Predicting choice-averse and choice-loving behaviors in a field experiment with actual shoppers

Ong, David

Jinan University-University of Birmingham Joint Institute, Jinan University

21 June 2021

Predicting choice-averse and choice-loving behaviors in a field experiment with actual shoppers

FORTHCOMING IN JOURNAL OF ECONOMIC BEHAVIOR AND ORGANIZATION, AUGUST 2021

DAVID ONG[1]

A large body of chiefly laboratory research has attempted to demonstrate that people can exhibit choice-averse behavior from cognitive overload when faced with many options. However, meta-analyses of these studies, which are generally of one or two product lines, reveal conflicting results. Findings of choice-averse behavior are balanced by findings of choice-loving behavior. Unexplored is the possibility that many consumers may purchase to reveal their tastes for unfamiliar products, rather than attempt to forecast their tastes before purchase. I model such 'sampling-search' behavior and predict that the purchases of unfamiliar consumers increase with the available number of varieties for popular/mainstream product lines and decrease for niche product lines. To test these predictions, I develop a measure of popularity based on a survey of 1,440 shoppers for their preferences over 24 product lines with 339 varieties at a large supermarket in China. 35,694 shoppers were video recorded after the varieties they faced on shelves were randomly reduced. As found in the meta-studies, choice-averse behavior was balanced by choice-loving behavior. However, as predicted, the probability of choice-loving behavior increases with the number of available varieties for popular product lines, whereas choice-averse behavior increases with available varieties for niche product lines. These findings suggest that increasing the number of varieties has predictable opposing effects on sales, depending upon the popularity of the product line, and opens the possibility of reconciling apparently conflicting prior results.

Keywords: field experiment, choice overload, choice-aversion, consumer search

JEL Codes: C93, D83, M31

---

[1] This study builds on a very preliminary theory and a very small sample experiment in Ong (2012), which is available on request. I thank Mengxia Zhang for her help getting access to the supermarket, experimental design, collection, coding of the data, and the preliminary statistical analysis, which became the basis for her MA thesis. I am grateful to United States National Science Foundation grant SES-08-51315 for financial support and to an anonymous referee, Aurelion Baillon, Miguel Vilas-Boas, Andrew Ching, Paolo Crossetti, Avi Goldfarb, Joseph K. Goodman, Liang Guo, Tanjim Hussain, Ralph Hertwig, Bing-Yi Jing, Emir Kamenica, Peter Landry, Jianpei Li, Yeshim Orhun, Benjamin Scheibehenne, Jason Somerville, Panayiota Touloupou, the participants at the 19th Annual International Industrial Organization Conference (virtual 2021), SABE Conference (Moscow, 2020), Econometric Society World Congress (Bocconi, 2020), ISMS Marketing Science Conference (Fuqua School, 2020), Marketing Dynamics Conference (Smith School, 2019), Second Winter Workshop: Behavioral and Experimental Economics of Food Consumption (France, 2019), Third Behavioral Industrial Organization and Marketing Symposium (Ross School, 2018), Advances with Field Experiments (University of Chicago, 2016), and the Johns Hopkins Carey Business School, Washington University Olin Business School, University of Basel Department of Psychology, Harbin Institute of Technology (Shenzhen), University of Birmingham and Huazhong University of Science and Technology, University Of International Business and Economics (Beijing) and Lingnan University (Hong Kong SAR) seminars for their helpful comments.
Ong: Jinan University- University of Birmingham Joint Institute, Jinan University, Guangzhou, email: dvdong@gmail.com.

1 Introduction

How people react to more options is of fundamental importance to the welfare implications of free markets and many government policies, e.g., the number of health insurance options the US Government should offer to Medicare recipients (Ketcham, Lucarelli, & Powers, 2015). Standard consumer theory in economics implies that more options cannot result in lower welfare. Among other advantages, more options offer the possibility of better matches, greater flexibility, and convenience by reducing search costs (Scheibehenne, Greifeneder, & Todd, 2010). All else being equal, rational decision-makers should exhibit choice-loving behavior (CLB). However, a growing body of research in psychology, marketing, and economics suggests that people can at least act *as if* more options decrease their welfare. Much of this literature was initiated by a now-iconic 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 ten times more likely to purchase at the shelves when only six varieties were at the display. Many follow-up laboratory studies provide supporting evidence.

A number of theories have been put forward to explain this choice-averse behavior (CAB). These include actual psychological "overload" from the presumed increased psychological cost of ranking a large number of options (Chernev, Böckenholt, & Goodman, 2015; Scheibehenne et al., 2010). This psychological overload can be formally modeled within a search context where consumers who are unfamiliar with the product line can, a) weigh the costs and benefits of searching and revealing the match quality between a specific variety (e.g., chocolate) in the product line (e.g., Oreo biscuits) and their tastes before purchasing, b) purchase randomly without knowing the match quality, or c) forgo purchasing at all (Kuksov & Villas-Boas, 2010).[2] CAB can also be modeled as the result of "contextual inference". Within this framework, consumers know the firm will prioritize the introduction of more popular varieties (Kamenica, 2008; Kuksov & Villas-Boas, 2010).[3] Unfamiliar consumers, who must also choose randomly, therefore, know that their odds of

---

[2] The tradeoff between incurring search cost and a potentially higher quality match predicts the existence of an optimal number of options, which has been confirmed in a recent study exploiting voter self-selection into more or less refined rankings of candidates in Australian elections (Nagler, 2015).

[3] Such inferences help to explain why additional funds in an individual's 401(k) plan are 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 can signal higher levels of surplus extraction by the firm, given that a more tailored fit can sustain higher prices (Villas-Boas, 2009).

success from randomly sampling untried varieties are greater for product lines with fewer varieties (e.g., chocolate flavor) than for product lines with more varieties.

Both of these theories of CAB also allow for CLB via a high share of consumers who have a low cost of ranking or searching among available varieties, i.e., those consumers who are familiar with the product line. Hence, in principle CAB is reconcilable with standard assumptions within economics and psychology through the introduction of a cost for cognition. However, predicting the CAB of subjects in experiments has still been problematic. Within the marketing literature, the evidence supporting CAB is roughly balanced by evidence supporting CLB and with no established predictor for either.

A meta-analysis of 63 marketing studies made up almost entirely of laboratory experiments, with a majority using hypothetical choices over one or two product lines, with *N*=5,036 subjects revealed a "mean effect size of virtually zero" (Scheibehenne et al., 2010). The net-zero effect finding was confirmed in a follow-up meta-study with *N*=7,202 (Chernev et al., 2015) by authors who challenged the validity of the initial meta-study on the grounds that the treatments across the different experiments were not identical (Chernev, Böckenholt, & Goodman, 2010). This last meta-study reveals four moderating factors: task difficulty, set complexity, preference uncertainty, and effort minimization goals that can explain nearly 70 percent of the variation in past treatment effects (Chernev et al., 2015). However, to my knowledge, none of these moderating factors have been used to actually predict CAB or CLB in the lab or 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 among a greater number of plans for a fixed decrease in the cost of the plans (Ketcham et al., 2015). The authors conclude that there is no such phenomenon as CAB. While this conclusion is perhaps premature, their finding does highlight the potential importance of field experimental studies of ordinary consumers in non-laboratory settings with real choices for corroborating laboratory studies.

Notwithstanding the importance of field experiments for establishing the external validity of a theory, field experiments in marketing are still rare because of the inherent difficulties of eliciting the cooperation of firms. Despite these difficulties, the use of field experiments is increasing due to growing concerns about the validity of laboratory experiments (Gneezy, 2017; Simester, 2017). However, field experiments on CAB with ordinary shoppers are subject to their own particular design challenges. For example, the original field experiment that initiated the literature used a special display to attract subjects. The display for the low-variety (six) treatment was smaller, and thus, may have attracted "motivationally different consumers" (in other words, having a greater willingness to purchase) than the high-variety (24) treatment (Iyengar & Lepper, 2000). Notably, subsequent attempts to replicate the main finding that more varieties lead to fewer purchases were unsuccessful. Scheibehenne (2008) found no negative effect of more varieties on the probability of purchase in a field experiment in a German supermarket. Boatwright and Nunes (2001, 2004) find that even large reductions in variety over many product lines by an online store did not affect sales.[4] These field studies vary only variety and do not test for a moderating factor for CAB. Most importantly, except for Boatwright and Nunes (2001, 2004), field experimental and empirical studies (Bertrand, Karlan, Mullainathan, Shafir, & Zinman, 2010; Iyengar, Jiang, & Huberman, 2004; Iyengar & Kamenica, 2010) in both the marketing and the economics literatures, similar to the laboratory studies, tested for only one, and exceptionally, two product lines or services at a time.

A crucial assumption in the prior literature is that CAB is driven by the consumer's attempt to forecast their preferences for unfamiliar products through costly introspection before purchasing. I refer to such attempts at forecasting (including contextual inference) as 'forecasting-search'. The present study is motivated by the possibility that the consumer finds it infeasible or not worthwhile to exert the cognitive costs of forecasting their tastes for unfamiliar products before purchasing. Instead, the consumer would rather choose non-deliberatively in order to reveal the match quality between their tastes and the product characteristics with certainty for future repeated purchases.[5] I refer to these searches as 'sampling-search'. I hypothesize that such sampling-search may be the better option for ordinary consumers when faced with unfamiliar varieties of relatively inexpensive household products within the ordinary cognitively non-pristine busy supermarket context, which may increase the cost of deliberation.

While both Kuksov and Villas-Boas (2010) (henceforth, KV-B) and Kamenica (2008) include the possibility of random purchases of varieties, neither theory derives the information value of

[4] Some caution is warranted in concluding CAB does not exist from this evidence because the criteria by which shoppers were assigned to the full- and reduced- variety treatments was not disclosed to the authors. Moreover, they do not identify or control for the behavior of familiar shoppers.

[5] Such non-deliberative searches may be regarded as random from the perspective of the shopper. Other factors exogenous to their tastes, e.g., proximity to a specific variety when they arrive at the shelf, may then determine their choice. Note that such non-deliberate randomization is distinct from the intentional randomization with a randomizing device proposed by the literature on a preference to randomize (Machina, 1989). Such deliberate use of a randomizing device may require some extra costs from the intentional purchase and utilization of an actual randomizing device.

such random purchases to guarantee positive surpluses for repeated future purchases. The present study develops a sampling-search framework to explain CAB\CLB based on the Hoteling model of product differentiation used in KV-B (See the Theoretical Appendix). It is shown that when consumers observe a larger number of varieties on shelves, they should infer greater heterogeneity in tastes and a greater probability of mismatch between their tastes and the characteristics of the product when they purchase randomly. This greater probability of mismatch decreases the expected value of sampling-search while, at the same time, increasing the optimal sample size, conditional on sampling-search being optimal at all. Thus, the more popular the product line, the less likely the decrease in expected value effect will prevent the consumer from purchasing at all, and the more likely the increased optimal sample size effect from the increased dispersion of utility outcomes will be realized as purchases. This comparative static result implies that when unfamiliar consumers face a greater number of varieties for a popular product line, they exhibit CLB. In contrast, when unfamiliar consumers face a greater number of popular varieties for a niche product line, they exhibit CAB.

To test these predictions, I developed an estimator of potential purchasers' perception of the popularity of a product line by surveying 1,440 shoppers for their "likes", "neutrals", "dislikes", and "untried" for hundreds of varieties (e.g., vanilla) across 24 product lines (e.g., Oreo biscuits) when they exited the large supermarket in China in which I conducted the experiment. I 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 perceived popularity of a product line. I interpret this percentage-of-likes measure of popularity as giving the odds that a random consumer receives a positive surplus from a variety that she chooses randomly.

To observe shoppers' behavior in front of shelves, mini video cameras were mounted innocuously on the ceiling where there were already security cameras. With these cameras, shoppers were recorded from behind and above as they passed by, stopped, or purchased items. The varieties shoppers faced on the shelves were reduced on randomly selected days across the two weeks of the experiment. This random assignment ensures that the distribution of the types of consumers is constant across the full- and reduced-variety treatments for each product line used in the experiment. I recorded approximately 12 hrs of color video of shoppers' behavior per product line per each of two experimental days (approximately 500 hrs in total). I supplement this video data of shopper behavior, which may not always show clearly what variety the shopper removed from shelves, with point-of-sale purchase data provided by the store with exact quantities of each variety purchased. For this study, I focus on the 35,694 shoppers observed between the maximum traffic hours during the evening 5:30 pm-9:30 pm.

Using only the data from video footage, I find, inconsistent with Iyengar and Lepper (2000) that the reduction in variety has no net effect on the probability of stopping in front of the shelf. However, the display for the product line was constant, ruling out the possibility that different display sizes induce selection for motivationally different shoppers as a confounder. Consistent with the meta-studies (Chernev et al., 2015; Scheibehenne et al., 2010) and with the field study of Boatwright and Nunes (2001, 2004), the number of options does not affect the average probability of purchase. The question of the comparability of treatments raised by Chernev et al. (2010) is not an issue for this finding because the treatments were identical across product lines.

These results based on data from video footage use the standard measure of CAB, which tests for the effect of the number of varieties on the probability of purchase *per product line*. While my experiment has many more product lines and varieties than prior studies using the standard measure, the sample size of 48 observations (24 product lines on 2 treatment days) is still modest. Moreover, this method of calculating the treatment effect of variety per product line does not control for the substitution behavior of familiar consumers. Thus, the lack of effect of the reduced-variety treatment on the probability of purchase can also be due to the CLB of familiar shoppers canceling out the CAB of unfamiliar shoppers.

The substitution behavior of familiar shoppers is an important confounder in field studies of CAB. Prior researchers have tried to reduce the influence of those subjects who are familiar with the product line on their data by either using exotic product lines or by dropping the data from these familiar subjects in laboratory studies in which they can be identified with a survey (Iyengar & Lepper, 2000). However, even if such strategies succeeded, restricting product lines to exotic brands and/or dropping data from familiar subjects limits the product characteristics or populations for which hypotheses can be tested and the generalizability of any results found.

One of the novelties of this study is to measure the probability of CAB/CLB *per variety* instead of per product line. Doing so has two important advantages. It greatly increases the statistical power when there are many varieties among many product lines, as in this study. It also permits the econometric control for the substitution behavior of familiar shoppers and the forecasting-search behavior of unfamiliar shoppers, which I discuss in detail in Section 2

Using this per-variety measure of CAB and controlling for the substitution behavior of familiar consumers, I find a highly significant relationship between the probability of CAB/CLB and my measure of popularity (percent of likes). Consistent with the theoretical predictions, the probability of CAB (CLB) decreases (increases) as the percentage of likes increases. For product lines with a percentage of likes exceeding 60 percent, more variety leads to a higher rate of purchase (CLB). By contrast, for product lines exhibiting a percentage of likes below 60 percent, more variety leads to a lower rate of purchase (CAB).

The effect of popularity on CAB/CLB behavior is robust to the inclusion of the number of available varieties in the product line. The available number of varieties serves as a measure of cognitive load/forecasting-search cost within the standard framework of CAB. Moreover, the number of available varieties is insignificant after I control econometrically for the substitution behavior of familiar consumers. In contrast, the coefficient for the percentage of likes is similar in magnitude and identical in significance with these additional controls. The robustness of the estimated effect of popularity suggests that sampling-search is more prevalent than forecasting-search for the shoppers in my sample.

