad \text{Eq.(7)}$$

Table IV reveals that the coefficient for the variety treatment is not significant in any model.

[Insert Table IV here.]

Thus, we replicate the finding in the meta-studies that the average effect of variety on the probability of purchase is zero (Chernev et al., 2014; Scheibehenne et al., 2010). We find a similar lack of significance for the coefficient of the dummy variable for treatments when we use the average percent of the change in sales of all varieties in a product line in each treatment, not only those varieties that were constant across both treatments, controlling for average prices for the product line. Because the sample size of this regression is 24, we omit these results from the main text and make it available on request. We again find a similar lack of significance in the coefficient for the treatment dummy when we use raw sales for each product line instead of percentage change in sales. (Again, these results are available on request.) One potential explanation is that the required rate of success per product line for individual shoppers are randomly distributed across shoppers and product lines. We leave the testing of this conjecture to future study. Observation II summarizes our finding.

Observation II. Exogenous numbers of varieties do not have a significant effect on the probability of purchase.

We now discuss how we test Hypothesis I for the possibility that rate of success, as measured by $\%like_i$, predicts CAB or CLB – our main result. Table V exhibits an example of the value of $CAB_{i,j}$ for Oreo biscuits. $buys_{i,j}$ refers to the purchases of variety j from product line i per treatment using store data. Again, we do not use data from the video footage here because purchases of individual varieties were often obscured by the purchasing shoppers in those.

[Insert Table V here.]

We collect the definitions of all of the variables for our main estimation of (Equation (9)) in Table VI for the convenience of the reader.

[Insert Table VI here.]

As mentioned, in order to maintain comparability of the data across our treatments, we had to drop some data points. For example, we dropped varieties that were not in the reduced-variety treatments. See the notes of Table VII for the specific rationals. The summary statistics in Table VII indicate that the statistical properties of our variables were preserved nonetheless.

[Insert Table VII here.]

We test Hypothesis I by estimating the probability of choice averse behavior ($COB_{i,j} = 1$) for variety *j* in product line *i* as a function of $\%like_i$, controlling for price changes[11], in Eq. (9).

$$Pr\left(CAB_{i,j} = 1 \middle| \%like_i, price\ change_{i,j}, D_2, \ \dots D_{23}\right)$$

$$= \Phi(\alpha_0 + \beta_0 \%like_i + \beta_1 \cdot price\ change_{i,j} + \textstyle\sum_{k=2}^{23} \beta_k \cdot D_k \cdot price\ change_{i,j}) \quad \text{Eq.(8)}$$

We add dummies D_k multiplied by $price\ change_{i,j}$ to control for the possibility of variations in the demand curves for different product lines. We use the $\%like_i$ measure, which is at the product line level, to predict relative rates of purchases ($CAB_{i,j}$) within a specific variety j of a product line i across the full-and reduced-variety treatments. Excluding varieties could induce

[11] Unbeknownst to us, the store changed the prices of some of the items across our two experiments.

shoppers to substitute to a variety in the reduce-variety treatment in product line *i* or not purchase at all. Moreover, the rate of substitution can vary across product lines according to $\%like_i$, if it is also interpreted as measure of substitutability, in so far as a higher $\%like_i$ indicates a greater share of products liked by everyone. We show how the sampling risk effect of $\%like_i$ is separable from the substitutability effect in our empirical analysis. We cluster standard errors by product lines in all models to allow different product lines to have different standard errors in our estimations.

Table VIII displays the main result from Probit regressions, which we use because $CAB_{i,j}$ is a binary variable. Consistent with our prediction, the coefficient for $\%like_i$ is negative. This shows that CAB is more likely when sampling risk is higher ($\%like_i$ is lower). None of the alternative measures: $\%dislike_i$, $\%tried_i$, $relativ\ size_i$ and $\%sales\ of\ the\ removed\ varieties$ (our measure of the possible influence of the removed varieties) are significant. Notably, $\%tried_i$, our measure of familiarity, is not significant. This lack of significance is likely because, as mentioned, we removed the least popular varieties in the experiment in order to minimize the imposition on familiar shoppers. Downward sloping demand predicts the positive coefficient for $price\ change_{i,j}$; shoppers purchase fewer units in the full-variety treatment when prices increased for that treatment.

[Insert Table VIII here.]

Table IX shows that both the magnitude and significance of the findings in Table VIII are preserved when we *Adjusted* CAB to take into account that the probability of purchase of any specific variety is lower in the full-variety treatment.

[Insert Table IX here.]

The number of varieties in the full-variety treatment show only marginal significance in model (7) and (8) in Table VIII but not in Table IX. Thus, our results are weakly consistent with the hypothesis that people are overwhelmed when faced with more choices. However, we also show that $\%like_i$ is a better predictor of when CAB or CLB occur.

Figure II also illustrates the main result from model (1) in Table VIII. CAB is decreasing on $\%like_i$, when prices and other control variables are held constant.

[Insert Figure II here.]

Figure II illustrates that this average required rate of success is $\%like$ = 0.60. Our main finding exhibited in Table VIII, Table IX, and Figure II is summarized in Observation III, which confirms the prediction in Hypothesis I.

Observation III. CAB decreases with shoppers' belief about the rate of success of finding a product that they like.

Note that the finding in Observation III goes in the opposite direction of potential substitution effects. Shoppers whose most preferred varieties are in the full-variety treatment but not in the reduced-variety treatment should be less likely to purchase anything at all, unless there are good substitutes. However, if we interpret $\%like$ as a measure of substitutability among varieties within the product line, we should have observed a greater decrease in sales in the reduced-variety treatment when $\%like$ is low, i.e., when the varieties that are missing from the reduced-variety treatment have more imperfect substitutes than when $\%like$ is high. Instead, we observe the opposite pattern, a decrease in sales in the full-variety treatment when $\%like$ is low and an increase in sales in the full-variety treatment when $\%like$ is high.

We now show findings which are consistent with our risk aversion hypothesis as summarized in Hypothesis II and Hypothesis III. First, we estimate

$$\frac{sales_{i,treatment}}{buyers_{i,treatment}} = \beta_0 + \beta_1 full + \beta_2 price_{i,treatment} + \beta_3 week + \varepsilon_{i,treatment} \qquad \text{Eq.(9)}$$

Table X shows that sales per buyer increases for the full-variety treatment ($full$) with respect to the reduced-variety treatment.

[Insert Table X here.]

Consistent with Hypothesis II,

Observation IV. The units purchased per shopper in the full-variety treatment is larger than in the reduced-variety treatment (for a fixed variety in a product line).

