Choice Frictions in Large Assortments

Olivia R. Natan*

April 18, 2022

Click here to download the most recent version

Abstract

This paper studies how the growth and evolution of product assortments impact consumer adoption, churn, and long run consumption. Most economic theories of product variety and the value of platforms suggest consumers at least weakly prefer larger product assortments. In contrast, the psychological literature on the phenomenon of choice overload finds that larger assortments overwhelm consumers with decision costs or induce more regret. I provide empirical evidence of how the size and contents of product assortments impact consumers over their lifetime in an online platform market that provides restaurant delivery. I find that assortment expansion increases the acquisition of new consumers but reduces the frequency of consumption among consumers who remain on the platform. I rationalize these impacts via a model of costly attention and choice under limited information. Counterfactual exercises show that targeting choice set reductions can improve revenue among existing customers.

*University of California, Berkeley Haas School of Business; olivia.natan@haas.berkeley.edu. I am grateful to Jean-Pierre Dubé, Pradeep Chintagunta, Anita Rao, and Ali Hortaçsu for their advice and support. Thanks go to Junhong Chu, Ali Goli, Joonhwi Joo, Uyen Tran, Walter Zhang, Stephan Seiler, and participants at various seminars for their thoughtful comments. I thank the data provider; all errors are my own.

1 Introduction

Consumers have access to more product variety now than ever before. Both online and offline, retailers offer consumers greater product variety through larger and more varied assortments.[1] Whether this increased variety is inherently beneficial to consumers is uncertain. Psychologists have documented that presenting more choice alternatives to consumers lowers the likelihood of purchase (Iyengar and Lepper, 2000). Choice overload - the negative effect of additional choice alternatives on consumer purchase and satisfaction - has been sometimes demonstrated in small assortments both in the lab and in small field tests (Chernev et al., 2012; Scheibehenne et al., 2010). Despite this evidence, economists approach product variety by using models that assume 'more is better' for individual consumers. Widely used models of consumer demand preclude the possibility that consumers may prefer fewer choice alternatives.

An empirical literature on retail assortments has found mixed results on the impact of assortment size on category sales (Broniarczyk et al., 1998; Drèze et al., 1994), interpurchase time (Borle et al., 2005), and store choice (Briesch et al., 2009). These studies have primarily focused on changes to the store-level variety and a single purchase occasion. Changes to product assortments may affect customers differently, depending on category familiarity, their taste for variety, and the match between new products and their tastes. Optimal assortment strategy will reflect this consumer heterogeneity.

In this paper, I empirically examine how larger product assortments affect individual consumption. I document the dynamic impact of changing product assortments on consumer acquisition, short-term retention, and long-term consumption frequency using quasi-experimental variation. I find that more variety is detrimental to the purchase frequency of existing customers, though

[1] The average grocery retailer carries nearly 50,000 SKUs in 2008, up from 9,000 in 1975. (Food Marketing Institute). Spotify offers over 40 million songs (Aguiar et al., 2021).

consumer adoption into the category grows due to greater product variety. Consumers engage in longer search when they face larger assortments, and experiencing these higher search costs reduces their future purchase frequency. To separate the cost of searching through variety from possible changes to consumer match value, I construct and estimate a model of attention allocation where consumer beliefs and attention costs are a function of assortment size. I show that consumers' expectations about product valuations change, but search costs are not directly affected. Additionally, I demonstrate how only individually-targeted assortment reductions can improve sales among existing consumers.

The setting for this paper is an online restaurant-to-consumer delivery platform where the assortment varies across both consumers and time. In this context, product variety is measured as the number of restaurants and breadth of cuisines that deliver via the platform to a consumer's location. As the platform expands over time, consumers face growing numbers of restaurants that will deliver to their location. Unlike many offline retail contexts, the variation in assortment size is observed exactly instead of inferred from purchase data. Additionally, the impact of assortment growth in online restaurant-to-consumer delivery markets can be isolated from other confounding factors, since prices are typically fixed and new consumers have experience in similar categories.

In the restaurant delivery platform data, I observe higher rates of consumer adoption of the platform in neighborhoods with larger local assortments of restaurants. Conditional on adoption, I observe higher consumer spending in neighborhoods with larger local assortments. Since households likely self-select into neighborhoods based in part on local amenities including restaurants, I cannot use this raw correlation to determine the causal effect of variety.

To account for the endogeneity generated by household location choices, I use staggered changes to consumers' choice sets across time and geographic space to identify the causal effect of the size of the assortment. This approach compares the within-household variation in choice sets and purchase behavior among households in the same neighborhood.

I find that larger assortments increase platform adoption but decrease the frequency of purchase among existing users. I test whether the decrease in consumer purchase frequency occurs due to lower rates of consumer conversion on search (as demonstrated in Iyengar and Lepper (2000)), or whether consumers engage with the platform less (as theorized in Kuksov and Villas-Boas (2010)). Consumers are less likely to engage with (e.g., search for restaurants) the platform due to assortment expansion, but consumers' conversion from search to purchase is unaffected. These findings are not predicted by choice overload experiments or by typical (static) demand models. I find suggestive evidence that larger assortments lead to consumers experience higher search costs, which in turn lead consumers to wait longer between purchases.

Next, in order to quantify the mechanism by which assortment expansion negatively impacts existing customers, I construct and estimate a structural model of consumer attention and demand where prior beliefs and information costs may vary with the assortment size. This Rational Inattention model proposes that consumers select how much costly information to acquire about choice alternatives before making a discrete purchase decision. The model, which builds off of Joo (2021), nests a test of how larger assortments increase information costs separately from expected consumption utility, unlike adjustments to simple discrete choice models. I reject that the cost of a unit of information changes as the assortment grows. Instead, I find that inter-purchase time is increasing in the assortment size because consumers' expectations of untried products are lower as the assortment grows.