This study makes empirical and theoretical contributions to the literature on consumer search. First, on the empirical side, this study is one of the few large-scale field experiments with actual shoppers facing a large number of varieties. Second, on the theoretical side, it introduces a model of sampling-search into theories of consumer search, which have focused on pre-purchase forecasting-search to explain CAB. Third, with regards to the operationalization of the theory, this study develops a per-variety measure of CAB/CLB to test for sampling-search behavior that allows for the econometric controls of both the substitution behavior of familiar consumers and the forecasting-search behavior of unfamiliar shoppers. Fourth, the substantive contribution of this study is to show that popularity predicts CAB for niche product lines and CLB for popular product lines. This finding that reductions in the number of varieties has opposing effects on purchases offers a potential channel by which to reconcile conflicting results between prior studies of individual product lines and null findings in the meta and large-scale field studies.

The theory and experiment presented here reveal the possibility that some consumers find sampling after purchase more optimal than forecasting their own taste experiences. The results of the theory and experiment apply to situations in which the consumer has a low level of familiarity with the product line, a high opportunity cost of introspection and forecasting their tastes relative to the cost of purchase. The findings here are likely not valid for predicting behavior where sampling is not applicable, e.g., to hypothetical choices used in many laboratory studies, to one-time purchases of expensive items, financial products, health insurance, and voting, and generally, when choices are not repeated or there are no varieties.[6]

The remainder of the paper is organized as follows. Section 2 presents the conceptual framework as it relates to the econometric specification. For readers interested in the theoretical basis of the econometric specification and predictions, the Appendix contains: Section A-1, the details of the Hoteling linear taste setup; Section A-2, the theoretical implications, and; Section A-3, connects these implications to the empirical strategy and predictions. The experimental design is in Section 3. The main results are in Section 4. Section 5 presents concluding remarks.

2 Outline of Conceptual Framework

Here I outline only the parts of the framework that are necessary for specifying the econometric models that I test. See the Appendix for the mathematical development of the conceptual framework based on the assumptions of standard consumer and search theories.

Let $stops_i$ be the number of shoppers who stop in front of the shelf of product line i (for more than three seconds in the video footage), and $buyers_i$ be the number of shoppers observed taking something off the shelf of product line i.[7] The buyers-per-stop ratio, which measures the probability of purchase for product line i, is $\frac{buyers_i}{stops_i}$. I define $buys_{i,j}$ as the number of purchases of variety j from product line i (using in-store sales data). The buys-per-stop ratio, which measures the probability of purchase for variety j of product line i, is defined as

$$BS_{i,j} = \frac{buys_{i,j}}{stops_i}. \qquad \text{Eq. 1}$$

[6] Among the real/non-hypothetical choice experiments are a number that allow subjects to sample unfamiliar varieties of exotic product lines in the first stage of the experiment. In the field experimental part of Iyengar and Lepper (2000), shoppers could taste the available varieties of jam before purchase at a tasting booth. In the lab experimental part of Iyengar and Lepper's study, subjects were given a free single-unit sample in the first stage of the experiment. In the second stage, they could choose to purchase a four-unit set. My model would apply to the sampling behavior in the first stage in such cases, not to the actual purchases in the second stage of these experiments. However, if subjects sampled multiple varieties, a model that allows for dynamic updating would be required.

[7] Note that in conception, $buyers_i = \sum_{sj} buys_{sij}$, where s is the index for a specific shopper, if I had data on the individual purchases of shoppers, which I do not.

I next define my alternative measure of CAB, which I use for the main results. To measure only increases in the purchase of varieties rather than increases in the number of items purchased, I use a dichotomous outcome variable $CAB_{i,j}$

$$CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,R} > BS_{i,j,F} \\ 0, if\ BS_{i,j,R} < BS_{i,j,F} \\ \text{dropped}, if\ BS_{i,j,R} = BS_{i,j,F} \end{cases} \qquad \text{Eq. 2}$$

that takes on the value 1 if and only if the purchase rate in the reduced-variety (*R*-) treatment is strictly higher than in the full-variety (*F*-) treatment for a given variety j in product line i. In other words, $CAB_{i,j}$ excludes information about the relative intensity of purchases across the two treatments. Such intensity information would confound the number of units purchased of a given variety or varieties with the mere probability of CAB that I seek to measure. The dichotomous functional form, combined with the probit regression described below, also allows me to avoid strong functional form assumptions (e.g., linearity) in my estimations unwarranted by my theory. I dropped data where the $BS_{i,j}$ ratio was constant across the reduced-and full-variety treatments.[8] I interpret CLB as the opposite of CAB: $CLB = 1 - CAB$.

A key feature of this measure of $CAB_{i,j}$ is it compares the relative rates of purchase across the reduced-and full-variety treatment for a *fixed* variety j in product line i. This comparison of the relative rates of purchase for the fixed varieties, which are constant across the reduced- and full-variety treatments, has important advantages for measuring the sampling-search behavior of uninformed shoppers, which I now enumerate.

As a benchmark, note that the standard measure of CAB counts the number of subjects choosing an option other than the default option, termed 'active' choice, and the number of subjects retaining the default option, termed 'passive' choice. In the laboratory, the number of active choices is often the same as the number of subjects who make such choices because each subject generally can make only one choice. In field experiments, the number of active or passive choosers, not the number of active or passive choices is usually analyzed. These subjects make their choice under

---

[8] My results are largely unaffected when I include the data in which $BS_{i,j,R} = BS_{i,j,F}$ and define a) $CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,R} > BS_{i,j,F} \\ 0, if\ BS_{i,j,R} \leq BS_{i,j,F} \end{cases}$ or b) $CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,R} \geq BS_{i,j,F} \\ 0, if\ BS_{i,j,R} < BS_{i,j,F} \end{cases}$. The absolute value of the magnitude of the coefficient of the predictor, $\%like_i$, increases by approximately 1 in the case of a) and decreases by 1 in the case of b), as compared to a baseline coefficient of approximately 2.7. The level of significance of the coefficient of $\%like_i$ increases to become uniformly 1% for all specifications in the case of a) and drops to the 10% level for some specifications in the case of b). I present the results for the intermediate case in the current specification in Table 9. However, a) seems more reasonable by including in the null hypothesis the possibility that $BS_{i,j,R} = BS_{i,j,F}$,i.e., non fixed varieties have no effect on purchasing behavior among the fixed varieties, all else being equalized econometrically. The regression results for a) and b) are available on request.

two conditions: when facing a large set of varieties or a small set of varieties from the same product line. The treatment effect of increased variety is per product line. Such studies generate four numbers per product line: a count of active and of passive choices in the lab or a count of the number of active or passive choosers in the field for each of the full- or reduced-variety treatments. These numbers yield a probability of an active choice or active chooser, respectively, within each treatment.[9]

In this field study, the active choice is purchasing, while the passive is not purchasing. To derive the probability of purchase for consumers within each treatment, I need to condition on the consumer stopping in front of the shelf, i.e., condition on considering purchasing among the product lines included in this study. A higher buyers-per-stop ratio in the reduced- as compared to the full-variety treatment: $\frac{buyers_{i,R}}{stops_{i,R}} > \frac{buyers_{i,F}}{stops_{i,F}}$, reflects CAB. To illustrate, in the full-variety treatment, the number of active choosers is the total number of buyers in Table 1, which is equal to 5.

[Insert Table 1]

The number of stops minus the total number of buyers for the same treatment is the number of passive choosers, which is also equal to 5. The buyers-per-stop ratio is $\frac{5}{10}$ in the full-variety treatment and $\frac{3}{6}$ in the reduced-variety treatment. Both treatments show a 50 percent probability of purchase, which indicates no CAB.

By contrast with the standard measure of CAB, this study measures CAB for each variety j by comparing the buys-per-stop ratio, $BS_{i,j}$, across the reduced- and full-variety treatments for each variety in a product line. For example, for variety z_1 among the fixed varieties $\{z_1, z_2, z_3\}$ in Table 1, the buys-per-stop ratio for the full-variety treatment is smaller than the buys-per-stop ratio for the reduced-variety treatment: $BS_{1,F} = \frac{1}{2} < BS_{1,R} = \frac{2}{2}$. Therefore, $CAB_1 = 1$. By similar reasoning, $CAB_2 = 1$, and $CAB_3 = 1$ for the other fixed varieties. The non-fixed varieties, $\{z_4, z_5\}$ in Table 1, are the treatment varieties, which are present in the full-variety treatment, but absent in the reduced-variety treatment. While the standard measure of CAB reveals no CAB in this example, my per-variety measure shows CAB behavior for all three fixed varieties in this stylized case.

[9] A Fisher exact test for one product line or a Chi-squared test for a number of product lines can then be used to test for the statistical significance of changes in the rate of active choices to passive choices across the reduced- and full-variety treatments.

In general, my per-variety measure of CAB is a more sensitive measure of CAB/CLB than the standard per-product line measure when there are many varieties in the fixed-variety treatment. First, to see that my measure captures the notion of CAB, note that if increasing the number of varieties increases the cognitive costs of choosing, it should do so among the fixed varieties. In other words, conceptually, CAB requires only that adding the nonfixed varieties $\{z_4, z_5\}$ to the choice set decreases the sum of the purchases among the fixed varieties $\{z_1, z_2, z_3\}$ in the full variety treatment, conditional on the same number of stops, and after adjusting for the lower odds of purchasing among the fixed varieties within the larger choice set in the full variety treatment. (See the discussion of the Adjusted $CAB_{i,j}$ in equation Eq. 4 below.) The concept of CAB does not require the reduction in the probability of purchase of varieties in the full-variety set $\{z_1, z_2, z_3, z_4, z_5\}$ to be of such a degree that the sum of the purchases from this full-variety set is smaller than the sum of the purchases from the reduced-variety set $\{z_1, z_2, z_3\}$, as is currently supposed in the standard measure. I use a probit regression for the main results to calculate relative odds of CAB across treatments to allow for a mixture of cases where some varieties reveal CAB while others do not.

The standard measure also loses information by using the ratio of the sum of the buyers over the sum of stops over all varieties. By aggregating over the purchases of all buyers in each treatment, it ignores the variation in the number of purchases per variety. In contrast, by counting the number of purchases per variety for a fixed set of varieties across the full- and reduced-variety treatments per product line, my $CAB_{i,j}$ measure generates $2{*}n$ numbers, where n is the number of fixed varieties per product line. Hence, when the number of available product lines tested and the number of varieties per product line is large, as is the case for my supermarket environment, the $CAB_{i,j}$ measure yields higher statistical power than the standard measure.

Another important limitation of the standard measure of CAB is, it does not allow econometric control for the influence of familiar shoppers. To mitigate this problem, the practice in field studies is to use exotic brands that are presumed to be unfamiliar to shoppers. With laboratory experiments, where the subjects' familiarity with the product line can be elicited through a survey, the data from subjects who reveal familiarity are dropped (Iyengar & Lepper, 2000). As mentioned, my test for the treatment effect of variety on the probability of purchase per product line with the video footage data using the standard measure of CAB suffers also from the weakness that I do not observe, and therefore, cannot control for the behavior of informed shoppers. However, the point-of-sales data obtained from the store, combined with my alternative per-variety measure of CAB, does allow me to control for the effect of familiar shoppers and the forecasting-search behavior of unfamiliar shoppers on the data econometrically.

To explain how I econometrically control for the non-random purchases of familiar shoppers, I must first characterize the anticipated differences in behavior between shoppers who are familiar with the product line and shoppers who are unfamiliar across the reduced- and full-variety treatments. Unfamiliar shoppers are those who do not know which varieties match their tastes, and thus, must choose randomly among what they would regard as the full set of varieties. The varieties missing from the reduced-variety treatment do not affect their behavior across my two treatments, except insofar as their odds of choosing any variety among the fixed varieties is reduced by the ratio of the number of fixed varieties over the number of fixed and non-fixed varieties. I specifically adjust for this reduction in odds below in the discussion of the Adjusted $CAB_{i,j}$. By contrast, familiar shoppers not only choose deterministically because they know what they want, but their purchases among the fixed-varieties may vary systematically according to the treatment because of their substitution behavior.[10]

I use the stylized example in Table 1 to illustrate the three cases in which familiar shoppers may behave deterministically across the reduced- and full-variety treatments and point out in which cases their purchases may affect my results. Suppose a familiar shopper prefers a variety among the fixed varieties, e.g., z_3, most of all. This shopper would not contribute to a difference in the data between the reduced- and full-variety treatments because she buys z_3 in both treatments. In the second case, suppose the familiar shopper prefers a variety among the non-fixed varieties, e.g., z_4, which is only in the full-variety treatment, to all other varieties. There are then two potential subcases. In the first subcase, she does not derive a positive surplus from any other variety and would not purchase at all in the reduced-variety treatment when z_4 is missing, but would purchase z_4 in the full-variety treatment. Again, she would have no effect on the relative rates of purchases among the fixed varieties across the treatments.[11] In the second subcase, where the familiar shopper purchases her second-favorite option, z_3, in the reduced-variety treatment, but her favorite

[10] With regards to the behavior of consumers who are familiar with some varieties and not with others, the correlation coefficient of the regressions captures the average behavior, i.e., convex combinations of both types of consumers, which is in effect the consumer with partial familiarity.

[11] Similarly, substitution to product lines outside of the ones I measure also do not affect the results (e.g., fruit preserves instead of the jams in this example) because such purchases do not affect the relative purchase rates among the fixed varieties in across the full- and reduced-variety treatments.

option, z_4, in the full-variety treatment, her behavior would *be observationally equivalent to* CAB. Therefore, without adequate controls, a decreased probability of purchase of the fixed varieties in the full-variety treatment relative to the reduced-variety treatment can be due to the prevalence of the substitution behavior of familiar shoppers and not to the sampling-search of unfamiliar shoppers. This analysis applies similarly to consumers who are familiar with only a part of a product line. In particular, this analysis applies to their substitution behavior across the reduced- and full-variety treatments for the part with which they are familiar.

Such substitution behavior was not a confounder in prior studies using the standard measure of CAB, because, as discussed above, the standard measure requires that consumers make fewer purchases *in total* across all available varieties in a product line for the large variety treatment than in the small variety treatment. However, my per variety measure requires a lower probability of purchase per variety among the fixed varieties, rather than a lower sum over all purchases across fixed and non-fixed varieties in the full-variety treatment as compared to the sum over all purchases among the fixed varieties in the reduced-variety treatment.

Whether a reduction in purchases among the fixed varieties within the full-variety treatment as compared to the reduced-variety treatment is due to the substitution of familiar shoppers is unobserved in this study. Nevertheless, the degree to which familiar shoppers substitute to the non-fixed varieties in the full-variety treatment excluded from the reduced-variety treatment should be correlated with the share of the non-fixed varieties among all varieties in the full-variety treatment towards which they substitute. Consequently, the substitution behavior of these familiar shoppers can be controlled for by including such measures of the availability of substitutable alternatives as the share of fixed varieties ($\%fixed_i$, which is the share of the complement of these substitutable alternatives) and the share of sales of the non-fixed ($\%sales\ non\ fixed_i$) varieties among all available varieties in the regressions. The effect of the forecasting-search behavior of unfamiliar shoppers can also be controlled econometrically by including the total number of available varieties in the regression, if cognitive load/forecasting-search cost increases on the total number of available varieties, as hypothesized by the current frameworks for understanding CAB.