To test Hypothesis III, we estimate

$$\frac{sales_{i,treatment}}{buyers_{i,treatment}} = \beta_0 + \beta_1 full + \beta_2 D\_like_i + \beta_3 full * D\_like_i + \beta_4\, price_{i,treatment} + \beta_5 week + \varepsilon_{i,treatment} \quad \text{Eq.(10)}$$

where $D\_\ \%like_i$ is a dummy that takes on the value of 1 when $\%like_i$ is above the average required rate of success of 0.60 in Figure II:

$$D\_like_i = \begin{cases} 1, if\ \%like_i > 0.60 \\ 0, if\ \%like_i \leq 0.60 \end{cases}$$

The estimation results are displayed in Table XI. The key finding is the significant coefficient for the interaction between the treatment variable and $\%like_i$ ($full * D\_like$) in models (4)-(6), where product line fixed effects are controlled for.

[Insert Table XI here.]

Consistent with Hypothesis III,

Observation V. The increase in units purchased in the full-variety treatment is larger when the shoppers' believed rate of success is higher, i.e., for product lines with higher than average $\%like_i$ (for a fixed variety in a product line).

Comparing Table XI to the above Table X, we can see in columns (1)-(6) that the dummy $full$ is no longer significant when D_like_i and the interaction term are included in the exogenous treatment while the interaction term is significant when we control for product fixed effects. This could be due to collinearity between $full$ and $full\ * D\_like_i$ (correlation= 0.7405 with *p*-value=0.00).

V. Discussion

We extend the contextual inference theory of CAB by introducing non-unit demand in an incomplete information framework[12], where shoppers face risk in not achieving a required rate of success in sampling untried (or unremembered) varieties. Our conceptual framework predicts that CAB is more likely to occur in product lines when heterogeneity in taste is high, while CLB is more likely to occur for product lines where heterogeneity in taste is low. We developed a simple survey measure of heterogeneity in taste across different product lines and tested it as a predictor of shopper CAB/CLB across these product lines simultaneously in a large supermarket. We replicate what appeared to be conflicting evidence for CAB found in meta-studies of the choice averse literature (Chernev et al., 2014; Scheibehenne et al., 2010) by demonstrating that the average effect of greater variety across many product lines is zero (Observation II). Our data confirms our main hypothesis (Hypothesis I) that beliefs about heterogeneity in taste, as measured by the probability of success in sampling (p_i) and as estimated by $\%like_i$, can predict shoppers' CAB and CLB (Observation III). Moreover, for product categories with low heterogeneity in taste (high $\%like_i$), we replicate the CLB finding in the large sample study of Medicare data (Ketcham et al., 2015), for which consumers had positive surveyed beliefs.

In addition to predicting when CAB/CLB occur (Hypothesis II), we also show that our theory can predict the degree to which they occur. Shoppers purchased more units given that they purchased any units at all, when they faced more varieties (Observation IV). The constancy of the probability of stopping across the full- and the reduced-variety treatments (Observation I) suggests that this increase in purchases was not due to a greater number of shoppers stopping in front of shelves. The constancy of the probability of purchases across the full- and reduced-variety treatments (Observation II) also suggests that this increase in purchases was not due to a greater probability of purchase by shoppers. The remaining possibility is that there were more purchases per shopper when the shelf contained more varieties. While it is possible that this effect was due to more expert shoppers making better matches when choosing among more varieties, we would expect that they would be more likely to do so when heterogeneity in taste is high ($\%like_i$ is low), and products are consequently less substitutable. But, as predicted

[12] An earlier version of Kuksov and Villas Boas (2009) also use an incomplete information framework. They focus on the firm's choice of the number of varieties in a product line given private information as to the importance of fit between the consumers' tastes and the product characteristics, rather than the risky decision of the consumer, as in our case.

(Hypothesis III), the increase in the number of purchases per shopper was greater when their belief of success was low (Observation V), which is consistent with unfamiliar shoppers being more willing to sample when their belief is relatively optimistic compared to their required rate of success. Though, our predictions that both the probability of purchase and the number of units purchased would be increasing on $\%like_i$ were born out by the data, there is much work to be done to develop a complete conceptual framework, based on an altogether different conception of the COB, which can be integrated with the prior marketing literature. Despite the shortcomings of our conceptual framework and our experimental design, we emphasize, however, our findings indicate that what has been regarded as conflicting results as to the existence of CAB could be due to in part to the interaction between shoppers' belief about their likely rate of success in sampling from untried varieties and their required rate of success across different product lines. In this respect, our results can be regarded as a useful step towards understanding the CAB and CLB phenomena.

References

Barberis, N. C. (2007). The Psychology of Tail Events : A Note on Progress and Challenges. *American Economic Review*, *103*(3), 611–616.

Breslin, P. A. S. (2013). An evolutionary perspective on food and human taste. *Current Biology*, *23*(9), R409-18.

Chernev, A., Böckenholt, U., & Goodman, J. (2010). Commentary on Scheibehenne, Greifeneder, and Todd Choice Overload: Is There Anything to It? *Journal of Consumer Research*, *37*(3), 426–428.

Chernev, A., Böckenholt, U., & Goodman, J. (2014). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, *25*(2), 333–358.

Coates, J. M., & Herbert, J. (2008). Endogenous steroids and financial risk taking on a London trading floor. *Proceedings of the National Academy of Sciences*, *105*(16), 6167–72.

Gigerenzer, G. (2007). *Gut feelings : the intelligence of the unconscious*. Penguin Publishing Group.

Iyengar, S. S., & Kamenica, E. (2010). Choice proliferation, simplicity seeking, and asset allocation. *Journal of Public Economics*, *94*(7–8), 530–539.

Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*(6), 995–1006.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan Publishers Limited.

Kamenica, E. (2008). Contextual Inference in Markets: On the Informational Content of Product Lines. *American Economic Review*, *98*(5), 2127–2149.

Ketcham, J. D., Lucarelli, C., & Powers, C. A. (2015). Paying Attention or Paying Too Much in Medicare Part D? *American Economic Review*, *105*(1), 204–233.

Kuksov, D., & Villas-Boas, J. M. (2010). When More Alternatives Lead to Less Choice. *Marketing Science*, *29*(3), 507–524.

Nagler, M. G. (2015). Trading off the benefits and costs of choice: Evidence from Australian elections. *Journal of Economic Behavior & Organization*, *114*, 1–12.

Palacios-Huerta, I., & Serrano, R. (2006). Rejecting small gambles under expected utility. *Economics Letters*, *91*(2), 250–259.

Rabin, M. (2000). Risk Aversion and Expected-Utility Theory: A Calibration Theorem. *Econometrica*, *68*(5), 1281–1292.

Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload. *Journal of Consumer Research*, *37*(3), 409–425.

Thaler, R., & Tversky, A. (1997). The effect of myopia and loss aversion on risk taking: An experimental test. *Quarterly Journal of Economics*, 647–661.