Finally, I use the structural model to test how much the platform can offset the downsides of assortment expansion by offering personalized choice sets. Revenues can be improved by offering different types of assortment reductions to different consumers. Testing counterfactual assortments is necessary to understand two countervailing forces in demand for heterogeneous products: (1) the negative effect of larger assortments on consumer expectations and (2) the improved possible match value generated by a larger assortment. I consider assortment restrictions that hold the supply and contents of restaurant fixed. For a potential restricted size of the assortment, I target the contents based on several metrics (e.g., probability of purchase, expected platform revenue). I simulate consumer choices and platform revenue under proposed assortment reductions, and I compare revenue across both the magnitude of the assortment reduction and across targeting methods. I find that the platform can improve weekly purchase frequency up to 40% among existing customers by offering assortment reductions that target based on consumer preferences; reductions at random make purchase even less likely.

1.1 Assortment Management in Practice

Retailers and intermediaries use reductions to the scope of their assortments—a practice known as SKU rationalization—to address operational concerns. Limiting the number of products stocked on shelves simplifies operations and stocking costs. However, results of this strategy has proven mixed in offline contexts (Boatwright and Nunes, 2001; Borle et al., 2005; Sloot et al., 2006). WalMart, for example, ultimately reversed course after trying a large-scale SKU rationalization by bringing back 8,500 SKUs to their stores.[2] This paper provides evidence suggesting another mechanism by which retailers can benefit from SKU rationalization. In particular, if reductions can be customized to individual customers, retailers can improve retention via SKU rationalization.

1.2 Related Literature

This paper relates closely to the marketing literature on product assortments, including work considering the possibility of making assortments strategically smaller. Much attention has been devoted to assortments at the store, category, and product line levels in grocery retailing (Boatwright and Nunes, 2001; Borle et al., 2005; Briesch et al., 2009; Broniarczyk et al., 1998; Draganska and

[2]Source: https://retailwire.com/discussion/walmart-reverses-course-on-sku-rationalization/

Jain, 2005; Drèze et al., 1994).[3] I add to this stream of literature by measuring the effect of assortment size on individual consumers in platform adoption and repeat-purchase settings. Existing empirical work finds mixed results—in some contexts, removing low-selling items improves in-store sales, while in others, more products on the shelf yield higher sales.

Consumer behavior research has documented robust instances of choice overload, starting with Iyengar and Lepper (2000). These papers (reviewed in Chernev et al. (2012)) find that showing consumers a larger variety of products induces more interest in browsing the products, but fewer overall purchases. However, Scheibehenne et al. (2010) suggest that the average empirical effect is close to zero, depending on the context studied. I build on these small-scale, primarily lab-based, empirical findings by providing evidence for choice overload effects in an empirical setting with large choice sets and repeated consumption. Because of the richness of my empirical setting, I test how larger assortments are differentially costly across consumers and choice contexts.

Choice overload can be rationalized by several economic theories.[4] Kamenica (2008) provides a theoretical account of the effect based on equilibrium behavior of firms selling multiple products. Consumers make inferences about product quality from the length of the product line. In the presence of search costs, the expected cost of searching can deter consumers from entering the market, since they expect to have to search too much to find the product they want when the number of products grow (Kuksov and Villas-Boas, 2010).

This paper also relates to the economic literature on product variety. Theoretical work on product variety models consumers who receive utility directly from variety when they purchase a basket of goods (Bronnenberg, 2015; Dixit and Stiglitz, 1977) or due to many goods meeting heterogeneous tastes better (Hotelling, 1929; Lancaster, 1975, 1990). Empirical approaches to product

[3] The structure of the assortment, such as shelf space allocation offline and product organization, can also influence consumer perceptions about the assortment's variety (Eisend, 2014; Kahn and Wansink, 2004). While changes to assortment size will be considered here, the fine-tuned adjustment of shelf facings is beyond the scope of this paper.

[4] Variety may negatively impact consumption levels for other reasons not explored by this paper. Examples include matching markets (Halaburda et al., 2018) or consumer learning (Kim, 2021).

variety have used the characteristics approach to measure whether markets provide socially optimal product variety, and they typically find positive returns to variety on market size(Berry et al., 2016; Berry and Waldfogel, 1999; Illanes and Moshary, 2020; Quan and Williams, 2018). However, these discrete choice models of demand suffer from a mechanical issue with the introduction of new products. Each new product generates a new characteristic from each products demand shock, which mechanically increases consumer welfare (see Ackerberg and Rysman (2005)). I extend work on Rational Inattention models of demand (Caplin and Dean, 2015; Joo, 2021; Matêjka and McKay, 2015) to estimate the returns to product variety in a manner that breaks this mechanical connection.

My work is also related to empirical work on platforms and retailers that has focused on how the size of the seller base impacts competitive and demand dynamics on platforms. The two most closely related works, Li and Netessine (2020) and Farronato et al. (2020), find evidence for limited-to-no cross network effects in online platforms, while others document positive cross-network effects (Chu and Manchanda, 2016; Lin, 2017). Reshef (2020) uses similar data and identifying variation to study the impact of assortment changes on how platform sellers price differently under increased competition. The findings of that paper - that new entries benefit ‘strong‘ incumbents and hurt ‘weak‘ ones - is consistent with the findings in this paper. Ershov (2018) similarly looks at how changes to search frictions changes entry quality; he finds a reduction in search costs on a platform spurs entry of low-quality products. I contribute to these works by studying (i) how entry of sellers differentially impacts individual consumers based what part of the consumer lifecycle they are in and (ii) how platforms can leverage the online nature of their business to offer an individually-targeted solution.

The rest of this paper is structured as follows. Section 2 describes the relevant details on the market context and the data. Section 3 presents the research design, reduced form results, and robustness checks. Section 4 presents the demand model, estimates, and counterfactual results. In

Section 5, I conclude and discuss possible extensions.