My controls for substitution behavior also control for the influence of the purchasing of multiple units of the same varieties in the reduced-variety treatment as a substitute for purchases of a greater number of unique varieties in the full-variety treatment. Consumers who make multiple purchases of the same varieties among the fixed varieties $\{z_1, z_2, z_3\}$ independently of the nonfixed $\{z_4, z_5\}$

would not affect my findings since my measure of CAB uses only the fixed varieties. Similarly, my results are not affected when consumers make these multiple purchases among the nonfixed or even some combinations of fixed and nonfixed varieties. As long as these multiple purchases across the fixed and non-fixed are independent of each other, they would not affect my measured CAB.

This assumption of the independence of purchases across fixed and non-fixed varieties is valid for unfamiliar consumers who purchase randomly. The only case where the consumers' multiple purchases of the same variety in the reduced-variety treatment and purchases of multiple varieties among the nonfixed in the full-variety treatment can be correlated is if the consumer is familiar with the product line and substitutes multiple units of her second favorite variety among the fixed, e.g., two units of z_3 in the reduced-variety treatment for her missing favorite varieties among the non-fixed, e.g., one unit each of z_4 and z_5 in the full-variety treatment. However, such substitution behavior amounts to substituting one unit of z_3 for one unit of z_4 and one unit of z_3 for one unit of z_5 in the, respectively, reduced-variety and full-variety treatments. Such substitution behavior is equivalent to that already dealt with above.

As mentioned, beyond controlling for the influence of familiar shoppers on the data, I also need to include a further control for the reduction in the odds of the unfamiliar shoppers choosing any particular fixed varieties in the full-variety treatment merely because the full-variety treatment has additional options that must, as a mathematical necessity, reduce the odds of shoppers choosing any of the fixed varieties randomly. To illustrate with the stylized example in Table 1, the unfamiliar shopper's odds of choosing z_1 randomly is $\frac{1}{5}$ in the full-variety treatment and $\frac{1}{3}$ in the reduced-variety treatment. Hence, even the random purchasing behavior of the uninformed shopper may be a confounder for CAB, if not properly controlled for.

To offset this lower probability of purchasing any fixed variety in the full-variety treatment, I multiply the buys-per-stops ratio ($BS_{i,j}$) 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 ratio:

$$Adjusted\ BS_{i,j,F} = BS_{i,j,F} \cdot \frac{\#\ varieties\ in\ full\ treatment}{\#\ varieties\ in\ reduced\ treatment}. \qquad \text{Eq. 3}$$

and

$$Adjusted\ CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,R} > Adjusted\ BS_{i,j,F} \\ 0, if\ BS_{i,j,R} < Adjusted\ BS_{i,j,F} \\ null, if\ BS_{i,j,R} = Adjusted\ BS_{i,j,F} \end{cases}. \qquad \text{Eq. 4}$$

The reader may realize that the mathematical necessity of unfamiliar shoppers being less likely to buy from the fixed varieties in the full-variety treatment should already be controlled for when I control for the share of fixed varieties. The $Adjusted\ CAB_{i,j}$ serves as a further robustness check. Given the adequacy of this adjustment and the econometric controls for the substitution behavior of familiar shoppers discussed above, I can now test for the treatment effect of the mere presence or absence of the non-fixed varieties on purchases among the fixed varieties.

Turning to predictions, I formulate hypotheses in terms of increased varieties rather than reduced varieties to be consistent with the conception of CAB in the literature. The definitions of all variables are collected in Table 8 for the convenience of the reader. The mathematical derivations of the predictions are in the Theoretical Appendix. Here, I merely summarize the main results and cite the relevant implication from the theory.

Within the standard Hoteling linear taste framework, firms only increase the number of varieties if consumers are heterogeneous in their tastes. Given that consumers understand this motivation at an intuitive level, increasing the number of available varieties in the product line increases the consumer's perception of the level of heterogeneity in tastes and their belief in the marginal mismatch cost between the tastes of the consumer and the product characteristics (Assumption I). Increases in the marginal mismatch cost decreases the expected value of the initially random purchases required for sampling-search (Implication I), while at the same time, it increases the optimal sample size (Implication II), conditional on the optimality of purchasing at all. These comparative statics results imply that the marginal mismatch cost increases the varieties unfamiliar consumers purchase, if the product line serves popular tastes. In contrast, the marginal mismatch cost decreases the odds that unfamiliar consumers purchase at all, if the product line serves niche tastes (Implication III). In sum, these comparative statics results imply that when unfamiliar consumers face a greater number of varieties for a popular product line, they exhibit CLB. In contrast, when unfamiliar consumers face a greater number of popular varieties for a niche product line, they exhibit CAB. Accordingly,

Prediction I. CAB decreases and CLB increases with the popularity of the product line.

More formally, for a variety j among the fixed varieties in product line i and product line i', I predict

$$Pr(CAB_{i',j} = 1 | popularity_{i'}) \geq Pr(CAB_{i,j} = 1 | popularity_i). \qquad \text{Eq. 5}$$

for $popularity_{i'} < popularity_i$, where popularity is measured by the percent of likes. I define popularity mathematically in equation A-Eq. 3 in the Appendix, operationalize this definition in Section A-3, and explain the measurement of percent of likes in detail in Section 3 Experimental Design.

Prediction I is about the aggregate purchases of all shoppers of a specific variety. It leaves open the possibility that instead of there being more purchases per shopper (when the number of options increases for popular product lines, as shown in Implication II) there can merely be a larger share of shoppers who purchase. Prediction II refines Prediction I by specifying that the increase in the purchases from the increase in variety for popular product lines is from an increase in the average number of purchases per shopper. Therefore,

Prediction II. The average units purchased per shopper among the fixed varieties increase on the interaction between the treatment and the popularity of the product line.

More formally, let $sales_i = \sum_j buys_{i,j}$. For product line i and i'

$$\begin{aligned} &Average\left(\frac{sales_{i,F}}{buyers_{i,F}}(popularity_i) - \frac{sales_{i,R}}{buyers_{i,R}}(popularity_i)\right) \geq \\ &Average\left(\frac{sales_{i',F}}{buyers_{i',F}}(popularity_{i'}) - \frac{sales_{i',R}}{buyers_{i',R}}(popularity_{i'})\right) \end{aligned} \qquad \text{Eq. 6}$$

where F=full- and R=reduced-variety treatments for $popularity_{i'} < popularity_i$.

3 Experimental Design

Previous studies tested *whether* CAB/CLB occurs for a small number of product lines. In contrast, I use shoppers' belief about the popularity of a product line to predict the relative likelihood of CAB/CLB occurring across a large set of product lines. To estimate shoppers' beliefs about the popularity of a product line, four student research assistants from Peking University asked shoppers as they exited the supermarket to take a survey for a small gift (e.g., colored pens). The research assistants were not informed about the hypotheses of the experiment. In this survey, the shopper's "likes", "neutrals", "dislikes", and "untried" for each variety (e.g., chocolate, vanilla,…) in each of the 24 product lines (e.g., Oreo biscuits) were elicited.[12] See Table 2 for an example of a survey for Oreo biscuits for one shopper.

[Insert Table 2 here]

Based on this survey, I calculate the $\%like_i$, which is equal to the count of likes divided by the sum of the count of likes, neutrals, and dislikes, i.e., tries:

$$\%like_{i,week} = \frac{\#likes_{i,week}}{\#likes_{i,week} + \#neutrals_{i,week} + \#dislikes_{i,week}} \qquad \text{Eq. 7}$$

for product line i where $i = 1, 2, \ldots, 24$ and $week = 1, 2$. The probability is the average of the $\%like_i$s for each of the two treatment weeks:

$$\%like_i = \frac{\%like_{i,1} + \%like_{i,2}}{2} \qquad \text{Eq. 8}$$

$\%like_i$ can be interpreted as an estimate of the probability that a person considering purchasing from product line i likes a random variety of product line i—among those varieties the shopper is likely to consider.[13]

For the survey, I chose to use the varieties listed on www.jd.com, a popular online store's website with a large selection of each product line used in the experiment. Using the more comprehensive listing of varieties per product line helps to measure unfamiliar shoppers' general perception of the product line. Such general perceptions seem more appropriate as estimators than surveying shoppers for the varieties which happen to be on the shelves at this particular store on a particular day and possibly because of a promotion or other idiosyncratic/transient factors. Moreover, I only consider varieties of qualities/tastes in the survey rather than quantities/sizes. Since sizes are observable to all consumers, all consumers choose them deterministically. I did not attempt to replicate Iyengar and Lepper's (2000) jam result because jam, similar to bread and toast, is still a novelty in China. There were only a few varieties of any brand of jams.

The survey used to derive $\%like$ was unincentivized. I do not suppose that shoppers would be embarrassed or have other disincentives to admit their preferences. At worst, shoppers may be

[12] I had unexpected problems with the mini-video cameras for six product lines and had to drop them from the study, which originally had 30 product lines.

[13] $\%like$, 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 my 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 I expect these two types of popularity converge in outcomes.

careless, which would add noise to the predictor. That would result in underestimating the actual correlation between CAB and $\%like$.

Figure 1 illustrates a stylized version of the survey results. $\%like_{i,week}$ is the area of 1s in the area of 1s, -1s and 0s. I use these numbers 0, 1, and -1 merely for visual clarity and not for any calculations. Each of the 48 actual tables summarizing the data for each of the 24 product lines used in this paper on two treatment days contain 30 columns, one for each shopper surveyed, and as many rows as there are varieties in the product line.

[Insert Figure 1 here.]

As an example of the actual data, for Want-want QQ gummies candy, $\%like = 0.70$, while for Oreo biscuit, $\%like = 0.59$. According to the conceptual framework, uninformed shoppers are more likely to experience CAB (less likely to experience CLB) when choosing Oreo biscuits than when choosing Want-want QQ gummies candy. Table 4 displays $\%like$s for each product line used in the experiment. While $\%like$ is not randomly assigned to product lines, the correlation between the number of varieties within a product line and its $\%like$ is insignificant (adjusted R^2=0.01, F-test p-value=0.62). The average $\%like$ for all product lines is 66 percent when equally weighted across all product lines and both treatment and control days.[14]

I use $\%like_i$ to estimate the perceived likely rate of success for those who were considering purchasing (since they stopped in front of the shelf) using the reported preferences of shoppers who had already purchased. Conceivably, $\%like_i$ may be a biased estimator when applied to those shoppers who consider purchasing but who may or may not actually purchase. However, I use $\%like_i$ to estimate the cross-product line variation in purchases between the full- and reduced-variety treatments. $\%like_i$ would be an unbiased estimator if the difference in the within-product line purchasing rates of those who merely consider and those who actually purchase is constant across the product lines in this study. Clearly, the within-product line cost of purchasing after stopping before the shelf is constant 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 cross-product line

[14] Due to an error, I was unable to use the Sept. 13 survey data for Lipton teas. Therefore, I calculate the *%like* using only the survey data from Sept. 27. The results are qualitatively identical and quantitatively nearly identical when I estimate the Sept. 13 *%like* for Lipton teas using the Sept. 27 *%like* and coefficients derived from the Sept. 13 and Sept. 27 results from other product lines.

cost of purchasing should also be constant because the aisles are nearly identically structured for all product lines, as they generally are in most supermarkets.

The total likes, neutrals, and dislikes are 2643, 1130, and 351 respectively, summing to a total of 4124 tried. The percent of likes, neutrals, and dislikes using these totals is 64, 27, and 9 percent, respectively. Among those who have tried, most liked, and very few were neutral or disliked the products they tried. Such a distribution is to be expected in a competitive market where disliked and neutral varieties would presumably be quickly dropped by stores. The percent of tried compared with the sum of tried and untried (4124+14116) is 23 percent. Hence, most shoppers at this store are unfamiliar with the varieties used in the experiment.[15]

The store opens Monday to Thursday 8:00 am-10:30 pm and Friday to Sunday 8:00 am-11:00 pm. 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.[16] Each of the 24 product lines was randomly selected for the reduced-variety treatment on a specific day of the week (e.g., Wednesdays for Oreo biscuits) across two weeks. Approximately eight product lines were tested per day.

In the reduced-variety treatments, the supermarket staff, who were uninformed of the hypotheses being tested, moved the non-fixed varieties off the shelf, filling the rest of the shelf with the remaining varieties from additional stock. Hence, the shelf space was fixed in size across the full- and reduced-variety treatments. The product lines were selected based on two criteria: a high number of varieties and easy re-shelving, e.g., not bottles of wine. The selected product lines are listed in in Table 4.

Between one-half to one-third of the varieties were removed, depending upon whether and wherein the product line there was a steep drop-off in popularity (as measured by store sales data from one day before the experiment on September 7, 2013). The least popular varieties were removed to minimize the imposition on shoppers familiar with the product line who might have sought missing varieties and to minimize the effect of the substitution behavior of these familiar

---

15 The prevalence of unfamiliar shoppers may be expected because Shenzhen was a fishing village 30 years ago. It has been one of the fastest-growing cities in China. Now, it is a teeming metropolis of 12 million with mostly recently arriving migrants to the city. The store where the experiment was conducted is in one of the newer districts of the city.

16 The Mid Autumn Festival and National Day of the People's Republic of China fell respectively on September 19 – 21 and October 1-7 in 2013. These holidays may have affected purchasing behavior for some product lines, e.g., dishwashing detergent. However, these holidays should not affect my experimental findings because I randomly assigned my treatments for each product line across the days of the experiment.

shoppers on the data. See Table 4 for the number of removed items for each product line, their share of the product line, and their percentage share of sales per product line.

Table 5 shows that none of the correlations between these variables and $\%like$ are significant. This lack of correlation suggests that the removed items, and therefore, the substitution behavior of familiar shoppers, would not affect the main results. However, I control for the share of the removed product lines and their percentage share of sales to further rule out the effect of possible substitution behavior in the main results.

All experimental setup for the next day was done in the previous night after the store closed. Cameras were installed on the ceilings above the shelf to record the $pass\ bys$, $stops$, and $buys$ from above and behind the shopper.[17] The floor manager may not have turned on the cameras precisely at the store opening time. Therefore, to avoid unobserved factors influencing the data, and also because I already have 35,694 shoppers between 5:30 pm-9:30 pm, I limit this study to the shoppers during this the highest traffic time.

Though the store management agreed to maintain constant prices across both weeks, some prices were found to have changed after the experiment was completed. However, random assignment of product lines, first to a specific day of the week, and then to a specific week across the two weeks of the experiment for the reduced-variety treatment, should counteract any systematic influence of these price changes on the results. Moreover, I control for prices or their changes in all regressions. Of the 35,694 shoppers who were observed to have passed by in the video footage, 3,291 stopped (for more than 3 seconds), and 607 purchased (presumed if an item was taken off the shelf). I use store provided sales data for the whole day instead of the purchases observed in the video footage because the store sales data allow me to reliably identify each of the 339 varieties within the 24 product lines purchased. The purchased varieties are sometimes obscured by the shopper, because the video footage was taken from above and behind shoppers. Identifying the variety is necessary for distinguishing whether a purchase was among the fixed varieties or the non-fixed varieties in the full-variety treatment. This store data recorded 1,530 sales.