Tversky, A., & Kahneman, D. (1974). Judjment under Uncertainty: Heuristics and Biasses. *Science*, *185*(4157), 1124–1131.

Villas-Boas, J. M. (2009). Product Variety and Endogenous Pricing with Evaluation Costs. *Management Science*, *55*(8), 1338–1346.

Wiuf, C., & Stumpf, M. P. H. H. (2006). Binomial subsampling. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *462*(2068), 1181–1195.

Tables and figures

TABLE I: A SAMPLE OF SURVEY FOR OREO BISCUIT PRODUCT LINE

	Not tried	Neutral	Like	Dislike
1. Chocolate	√			
2. Milk	√			
3. Mocha	√			
4. Strawberry		√		
5. Grape and peach	√			
6. Raspberry and blueberry	√			
7. Orange and mango		√		
8. Vanilla			√	
9. Cake				√
10. Strawberry cream	√			
11. Chocolate cream	√			
12. Green tea	√			

Notes: Each survey represents the preferences of one shopper.

Product line i

Variety	Shoppers: 1	2	3	4	5	6	7	8	9	10
1	1	0	1	1		1	0		1	
2	1	1	1	1		1	0		1	-1

FIGURE I: A SIMPLIFIED EXAMPLE OF SURVEY DATA

Notes: The 2 rows are 2 varieties of product line i. The 10 columns are for the 10 shoppers who took the survey. For this example, the $\%like = \frac{11}{15}$.

TABLE II: NAMES, NUMBER (#) AND $\%like$ FOR VARIETIES IN THE FULL-VARIETY TREATMENT FOR 24 PRODUCT LINES

Product line	# of varieties	$\%like$ (%)	Product line	# of varieties	$\%like$ (%)	Product line	# of varieties	$\%like$ (%)
Glico biscuits	19	32.62	Häagen-Dazs ice creams	24	61.36	Huaweiheng dried fruits	8	69.91
Huashengtang vinegars	5	53.01	Store made sushi	26	61.39	Want-want QQ gummies	9	70.43
Nestle milk powders	6	56.06	Alpine candies	7	61.93	Nissin cup noodles	19	70.79
Liby clothes detergents	15	56.81	Tongyi100 instant noodles	11	64.81	Vinda small pack tissues	7	73.17
Kangshifu instant noodles	30	58.69	Liby dish detergents	21	65.30	Comfort detergents	22	74.35
Oreo biscuits	10	58.73	Dove chocolates	11	68.29	Laoganma sauces	9	74.58
Lee Kum Kee soy sauces	19	59.15	Lipton teas	13	68.33	Haitian sauces	11	76.59
U. loveit milk teas	9	60.09	Knorr soup bases	10	68.97	Heinz rice powders	18	83.56

TABLE III: REGRESSION OF PROBABILITY OF STOP ON REDUCED NUMBER OF VARIETIES TREATMENT

Independent variables	$\frac{stops_{i,treatment}}{passbys_{i,treatment}}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment (Reduced-variety = 1)	-0.00249	-0.00297	-0.00782	-0.00249	-0.00255	-0.00657
	(0.0189)	(0.0189)	(0.0188)	(0.00973)	(0.0101)	(0.00792)
$price_{i,treatment}$		-0.00106	-0.00101		-0.000134	0.00266
		(0.000916)	(0.000903)		(0.00371)	(0.00298)
Week (Second = 1)			0.0292			0.0316***
			(0.0188)			(0.00809)
Constant	0.0990***	0.113***	0.100***	0.0990***	0.101*	0.0503
	(0.0134)	(0.0181)	(0.0196)	(0.00688)	(0.0500)	(0.0410)
Product line				Yes	Yes	Yes
R^2	0.000	0.029	0.080	0.003	0.003	0.423
Observations	48	48	48	48	48	48

Notes: $stops_i$ counts the shoppers who stopped in front of experimental product line i for at least three seconds. $passbys_i$ counts the number of shoppers who passed by the shelf but did not stop. Treatment is a dummy for the full- and reduced-variety treatments. Price is the average price of each product line in each treatment. Week is a dummy for the first and second weeks. Models (1)-(3) do not control for product line fixed effect, while models (4)-(6) do. Models (2) and (5) control the average price of each product line in each variety treatment. Models (3) and (6) control for week fixed effects. The dependent variable is $\frac{stops_{i,treatment}}{pass\,bys_{i,treatment}}$ of each product line i in each treatment, so there are 48 observations in total. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

TABLE IV: REGRESSION OF PROBABILITY OF PURCHASE ON NUMBER OF VARIETIES TREATMENT

Independent variables	$\frac{buyers_{i,treatment}}{stops_{i,treatment}}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment (Reduced-variety = 1)	0.00803	0.00652	0.00244	0.00803	0.00638	0.00310
	(0.0316)	(0.0302)	(0.0308)	(0.0201)	(0.0208)	(0.0207)
$price_{i,treatment}$		-0.00335**	-0.00331**		-0.00365	-0.00137
		(0.00147)	(0.00148)		(0.00765)	(0.00779)
Week (Second = 1)			0.0246			0.0259
			(0.0308)			(0.0212)
Constant	0.187***	0.232***	0.221***	0.187***	0.236**	0.194*
	(0.0223)	(0.0290)	(0.0321)	(0.0142)	(0.103)	(0.107)
Product line				Yes	Yes	Yes
R^2	0.001	0.105	0.117	0.007	0.017	0.082
Observations	48	48	48	48	48	48

Notes: $stops_i$ counts the shoppers who stopped in front of experimental product line *i* for at least three seconds. Treatment is a dummy for the full- and reduced-variety treatments. Price is the average price of each product line in each treatment. Week is a dummy for the first and second weeks. Models (1)-(3) do not control for product lines fixed effect, while models (4)-(6) do. Models (2) and (5) control the average price of each product line in each variety treatment. Models (3) and (6) control for week fixed effects. The dependent variable is $\frac{buyers_{i,treatment}}{stops_{i,treatment}}$ of each product line *i* in each treatment, so there are 48 observations in total. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

TABLE V: EXAMPLE OF EXPERIMENTAL DATA FOR OREO BISCUITS

Product lines (# Varieties in full)	BS_F	BS_H	$BS_H > BS_F$	$\%like_i$	$price\ change_{i,j}$
Oreo biscuit (10)				55%	
Milk	0.11	0.13	Yes	55%	0.00
Strawberry	0.02	0.00	No	55%	0.00
Chocolate	0.00	--	--	55%	--
Mocha	0.02	0.02	--	55%	-0.10