2 Context and Data

2.1 Context

This paper studies product variety in the U.S. restaurant-to-consumer delivery market. This is a large market, with annual revenue over $22 billion (Statista 2019). The two main channels for purchasing restaurant delivery are directly from individual restaurants (primarily offline, such as over the phone) and online ordering and delivery platforms. The former channel is older and remains the largest segment (Statista 2019). In this traditional model, consumers place orders by calling local restaurants and waiting for a restaurant-employed driver to deliver food to the specified location. Some restaurants offer direct online ordering.[5] The prevalence of online platforms has taken off since 2010. In this channel, consumers may access many restaurants through a single website or mobile application, where they see menu options, place and pay for their order, and receive delivery through the service.

The online restaurant ordering and delivery market is useful for studying product variety and assortments for both practical and substantive reasons. Typical assortment studies have focused on grocery retail, but their assortments are varying in often unobserved ways (due to product stock-outs). Online restaurant ordering assortments are observed based on restaurant entry, exit, and delivery zones on the platform. Moreover, unlike offline retail assortments or other types of online assortments, these platforms' assortments vary across individuals within the same time period based on their location. Much like the offline restaurant delivery market, each restaurant may choose which addresses they will deliver to. Neighbors can face different assortments on the platform based on the restaurants' delivery zone decisions.

[5]For example, see Domino's Pizza online ordering system. Domino's alone sold $9.8 Billion in delivery pizza in the US in 2018 (2018 Domino's Annual Report)

In addition to the observed and widespread variation in assortments in this market, prices are typically fixed,[6] and new consumers have experience in similar categories (e.g. ordering restaurant food over the phone or in person). These context-specific factors allow me to rule out alternative explanations for why assortment expansion could be detrimental. Consumers who are new to this market are not new to the broader category of prepared food. Consumption through the platform may allow them to learn about ordering online, but it should not cause them to update their beliefs or preferences for food characteristics.

Data was generously and anonymously shared by a company whose business includes the operation of an online (desktop, mobile, and app) restaurant food ordering business in the U.S. (henceforth, "the platform"). On the platform, consumers find restaurants that operate initially offline, but then start selling additionally via this online channel. I study the Los Angeles metro area market[7] from 2015 - mid 2018. The platform faced competition from several online competitors during this time period. I will largely abstract from competitive dynamics.

Entry onto the platform by existing restaurants comprises much of the variation in assortments in this data.[8]. Based on discussions with both restaurants and the platform, the typical entrant during this time period was not a new restaurant, but instead was a restaurant already operating offline. New products in this context expand the online assortment but don't also grow the consumer's offline choice set. This will allow us to study how consumers value variety in a specific channel, rather than across all channels or across all related markets.

[6] I will discuss prices further in section 2.2. Changes to the price of a meal happen for several reasons. First, prices of a meal differ across restaurants, including within cuisine. An upscale sushi restaurant may charge more for a single meal than both a local pizzeria and a casual neighborhood sushi restaurant. Second, the meal price can differ within a restaurant based on consumer choices of menu items. Third, the price can differ because the menu prices change. Finally, the total price can differ when consumers use platform promotional discounts.

[7] This includes five counties: Los Angeles, Ventura, Orange, Riverside, San Bernardino.

[8] Exit does occur, but it is much less common than entry

2.2 Data

I combine several data sources from the platform: (1) two consumer search and purchase panels, (2) restaurant delivery zones, (3) restaurant entry and exit data, and (4) additional restaurant characteristics data. The first sample is the set of all new users in the relevant geographic area from January 2015 until June 2016. I use this large, repeated cross-section to study customer acquisition and immediate retention, as it records the timing and contents of the first order on the platform and how many subsequent orders the consumer made on platform. The second panel is a subset of the initial cross-section, where I follow a cohort of new users from the first 6 months of 2015 in LA. I keep users who order a second time in the first 60 days of their 'lifetime' and remain in Los Angeles area. I then follow their activity in full from 60 days after their first order until September 2018. I make this second panel more restricted to measure any effects on returning or active customers. I supplement these data by matching all consumers to Census demographic data at the tract level.

To supplement the consumer order data, I add restaurant characteristics data. These include cuisine, price and fee measure, location, entry and exit dates, and matched Yelp data. These data are fixed across time for each restaurant. In particular, I use a measure of price that is not time-varying. Restaurant prices are high-dimensional: each menu item has a price, and the composition of items or the prices attached to them may change or remain constant over time. Item-level pricing is not available. For this reason, I use the average total spending on food and beverage at the restaurant to capture the typical price of buying food at that restaurant. Delivery fees are measured somewhat noisily, so I take the average over time for each restaurant. To introduce additional price variation, I also construct a city-level panel of sales tax rates, which vary over the course of this time period for this region.

A novel component of the data is delivery zones. For the time period studied, the platform understands that restaurants set their delivery zones to be similar to what a restaurant would offer for their phone-based delivery orders. These zones are relayed in terms of a geographic polygon; a location can receive delivery from the restaurant only if its point contained by the polygon. For the customer data, I create the realized choice sets at each point in time based on whether the location is in the delivery zone and whether the restaurant is available on the platform.

I document several key facts about the consumer and restaurant data. Table OA2 shows the consumer panel summary stats. The median consumer orders 8 times from 4 unique restaurants in the 3.5 year period studied. However, the means of these measures are more than double the median, since there is a substantial right tail of high-consumption users. In Figure 9, I plot the distribution of choice set sizes for all area census tracts that experience any adoption in the 18 month period for which I have complete adoption data. Across the entire geographic area, the median consumer has relatively little choice - the median of this cross-section is 21 restaurants, and the mean is 44 restaurants. These relatively low summary stats reflect the construction of the data: I include census tracts even if they only have a single adopter over the time period. These may be outerlying areas with little restaurant availability. In contrast, the selection of choice set sizes for the consumer-level panel (shown in Figure 10), a selected sample of consumers, has a median of 85 restaurants. By the end of the panel, this median has grown to 243 restaurants (see Figure 11).