17 Hence, the shopper's privacy is preserved.

4 Results

The first set of results is motivated by previous studies rather than by my theoretical framework. First, I test whether the reduction in the numbers of varieties has any effect on stops at the shelf for longer than 3 seconds ($stops_i$). Recall that Iyengar and Lepper (2000) find that significantly more people stopped in front of their full-variety display than for their limited-variety display. I estimate the probability of stops for the full- and reduced-variety treatments with Eq. 9.

$$\frac{stops_{i,treatment}}{passbys_{i,treatment}} = \alpha_0 + \beta_0 treatment + \beta_1 price_{i,treatment} + \beta_2 week + \varepsilon_{i,treatment} \quad \text{Eq. 9}$$

$price_{i,treatment}$ is the average price of each product line i in each treatment. Table 6 reveals that the coefficient for the treatment is not significant for any model.

[Insert Table 6 here.]

This result is summarized in Observation I.

Observation I. The probability of stops in front of the shelf does not increase with the number of varieties displayed, fixing display size.

One possible reason for the difference between the finding in Observation I and Iyengar and Lepper's (2000) is that, as mentioned in the introduction, they used a separate display to attract shoppers and hand out discount coupons that was larger for the large variety treatment, as well as the regular shelf space, where shoppers can remove products for purchase. By contrast, I used only the regular shelf space which was fixed in size for both the full- and reduced-variety treatments. Thus, the inconsistency between the finding here and Iyengar and Lepper's can be explained by their conjecture that physically larger displays draw more shoppers.

I next test the effect of the reduction in the number of varieties on purchasing behavior. Using the video footage, I count the number of buyers ($buyers_{i,treatmeant}$) for each product line in each treatment. I use this to estimate the probability of purchase per individual shopper per product line i in Eq. 11.

$$\frac{buyers_{i,treatment}}{stops_{i,treatment}} = \alpha_0 + \beta_0 treatment + \beta_1 price_{i,treatment} + \beta_2 week + \varepsilon_{i,treatment} \quad \text{Eq. 10}$$

Table 7 reveals that the coefficient for the variety treatment is not significant in any model.[18] This result also contrasts with Iyengar and Lepper's finding that shoppers faced with more varieties were less likely to purchase. However, as they acknowledged, their larger display for their large variety treatment may not only have drawn more shoppers, but may also have drawn shoppers who were less motivated to purchase. This lack of significance is, nonetheless, consistent with the finding of an average zero effect of the number of varieties found in the meta and field studies.

[Insert Table 7 here.]

Accordingly,

Observation II. Changes in the number of varieties do not have a significant effect on the probability of purchase.

Columns (4)-(6) of Table 7 shows that the interaction of the treatment with $\%like_i$ is also insignificant in explaining the probability of purchase. The lack of significance of $\%like_i$ in predicting the probability of purchase can be due, first, to the purchases of familiar shoppers not being controlled for, and second, to the modest sample size of 48 observations. I deal with these issues next using my alternative measure of CAB/CLB.

To prepare for the test of Prediction I, that the rate of success, as measured by $\%like_i$, predicts CAB or CLB, I collect the definitions of all of the variables that I use in the main estimation of Eq. 11 with a summary of their interpretation in Table 8 for the convenience of the reader.

[Insert Table 8 here.]

I test Prediction I by estimating the probability of choice-averse behavior ($CAB_{i,j} = 1$) for variety j in product line i as a function of $\%like_i$, controlling for changes in prices across treatments. Controls for the level of prices are unnecessary because the dependent variable $CAB_{i,j}$

[18] I find a similar lack of significance for the coefficient of the dummy variable for treatments when I 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 price per product line. Because the sample size of this regression is 48, the regression result is merely suggestive. I also find a lack of significance in the coefficient for the treatment dummy when I use raw sales for each product line instead of the percentage change in sales. These results are available upon request.

records treatment effects for a fixed variety j. The basic model is in Eq. 11, which I estimate using the probit regression is

$$
\begin{aligned}
&Pr\left(CAB_{i,j} = 1 \middle| \%like_i, price\ change_{i,j}, D_2,\ \ldots D_{23}\right) \\
&= \Phi(\alpha_0 + \beta_0 \%like_i + \beta_1 \cdot price\ change_{i,j} \\
&\qquad + \sum_{k=2}^{23} \beta_k \cdot D_k \cdot price\ change_{i,j,}, control\ variables)
\end{aligned}
\qquad \text{Eq. 11}
$$

Here, $price\ change_{i,j} = \frac{price_{i,j,F} - price_{i,j,R}}{price_{i,j,R}}$ measures the percent price change between the reduced-variety and full-variety treatments for variety j in product line i based on point-of-sale data. The dummies D_k multiplied by $price\ change_{i,j}$ control for the possibility of variations in the demand curves for k different product lines. With regards to the sample size of 152-fixed varieties, from the total 339 varieties, 182 are fixed varieties. 30 that have ties in their buys-per-stops ratio ($BS_{i,j}$) between the full- and reduced-variety treatments. That leaves 152 fixed varieties for the dependent variable of the regression, $CAB_{i,j}$, to take on the value of either 0 or 1, as defined in Eq. 2.

Table 9 displays the main result from the probit regression. Consistent with Prediction I, the coefficient for $\%like_i$ is negative. This shows that CAB is more likely for low popularity product lines. A downward sloping demand curve is consistent with the positive coefficient for $price\ change_{i,j}$, which implies that shoppers purchase fewer units in the full-variety treatment when prices increased relative to the reduced-variety treatment.

While the correlation matrix in Table 3 reveals that $\%like_i$ and $\%dislike_i$ and $\%like_i$ and $\%neutral_i$ are highly correlated, the lack of significance for $\%neutral_i$ and $\%dislike_i$ in the regression in column (1) of Table 9 might still be expected because I removed the least popular varieties. Similarly, removing unpopular varieties may also have diminished the significance of the coefficient for $\%tried_i$, which includes $\%neutral_i$ and $\%dislike_i$, as well as $\%like_i$. In addition, as mentioned in Section 3, the sample sizes for $\%neutral_i$ and $\%dislike_i$ in the survey were fractions of that for $\%like_i$, which would further reduce their significance.

It is interesting to note that the coefficient for $\%like_i$ in column (2) of Table 9 multiplied by the average $\%like_i$: -2.781·0.65=-1.80, is similar in magnitude and opposite in sign to the value of the constant term, 1.87. This is consistent with the CAB balancing CLB at the average $\%like_i$. Such balancing of CAB and CLB at the average $\%like_i$ in large supermarkets would help explain the zero average effect of decreases in variety found in Observation II, Boatwright and Nunes (2001, 2004) and in the meta-studies.

Column (9) shows that the constant term (-1.344) is itself insignificantly different from zero when $\%fixed_i$ is included in the specification. The coefficient for $\%fixed_i$ is the opposite sign and twice the magnitude of the coefficient of $\%like_i$, but of lower significance. The opposing signs and similar magnitudes of $\%like_i$ and $\%fixed_i$ suggests that the substitution behavior of familiar shoppers may indeed be cancelling the sampling-search behavior of unfamiliar shoppers for some product lines.

Column (7) of Table 9 reveals that the number of varieties in the full-product line ($\#full_i$) is significantly positively correlated with $CAB_{i,j}$. This positive coefficient of $\#full_i$ is consistent with the hypothesis that a larger number of varieties increases the cognitive load for unfamiliar shoppers, and hence, CAB in the standard framework. However, the positive coefficient of $\#full_i$ is also consistent with my hypothesis that the substitution behavior of familiar shoppers increases with a greater number of non-fixed alternatives increases the substitution behavior of familiar shoppers. Indeed, consistent with this hypothesis that $\#full_i$ is capturing substitution behavior, column (10) shows that $\#full_i$ loses significance when the share of fixed varieties ($\%fixed_i$) is included. $\%like_i$ remains similar in magnitude and identical in significance. Moreover, columns (8), (9), and (10) show that coefficient for $\%like_i$ is similarly significant with the inclusion of any combination of these controls.

Column (11) shows that the percent of sales of the non-fixed varieties ($\%sales\ non\ fixed_i$), which controls for the substitution behavior of familiar shoppers weighted by sales, is not significantly correlated with $CAB_{i,j}$. The lack of significance of the coefficient for $\%sales\ non\ fixed_i$ is likely due to the large standard error, which could have arisen from the fact that the least popular varieties by sales volume were removed in the reduced-variety treatments. The insignificance of substitution behavior, in particular, is to be expected given that the rate of familiarity among shoppers for the varieties included in the experiment is 23 percent. In any case, column (12) shows that the coefficient for $\%like_i$ gains in significance when the sales of the non-fixed varieties is included.

[Insert Table 9 here.]

Table 10 shows that both the magnitude and significance of the coefficient for $\%like_i$ in Table 9 are preserved when I use the $Adjusted\ CAB_{i,j}$ to take further into account that the probability of purchase of any specific variety is lower in the full-variety treatment than in the reduced-variety treatment for unfamiliar shoppers' random choice among available varieties, as discussed in Section 2. However, as shown in columns (7)-(12), all of the control variables are now insignificant.

[Insert Table 10 here.]

The coefficient for $\%like_i$ remains roughly constant across these different specifications and across Table 9 and Table 10, increasing the confidence that $\%like_i$ is the key factor influencing unfamiliar shoppers purchasing behavior.

Figure 2 illustrates the main result from model (1) in Table 9. CAB is decreasing on $\%like_i$ when prices and other control variables are held constant.

[Insert Figure 2 here.]

Figure 2 also reveals that shoppers are more likely to exhibit CLB instead of CAB when variety is increased and $\%like_i$ is above 0.60. The main finding exhibited in Table 9, Table 10, and Figure 2 is summarized in Observation III, which confirms the prediction in Prediction I.

Observation III. CAB decreases and CLB increases with shoppers' belief about the popularity of the product line.

While the substitution behavior of familiar shoppers should already be controlled for in Table 9 and Table 10 with the inclusion of $\%fixed_i$ and $\%sales\ non\ fixed_i$, within the Hoteling framework of Section 2, $\%like_i$ can itself be interpreted as a proxy for the degree of substitutability for familiar shoppers among varieties in the product line, rather than as a measure of sampling-search behavior engaged in by unfamiliar shoppers.[19] To see this alternative interpretation, recall that $\%like_i$ estimates the popularity of product line i, which is modeled by $1 - F_i(p_i)$ in Section A-2 of the Appendix. The increase in $\%like_i$ corresponds to increases in $1 - F_i(p_i)$ in the theoretical model. As also shown there in equation A-Eq. 13, increases in $1 -$

[19] I thank Miguel Vilas-Boas for pointing this out.

$F_i(p_i)$ decreases $F_i(u_i) = \frac{u_i - \bar{u}_i}{v_i - \underline{u}_i}$ and can increase the lower bound of the support of the distribution, $\underline{u}_i$, which decreases the dispersion in utility outcomes if the upper bound of the distribution, v_i, is fixed, as is assumed here.

The decrease in the dispersion of utility outcomes can be from a decrease in marginal mismatch cost t_i within the Hoteling framework discussed in Section 2. Such a decrease in mismatch cost entails a decrease in the utility gap between the shopper's favorite variety and her second favorite variety. Such a decrease in the utility gap could affect my results if the shopper's favorite variety yielded her a positive surplus from purchasing and was among the non-fixed varieties in the full-variety treatment while her second favorite variety among the fixed-varieties in the reduced-variety treatment yielded her negative surplus. This shopper would only purchase her favorite variety among the non-fixed varieties in the full-variety treatment and not purchase at all in the reduced-variety treatment, which contains only her second favorite variety, when $\%like_i$ is low. By contrast, when $\%like_i$ is high, she would not only purchase her favorite variety among the non-fixed varieties in the full-variety treatment, but she would also purchase her second favorite variety in the reduced-variety treatment. Thusly, the reduction in the dispersion of utility outcomes implied by a high $\%like_i$ could indicate an increase in the shopper's utility from her second favorite variety such that she receives a positive surplus from purchasing it.

This substitution behavior would not increase observations of CAB uniformly across all $\%like_i$, as would the substitution behavior of familiar shoppers discussed in Section 2. Rather, this substitution behavior varies across different values of $\%like_i$. Such substitution behavior would specifically be less likely for low $\%like_i$ and more likely for high $\%like_i$. In other words, such substitution behavior would predict a positive sign for the coefficient of $\%like_i$: a higher $\%like_i$ is associated with higher odds of CLB, while a lower $\%like_i$ is associated with higher odds of CAB, which is the opposite of what I find in Table 9. Hence, I can rule out this second form of substitution behavior as an interpretation of Observation III.

To test Prediction II, that purchases per buyer in the full-variety treatment increase with high $\%like_i$, I estimate

$$\frac{sales_{i,treatment}}{buyers_{i,treatment}} = \beta_0 + \beta_1 full + \beta_2 D\_like_i + \beta_3 full \cdot D\_like_i + \beta_4\, price_{i,treatment} + \beta_5 week + \varepsilon_{i,treatment} \quad \text{Eq. 12}$$

where $D\_\ \%like_i$ is a dummy that takes on the value of 1 when $\%like_i$ is above 0.60 in Figure 2:

$$D\_like_i = \begin{cases} 1, if\ \%like_i > 0.60 \\ \\ 0, if\ \%like_i \leq 0.60 \end{cases} \qquad \text{Eq. 13}$$

The estimation results are displayed in Table 11. The significance of the dummy $full$ in columns (1)-(3) reveals that the units purchased of the fixed varieties per shopper in the full-variety treatment is larger than in the reduced-variety treatment. Columns (4)-(9) show that the treatment dummy $full$ is no longer significant when the interaction between the treatment dummy and $\%like_i$ ($full \cdot D\_like$) is included. This interaction term itself is insignificant when I do not control for product line fixed effects in columns (4)-(6). This lack of significance can be due to collinearity between $full$ and $full \cdot D\_like_i$ (correlation= 0.7405 with p-value=0.00). However, the interaction is significant when I include fixed effects in columns (7)-(9).

[Insert Table 11 here.]

Consistent with Prediction II,

Observation IV. The units purchased per shopper among the fixed varieties increase on the interaction between the treatment and the popularity of the product line.

5 Discussion

In this study, I find that greater variety does not attract more shoppers, fixing the size of the display (Observation I). I also find that greater variety does not increase the probability of purchase using video footage data (Observation II). This latter result supports the finding in the meta and field studies that the average effect of variety on the probability of purchase is zero (Boatwright & Nunes, 2001, 2004; Chernev et al., 2015; Scheibehenne et al., 2010). However, as acknowledged already, these parallel null effects of variety on active choices/purchases in prior studies and purchases in my study may merely be coincidences. To test that the conflicting/canceling results across prior studies are driven by heterogeneous levels of popularity of the product lines used, as hypothesized in the present study, one would need to collect data on the popularity of the product lines used, since such data was not collected in the original studies.