Notes: $stops_i$, which counts the shoppers who stopped in front of experimental product line *i* for at least three seconds. $buys_{i,j}$ refers to the purchase of variety j from product line *i* using store data. Buyers per $stops_i$ is defined as $BS_{i,j} = \frac{buys_{i,j}}{stops_i}$ for variety j of product line i. Subscript *F* refers to full-variety treatment while *H* refers to reduced-variety. $price\ change_{i,j} = \frac{price_{i,j,F}}{price_{i,j,H}} - 1$ for variety j of product line i. "Oreo biscuit" is a product line with ten varieties in the full-variety treatment. If $BS_{i,j,H}$ is greater than $BS_{i,j,F}$, we record a "Yes" for CAB. $\%like_{Oreo\ biscuit} = 55\%$. Variety "Milk" and variety "Strawberry" are a part of the same product line "Oreo biscuit". In addition, both have $price\ change_{Oreo\ biscuit,Milk}$ and $price\ change_{Oreo\ biscuit,Strawberry}$, because they were in both full- and reduced-variety treatments, like "Mocha", while "Chocolate" was not. Variety "Chocolate" only has $BS_{i,j,F}$ because it was only in the full-variety treatment. For variety "Mocha", for illustration, $BS_{i,j,F} = BS_{i,j,H}$.

TABLE VI: DEFINITIONS OF VARIABLES

Variable	Meaning	Definition
i	Product line index	
D_i	Dummy for product line i	$D_{\text{i}} = \text{I}_{(\text{product line}=i)}$
j	Variety index	
$stops_i$	# of shoppers who stop for more than 3 seconds using video data	
$buys_{i,j}$	purchases of the variety *j* in product line *i* in the reduced-variety treatment using store data	
F	full-variety treatment	
H	reduced-variety treatment	
$BS_{i,j} = \frac{buys_{i,j}}{stops_i}$	Probability of purchase for variety j of product line i in the reduced-variety treatment based on store data	
$CAB_{i,j}$	An indicator of choice averse behavior for the variety *j* in product line in the reduced-variety treatment	$CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,H} > BS_{i,j,F} \\ 0, if\ BS_{i,j,H} < BS_{i,j,F} \\ \text{dropped}, if\ BS_{i,j,H} = BS_{i,j,F} \end{cases}$
$\%like_i$	Estimate of success rate p_i probability of success when sampling randomly	$\frac{\#1s}{\#1s + \#0s + \# - 1s}$
$\%dislike_i$	A measure for percentage of dislike	$\frac{\# - 1s}{\#1s + \#0s + \# - 1s}$
$\%tried_i$	A measure for familiarity	$\frac{\#1s + \#0s + \# - 1s}{\#1s + \#0s + \# - 1s + \#blanks}$
$relativ\ size_i$	A measure of the reduced-variety treatment's share of the full-variety treatment for each product line	$\frac{\#varieties\ in\ reduced_i}{\#varietiesin\ full_i}$
%sales of removed varieties	A measure of the popularity of the removed varieties for each product line	$\frac{Sales\ of\ varieties\ not\ in\ reduced\ treatment\ in\ full\ teatment}{Sales\ of\ all\ varieties\ in\ full\ teatment}$
$price\ change_{i,j}$	Relative price change between reduced-variety and full-variety treatments for the variety *j* in product line *i* in the reduced-variety treatment using store data	$\frac{price_{i,j,F}}{price_{i,j,H}} - 1$

TABLE VII: SUMMARY STATISTICS FOR VARIABLES IN EQUATION (9)

Data	# Observations	Mean	Median	Max	Min	Standard Deviation
Product line	24 (23 have variety information)					
$\%like_i$	24	0.65	0.65	0.84	0.33	0.10
Total varieties	339					
Varieties in reduced-variety treatment	182					
$price\ change_{i,j}$	182	0	0	0.29	-0.20	0.08
$BS_{i,j,H} \neq BS_{i,j,F}$	150					
$CAB_{i,j}$	150					
$price\ change_{i,j}$ in regression	150	0	0	0.29	-0.20	0.08
$\%like_i$ in regression	23	0.65	0.65	0.84	0.33	0.10

Notes: We went from 24 product lines to 23 because we were not able to obtain individual variety sales data for store-made sushi. There are 339 varieties among the 24 product lines. Since we need to compare purchase behavior in the full- and reduced-variety treatments, we drop 157 varieties that are not in the reduced-variety treatment and varieties of store-made sushi, leaving 182 varieties in both. We, furthermore, drop 30 varieties where $BS_{i,j,H} = BS_{i,j,F}$, resulting in 152 observations.

TABLE VIII: PROBIT REGRESSIONS OF PROBABILITY OF CAB

Independent variables	$COB_{i,j}$											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\%like_i$	-2.339**	-2.781***		-2.641***		-2.833***		-2.354**	-2.123**	-2.074**		-2.829***
	(1.029)	(0.965)		(0.947)		(0.933)		(0.941)	(0.977)	(0.946)		(0.907)
$\%dislike_i$	0.394		2.908	2.162								
	(3.797)		(3.074)	(3.700)								
$\%tried_i$	1.226				0.925	1.209						
	(0.970)				(1.126)	(0.804)						
# of varieties in full-treatment	0.0131						0.0341**	0.0260*		0.0120		
	(0.0213)						(0.0152)	(0.0156)		(0.0198)		
$relativ\ size$	3.751								4.664*	3.605		
	(2.701)								(2.388)	(2.997)		
%sales of removed varieties	0.565										-0.282	0.136
	(0.797)										(0.960)	(0.797)
$price\ change_{i,j}$	18.72***	20.27***	20.83***	20.85***	21.19***	20.84***	19.41***	19.79***	19.68***	19.48***	21.27***	19.89***
	(2.297)	(0.742)	(0.895)	(0.833)	(1.087)	(0.827)	(0.776)	(0.746)	(0.763)	(0.746)	(2.408)	(2.067)
Control for product line fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-1.406	1.874***	-0.130	1.603**	-0.147	1.561***	-0.474	1.155*	-1.344	-0.944	0.179	1.872***
	(1.891)	(0.549)	(0.288)	(0.727)	(0.384)	(0.491)	(0.327)	(0.696)	(1.818)	(1.970)	(0.269)	(0.548)
Observations	150	150	150	150	150	150	150	150	150	150	150	150
Pseudo R^2	0.225	0.194	0.169	0.198	0.169	0.204	0.185	0.206	0.211	0.213	0.163	0.195

Notes: See Table VI for definition of terms. This table shows that the significance of $\%like_i$ is robust under many specifications. Column (1) includes all control variables. Column (2) includes no control variables except $price\ change_{i,j}$. Columns (3) and (4) control for $\%dislike_i$. Column (5) and (6) control for $\%tried_i$. Columns (7) and (8) control for the size of full-product line. Columns (9) - (12) controlled for $relative\_size$ and $\%sales\ of\ removed\ varieties$. We cluster standard errors by product lines in all models to allow different product lines to have different standard errors in our estimations. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