3 Causal Impact of Assortment Expansion

3.1 Research Design

The assortments consumers face on the platform are not randomly assigned in size or contents. Instead, they are drawn from the equilibrium availability of local, offline-operating restaurants. Additionally, these local restaurants choose whether and when to enter the online delivery market. The timing of entry may also coincide with platform-level promotional activity. In order to identify how a change to the platform assortment impacts consumers, I will need to address the effects of local geographic offline market equilibria, restaurant strategy, and platform promotions.

The number of local offline restaurant markets reflect the tastes of local consumers, and more restaurants can be sustained by a local market with greater preferences for consuming restaurant food. This equilibrium reflects two sources of selection. First, restaurants choose to open in neighborhoods where, all else equal, they expect higher demand. Second, consumers choose to live in neighborhoods, all else equal, with local amenities that match their tastes. These amenities include local restaurants. As a result, I expect that consumers who live in areas with higher offline restaurant availability have a higher average preference for ordering from restaurants. This would generate positive (spatial) correlation between the size of the choice set and unobserved consumer or neighborhood heterogeneity in platform behavior.

The entry timing of restaurants may coincide with unobserved demand shocks, as restaurant may choose to enter the online market when they expect particularly high demand. Restaurants may also consider potential demand when selecting their delivery zones. If this is the case, I expect consumers who live within the delivery boundary to have higher demand for the entering restaurant on the platform than those outside the boundary. Additionally, I cannot rule out that the platform engages in promotional activity that corresponds to time periods with high degrees of new restaurant entry. These forces could also generate a positive correlation between changes to the choice set on the platform and unobserved platform demand shocks for a given time period.

To address these potential sources of positive correlation between assortment size and consumer outcomes (platform adoption, churn, and purchase frequency), I consider a staggered differences-in-differences design. I will first discuss the intuition behind my identification strategy, and then I will present the specific assumptions used.

Consumers (i) living in a neighborhood (z) face a platform choice set of restaurants at time t of S_{it}. As discussed in the prior section, this assortment of restaurants varies across consumers and time, but the across-consumer variation is driven only by consumers' order location. My identification approach will use the variation generated by differential entry of restaurants, controlling for consumer or neighborhood unobserved heterogeneity (through fixed effects). The simplest strategy will also control for granular time effects, which will absorb confounding variation generated by restaurant timing selection or platform promotions. However, such an approach will consider consumers who live extremely far apart to be comparable, comparing the 'within-user' variation of a downtown resident to a suburban household. There may be other time-varying differences between these consumers, such as the availability of offline options.

I will additionally consider only within-neighborhood variation in the size of the choice set. Using neighborhood-time fixed effects will isolate variation in restaurant entry to the platform within local areas. The treatment effect will average across these local comparisons, but will leverage only comparisons between consumers in the same neighborhood who receive different assortment sizes. This variation is comparing consumers who live on either side of the delivery zone specified by the entering restaurant.

Identification Example 1. *To better understand the variation that will generate these 2 sets of estimates, consider a simple 2-period example. A restaurant ("Z") enters at the beginning of the second period. Half of the consumers, as noted below, are now granted an additional choice on the platform.*

Consumer	*Neighborhood*	*Restaurant Z Availability*
1	*A*	*0*
2	*A*	*0*
3	*B*	*0*
4	*B*	*1*
5	*C*	*1*
6	*C*	*1*

In the main specifications (with consumer and time fixed effects), the variation used to identify the effect of more choices is comparing all treated users (Consumers 4, 5, 6) to all untreated users

(Consumers 1, 2, 3). If orders are the outcome of interest, the main specification will compare changes in orders between treated and untreated users.

In the specification with neighborhood-time fixed effects, the regression will only use variation in entry that varies within a neighborhood-time period. In this example, only the variation in choice sets in Neighborhood B will be used, since the neighborhood-time fixed effects will soak up any variation from Neighborhood A and Neighborhood C.

Before discussing the assumptions required to interpret this specification as causal, I want to highlight the residual variation used in this context under these strategies. The two-way fixed effects approach soaks up nearly all of the variation in choice set size in this data. To describe this, I present the R-squared and F-statistics from regressions of different two-way fixed effect regressions on the treatments of interest: the assortment size in levels and changes to the assortment size. Table OA3 presents these results. In levels, the fixed effects explain nearly all of the differences across consumers and time in the size of the choice set. However, this is slightly misleading. The marginal effect of the assortment size is identified here from changes in assortments. The fixed effects explain a considerable share of the changes to assortments: over 90 percent of the entry is explained by neighborhood-week fixed effects alone. This highlights two notes for future results. First, despite the size of the data, I should expect these results to be relatively low power given the share of actual variation used to identify this main effect. Second, the set of residual variation used is quite small, so endogeneity concerns should be addressed specifically with this variation in mind.

3.2 Identification Assumptions

I estimate models of the form:

$$y_{it} = \alpha_i + \alpha_{zt} + f(S_{it}, \beta) + \epsilon_{it}$$

$$y_{ct} = \alpha_c + \alpha_{zt} + f(S_{ct}, \beta) + \epsilon_{ct}$$

As described above, I will also consider a version with $\alpha_{zt} = \alpha_t \forall z$. The outcomes y_{it} include adoption rate (A_{ct}, measured at the census tract c level), churn rate (C_{ct}), and weekly spending and orders (b_{it},o_{it}). Consider a linear effect of assortment size, so that $f(S_{it}, \beta) = \beta|S_{it}|$.

I assume that $|S_{it}| \perp \epsilon_{it}|\alpha_i, \alpha_t$: the assortment size is conditionally independent of unobserved determinants of adoption and ordering. Further, I assume that in the absence of changes to the assortment, consumer behavior follows parallel trends.