The main contribution of this study is to develop and test a model of consumers' sampling-search over unfamiliar varieties. This model predicts that the probability of CAB decreases and the probability of CLB increases with the popularity of the product line. This prediction was tested in a field experiment using a simple survey measure of popularity as a predictor and the relative purchasing rates of a large number of shoppers among a large number of varieties across the reduced-and full-variety treatments as the outcome. As expected, the popularity of the product line predicts the relative rates of shoppers' CAB and CLB (Observation III). This model is applicable when the expected surplus of purchasing is high enough to motivate consumers to purchase, but the returns to forecasting their own taste before a purchase are not sufficient to offset the cost. To my knowledge, such a model has been absent in previous models of consumer search.

This study overcame a set of important technical problems by developing a method of measuring CAB/CLB per variety instead of using the standard measure, which is per product line. The per-variety measure of CAB/CLB greatly increases statistical power when each product line has a large number of varieties. Moreover, the per-variety measure also allows econometric control for the substitution of familiar shoppers that prior field studies attempted to do by restricting their product lines to those of exotic brands or by dropping data in laboratory settings. The main regression in Table 9 shows that the substitution behavior of familiar shoppers and the forecasting-search of unfamiliar shoppers do not drive the findings. Further supporting the sampling-search hypothesis, I find that there were specifically more purchases per unfamiliar shopper for popular varieties when the shelf contained more varieties, as predicted by the theory in Section 2, rather than from a larger share of shoppers who purchase (Observation IV).

The validity of the two-period model of consumers trading off the cost and benefit of the information gained from random purchases on across-subject data proposed here relies upon several assumptions which I now spell out. The assumption that unfamiliar shoppers gain information from initially random purchases requires only that consumers have a normal memory combined with a modicum of rationality. Given this assumption, unfamiliar shoppers face a tradeoff between the cost and the benefit of the information gained. Hence, the validity of the model then depends on whether the observed opposing treatment effect on sales from a reduction in variety is due to this tradeoff or to some other motivations. However, the alternative motivations to sampling-search, namely the forecasting-search of unfamiliar shoppers and the deterministic purchases of familiar shoppers, were controlled for econometrically in the main results and shown to be insignificant.

That said, this study does not observe consumers actually calculating the cost and benefit of the information gained from the initially random purchases. In particular, it is not clear whether and how consumers can prospectively optimally tradeoff the monetary costs and benefits of information gained from random purchases, as predicted by my theory. Nonetheless, there is an intuitive basis to believe that consumers make such tradeoffs optimally.

Recall that for my model to be a good approximation of shoppers' behavior, shoppers would need to associate a greater risk of mismatch in tastes from observing more varieties. This is intuitive, because greater variety implies greater heterogeneity in tastes, and consequently, lower popularity of any particular variety. Such an inference would seem both natural and supported by the well-known result from the standard Hoteling model of product differentiation. Given the greater perceived risk of mismatch from observing more varieties, consumers should have diminished expectations of their surplus as well a larger sample size, if the expected value of sampling-search is greater than the value of their outside option. The diminished expectations lead to lower average rates of purchase, while the need for a larger sample size leads to higher average rates of purchase. Which effect dominates depends upon the level of risk, which in this experiment is determined by the popularity of the product line.

Thus, the remaining question is whether consumers can, at low cognitive cost (e.g., non-deliberatively), optimally take into account the greater risk of mismatch. In support of this possibility, there is a well-established stream of research in psychology showing that subjects in risky hypothetical situations in laboratory settings may decide advantageously through mere repeated feedback before explicitly knowing in the sense of 'system 2' the advantageous strategy (Bechara, Damasio, Tranel, & Damasio, 1997; Chiu, Huang, Duann, & Lin, 2018). Shoppers in my experiment may "learn" in the sense of non-deliberative 'gut-level' (Gigerenzer, 2007) or 'system 1' reactions (Kahneman, 2011) even more efficiently when the feedback is based on repeated experiences with supermarket products, where the stakes are real, rather than merely hypothetical. Thus, while this study does not observe consumers actually calculate the tradeoff between the cost and benefit of the information gained from random purchases, there is some basis beyond this study in the social science literature to believe that they do.

The field experiment in this study itself has important limitations. I use $\%like_i$ to predict *the relative likelihood* of CAB/CLB for a given product line for shoppers in aggregate rather than their actual occurrence for a given individual shopper. Moreover, while I do have random assignment of the treatments to product lines for customers across particular days across the two weeks of the experiment (which ensures a constant distribution of consumer types across full- and reduced-variety treatments and product lines), I do not have random assignment of $\%like_i$. Accordingly, it is possible that other unobserved factors correlated with $\%like_i$, e.g., unobserved cross-product line differences in product complexity, quality…and advertising expenditure can be driving the correlation between CAB/CLB and $\%like_i$. I leave for future research the investigation of the relation of $\%like_i$ to other potential factors.

I find that neither the substitution behavior of familiar shoppers nor the forecasting-search behavior of unfamiliar shoppers is significant. The former can be due to the low level of familiarity of shoppers for the product lines used in my experiment, while the latter may be due to the fact that I record ordinary shoppers within the cognitively non-pristine environment of a supermarket. Unlike the college students in laboratories who are the usual subject of choice experiments, such shoppers may be too distracted, fatigued, or in a rush to engage in attempts at forecasting their tastes for unfamiliar varieties.

The low level of familiarity I find among the shoppers in the store for the product lines in the experiment limits the scope of the implications of the findings for store managers. The results suggest that store managers can potentially increase the purchases of shoppers by increasing the varieties that customers face for products with higher $\%like_i$—should their shoppers be as unfamiliar as those in my experiment.

On the other hand, the results here would seem to suggest that there is no optimal level of variety for all product lines for all stores. Instead, they suggest that the optimal level would depend on $\%like_i$—given that the share of unfamiliar shoppers is large enough. In this respect, my findings may be helpful even in mature markets for helping to optimize the number of varieties to introduce at once for a new product line.

Despite the shortcomings of the conceptual framework and the experimental design, this study contributes significantly to the economics and marketing literatures on CAB by providing evidence that what has been regarded as unpredictable and conflicting results can be explained by a relatively straightforward theory of sampling-search. This theory was tested in a large-scale field experiment that was exceptional in the number and representativeness of subjects observed, and the realism of the conditions under which they were observed. The theory and the field experimental corroboration suggest that these conflicting prior results may be a part of a spectrum of predictable findings.

References

Bechara, Antoine, Damasio, Hanna, Tranel, Daniel, & Damasio, Antonio R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science (New York, N.Y.)*, *275*(5304), 1293–5.

Bertrand, Marianne, Karlan, Dean, Mullainathan, Sendhil, Shafir, Eldar, & Zinman, Jonathan. (2010). What's advertising content worth? Evidence from a consumer credit marketing field experiment. *Quarterly Journal of Economics*, *125*(1), 263–306.

Boatwright, Peter, & Nunes, Joseph C. (2001). Reducing Assortment: An Attribute-Based Approach. *Journal of Marketing*, *65*(3), 50–63.

Boatwright, Peter, & Nunes, Joseph C. (2004). Correction note for "reducing assortment: An attribute-based approach." *Journal of Marketing*, *68*(July).

Bronnenberg, Bart J., Dubé, Jean Pierre H., & Gentzkow, Matthew. (2012). The evolution of brand preferences: Evidence from consumer migration. *American Economic Review*.

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

Chernev, Alexander, Böckenholt, Ulf, & Goodman, Joseph. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, *25*(2), 333–358.

Chiu, Yao-Chu, Huang, Jong-Tsun, Duann, Jeng-Ren, & Lin, Ching-Hung. (2018). Twenty Years After the Iowa Gambling Task: Rationality, Emotion, and Decision-Making. *Frontiers in Psychology*, *8*.

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

Gneezy, Ayelet. (2017). Field Experimentation in Marketing Research. *Journal of Marketing Research*, *54*(1), 140–143.

Iyengar, Sheena S., Jiang, Wei, & Huberman, Gur. (2004). How Much Choice is Too Much ?: Contributions to 401 (k) Retirement Plans. *Pension Design and Structure: New Lessons from Behavioral Finance*, 83–95.

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

Iyengar, Sheena S., & Lepper, Mark 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, Emir. (2008). Contextual Inference in Markets: On the Informational Content of Product Lines. *American Economic Review*, *98*(5), 2127–2149.

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

Kuksov, Dmitri, & Villas-Boas, J. Miguel. (2010). When More Alternatives Lead to Less Choice.

Marketing Science, *29*(3), 507–524.

Liu, Lin, & Dukes, Anthony. (2016). Consumer Search with Limited Product Evaluation. *Journal of Economics and Management Strategy*, *25*(1), 32–55.

Machina, Mark J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature*, *27*(4), 1622–1668.

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

Ong, David. (2012). Can there ever be too many flavors of Häagen-Dazs? Product stereotypes and choice overload behavior. *Working Paper*.

Scheibehenne, Benjamin. (2008). The Effect of Having Too Much Choice. *Working Paper*.

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

Simester, Duncan. (2017). Field Experiments in Marketing. *Handbook of Economic Field Experiments*, *1*, 465–497.

Sun, Monic. (2012). How does the variance of product ratings matter? *Management Science*, *58*(4), 696–707.

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

Waldman, Don E., & Jensen, Elizabeth J. (2019). *Industrial Organization: Theory and Practice*. *Textbook*. Routledge.

Tables and figures

TABLE 1: STYLIZED FULL- AND REDUCED-VARIETY TREATMENTS

	z_1	z_2	z_3	z_4	z_5	Total
Full-Variety Treatment						
$buyers_i$	1	1	1	1	1	5
$buys_{ij}$	1	1	2	1	1	7
$stops_i$	2	2	2	2	2	10
$BS_{i,j,F}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{2}$			
Reduced-Variety Treatment						
$buyers_i$	1	1	1			3
$buys_{ij}$	2	2	3			7
$stops_i$	2	2	2			6
$BS_{i,j,R}$	$\frac{2}{2}$	$\frac{2}{2}$	$\frac{3}{2}$			
$CAB_{i,j}$	1	1	1			

Notes: See Table 8 for definitions of these terms.

TABLE 2: 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.

Shoppers

Product line i	Variety	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 1: A STYLIZED EXAMPLE OF SURVEY DATA

Notes: Rows represent survey responses for different varieties of product line i. Columns represent responses for different shoppers who took the survey. Likes correspond to 1, neutrals to 0, and dislikes to -1. These numbers are merely for visual clarity and are not used within the analysis. Untried varieties are assigned no value. $\%like_i \; = \frac{\#likes_i}{\#likes_i + \#neurtrals_i + \#dislikes_i}$ For this example, the $\%like = \frac{11}{15}$.

TABLE 3: CORRELATION MATRIX OF $\%LIKE$, $\%DISLIKE$, $\%NEUTRAL$, AND $\%UNTRIED$

	$\%like_i$	$\%disike_i$	$\%neutral_i$	$\%untried_i$
$\%like_i$	1			
$\%disike_i$	-0.2048	1		
	(0.0056)			
$\%neutral_i$	-0.9008	-0.2406	1	
	(0)	(0.0011)		
$\%untried_i$	-0.0818	-0.1413	0.1439	1
	(0.2722)	(0.0571)	(0.0527)	

Notes: See Table 8 for definitions of these terms. There are 24 observations for each variable.

TABLE 4: VARIETIES AND SALES FOR 24 PRODUCT LINES

	Product line Name (Chinese Name)	# varieties 1st week	# varieties 2nd week	$\%fixed_i$	Sales in full-variety treatment	Sales in reduced-variety treatment	$\%\,sales\ non\ fixed_{\ i}$	$\%like$ (%)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	Glico biscuits(格力高)	19	12	0.63	32	50	0.09	32.6
2	Huashengtang fruit vinegars(华生堂果醋)	3	5	0.60	18	14	0.11	53.0
3	Nestle milk powders(雀巢奶粉)	3	6	0.50	13	9	0.31	56.1
4	Liby clothes detergents(立白洗衣粉)	12	21	0.52	27	17	0.33	56.8
5	Master Kong instant noodles(康师傅方便面)	30	20	0.67	154	147	0.12	58.7
6	Oreo biscuits(奥利奥)	12	5	0.50	23	15	0.22	58.7
7	Lee Kum Kee soy sauces(李锦记酱油)	19	12	0.63	40	23	0.40	59.2
8	You Lemei milk teas(优乐美奶茶)	9	5	0.56	14	8	0.21	60.1
9	Häagen-Dazs ice creams(哈根达斯)	24	15	0.63	1	5	1.00	61.4
10	Store made sushi	26	13	0.50	40	37	NA	61.4
11	Alpine candies(阿尔卑斯糖)	7	4	0.57	11	14	0.00	61.9
12	Tongyi instant noodles(统一方便面)	11	6	0.55	28	53	0.29	64.8
13	Liby dish detergents(立白洗洁精)	9	15	0.60	44	30	0.23	65.3
14	Dove chocolates(巧克力)	7	11	0.64	33	17	0.18	68.3
15	Lipton teas(立顿)	6	13	0.46	5	1	0.80	68.3
16	Knorr soup bases(家乐浓汤宝)	10	6	0.60	12	8	0.08	69.0
17	Huaweiheng dried fruits (华味亨蜜饯)[20]	4	8	0.50	24	6	0.46	69.9
18	Want-want QQ gummies candies(旺仔 QQ 糖)	9	5	0.56	31	25	0.45	70.4
19	Nissin cup noodles(合味道方便面)	11	19	0.58	67	53	0.34	70.8
20	Vinda small pack tissues(维达手帕纸)	4	7	0.57	156	38	0.00	73.2
21	Comfort fabric softeners(金纺衣物护理剂)	22	13	0.59	15	14	0.20	74.4
22	Laoganma sauces(老干妈)	5	9	0.56	67	39	0.43	74.6
23	Haitian sauces(海天酱)	11	6	0.55	15	8	0.47	76.6
24	Heinz rice powders(亨氏米粉)	18	10	0.56	12	17	0.00	83.6

Notes: $\%fixed_i = \frac{\#fixed\ varieties_i}{l\ \#varieties\ full\ varieties_i}$ for each product line i. $\%sales\ non\ fixed_i = \frac{sales\ non\ fixed\ varieties_i}{sales\ of\ full\ varieties_i}$. See Table 8 for the definitions of other terms. NA means not applicable. In the case of Sushi, I was not able to obtain sales data per variety from the store. With Alpine candies, Heinz rice powders Vinda tissues, I had no sales among the non-fixed varieties. With Häagen-Dazs, I had sales only among the non-fixed varieties.

[20] One variety of Huaweiheng dried fruit was present in the reduced-variety treatment (which had a total of 5 varieties including this variety) that was not sold in the full-variety treatment. However, the loss of data for this single variety with zero sales had a negligible influence on the regression results regardless of their inclusion or exclusion.

TABLE 5: CORRELATION BETWEEN %LIKE, %SALES OF NON-FIXED VARIETIES, AND SHARE OF FIXED VARIETIES (P-VALUES IN PARENTHESES)

	$\%like_i$	$\%sales\ non\ fixed_i$	$\%fixed_i$
$\%like_i$	1		
$\%sales\ non\ fixed_i$	0.112 (0.6108)	1	
$\%fixed_i$	-0.2172 (0.3194)	-0.1968 (0.3681)	1

Notes: There are 24 observations for each variable. The table reveals $\%like_i$, $\%sales\ non\ fixed_i$, and $\%fixed_i$ are not significantly correlated. $\%sales\ non\ fixed_i = \frac{\#Sales\ non\ fixed\ varieties_i}{\#Sales\ full\ variety\ teatment_i}$. $\%fixed_i = \frac{\#varieties\ in\ reduced_i}{\#varieties\ in\ full_i}$. See Table 8 for the definition of other terms.