TABLE IX: PROBIT REGRESSIONS OF PROBABILITY OF CAB WITH ADJUSTED CAB

Independent variables	$Adjusted\ COB_{i,j}$												
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
$\%like_i$	-2.181**	-2.440**		-2.469**		-2.326**		-2.179*		-2.050*	-2.030*		-2.636**
	(1.006)	(1.084)		(1.067)		(0.953)		(1.130)		(1.149)	(1.155)		(1.131)
$\%dislike_i$	-2.113		-0.453	-0.924									
	(2.780)		(2.218)	(2.885)									
$\%tried_i$	-0.515				-0.986	-0.705							
	(0.790)				(1.016)	(0.620)							
# of varieties in full-treatment	0.0152						0.0281	0.0206			0.0137		
	(0.0199)						(0.0201)	(0.0202)			(0.0213)		
$relativ\ size$	2.645								4.667*	3.006	1.799		
	(2.912)								(2.570)	(2.341)	(2.343)		
%sales of removed varieties	0.819											0.0278	0.502
	(0.749)											(0.983)	(0.765)
$price\ change_{i,j}$	18.64***	22.25***	22.28***	22.07***	21.78***	21.82***	21.68***	21.77***	21.62***	22.01***	21.78***	22.29***	20.85***
	(2.304)	(0.675)	(0.868)	(0.800)	(1.050)	(0.702)	(0.948)	(0.876)	(0.829)	(0.727)	(0.898)	(2.287)	(1.818)
Control for product line fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.690	1.205*	-0.285	1.300*	-0.0533	1.330**	-0.812**	0.682	-3.131**	-0.847	-0.373	-0.329	1.205*
	(1.854)	(0.678)	(0.276)	(0.745)	(0.373)	(0.599)	(0.330)	(0.828)	(1.487)	(1.711)	(1.671)	(0.287)	(0.700)
Observations	153	153	153	153	153	153	153	153	153	153	153	153	153
Pseudo R-squared	0.248	0.230	0.201	0.230	0.208	0.233	0.215	0.237	0.218	0.236	0.238	0.201	0.232

Notes: See Table VI for definition of terms. As robustness check, we "adjusted" $COB_{i,j}$ in all models because the probability of being purchased for a specific variety in a larger choice set is diluted by more varieties. $Adjusted\ CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,H} > Adjusted\ BS_{i,j,F} \\ 0, if\ BS_{i,j,H} < Adjusted\ BS_{i,j,F} \\ null, if\ BS_{i,j,H} = Adjusted\ BS_{i,j,F} \end{cases}$. $Adjusted\ BS_{i,j,F} = BS_{i,j,F} \cdot \frac{\#\ varieties\ in\ full\ treatment}{\#\ varieties\ in\ reduced\ treatment}$. Again, this table shows that the significance of $\%like_i$ is robust under many specifications. Column (1) includes all control variables. Column (2) includes no control variables except $price\ change_{i,j}$. Columns (3) and (4) control for $\%dislike_i$. Column (5) and (6) control for $\%tried_i$. Columns (7) and (8) control for the size of full-product line. Columns (9) - (12) controlled for $relative\_size$ and $\%sales\ of\ removed\ varieties$. We cluster standard errors by product lines in all models to allow different product lines to have different standard errors in our estimations. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

.

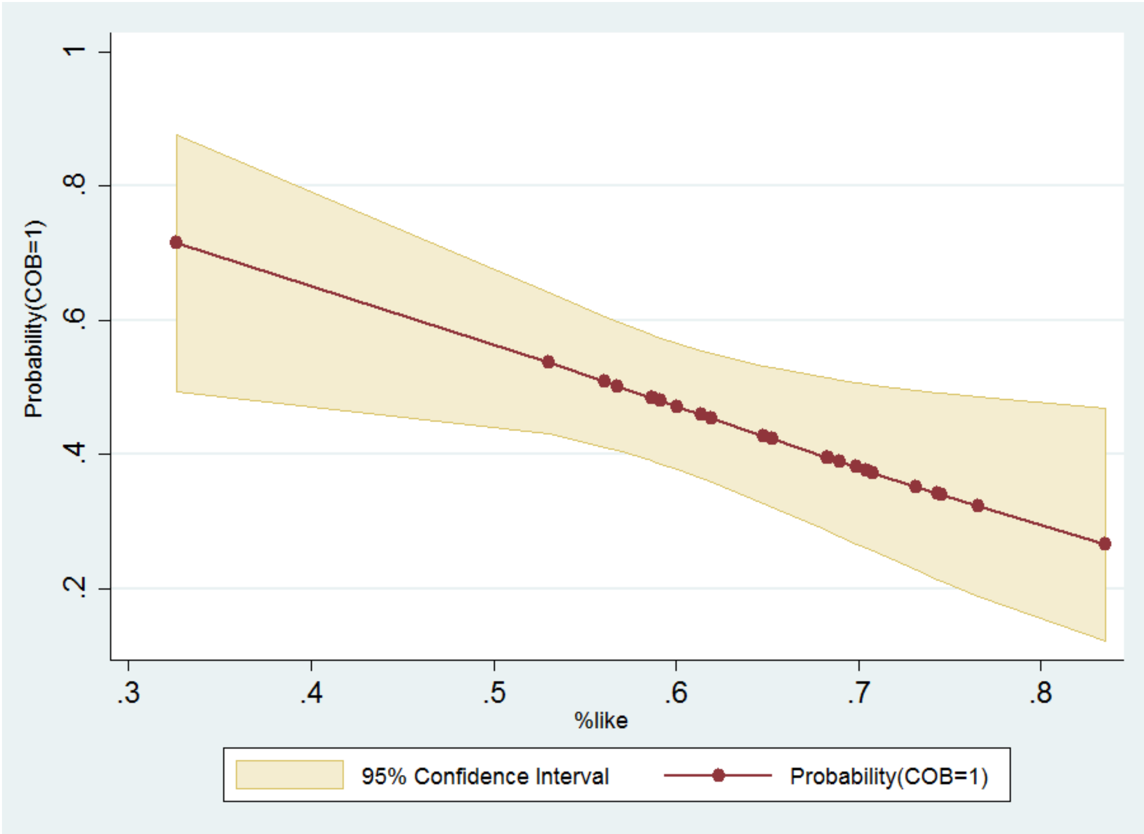


FIGURE II: PREDICTED PROBABILITY AS A FUNCTION OF $\%like_i$

Notes: This figure reveals how the probability of CAB decreases on $\%like_i$ for product line *i*. The dots are product lines. These correspond to those listed in Table II. The horizontal axis measures $\%like_i$. The vertical axis measures $\Pr(CAB_{i,j} = 1|\%like_i, other\ control\ variables)$ holding $other\ control\ variables$ at their means in model (1) of Table VIII. The yellow area surrounding the trend line is the 95% confidence interval in model (1) of Table VIII. The slope above, which looks straight, is not necessarily linear. It is in fact slightly curved downwards.