After controlling for consumer geographic selection (through the fixed effects), the main challenge to identification is strategic behavior on the part of restaurants. In my empirical context, restaurants cannot control their exact entry timing precisely, so they cannot select entry timing to coincide with positive demand shocks. However, the shape and size of their choice of delivery zones could potentially violate the identification assumption. The overall size of these zones is fairly uniform, with typical radii around the physical restaurant location of 3 to 4 kilometers. A larger concern would be strategic behavior in drawing the exact boundary, conditional on approximate size. In particular, restaurants may choose to select their boundaries by including blocks where they expect to sell and excluding blocks with low demand, on the margin. If this is the case, then the estimates here will be an upper bound on the true effect, since such strategic behavior would generate positive selection.

These are strong assumptions and worth discussing in practical detail. In particular, I want to emphasize what strategic behavior by restaurants and the platform is *ruled out* by this design. I will additionally discuss assumptions about dynamic treatment effects that are testable and ruled out.

I am implicitly assuming that platform promotional activity (advertising or discounts) only varies across consumers independent of changes to the assortment. In particular, this rules out that the platform engages in targeting resulting from past treatment effects. For example, if β is positive, I assume that the platform does not send promotions to remedially improve adoption or retention in areas with low assortment size. If β is negative, I assume that the platform doesn't remedially target areas or consumers with high assortment growth with promotions. However, conditional on past entry (and any consumer dynamics), I am ruling out that *entry itself* alters platform promotions.

I am further assuming that restaurant entry timing is not a function of time-specific demand shocks for covered households (versus non-covered households). Given the level of granularity studied (weekly), I am skeptical that restaurants are timing entry in such a strategic manner. The greater concern is that restaurants are selecting their delivery zone (on the margin) in a manner that would violate my identification assumption. The overall size of these zones is fairly uniform, with typical radii around the physical restaurant location of 3 to 4 kilometers. Restaurants may choose to select their boundaries by including blocks where they expect to sell and excluding blocks with low demand, on the margin. If this is the case, then the estimates here will be an upper bound on the true effect, since such strategic behavior would generate positive selection.

Finally, the difference-in-differences strategy rules out dynamic treatment effects. However, some of these dynamic effects can be included by testing estimating equations that include treatment lags or cumulative measures of past treatment changes. In light of recent work highlighting potential pitfalls of two-way fixed effects for estimating difference-in-differences research design (Callaway and Sant'Anna, 2020; De Chaisemartin and d'Haultfoeuille, 2020; Sun and Abraham, 2020), I also test robustness to alternative estimators in Appendix OA5.

3.3 Results

Larger platform assortments increase the number of adopting (first-time) consumers but reduce the frequency with which returning consumers order on the platform. Conditional on adoption, the size of the assortment at the point of adoption does not significantly alter the probability of churn after the first order. Table 1 summarizes the direction of the estimated effects.

Table 1: Summary of Effects

	Sign	Effect
$\frac{\partial A_{ct}}{\partial \lvert S_{ct} \rvert}$	(+)	Large assortments increase adoption
$\frac{\partial C_{ct}}{\partial \lvert S_{ct} \rvert}$	(0)	Larger assortments at adoption don't impact immediate churn
$\frac{\partial o_{it}}{\partial \lvert S_{it} \rvert}$	(-)	Larger assortments reduce order frequency among returning users
$\frac{\partial s_{it}}{\partial \lvert S_{it} \rvert}$	(-)	Larger assortments reduce search sessions
$\frac{\partial o_{it}}{\partial \lvert S_{it} \rvert} \mid s_{it} > 0$	(0)	Larger assortments don't impact orders conditional on search

Table 2: Main Effects of Assortment Size

	Dependent variable:					
	Adoption Rate		Churn Rate		Orders	
	(1)	(2)	(3)	(4)	(5)	(6)
Restaurant Count (tract)	0.0002***	0.0002	0.020**	−0.023		
	(0.00002)	(0.0001)	(0.006)	(0.028)		
Restaurant Count (household)					−0.0001***	−0.0001
					(0.00003)	(0.0001)
ZCTA-Week FEs?	N	Y	N	Y	N	Y
Observations	204,798	204,798	104,218	104,218	2,058,406	2,058,406
R^2	0.718	0.776	0.118	0.331	0.211	0.237
Adjusted R^2	0.713	0.740	0.091	0.114	0.207	0.207

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include week and unit FEs
Standard Errors clustered at tract or household level
Churn Rate results omit promotion use control
Rates range from 0 to 100

Table 2 shows the results for these three outcome measures for the main specifications. Columns 1 and 2 show the estimated marginal impact of an additional restaurant available on the platform on the census tract adoption rate. The effect size is small in levels: the addition of 10 additional restaurants on the platform increases the adoption rate by 0.002 percentage points. This corresponds to a 5-7 percent increase in the adoption rate. I also present the effect of platform assortment size *at adoption* on subsequent churn behavior in columns 3 and 4. These effects are inconsistent in sign across specifications, and the effect is not statistically significant in the preferred specification. The scope of restaurant variety at adoption does not drive consumers to immediately churn from the platform. These results omit the coefficients from a control: the average share of the initial order purchased under promotion. The estimated coefficients on this promotion usage variable are large, positive and significant—consumers who use coupons when adopting are considerably more likely to churn after their first order. Columns 5 and 6 report the effect of assortment size on weekly ordering from consumers who remained on the platform beyond their first order. Adding 10 restaurants to the platform assortment reduces weekly orders by 0.001, or about 1 percent.