TABLE 6: REGRESSION OF PROBABILITY OF STOP

Dependent variable	$\frac{stops_i}{passbys_i}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment (Reduced-variety = 1)	-0.00249	-0.00297	-0.00782	0.0461	0.0430	0.0298
	(0.0189)	(0.0189)	(0.0188)	(0.0898)	(0.0896)	(0.0888)
$Treatment \cdot \%like_i$				-0.000753	-0.000712	-0.000581
				(0.00136)	(0.00136)	(0.00134)
$price_{i,treatment}$		-0.00106	-0.00101		-0.00104	-0.00100
		(0.000916)	(0.000903)		(0.000924)	(0.000911)
Week (Second = 1)			0.0292			0.0287
			(0.0188)			(0.0190)
Constant	0.0990***	0.113***	0.100***	0.0990***	0.113***	0.100***
	(0.0134)	(0.0181)	(0.0196)	(0.0135)	(0.0182)	(0.0198)
R^2	0.000	0.029	0.080	0.007	0.035	0.084
Observations	48	48	48	48	48	48

Notes: $stops_i$ counts the shoppers who stopped in front of the experimental product line *i* for at least three seconds in the video footage. $passbys_i$ counts the number of shoppers who passed by the shelf but did not stop in the video footage. 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. The dependent variable is $\frac{stops_i}{pass\ bys_i}$ of each product line i in each treatment. Hence, there are 48 observations in total, consisting of 24 product lines and two treatments. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

TABLE 7: REGRESSION OF PROBABILITY OF PURCHASE

Dependent variable	$\frac{buyers_i}{stops_i}$					
	(1)	(2)	(3)	(4)	(5)	(6)
Treatment (Reduced-variety = 1)	0.00803	0.00652	0.00244	0.0908	0.0808	0.0697
	(0.0316)	(0.0302)	(0.0308)	(0.150)	(0.143)	(0.145)
$Treatment \cdot \%like_i$				-0.00127	-0.00114	-0.00103
				(0.00227)	(0.00217)	(0.00218)
$price_{i,treatment}$		-0.00335**	-0.00331**		-0.00335**	-0.00332**
		(0.00147)	(0.00148)		(0.00148)	(0.00148)
Week (Second = 1)			0.0246			-0.00103
			(0.0308)			(0.00218)
Constant	0.187***	0.232***	0.221***	0.187***	0.231***	0.221***
	(0.0223)	(0.0290)	(0.0321)	(0.0224)	(0.0291)	(0.0322)
R^2	0.001	0.105	0.117	0.009	0.112	0.125
Observations	48	48	48	48	48	48

Notes: $stops_i$ counts the shoppers who stopped in front of product line *i* for at least three seconds in the video footage. $buyers_i$ counts the shoppers observed taking something off the shelf of product line *i* in the video footage. *Treatment* is a dummy for the full- and reduced-variety treatments. Price is the average price of each product line in each treatment based on point-of-sale data provided by the store. Week is a dummy for the first and second weeks. The dependent variable $\frac{buyers_i}{stops_i}$ is interpreted as the probability of purchase for each product line i in each treatment. Hence, there are 48 observations in total, consisting of 24 product lines and two treatments. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

TABLE 8: DEFINITIONS OF VARIABLES DESCRIBING EXPERIMENTAL RESULTS

Variable	Definition	Interpretation
i		• Product line index.
D_i	$D_{\mathrm{i}} = \mathrm{I}_{(\text{product line}=i)}$	• Dummy for 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 provided data.
$buyers_i$		• Number of buyers from video data of product line i.
$sales_i$	$sales_i = \sum_j buys_{i,j}$	• Sum $buys_{i,j}$ over all of the varieties j among the fixed varieties in product line i for each treatment.
F		• Full-variety treatment.
R		• Reduced-variety treatment.
$BS_{i,j} = \frac{buys_{i,j}}{stops_i}$		• Probability of purchase for variety j among the fixed-varieties of product line i.
$CAB_{i,j}$	$CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,R} > BS_{i,j,F} \\ 0, if\ BS_{i,j,R} < BS_{i,j,F} \\ \text{dropped}, if\ BS_{i,j,R} = BS_{i,j,F} \end{cases}$	• An indicator of choice-averse behavior among fixed-varieties j of product line i. See Eq. 2 for development.
$Adjusted\ BS_{i,j,F}$	$Adjusted\ BS_{i,j,F} = BS_{i,j,F} \cdot \frac{\#\ varieties\ full\ treatment}{\#\ varieties\ reduced\ treatment}$.	• $BS_{i,j,F}$ adjusted for the lower probability of purchase of any variety in the full-variety treatment compared with the reduced-variety treatment. See Eq. 3 for development.
$Adjusted\ CAB_{i,j}$	$Adjusted\ CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,R} > Adjusted\ BS_{i,j,F} \\ 0, if\ BS_{i,j,R} < Adjusted\ BS_{i,j,F} \\ null, if\ BS_{i,j,R} = Adjusted\ BS_{i,j,F} \end{cases}$.	• $CAB_{i,j}$ adjusted for the lower probability of purchase of any variety in the full-variety treatment compared with the reduced-variety treatment. See Eq. 4 for development.
$\%like_i$	$\frac{\#likes_i}{\#likes_i + \#neutrals_i + \#dislikes_i}$	• See Section 3 Experimental Design for definitions.
$\%dislike_i$	$\frac{\#dislikes}{\#likes_i + \#neutrals_i + \#dislikes_i}$	
$\%neutral_i$	$\frac{\#neutrals_i}{\#likes_i + \#neutrals_i + \#dislikes_i}$	
$\%tried_i$	$\frac{\#likes + \#neutrals + \#dislikes}{\#likes_i + \#neutrals_i + \#dislikes_i + \#untried_i}$	
$\%fixed_i$	$\frac{\#varieties\ in\ reduced_i}{\#varieties\ in\ full_i}$	• Share of varieties that were constant across the reduced- and full-variety treatments.
$\%sales\ non\ fixed_i$	$\frac{\#Sales\ non\ fixed\ varieties_i}{\#Sales\ full\ variety\ teatment_i}$	• Share of sales of the varieties found only in the full-variety treatment.
$\#full_i$		• The number of varieties in the full variety treatment.
$price\ change_{i,j}$	$\frac{price_{i,j,F}}{price_{i,j,R}} - 1$	• Relative price change between the reduced-variety and full-variety treatments for variety j in product line i based on point-of-sale data.

TABLE 9: PROBIT REGRESSIONS OF PROBABILITY OF CAB

Dependent variable	$CAB_{i,j}$											
	(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)						
$\#full_i$	0.0131						0.0341**	0.0260*		0.0120		
	(0.0213)						(0.0152)	(0.0156)		(0.0198)		
$\%fixed_i$	3.751								4.664*	3.605		
	(2.701)								(2.388)	(2.997)		
$\%sales\ non\ fixed_i$	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	152	152	152	152	152	152	152	152	152	152	152	152
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: $CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,R} > BS_{i,j,F} \\ 0, if\ BS_{i,j,R} < BS_{i,j,F} \\ \text{dropped}, if\ BS_{i,j,R} = BS_{i,j,F} \end{cases}$. $\#full_i$= number of varieties in the full-variety treatment. $\%fixed_i = \frac{\#varieties\ in\ reduced_i}{\#varieties\ in\ full_i}$. $\%sales\ non\ fixed_i = \frac{\#Sales\ non\ fixed\ varieties_i}{\#Sales\ full\ variety\ teatment_i}$. $price\ change_{i,j} = \frac{price_{i,j,F}}{price_{i,j,R}} - 1$. See Table 8 for the definition of other terms. This table reveals 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 $\%fixed_i$ and $\%sales\ non\ fixed_i$. Robust standard errors clustered by product lines in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

TABLE 10: PROBIT REGRESSIONS OF PROBABILITY OF ADJUSTED CAB

Dependent variable	$Adjusted\ CAB_{i,j}$											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\%like_i$	-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)						
$\#full_i$	0.0152						0.0281	0.0206		0.0137		
	(0.0199)						(0.0201)	(0.0202)		(0.0213)		
$\%fixed_i$	2.645								3.006	1.799		
	(2.912)								(2.341)	(2.343)		
$\%sales\ non\ fixed_i$	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***	22.01***	21.78***	22.29***	20.85***
	(2.304)	(0.675)	(0.868)	(0.800)	(1.050)	(0.702)	(0.948)	(0.876)	(0.727)	(0.898)	(2.287)	(1.818)
Control for product line fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	-0.690	1.205*	-0.285	1.300*	-0.0533	1.330**	-0.812**	0.682	-0.847	-0.373	-0.329	1.205*
	(1.854)	(0.678)	(0.276)	(0.745)	(0.373)	(0.599)	(0.330)	(0.828)	(1.711)	(1.671)	(0.287)	(0.700)
Observations	152	152	152	152	152	152	152	152	152	152	152	152
Pseudo R-squared	0.248	0.230	0.201	0.230	0.208	0.233	0.215	0.237	0.236	0.238	0.201	0.232

Notes: $\#full_i$=number of varieties in the full-variety treatment. $\%fixed_i = \frac{\#varieties\ in\ reduced_i}{\#varieties\ in\ full_i}$. $\%sales\ non\ fixed_i = \frac{\#Sales\ non\ fixed\ varieties_i}{\#Sales\ full\ variety\ teatment_i}$. $price\ change_{i,j} = \frac{price_{i,j,F}}{price_{i,j,R}} - 1$. See Table 8 for the definition of other terms. As robustness check, I "adjusted"$CAB_{i,j}$ in all models because the probability of being purchased for a specific variety in a larger choice set is diluted by more varieties. $Adjusted\ CAB_{i,j} = \begin{cases} 1, if\ BS_{i,j,R} > Adjusted\ BS_{i,j,F} \\ 0, if\ BS_{i,j,R} < Adjusted\ BS_{i,j,F} \\ null, if\ BS_{i,j,R} = Adjusted\ BS_{i,j,F} \end{cases}$. $Adjusted\ BS_{i,j,F} = BS_{i,j,F} \cdot \frac{\#\ varieties\ full\ treatment}{\#\ varieties\ reduced\ treatment}$. Again, as in Table 9, this table reveals 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 $\%fixed_i$ and $\%sales\ non\ fixed_i$. Robust standard errors clustered by product lines in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

.

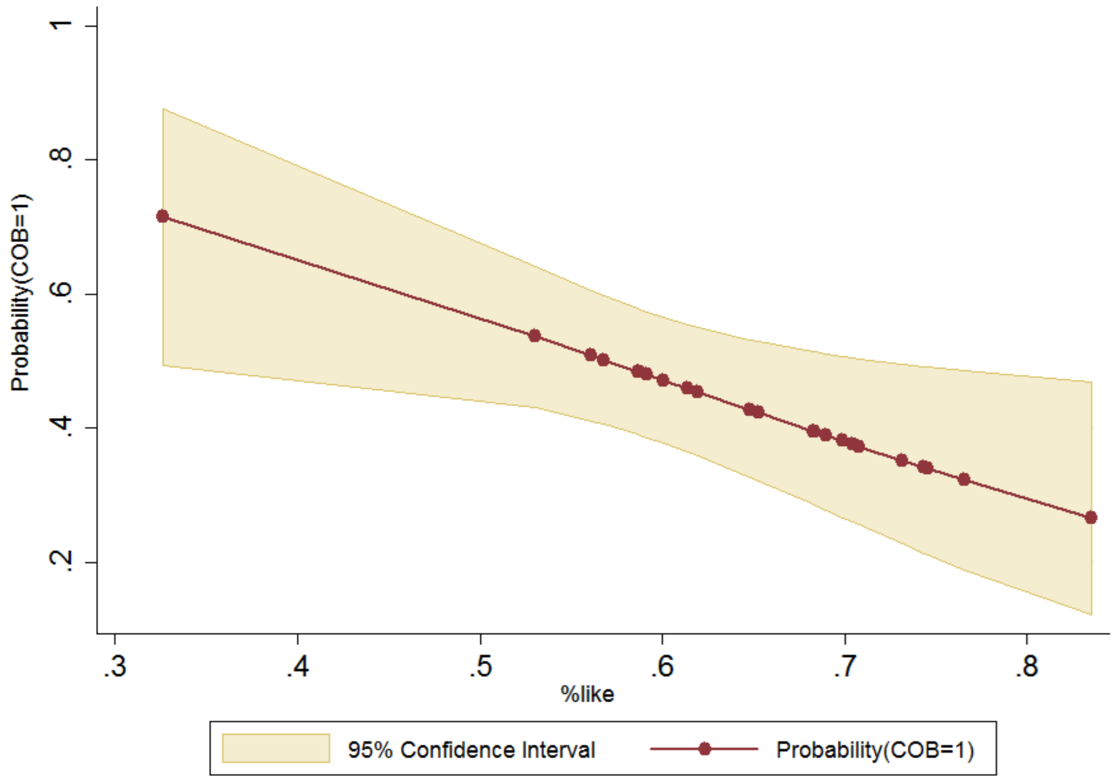


FIGURE 2: PREDICTED PROBABILITY OF CAB AS A FUNCTION OF $\%like_i$

Notes: See Table 8 for the definition of terms. This figure reveals how the probability of CAB decreases on $\%like_i$ for product line *i*. The dots are product lines. Their $\%like_i$ correspond to those listed in Table 4. The horizontal axis represents $\%like_i$. The vertical axis measures $\Pr\left(CAB_{i,j} = 1 \middle| \%like_i, other\ control\ variables\right)$ holding other control variables at their means in model (1) of Table 9. The yellow/shaded area surrounding the trend line is the 95% confidence interval in model (1) of Table 9.