TABLE X: SALES PER BUYER

Independent variables	$\frac{sales_{i,treatment}}{buyers_{i,treatment}}$					
	(1)	(2)	(3)	(4)	(5)	(6)
$full$ (full-variety = 1)	0.728*	0.742*	0.716*	0.728*	0.804**	0.771**
	(0.403)	(0.398)	(0.382)	(0.365)	(0.358)	(0.329)
$price_{i,treatment}$		-0.0318*	-0.0320*		-0.169***	-0.192***
		(0.0160)	(0.0164)		(0.0347)	(0.0434)
$week$ (second=1)			-0.158			-0.262
			(0.383)			(0.338)
Constant	2.248***	2.657***	2.752***	2.248***	4.422***	4.873***
	(0.263)	(0.364)	(0.416)	(0.182)	(0.463)	(0.594)
Observations	48	48	48	48	48	48
R-squared	0.066	0.119	0.122	0.150	0.208	0.226
Number of product lines				24	24	24

Notes: $full$ is a dummy for the full- and reduced-variety treatments, and $full$ is equal to 1 for full-treatments. Price is the average price of each product line in each treatment. Week is a dummy for the first and second weeks. Models (1)-(3) do not control for product line fixed effects, while models (4)-(6) do. Models (2), (3), (5), and (6) control the average price of each product line in each variety treatment. Models (3) and (6) control for week fixed effects. $sales_{i,treatment}$ is the sum of $buys_{i,j}$ over all of the varieties *j* in product line *i* in the reduced-variety treatment. $buyers_{i,treatment}$ is the number of buyers of each product line *i* in each treatment from 5:30pm-9:30pm. The dependent variable is $\frac{sales_{i,treatment}}{buyers_{i,treatment}}$ of each product line *i* in each treatment, so there are 48 observations in total. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

TABLE XI: SALES PER BUYER GIVEN HIGH OR LOW $\%LIKE$

Independent variables	$\frac{sales_{i,treatment}}{buyers_{i,treatment}}$					
	(1)	(2)	(3)	(4)	(5)	(6)
$full$ (full-variety = 1)	-0.118	-0.141	-0.163	-0.118	-0.291	-0.354
	(0.816)	(0.824)	(0.823)	(0.451)	(0.412)	(0.443)
D_like_i (D_like_i=1 if $\%like_i$>0.6)	-0.463	-0.466	-0.464			
	(0.745)	(0.750)	(0.775)			
$full$ * D_like_i	1.194	1.248	1.244	1.194*	1.593**	1.629**
	(0.937)	(0.941)	(0.958)	(0.643)	(0.643)	(0.676)
$price_{i,treat}$		-0.0330*	-0.0332*		-0.244***	-0.272***
		(0.0169)	(0.0175)		(0.0471)	(0.0594)
$week$ (second=1)			-0.150			-0.303
			(0.386)			(0.316)
Constant	2.576***	3.003***	3.092***	2.248***	5.387***	5.931***
	(0.707)	(0.773)	(0.778)	(0.175)	(0.561)	(0.789)
Observations	48	48	48	48	48	48
R-squared	0.105	0.162	0.165	0.234	0.346	0.370
Number of product lines				24	24	24

Notes: $full$ is a dummy for the full- and reduced-variety treatments, and $full$ is equal to 1 for full-treatments. Price is the average price of each product line in each treatment. Week is a dummy for the first and second weeks. Models (1)-(3) do not control fixed effect of product lines, while models (4)-(6) do. Models (2), (3), (5), and (6) control the average price of each product line in each variety treatment. Models (3) and (6) control for week fixed effects. D_like_i is omitted in models (4)-(6) because of collinearity. $sales_{i,treatment}$ is the sum of of $buys_{i,j}$ over all of the varieties j in product line i in the reduced-variety treatment. $buyers_{i,treatment}$ is the number of buyers of each product line in each treatment from 5:30pm-9:30pm. The dependent variable is $\frac{sales_{i,treatment}}{buyers_{i,treament}}$ of each product line i in each treatment, so there are 48 observations in total. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendices

Appendix I. Conceptual framework

For a given product line i, we assume that shoppers have a constant utility of consumption U for each variety when they get a product they like and 0 otherwise. The cost of purchasing any unit of any variety is c. We further limit ourselves to predicting the shopper's static choices, given their information at the time of choice[13].

Shoppers have a belief p about their probability of success, which is identical with the sample proportion of success, for any given variety randomly chosen from m varieties of product line i on the shelf. The probability of k successes on the shelf is binomial $Pr(k|\, m, p)$. Because the shopper randomly takes an n-size subsample of the m varieties on the shelf, their odds of getting k successes follows a binomial $Pr(k|\, n, p)$ distribution (Wiuf & Stumpf, 2006). Let X be the count of their successes and $\hat{p} = \frac{X}{n}$ be their sample proportion of success. We assume that the shopper has a required rate of success $\bar{r}$, which we take as exogenous because it reflects an optimal decision across all product categories given their entire budget, a decision which is outside the scope of our study to explain. Shoppers discount future periods at δ and repeat the purchase of what they end up liking for T periods. Thus, we conceive of the decision of shoppers to sample from one product line as a part of a T period consumption plan, where they choose

$$max\{\sum_{t=0}^{T} \delta^t k(U-c)(Pr(\hat{p} > \bar{r})) - nc, 0\} \qquad \text{Eq.(11)}$$

We refer to the LHS of this pair of choices as the decision to sample. Clearly, fixing the probability of sampling, this required rate $\bar{r}$ is decreasing on the shopper's consumption as reflected in the number of repeated purchases T in that product category. For example, heavy coffee drinkers should be more tolerant of getting $\frac{n-k}{n}$ share of coffee that they do not like when they sample n untried varieties, because they will buy many more units (i.e., T is large) of the coffee that they do like. The increase in lifetime utility from the

13 A fuller framework would allow consumers to employ a dynamic sampling strategy, where consumers choose an initial sample with a view to potential updates and further sampling. Our experimental results suggest that the explicit modeling of such updating is not necessary, at the level of a first approximation, to predict COB/CLB, though no doubt, the predictive power of beliefs about sampling risk can be improved if the possibility of updating could be incorporated into the model and into the experimental design.

many anticipated purchases of what they do like (the $\frac{k}{n}$ share of successes) will better compensate them for the lost surplus from what they do not like (the $\frac{n-k}{n}$ share of failures that each cost c). Thus, we introduce the following assumption.