Table 3: Effects of Assortment Size and Variety

	Dependent variable:					
	Adoption Rate		Churn Rate		Orders	
	(1)	(2)	(3)	(4)	(5)	(6)
Restaurant Count (tract)	0.0002^{***}	0.0002	0.021^{**}	-0.020		
	(0.00002)	(0.0001)	(0.006)	(0.028)		
Cuisine Count (tract)	-0.0002^{**}	-0.0002	-0.095	0.347^{*}		
	(0.0001)	(0.0001)	(0.090)	(0.165)		
Restaurant Count (household)					-0.0001^{***}	-0.0001
					(0.00003)	(0.0001)
Cuisine Count (household)					0.001^{**}	0.002
					(0.0005)	(0.001)
ZCTA-Week FEs?	N	Y	N	Y	N	Y
Observations	204,798	204,798	104,218	104,218	2,058,406	2,058,406
R^2	0.718	0.776	0.118	0.331	0.211	0.237
Adjusted R^2	0.713	0.740	0.091	0.114	0.207	0.207

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include week and unit FEs
Standard Errors clustered at tract or household level
Churn Rate results omit promotion use control
Rates range from 0 to 100

Beyond this specification, I also consider controlling for the assortment's variety. In Table 3, I present the results that additionally control for the number of cuisines offered in the platform's assortment for that household or census tract. The sign, magnitude, and significance of the estimated main effects of the number of restaurants is not impacted by the inclusion of cuisine variety controls. Conditional on the size of the assortment, there are negative returns to the cuisine variety on adoption (Columns 1-2). In contrast, cuisine variety is positively associated with the frequency of consumption for returning customers.

Table 4: Effect of Assortment Size on Spending

	Dependent variable:			
	Weekly Spend (USD)		Average Order Size (USD)	
	(1)	(2)	(3)	(4)
Restaurant Count	-0.002^{**}	0.001	0.004	0.022^{**}
	(0.001)	(0.003)	(0.002)	(0.008)
ZCTA-Week FE?	N	Y	N	Y
Observations	2,058,406	2,058,406	139,230	139,230
R^2	0.193	0.221	0.538	0.686
Adjusted R^2	0.188	0.190	0.508	0.523

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include Week and Individual FEs
Standard Errors clustered by individual
Order Size result conditions on purchase
Omits selection controls in columns 3 and 4

The negative effect of variety on purchase frequency for returning customers is mirrored in the effect on weekly spending. Conditional on purchase, consumers spend slightly more per order when faced with more restaurants. However, I cannot interpret this result as causal. Since a larger assortment reduces the probability of order, it is likely that the consumers who order nonetheless face larger unobserved demand shocks than those who are deterred from ordering. I interpret the small positive effect on order size as either reflecting a true marginal effect or selection.

An increase in assortment size may impact the consumer after having placed an order—in particular, interpurchase time. Time until the next purchase increases as a function of the size of the current assortment (see Table 5).

Table 5: Assortment Size and Time until next Purchase

	Dependent variable:			
	Interpurchase Time (Weeks)			
	(1)	(2)	(3)	(4)
Restaurant Count	0.004**	0.013*	0.004**	0.013*
	(0.001)	(0.006)	(0.001)	(0.006)
Search Duration (Mins)			0.013***	0.013**
			(0.003)	(0.004)
Count Elasticity	0.1329	0.4878	0.1380	0.4698
Duration Elasticity			0.1783	0.1696
ZCTA-Week FE?	N	Y	N	Y
Observations	100,310	100,310	100,310	100,310
R^2	0.319	0.582	0.322	0.584
Adjusted R^2	0.280	0.319	0.283	0.322
Note:	*p<0.05; **p<0.01; ***p<0.001			
	All specs include Cuisine and Individual FEs			

To understand the negative effect of variety on returning customers further, I consider several sources of treatment heterogeneity and several additional outcome measures. First, I document that nonlinearities in effects across the choice set size are limited. Table OA9 presents regression results that allow the marginal effect of an additional platform restaurant to differ across five assortment size bins. These results are consistent with the uniform effect - the marginal negative effect of additional restaurants is similar across choice set sizes.

Changes to the assortment may drive changes in which products consumers choose. If consumers seek to avoid the challenge of distinguishing between unfamiliar products, their choices may skew towards familiar, previously consumed products. I test which types of purchases are most affected by assortment changes. Table 6 shows these effects. I find that the effect of assortment size differs across orders which are repeat consumption (i.e. previously-ordered restaurants) and orders which are trying a new-to-the-consumer product (i.e. never-previously-ordered restaurants). The negative average effect on orders is coming from a large reduction in the probability of trying previously

Table 6: Effect of Assortment Size on Orders by Type

	Dependent variable:			
	First-Time Orders		Repeat Orders	
	(1)	(2)	(3)	(4)
Restaurant Count	-0.0001^{***}	-0.0001^{*}	0.00001	0.00001
	(0.00001)	(0.00003)	(0.00002)	(0.0001)
ZCTA-Week FE?	N	Y	N	Y
Observations	2,058,406	2,058,406	2,058,406	2,058,406
R^2	0.071	0.100	0.201	0.228
Adjusted R^2	0.066	0.064	0.197	0.197

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include Week and Individual FEs
Standard Errors clustered by individual

unsampled restaurants, while there appears to be no significant impact on repeat consumption. This is consistent with a higher information cost of trying new products.

Consumers may also purchase less frequently if large choice sets lead them to buy lower quality or more expensive products due to the difficulty of search. Table 7 shows how the average characteristics of restaurants ultimately chosen change with the size of the assortment. When customers do ultimately purchase from a larger assortment, they may purchase from more popular on-platform restaurants.[9] However, these chosen restaurants are, on average, less popular overall, as measured by the total number of Yelp reviews. There is not a consistent effect across specifications on the quality of the restaurant (measured as 4.5 or 5 stars on Yelp) or the average price of a basket of food at the restaurant. The selection of restaurants that are less popular offline may not necessarily reflect lower quality - successful offline restaurants may be popular due to the quality of in-person service, which does not translate to the quality of service online.

I next consider heterogeneous effects of different types of restaurant entry on order frequency. There are three dimensions on which I test differential effects: restaurant chain status, restaurant

[9] This popularity measure is constructed as the sales quantile of the restaurant among this cohort from the entire panel.