TABLE 11: OLS REGRESSION OF SALES PER BUYER GIVEN HIGH OR LOW $\%LIKE$

Dependent variable	$\frac{sales_i}{buyers_i}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$full$ (full-variety = 1)	0.728*	0.742*	0.716*	-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)
$price_{i,treatment}$	2.248***	2.657***	2.752***		-0.0330*	-0.0332*		-0.244***	-0.272***
	(0.263)	(0.364)	(0.416)		(0.0169)	(0.0175)		(0.0471)	(0.0594)
$week$ (second=1)			-0.158			-0.150			-0.303
			(0.383)			(0.386)			(0.316)
Constant	2.248***	2.657***	2.752***	2.576***	3.003***	3.092***	2.248***	5.387***	5.931***
	(0.263)	(0.364)	(0.416)	(0.707)	(0.773)	(0.778)	(0.175)	(0.561)	(0.789)
Control for product line fixed effects	No	No	No	No	No	No	Yes	Yes	Yes
Observations	48	48	48	48	48	48	48	48	48
R-squared	0.066	0.119	0.122	0.105	0.162	0.165	0.234	0.346	0.370

Notes: $full$ is a dummy for the full- and reduced-variety treatments. Price is the average price of each product line in each treatment provided by the store. Week is a dummy for the first and second weeks. I control the average price of each product line in each variety treatment except for models (4) and (7). D_like_i is omitted in models (7)-(9) because of collinearity. $sales_i$ is the sum of $buys_{i,j}$ for all of the varieties j in product line i in the reduced-variety treatment using store provided data. $buyers_i$ is the number of buyers of each product line in each treatment from 5:30 pm-9:30 pm in the video footage. The dependent variable is $\frac{sales_i}{buyers_i}$ of each product line i in each treatment. Thus, there are 48 observations in total, consisting of 24 product lines and two treatments. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Theoretical Appendix

A-1. Hoteling Linear Taste Setup

A population of consumers considers 'sampling' m_i varieties from n_i unfamiliar varieties in product line i in the first period of a two-period model to repeat the satisfactory purchases in the second period.[21] These varieties occupy positions on a line $z_{ij} \in [0,2]$, $j \leq n_i$, which spans the preference space for the product category to which the product line belongs. The consumer's preference for these varieties is defined by the pair (w_i, v_i). $w_i \in [0,2]$ defines the location of the consumer. $v_i \in (0,\bar{v}]$ is the utility level of the consumer at the same location as the variety.

The distance between the consumer's location and that of the variety, $|w_i - z_{ij}|$, measures the degree of mismatch. The cost of mismatch multiplies the degree of mismatch by the 'marginal mismatch cost' is t_i. t_i is high for niche tastes, where each consumer enjoys only a few varieties, and low for popular tastes, where each consumer enjoys many varieties. Hence, a consumer w_i's utility for variety z_{ij} is $u_{wij} = v_i - t_i \cdot |w_i - z_{ij}|$. I refer to $-t_i \cdot |w_i - z_{ij}|$ as the cost of mismatch. For simplicity, prices are identically p_i for the varieties in product line i. Thus, each consumer's *ex-post* surplus for varieties from product line i is $u_{wij} - p_i$.

With regards to information, there are two types of consumers: the informed (familiar), who know their own and the variety's location and the uninformed (unfamiliar) who knows neither. Both informed and uninformed types of consumers are uniformly distributed over $[0,2]$. The familiar consumer chooses deterministically according to the known match quality between their tastes and the variety in each period. The unfamiliar consumer chooses randomly without knowing their own location or the location of the variety, i.e., according to their uninformed *ex-ante* utility, in the first period and deterministically according to their informed *ex-post* utility in the second period, u_i. I now specify the uninformed consumer's *ex-ante* utility in the model.

Due to the unfamiliar consumers' lack of information as to the degree of mismatch between their tastes and any given variety, each unfamiliar consumer is, in effect, identically repeating an identical gamble when they purchase multiple varieties in the first period. The varieties are, prior to purchase, interchangeable for this type of consumer: $E(u_{wij}) = E(u_{w'ij'})$. Thus, there is no

[21] The choice of the consumers to purchase more than one variety at a time can be justified by a travel cost to the store, which I do not model. I model the choice of the consumer at the store, at which point, the travel cost is already sunk, and leave the modeling of how the travel cost affects the optimal sample size for future work.

loss of generality in dropping the location index as well as the index of the variety from the consumer's utility in the subsequent discussion about the properties of the expectations and the distribution of utility outcomes. [22] Hence, $E(u_i)$ now represents the *ex-ante* expected utility of any single variety for any consumer. Moreover, since I consider only the *ex-post* utility for a fixed consumer for a fixed variety, I also dispense with the consumer and location indices of the utility function, u_i, in the following.

This specification of the unfamiliar consumer is now similar to Sun's (2012) model in which unfamiliar consumers use familiar consumers' ratings to determine their own purchasing decisions. As Sun pointed out, the distribution of the utility of a uniform mass of consumers is also uniformly distributed.[23] Let $\underline{u_i}$ represent the lowest and $\bar{u}_i$ the highest achievable *ex-post* utility. I define the probability distribution function $F(u_\mathrm{i}) = \int_{\underline{u_i}}^{\bar{u}_i} f(u_i)du_i = \frac{u_i - \bar{u}_i}{\bar{u}_i - \underline{u_i}}$ and the probability density function $f(u_i) = \frac{1}{\bar{u}_i - \underline{u_i}}$ over this support.[24] Since $\bar{u}_i = v_i$, $F(u_\mathrm{i}) = \frac{u_i - \bar{u}_i}{v_i - \underline{u_i}}$ and $(u_i) = \frac{1}{v_i - \underline{u_i}}$.

To avoid unnecessary complexity in the analysis of the firm's decisions (which is not the focus of this paper), I assume a monopolist firm that chooses the number of varieties to produce in product line i based only on the distribution of the informed consumers.[25] Therefore, I can use the standard result for the Hoteling linear taste model that the monopoly firm chooses the location of varieties at equal intervals to minimize the mismatch cost of consumers and maximize its own profit margin.[26] In this case, the number of varieties is inversely related to the marginal mismatch cost. I assume that consumers know this based on the intuition that more varieties indicate greater heterogeneity in tastes, and can, therefore, infer that a larger number of varieties n_i implies higher marginal mismatch cost; $n'_i > n_i$ implies $t'_i > t_i$.[27] Thus,

---

[22] These assumptions are standard to the search literature (see, for example, Liu and Dukes (2016)) and greatly simplify the analysis by allowing the modeling of consumer sampling as sampling of one variety with replacement rather than of many varieties without replacement.

[23] Suppressing the product line index i, let $x = |w - z|$. Given that $u(x) = v - tx$ and x is distributed $F(x) = P\{X < x\}$, then $x = \frac{v-u}{t}$ and $P\left\{X < \frac{v-u}{t}\right\} = P\{v - Xt > u\} = 1 - P\{v - Xt < u\} = 1 - G(u)$. Therefore, if $F(x)$ is uniform, then so is $G(u)$.

[24] In the interest of parsimony and to minimize notation, I follow Sun (2012) and abuse notation somewhat by not distinguishing between the value of a random variable u_i and the random variable which takes on this value.

[25] However, the firm would still choose the same locations even if based on the distribution of uninformed consumers because I assume that they are also uniformly distributed.

[26] See, for example, Section 10.2.1 of Waldman and Jensen (2019).

[27] This result is directly assumed in Kamenica (2008); more varieties implies lower average popularity.

Assumption I. Increasing the number of available varieties, n_i, in product line i increases the consumer's perceived marginal mismatch cost, t_i.

Hence, an exogenous increase in n_i, as in my experiment, decreases the lower bound of the consumer's beliefs about utility outcomes $[\underline{u}_i, v_i]$. I discuss the effects of this decrease in the lower bound on the expected value of sampling and the consumer's decision to sample below when I explain the comparative statics of the model.

The choice of the unfamiliar consumer is as follows. In the first period, she observes the number of varieties n_i and infers the cost of mismatch. Within this simplified version of KV-B's framework, she can choose to not make a purchase and receive 0, or to purchase m_i out of n_i varieties randomly and receive an expected surplus of $\delta_i(m_i)E(u_i) - m_i \cdot p_i$ in the first period.[28] $E(u_i)$ is the expected value of a single variety from product line i purchased at random at the price of p_i. $\delta_i(m_i)$ is a positive strictly concave function where $\delta_i(1)$=1. The strict concavity of this utility for variety models the substitutability of varieties; The marginal utility of each successful purchase of a distinct variety decreases as it would for multiple units of the same variety.

In the second period, the consumer repeats the purchases of those varieties sampled that realized a positive surplus: $u_i - p_i > 0$. Crucially in this analysis, the consumer's choice of the number of varieties to purchase in the first period, m_i, anticipates the potential surplus for repeated purchases of varieties in the second period. I formalize the consumer's expected utility of repeated purchases in the second period next.

A-2. Determining the Optimal Sample Size

Let $B_i(m_i)$ be the expected benefit in the second period based on the information gained from random purchases in the first period from a sample size of m_i. Let N_i be the number of successful first-period purchases that the consumer anticipates repeating in the second period. The consumer's two-period expected 'shopping' utility is, thus,

$$V_i(m_i) = max\{0, \delta_i(m_i)E(u_i) - m_i \cdot p_i + N_i \cdot B_i(m_i)\}. \qquad \text{A-Eq. 1}$$

[28] I have normalized the value of the alternative to sampling to zero. However, multiple non-zero values can be introduced to allow consumers to differ in terms of their outside option to sampling new brands, e.g., because older consumers have higher levels of 'consumption capital' and have less to gain in terms of lifetime surplus from sampling new products than younger consumers (Bronnenberg, Dubé, & Gentzkow, 2012).

To find the optimal sample size, m_i^*, I note that if $\delta_i(m_i^*)E(u_i) - m_i^* \cdot p_i > 0$, then the information gained for second period purchases makes no difference to first period purchases. The consumer maximizes her surplus by setting the number of varieties to sample equal to the available number of varieties ($m_i^* = n$), and buys all available varieties. Accordingly, I analyze the case where the consumer suffers an expected loss in surplus from uniformed random purchases of m_i^* varieties in the first period: $\delta_i(m_i^*)E(u_i) - m_i^* \cdot p_i < 0$. This loss needs to be offset by the consumer's expected benefit from informed second-period purchases for the consumer to want to purchase at all in the first period. To facilitate analysis, I define

Definition I. The net expected value of sampling m_i varieties from product line i is

$$\delta_i(m_i)E(u_i) - m_i \cdot p_i + N_i \cdot B_i(m_i). \qquad \text{A-Eq. 2}$$

To derive the expected benefit $B_i(m_i)$ from sampling m_i varieties, I first enumerate all possible outcomes in the first period with their payoffs and probabilities. The consumer repeats a 'successful' random purchase from the first period in the second period if the realized utility leads to a positive surplus, given price: $u_i > p_i$. The probability of a successful purchase is the complement of an unsuccessful purchase, where $u_i < p_i$, which occurs with probability $F(p_i)$. Accordingly, the probability of a successful purchase is $1 - F(p_i)$. I use $1 - F(p_i)$ to define 'popularity' from here onwards. More formally,

Definition II. Product line i is more popular than product line i' if

$$1 - F_i(p_i) \geq 1 - F_{i'}(p_{i'}), \text{ given } p_i \text{ is set equal to } p_{i'}. \qquad \text{A-Eq. 3}$$

In other words, the product line is more popular if a greater share of consumers derive a positive surplus from it. I suppress the subscript for the distribution of consumer utilities, F, where I am not comparing the popularity of different product lines.

The expected value of the second-period purchase of a given variety in the first period, conditional on a positive surplus ($u_i > p_i$) is

$$\int_{p_i}^{v_i} u_i \frac{f(u_i)}{1 - F(p_i)} du_i \qquad \text{A-Eq. 4}$$

This expectation is conditional on $u_i > p_i$ because only then is the first-period purchase repeated in the second period. In this case, the consumer's payoff is her utility above the price she pays p_i. If the consumer has an unsuccessful purchase, she does not repeat it. In that case, she receives a negative surplus in the first period and a zero surplus in the second period from the purchase of this variety.

This binary outcome of successful and unsuccessful purchases can be modeled with a Bernoulli random variable, the sum of which is binomially distributed. Therefore, the consumer's expected benefit in the second period from randomly purchasing m_i varieties in the first period is

$$N_i \cdot B_i(m_i) = \sum_{j=0}^{m_i} \delta_i(j) \int_{p_i}^{v_i} u_i \frac{f(u_i)}{1 - F(p_i)} du_i \binom{m_i}{j} \left(\left(1 - F(p_i)\right)^j \right) F(p_i)^{m_i - j} \qquad \text{A-Eq. 5}$$

The expression

$$\sum_{j=0}^{m_i} \binom{m_i}{j} \left(\left(1 - F(p_i)\right)^j \right) F(p_i)^{m_i - j} \qquad \text{A-Eq. 6}$$

in A-Eq. 5 is merely the cumulative distribution function of the binomial distribution, which I use to make an exhaustive probability weighted count of each combination of successful and unsuccessful purchases. In particular, j counts the number of successful purchases to be repeated N_i times. Thus, the consumer's utility for these purchases is the surplus from each of these purchases shown in A-Eq. 4 scaled by $\delta_i(j)$, which is strictly concave in the number of varieties, j, that yields a positive surplus.

I next show that there exists an m_i that maximizes the net expected benefit of sampling. The optimality condition for the choice of m_i can be succinctly stated in terms of the consumer's two-period expected shopping utility as

$$V_i(m_i^*) - V_i(m_i^* - 1) > 0 > V_i(m_i^* + 1) - V_i(m_i^*), \qquad \text{A-Eq. 7}$$

which is a discrete analog of the standard first-order condition for the maximization of $V_i(m_i)$ if m_i were a continuous variable. I divide the discussion of the optimization into two parts. First, I show that A-Eq. 7 is satisfied for the first-period random purchase, and then, also for the second period expected benefit from sampling.

For any fixed p_i, by the concavity $\delta_i(m_i)$, $\delta_i(m_i + 1)E(u_i) > \delta_i(m_i)E(u_i)$ and the difference decreases on m_i. Moreover, the sum of -$(m_i + 1) \cdot p_i + (m_i) \cdot p_i$ is -p_i. Therefore, the constant incremental decrease in surplus from the price of the next variety purchased, -p_i, eventually dominates the decreasing incremental increase in surplus from the expected value, $\left(\delta_i(m_i + 1) - \delta_i(m_i)\right)E(u_i)$, as m_i increases. Thus, there exists some finite m_i such that the terms for the surplus from random purchases in the first period in $V_i(m_i)$ meets the condition in A-Eq. 7:

$$\delta_i(m_i)E(u_i) - m_i \cdot p_i - (\delta_i(m_i - 1)E(u_i) - (m_i - 1) \cdot p_i) > 0$$
$$> \delta_i(m_i + 1)E(u_i) - (m_i + 1) \cdot p_i - (\delta_i(m_i)E(u_i) - m_i \cdot p_i) \qquad \text{A-Eq. 8}$$

In regards to the expected benefit term $B_i(m_i)$ in the second period shown in A-Eq. 5, note that the term

$$\int_{p_i}^{v_i} u_i \frac{f(u_i)}{1 - F(p_i)} du_i \qquad \text{A-Eq. 9}$$

is constant with respect to m_i, and hence, does not affect the concavity of $B_i(m_i)$. I, therefore, focus on showing that the rest of the sum of $B_i(m_i)$, namely $\sum_{j=0}^{m_i} \delta_i(j) \binom{m_i}{j} \left(\left(1 - F(p_i)\right)^j \right) F(p_i)^{m_i - j}$ is concave. I rely upon an intuitive argument, the key part of which is that increases in m_i shift the probability mass away from the first terms of $B_i(m_i)$, namely $j = 0$ or 1, for which the concavity of $\delta_i(j)$ makes no difference, i.e., $\delta_i(0) = 0$ and $\delta_i(1) = 1$, to later terms of $B_i(m_i)$, for which the concavity of $\delta_i(j)$ makes increasing differences, i.e., $\delta_i(j) < j$ for each $j = 2,3,4, \dots, m_i$.