Assumption 1: Shoppers who anticipate purchasing more units in a given product category have a lower required rate $\bar{r}$ of success in sampling untried varieties in that product category.

In our experiment, the number of varieties on the shelf m is determined exogenously by our treatments. We assume that m influences n monotonically, so that when m increases to m', n also increases to n', should the consumer choose to sample at all. Consumers may sample not only to get something they like, but also in the anticipation of a decision of whether to sample again, conditional on finding something they like, or not. Therefore, they might want a representative sample when they sample. By contextual inference, they should sample more units when they face more variety, because more varieties signal greater variation in tastes, and therefore, a lower representativeness of a fixed number of varieties. There are also likely many potential psychological explanations of how an individual shopper's choice of n might be so influenced by the m varieties on the shelf, e.g., anchoring (Tversky & Kahneman, 1974). We leave the systematic investigation of such influence for future work. For now, we assume,

Assumption 2: The size of the subsample that shoppers choose n, should they sample any units at all, is positively influenced by the sample size m on the shelf.

To avoid unnecessarily complicating our exposition, we will no longer distinguish between a treatment that changes the number of varieties on the shelf m and the number of varieties that the consumer chooses n. Hence, we will only refer to the effect of exogenous changes in n from the treatments (i.e., reduced- or full-variety) in the decision to sample.

Lemma 1 summarizes how the required rate of success of shoppers interacts with their belief about their likelihood of success p can predict CAB and CLB. For a given product line i,

Lemma 1: When $p > \bar{r}$, shoppers are more likely to sample (i.e., more likely to exhibit choice loving behavior), when the number of varieties increases from n to n'; when

$p < \bar{r}$, shoppers are less likely to sample (i.e., less likely to exhibit choice averse behavior) when the number of varieties increases from n to n'.

Proof:

The probability that a shopper achieves a $\bar{r}$ rate of success when taking a sample of size n

$$Pr(\hat{p} > \bar{r}) = 1 - Pr(n \leq \bar{r}) \qquad \text{Eq. (1)}$$

By the Central Limit Theorem, the distribution of this average success rate is approaches the normal distributed with mean p and standard deviation $\frac{\sqrt{p(1-p)}}{n}$. By the Law of Large numbers, the mean of the $\hat{p}$'s converges to p. The joint effect is to concentrate the probability mass for the distribution of sample means around the population mean p. Hence, as n gets large, the probability in Eq. (1) will either go to 1, if $\bar{r} > p$, or 0, if $\bar{r} < p$.

Thus, when shoppers' belief is optimistic, more convergence (large n') is better; $n' - n$ more repetitions of the gamble moves mass from below their required rate to above: $Pr(\hat{p} > \bar{r} | n') > Pr(\hat{p} > \bar{r} | n)$ for $n' > n$. When shoppers' belief is relatively pessimistic, less convergence (small n') is better; $n - n'$ less repetitions of the gamble moves mass from above $\bar{r}$ to below: $Pr(\hat{p} > \bar{r} | n') > Pr(\hat{p} > \bar{r} | n)$ for $n' < n$.

Note that by Assumption 1, absolutely pessimistic belief (low p) about the success rate in sampling need not necessarily deter the purchasing of those n units. Nor, for the same reason, would even relatively pessimistic belief ($p < \bar{r}$). The probability of getting their required rate of success, even if low, could still be *high enough* such that the choice to sample in Equation (1) yields a positive surplus, i.e., if the surplus from the anticipated purchases of T many units of even a few successes compensate for the lost surplus from even many failures[14].

[14] Hence, the mass of the probability function above the required rate of success rather than the mean belief determines behavior. See the Appendix II for an illustration.

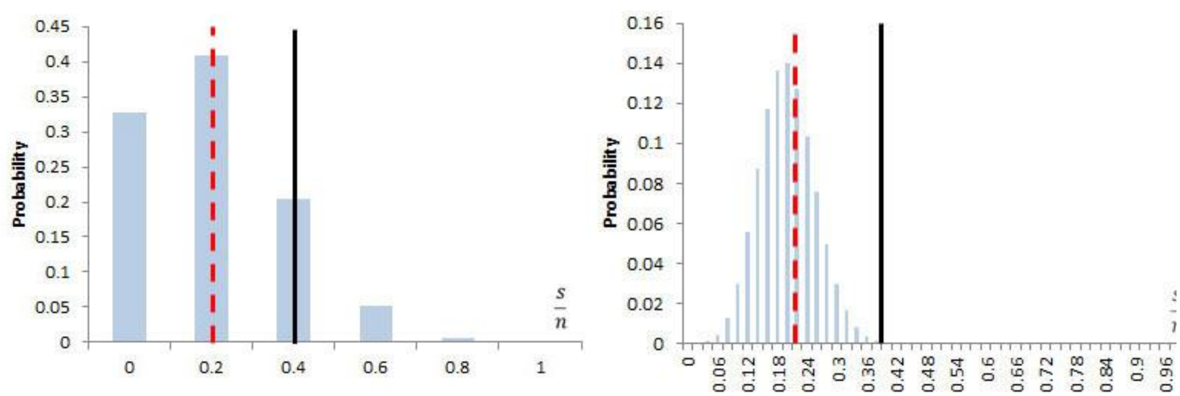
Appendix II. Numerical illustration of Lemma 1

We illustrate Lemma 1 with two special cases, first when shoppers have a pessimistic belief relative to their $\bar{r}$ ($p < \bar{r}$) and second, when shoppers have relatively optimistic belief. Suppose that for budgetary or taste reasons, a shopper decides to sample n =5 varieties at a time when there are m =10 varieties. However, when he faces a large product line with $m = 100$ varieties, he decides to sample $n = 50$ varieties at a time. Let this shopper believe that the population rate of success $p = 0.2$. Let him also have a required rate of success $\bar{r} = 2/5 = 40\%$. His sampling rate of success is $\hat{p} = \frac{X}{n}$, where X is the count of his successes. Then, the shopper should believe that probability of his sampling rate of success $\hat{p}$ exceeds the 40% required rate of success is 0.26 for the small product line and 0.000932 for the large product line. We organize these facts in A-Table XII.

A-TABLE XII: EXAMPLE WHERE $p < \bar{r}$

	# Varieties (m)	Sample (n)	$\bar{r}$	Success rate (p)	$Pr(\hat{p} \geq \bar{r})$
Small product line	10	5	$\frac{2}{5}$	$\frac{1}{5}$	0.26
Large product line	100	50	$\frac{2}{5}$	$\frac{1}{5}$	9.32×10^{-4}

We illustrate the findings in A-Figure III.