Table 7: Effect of Assortment Size on Ordered Restaurant

	Characteristics of Ordered Restaurant							
	Platform Popularity		Yelp Review Count		Yelp Rating over 4		Avg. Price	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Restaurant Count	0.0003***	−0.0001	−0.626***	−0.653	0.0002**	−0.0001	−0.0004	0.003
	(0.00003)	(0.0001)	(0.128)	(0.417)	(0.0001)	(0.0002)	(0.001)	(0.004)
ZCTA-Week FEs?	N	Y	N	Y	N	Y	N	Y
Observations	121,820	121,820	121,820	121,820	121,820	121,820	121,820	121,820
R^2	0.350	0.628	0.367	0.555	0.312	0.555	0.419	0.616
Adjusted R^2	0.302	0.408	0.320	0.293	0.261	0.292	0.377	0.390

Note: *p<0.1; **p<0.05; ***p<0.01
All specs include week and individual FEs
Standard Errors clustered by individual

vertical quality rating, and restaurant match with individual consumers. Consumers may have very different information or consideration costs from chain restaurants than local independent restaurants. In particular, I expect that there consumers would have little need to search over national or prominent local chains. Consistent with this story, the number of independent restaurants significantly reduces order frequency, but the number of chain restaurants has no significant effect (shown in Table 8).[10]

If new restaurants are lower-quality than incumbent restaurants, consumers expectations of the value of ordering on the platform may be diluted, driving lower return frequency. Given reduction in purchases from novel alternatives, I expect that adding low-quality entrants would reduce consumption frequency more than high-quality entrants. I test this by breaking up the assortment by Yelp star ratings. Table 9 presents the effect of assortment size binned by Yelp ratings. The results are noisy, but they suggest this effect is not ameliorated by high-rated restaurants entering the assortment. In particular, adding very-highly rated restaurants to the platform still reduces consumption frequency. These results are imprecise - I cannot rule out a small positive effect. I also

[10]Chain restaurants include large, national quick-serve and fast casual restaurants, regional chains, and local chains with at least 5 outlets.

Table 8: Effect of Chain Restaurants on Weekly Orders

	Dependent variable:	
	Weekly Orders	
	(1)	(2)
Independent Restaurant Count	-0.0001^{**}	-0.0001
	(0.00004)	(0.0001)
Chain Restaurant Count	-0.00003	0.0001
	(0.0002)	(0.0003)
ZCTA-Week FE?	N	Y
Observations	2,016,651	2,016,651
R^2	0.214	0.242
Adjusted R^2	0.209	0.214
Note:		*p<0.05; **p<0.01; ***p<0.001
		All specs include Week and Individual FEs
		Standard Errors clustered by individual

cannot rule out that restaurants of all vertical quality ratings reduce the probability of consumption on the platform by returning users.

Table 9: Effect of Assortment Quality on Weekly Orders

	Dependent variable:	
	Weekly Orders	
	(1)	(2)
Restaurant Count (4.5 or 5 stars)	-0.0002	-0.001
	(0.0002)	(0.001)
Restaurant Count (3.5 or 4 stars)	-0.00004	0.0002
	(0.0001)	(0.0002)
Restaurant Count (3 stars or less)	-0.0001	-0.0003
	(0.0003)	(0.001)
ZCTA-Week FE?	N	Y
Observations	1,991,134	1,991,134
R^2	0.214	0.239
Adjusted R^2	0.210	0.209
Note:		*p<0.05; **p<0.01; ***p<0.001
		All specs include Week and Individual FEs
		Standard Errors clustered by individual

Even if there is no effect of vertical quality, entry could still dilute individual consumer expectations about match value if the changes to the assortment are mostly low-match-value products for their particular tastes. To proxy for this, I distinguish between relevant (ever consumed) and irrelevant (never consumed) cuisines for each consumer. This proxy may be noisy. For a consumer who orders pizza, the addition of more pizza restaurants may be irrelevant, as they already have found a preferred pizza restaurant. Conversely, a consumer who never orders pizza on the platform may still consider it for purchase. Table 10 shows the marginal effect of relevant- (consumed) versus irrelevant- (never consumed) restaurants added into the assortment. I find that the negative effect on purchase frequency is driven by growth in relevant restaurants. The addition of restaurants which are less relevant, in contrast, increases the probability of purchase.

Table 10: Impact of Relevant Restaurant Entry on Returning Customers

	Dependent variable:			
	Weekly Orders			
	(1)	(2)	(3)	(4)
Ever Consumed Cuisine Restaurants	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.001^{***}
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Never Consumed Cuisine Restaurants			0.0003^{***}	0.0003^{**}
			(0.00004)	(0.0001)
ZCTA-Week FE?	N	Y	N	Y
Observations	2,058,406	2,058,406	2,058,406	2,058,406
R^2	0.213	0.239	0.213	0.239
Adjusted R^2	0.208	0.209	0.208	0.209

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include individual and Week FEs
Standard Errors clustered by individual

Finally, I consider heterogeneous effects across consumers. Consumers may value variety differently, and they may realize any costs of sifting through many products differently. I allow the effect of assortment size to differ by the degree of observed variety consumption in the panel.[11] Table 11 shows these estimates. Consumers in the bottom quartile of variety in consumption (in this

[11] These results condition on outcomes and should be taken as descriptive only.

case, those who only try 1 or 2 unique restaurants over 3 years) do not reduce their consumption frequency when the assortment expands. Consumers in the top quartile of variety in consumption have a larger negative effect of assortment expansion. Note that these consumers, by virtue of being in the top 25 percent of varied consumption, are above median in overall order frequency. However, their purchase frequency is reduced at a higher rate than other users when restaurants are added to their assortment. This may reflect that in order to access more varied consumption, consumers by definition must engage in more search. Low-variety consumers, on the other hand, can avoid searching entirely, since the platform's home page present recently purchased options prominently.