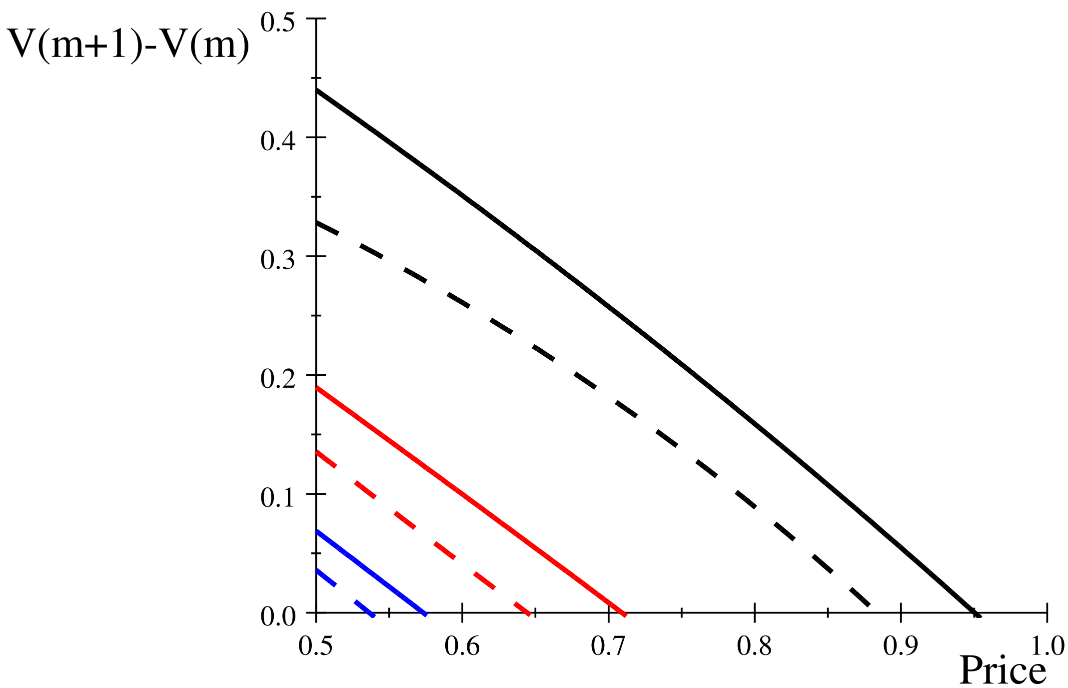
The movement of mass from increases in m_i is evident in the measure of the complement of the event measured by the first term: $1 - \left(F(p_i)\right)^{m_i}$, which is clearly increasing in m_i. Hence, as m_i increases, probability moves to the later terms in the sum, which are weighted by the concave function $\delta_i(j) < j$ for each $j = 2,3,4, \dots, m_i$. This shifted mass among the j th, $2 \leq j \leq m_i$, terms is itself uniformly divided among each of the m_i terms in the sum for $j = 1, 2,3, \dots, m_i$. This reweighting of mass towards higher values of $\delta_i(j)$, combined with the increasing number of terms as m_i increases, implies that $B_i(m_i)$ increases on m_i.

To show concavity, I also need to show that the rate of increase of $B_i(m_i)$ is decreasing. Note that because of the concavity of $\delta_i(m_i)$,

$$\delta_i(m_i + 1) - \delta_i(m_i) < \delta_i(m_i) - \delta_i(m_i - 1) \qquad \text{A-Eq. 10}$$

for each value of m_i, the increase of $B_i(m_i)$ from the increase from $\delta_i(m_i)$ to $\delta_i(m_i + 1)$ is smaller than the increase from $\delta_i(m_i - 1)$ to $\delta_i(m_i)$. As a consequence of the decreasing rate of increase of $\delta_i(m_i)$ from an increase in m_i, the increase in $B_i(m_i)$ from the increase in m_i is also increasing at a decreasing rate, making $B_i(m_i)$ also strictly concave in m_i. However, the loss from purchasing one more variety in the first period is always p_i. Consequently, as m_i increases, the constant incremental cost p_i from the cost of purchasing randomly in the first period will also dominate the decreasing incremental benefit of informed purchases in the second period for some finite m_i. Thus, the condition in A-Eq. 7 is satisfied.

The decreasing marginal rate of increase from the concavity of $V_i(m_i)$ in m_i necessary to meet the discrete first-order condition in A-Eq. 7 is illustrated by the decreasing rate of downward shift of the graphs of $V_i(m_i+1) - V_i(m_i)$ as the value of m_i increases from 1 to 3 in A-Figure 3. The dashed lines graph these marginal changes for the case of low-dispersion case of $u_i \in \left[\frac{1}{2}, \frac{3}{2}\right]$, while the solid lines graph the changes for the high-dispersion case of $u_i \in [0,2]$. (Graphs of higher values of m_i and other levels of dispersion showing similar results are available on request.)



A-FIGURE 3: $V_i(m_i+1, N_i) - V_i(m_i, N_i)$ FOR LOW-AND HIGH-DISPERSION OF UTILITIES

Notes: This graph is from a simulation of $V_i(m_i+1, N_i) - V_i(m_i, N_i)$ for $\delta_i(j) = j^{\frac{1}{2}}, N_i = 1$, and $m_i = 1, 2, 3$, where where the dashed line is for low utility dispersion ($u_i \in \left[\frac{1}{2}, \frac{3}{2}\right]$) and the solid line is for high utility dispersion ($u_i \in [0,2]$) over the price range $p_i \in \left[\frac{1}{2}, 1\right]$. Consistent with a diminishing marginal value of sampling and the concavity of $V_i(m_i, N_i)$, the left panel shows that the height of the lines decreases with m_i for all values of $p_i \in \left[\frac{1}{2}, 1\right]$. Within each set of either solid or dashed lines, $m_i = 1$ is represented by the top black lines, $m_i = 2$ by the middle red lines, and $m_i = 3$ is represented by the lowest blue lines. Consistent with the marginal value of sampling increasing with the dispersion of utility outcome, the height of each solid line is greater than the corresponding dashed line for the same value of m_i.

Now, I consider the effect of an increase in the marginal mismatch cost t_i. An increase in t_i affects the expected surplus in both the first and the second periods. In the first period, the increase in t_i decreases the lower bound $\underline{u}_i$ of the support of the probability distribution $F_i(p_i)$ on $[\underline{u}_i, v_i]$, which increases $F_i(p_i)$ and decreases the probability of success $1 - F_i(p_i)$ at a given price p_i. This decrease in $\underline{u}_i$ decreases $E(u_i)$ by moving mass towards the lower end of the support of the pdf. The decrease is evident for the uniform distribution: $E(u_i) = \int_{\underline{u}_i}^{v_i} u_i \frac{1}{v_i - \underline{u}_i} du_i$. Decreasing $\underline{u}_i$ moves mass lower for any fixed value of u_i in the interval $[\underline{u}_i, v_i]$ for the uniform pdf $f_i(p_i) = \frac{1}{v_i - \underline{u}_i}$.

In the second period, the decrease in $\underline{u}_i$ decreases $B_i(m_i)$ by moving mass from the positive surplus outcomes measured by the expected surplus conditional on positive surplus, $\int_{p_i}^{v_i} u_i \frac{f(u_i)}{1-F(p_i)} du_i$, to the zero surplus outcomes. The $\int_{p_i}^{v_i} u_i \frac{f(u_i)}{1-F(p_i)} du_i$ term is itself not affected, because the conditioning drops the support of $F_i(p_i)$ below p_i. The lack of effect on this expected surplus is evident in case of the uniform distribution where the expected surplus of a successful purchases is

$$\int_{p_i}^{v_i} u_i \frac{f(u_i)}{1-F(p_i)} du_i = \int_{p_i}^{v_i} u_i \frac{\frac{1}{v_i - \underline{u}_i}}{1 - \frac{p_i - \underline{u}_i}{v_i - \underline{u}_i}} du_i = \frac{1}{\bar{u}_i - p_i} \int_{p_i}^{v_i} u_i du_i = \frac{{v_i}^2 - {p_i}^2}{v_i - p_i} = v_i + p_i \qquad \text{A-Eq. 11}$$

which is not a function of the lower bound of the support, $\underline{u}_i$. Thus, for a fixed m_i^*, an increase in t_i decreases the expected utility of random purchases, $E(u_i)$, in the first period, and the expected benefit of sampling, $B_i(m_i)$, in the second period. Hence, the first implication is

Implication I. Increasing the cost of mismatch t_i decreases the lower bound of the utility outcomes $\underline{u}_i$, and therefore, decreases the net expected benefit from sampling $\delta_i(m_i)E(u_i) - m_i \cdot p_i + N_i \cdot B_i(m_i)$.

On the other hand, this decrease in $\underline{u}_i$ also increases the spread of the mass of the probability distribution over a larger support, which increases the dispersion of utility outcomes. This larger dispersion of utility outcomes increases the marginal expected utility of sampling, $V_i(m_i)$, within the expected shopping utility at a given m_i, and therefore, increases the optimal number of samples, m_i^*. A-Figure 3 illustrates how the marginal difference in the shopping utility is larger for the high dispersion case for each value of $m_i = 1, 2, 3$. Thus, I draw the next implication,

Implication II. Decreasing the lower bound of utility outcomes increases the dispersion of utility outcomes, which increases the marginal expected utility of sampling in the second period of the shopping utility $V_i(m_i)$.

Intuitively, when consumers observe more varieties, they infer greater heterogeneity in tastes and want to sample more, given that they want to sample at all.

In the experiment, I manipulated the number of available varieties the unfamiliar consumer observes on the shelf. By Assumption I, an increase in the number of available varieties, n_i, increases the consumer's perception of the marginal mismatch cost, t_i. By Implication I, increases in t_i decreases the lower bound of support of the $F(u_i)$ distribution, $\underline{u}_i$, along with the net expected benefit from sampling in A-Eq. 2 for a fixed m_i^*. By Implication II, such a decrease in $\underline{u}_i$ also increases $B_i(m_i^*)$ by increasing m_i^*. However, whether the increase in m_i^* is realized in terms of increased sales depends upon whether A-Eq. 2 is positive for a given increase in t_i.

To show how A-Eq. 2 is more likely to be positive for more popular product lines, I next show that the expression in A-Eq. 2 increases on the popularity of the product line as measured by $1 - F(p_i)$.

First, if $1 - F(p_i)$ increases, then $F(p_i)$ decreases. For the uniform distribution, $F(p_i) = \frac{p_i - \underline{u}_i}{v_i - \underline{u}_i}$ and $f(p_i) = \frac{1}{v_i - \underline{u}_i}$. Hence, the effect of the decrease in $F(p_i)$ can be determined by the derivative of $F(p_i)$ with respect to $\underline{u}_i$ and v_i, which are both negative:

$$\frac{d}{d\underline{u}_i}\left(\frac{p_i - \underline{u}_i}{v_i - \underline{u}_i}\right) = \frac{p_i - \bar{u}_i}{(v_i - \underline{u}_i)^2} < 0 \qquad \text{A-Eq. 12}$$

$$\frac{d}{dv_i}\left(\frac{p_i - \underline{u}_i}{v_i - \underline{u}_i}\right) = -\frac{p_i - \underline{u}_i}{(v_i - \underline{u}_i)^2} < 0 \qquad \text{A-Eq. 13}$$

Consequently, decreases in $F(p_i)$ increases $\underline{u}_i$ and v_i, which must increase $E(u_i)$, and also, $B_i(m_i^*)$ through the expected value of success, the $\int_{p_i}^{v_i} u_i \frac{f(u_i)}{1-F(p_i)} du_i$ term. Moreover, increases in $1 - F(p_i)$ increase the probability weight placed on $\int_{p_i}^{v_i} u_i \frac{f(u_i)}{1-F(p_i)} du_i$, which has a positive value in $B_i(m_i^*)$, and less on the value of failure, which has a value of zero. (A similar argument applies for non-uniform distributions because increases in the lower or upper bound of the support of a CDF must increase the mass above the argument of the CDF for a fixed value of the argument.) Hence, the more popular the product line is, the less likely that an increase in the number of varieties, n_i, would decrease the sum in A-Eq. 2 so much that the consumer chooses not to sample at all. In the case that the consumer still samples, by Implication II, the increase in n_i would result in an increase in the optimal sample size, m_i^*. This analysis leads to the final implication

Implication III. Increasing the number of available varieties, n_i, increases the observed purchases, m_i^*, of product lines with high popularity and decreases the observed purchases of product lines with low levels of popularity.

A-3. Applying the Theoretical Framework to Data Outcomes

I proceed now to discuss how these theoretical predictions are operationalized in the experiment. Regarding the independent variable, my measure of popularity is the share of consumers who reported 'liking' a variety divided by those who 'like', 'dislike', or are 'neutral' to the variety in a survey as they were exiting the store. I use like to capture a discrete level of satisfaction at purchasing. If a shopper reports liking a product, I infer that they received a positive surplus from purchasing, given the price.[29] From the 'neutral' response, I infer that the shopper received zero surplus, and from 'dislike', a negative surplus.

Hence, within my framework, the share of likes among the population of consumers or 'percentage of likes' models popularity defined as $1 - F_i(p_i)$, which is probability that a random consumer receives a positive surplus from purchasing a random variety within product line i at price p_i. I assume that, on average, the beliefs of unfamiliar consumers about popularity correspond to the actual levels of popularity. I use a ternary measure (of likes, neutrals, and dislikes) to derive a distribution of popularity rather than one with more values that can potentially capture the intensity of preferences because the focus of this study is on the interaction of the popularity of a product line with the number of varieties rather the underlying utility parameter, v_i, which determines the popularity of the product line, with the number of varieties.[30] This measure of popularity is otherwise conceptually straightforward and is discussed in Section 3 Experiment Design along with other measurement issues.

Developing the dependent variable is conceptually more challenging. To begin, I discretize the continuum of consumers which had been defined by a real number $x \in [0,1]$ for convenient discussion within the Hoteling linear taste context with a natural number index s. I now refer to these discretized consumers as 'shoppers'. Each shopper s who stops in front of the product line i

[29] In effect, I am eliciting consumer ratings. Like, being binary, omits intensity information about the consumer's surplus.

[30] I regard the elicitation of the preferences of consumers who have tried the varieties as more straightforward than the elicitation of the beliefs of consumers about the average level of popularity of other consumers a given product line. I also conjecture that this simple ternary measure may minimize the cognitive demands on survey respondents.

has an expected shopping utility $V_{si}(m_{si}^*)$ with a potential optimal sample size m_{si}^* . I take the distribution of N_{si} and the beliefs $F_{si}(u_{si})$ that influence the expectations in V_{si} as exogenously given. The shopper s *actually* purchases m_{si}^* varieties if $V_{si}(m_{si}^*) > 0$. Let $buys_{sij}$ be the number of purchases (i.e., realized m_{si}^*) of each shopper s of variety j in product line i, so that $buys_{sij} = m_{sij}^*$.

I observe only the number of varieties purchased, $buys_{i,j}$, based on point-of-sales data provided by the store, and not the number of varieties purchased per shopper $buys_{sij}$. Thus, I use the aggregate number of purchases of a given variety, $buys_{ij} = \sum_s buys_{sij}$, in the analysis of the experimental data. Continuing from this theoretical background, the main discussion of the dependent variable is in Section 2 of the main text.