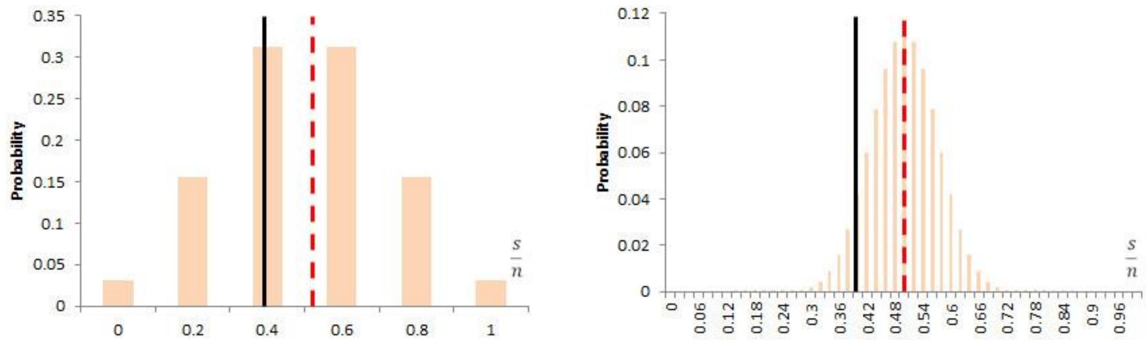


(a) Distribution of $\frac{X}{n}$ when $m = 10$ (b) Distribution of $\frac{X}{n}$ when $m = 100$

A-FIGURE III: EXAMPLE WHERE $p > \bar{r}$

Notes: The red dotted line denotes $p = 0.2$, the population rate of success. The black solid line denotes $\bar{r} = \frac{2}{5}$, the required rate of success. (a) and (b) are probability mass functions of $\hat{p} = \frac{X}{n}$, the sampling rate of success, where *X* is the count of successes. The blue bars represent the probability of $\frac{X}{n}$ equaling different values. For (a), $m = 10, n = 5$, thus $X \in \{1, 2, 3, 4, 5\}$ and $\frac{X}{n} \in \left\{\frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{5}{5}\right\}$. For (b), $m = 100, n = 50$, thus $X \in \{1, 2, \ldots, 50\}$ and $\frac{X}{n} \in \left\{\frac{1}{50}, \frac{2}{50}, \ldots, \frac{50}{50}\right\}$. The figure reveals that when the number of varieties (m) increases, the distribution of $\frac{X}{n}$ becomes more concentrated around the red dotted line (p). Thus, if $p < \bar{r}$, $Pr\left(\frac{X}{n} \geq \bar{r}\right)$ decreases when the number of varieties increases. This leads to CAB.

The reasoning is similar if we had assumed that the shopper had relatively optimistic belief, e.g., that the probability of success was $p = 0.5$ instead of 0.2. In this case, the shopper should believe that the probability of getting the 40% required rate of success is 0.81 for the small product line and 0.94 for the large product line. This can be seen in A-Figure IV.



(a) Distribution of $\frac{X}{n}$ when $m = 10$ (b) Distribution of $\frac{X}{n}$ when $m = 100$

A-FIGURE IV: EXAMPLE 2 WHERE $p > \bar{r}$

Notes: The red dotted line denotes $p = 0.50$, the population rate of success. The black solid line denotes $\bar{r}$, the required rate of success. (a) and (b) are probability mass functions of $p = \frac{X}{n}$, the sampling rate of success, where *X* is the count of successes. The pink bar represents the probability of $\frac{X}{n}$ equaling different values. For (a), $m = 10, n = 5$, thus $X \in \{1, 2, 3, 4, 5\}$ and $\frac{X}{n} \in \left\{\frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{5}{5}\right\}$. For (b), $m = 100, n = 50$, thus $X \in \{1, 2, \ldots, 50\}$ and $\frac{X}{n} \in \left\{\frac{1}{50}, \frac{2}{50}, \ldots, \frac{50}{50}\right\}$. A-Figure IV reveals that when the number of varieties (m) increases, distribution of $\frac{X}{n}$ is more concentrated around the red dotted line (p). Thus, if $p > \bar{r}$, $Pr\left(\frac{X}{n} \geq \bar{r}\right)$ increases when the number of varieties increases. This leads to CLB.

Appendix III. Removed varieties

TABLE A-XIII: LIST OF $\%sales\ of\ removed\ varieties$

ID	Product line	$relative\_size_i$	$\%sales$ of removed half	$\%like_i$
1	Häagen-Dazs ice cream	0.6250	1.0000	0.6136
2	Alpine candies	0.5714	0.0000	0.6193
3	Knorr soup bases	0.6000	0.0833	0.6897
4	Want-want QQ gummies	0.5556	0.4516	0.7043
5	Glico biscuits	0.6316	0.0938	0.3262
6	Liby dish detergents	0.6000	0.2273	0.5681
7	Liby clothes detergents	0.5238	0.3333	0.6530
8	Oreo biscuits	0.5000	0.2174	0.5873
9	Tongyi100 instant noodles	0.5455	0.2857	0.6481
10	U. loveit milk teas	0.5556	0.2143	0.6009
11	Dove chocolates	0.6364	0.1818	0.6829
12	Comfort detergents	0.5909	0.2000	0.7435
13	Haitian sauces	0.5455	0.4667	0.7659
14	Laoganma sauces	0.5556	0.4328	0.7458
15	Kangshifu instant noodles	0.6667	0.1234	0.5869
16	Lee Kum Kee soy sauces	0.6316	0.3250	0.5915
17	Heinz rice powders	0.5556	0.0000	0.8356
18	Huashengtang fruit vinegars	0.6000	0.1111	0.5301
19	Huaweiheng preserved fruits	0.6250	0.4583	0.6991
20	Lipton teas	0.4615	0.8000	0.6833
21	Nestle milk powders	0.5000	0.3077	0.5606
22	Nissin cup noodles	0.6316	0.3433	0.7079
23	Vinda small package tissues	0.5714	0.0000	0.7317
24	Store made sushi	0.5000	none	0.6139

Notes: $relative_{size_i} = \frac{\#varieties\ in\ half}{\#varietiesin\ full}$for each product line. $\%sales\ of\ removed\ varieties$ is estimated based on sales.

Table A-XIV: CORRELATION BETWEEN %LIKE, %SALES OF REMOVED HALF, AND RELATIVE SIZE (P-VALUE IN PARENTHESES)

	$\%like_i$	$\%sales\ of\ removed\ half$	$relative\_size_i$
$\%like_i$	1		
$\%sales\ of\ removed\ half$	0.112 (0.6108)	1	
$relative\_size_i$	-0.2172 (0.3194)	-0.1968 (0.3681)	1

Notes: The table shows $\%like_i$, $\%sales\ of\ removed\ half$, and $relative\_size_i$ do not have significant correlations.