Table 11: Effect of Assortment Size on Weekly Orders by User Consumption Variety

	Dependent variable:	
	Weekly Orders	
	(1)	(2)
Rest Ct : High Variety User	-0.001^{***}	-0.001^{***}
	(0.0001)	(0.0001)
Rest Ct : Low Variety User	0.0003^{***}	0.0003^{**}
	(0.00002)	(0.0001)
Rest Ct : Medium Variety User	-0.00004	-0.00002
	(0.00003)	(0.0001)
ZCTA-Week FE?	N	Y
Observations	2,058,406	2,058,406
R^2	0.214	0.240
Adjusted R^2	0.210	0.210

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include Individual and Week FEs
Standard Errors clustered at individual level
High and Low Variety Defined as
Top and Bottom Quartile of Restaurants Ordered

Poor experiences purchasing from a large assortment could potentially drive consumers to churn from the platform, though they remain observed in the panel. To consider whether this is driving my results, I look at three additional analyses. First, I control for individual-year fixed effects, and I find consistent results. As consumers churn, their lack of orders will be full absorbed by these fixed effects. Second, I look at the effect of assortment size for consumers who still make an order in the final year of data, and I find that assortment size decreases order frequency. Finally, I look descriptively at when consumers churn from the panel. Most churn occurs in the first year panel. Given this is the case, the results using more granular user-time fixed effects should control for this exodus of users.

To disentangle the negative effect of variety on returning customers further, I supplement the order data with summary data on search behavior for a subset of consumers.[12] Does the growth of assortments lead consumers to search more? Does longer search lead consumers to learn about the cost of finding a good option in large choice sets?

I observe weekly counts of search sessions on the platform, which allows me to construct conversion rates conditional on search. Using this subset of about half of the consumers, I document that the elasticity of searching with respect to assortment size is about -0.5: the addition of 1% more restaurants reduces weekly search by 0.5%. Conditional on searching, however, there appears to be no or a small positive impact of assortment size on search conversion into ordering. Table 12 shows the effect of assortment size on weekly search sessions and search duration. This reduction at the 'top of the funnel' is inconsistent with the in-person choice overload experiments. Table OA11 shows the effect of the number of restaurants on search duration and purchase, conditional on searching in the first place. Conditional on searching, the assortment size does not have a clear effect on duration and purchase.

This inconsistent average effect hides two distinct patterns in how the assortment size changes the intensity of search, conditional on searching. Table 13 shows how assortment size alters search duration by whether consumers ultimately purchase, conditional on search. Consumers who ultimately purchase spend longer searching prior to purchase in large assortments, but consumers who

[12] Appendix OA6.3 details selection of users into the search data.

Table 12: Effect of Assortment Size on Consumer Search Behavior

	Dependent variable:			
	Weekly Search Sessions		Weekly Search Duration	
	(1)	(2)	(3)	(4)
Restaurant Count	-0.0003^{***}	-0.0002	-0.003^{***}	-0.002
	(0.0001)	(0.0002)	(0.001)	(0.002)
ZCTA-Week FE?	N	Y	N	Y
Observations	1,436,956	1,436,956	1,436,956	1,436,956
R^2	0.222	0.262	0.128	0.170
Adjusted R^2	0.217	0.222	0.123	0.126
Note:	*p<0.05; **p<0.01; ***p<0.001			
	All specs include Week and Individual FEs			
	Standard Errors clustered by individual			
	Search duration measured in minutes			

ultimately don't purchase search less time searching. These effects are consistent with consumers avoiding extensive search unless they have a strong contemporaneous demand shock.

Table 13: Effect of Assortment Size on Search Duration by Purchase Status

	Dependent variable:			
	Search Duration (Minutes)			
	(1)	(2)	(3)	(4)
Rest Ct: No Purchase	-0.047^{***}	-0.040^{***}	-0.015	-0.016
	(0.007)	(0.007)	(0.025)	(0.024)
Rest Ct: Purchase	0.020^{**}	0.022^{**}	0.047	0.043
	(0.007)	(0.007)	(0.025)	(0.025)
ZCTA-Week FE?	N	N	Y	Y
Selection Controls?	N	Y	N	Y
Observations	84,630	84,630	84,630	84,630
R^2	0.251	0.272	0.556	0.569
Adjusted R^2	0.194	0.217	0.226	0.248
Note:	*p<0.05; **p<0.01; ***p<0.001			
	All specs include Week and Individual FEs			
	Omits search selection controls			
	Standard Errors clustered by individual			

3.4 Robustness Checks

I conduct three main robustness checks to ensure these causal effects are robust to alternative explanations: strategic restaurant entry timing, limited updating by consumers, and variation in entrant quality.

3.4.1 Merger Natural Experiment

I use a platform merger as a natural experiment to check that the effects are robust to otherwise endogenous entry timing. The platform studied in this paper, during the sample, acquired several small competing platforms. After the legal portion of the merger was completed, the platform on-boarded the restaurants from the acquired firms and released them online on a handful days during this time period. These discrete jumps serve as a natural experiment, since the entry timing of these restaurants was not based on restaurant strategic behavior or on platform strategic behavior – it was driven by availability of platform staff. The results from these natural experiments will be less precise, since the individual fixed effects cannot be estimated with precision, and the size of the data is much reduced. The results of this robustness check do not contradict the previous findings, but they are noisy null results. Since the sample size is less than 3% of the original analysis, the standard errors are consistent with magnitudes from power simulations for this sample reduction. The effect on adoption remains positive, and the effect on consumption among existing users remains consistently negative across all individual natural experiments.

3.4.2 Alternative Assortment Measures

I check the robustness of the causal effect to consumer limited information by considering the extreme where consumers only update their understanding of assortment size when they interact with the platform. I do this in two ways. First, I construct the size of the assortment at the lagged choice time period, and I carry it forward. Second, given the average growth trajectory on the platform, I allow an approximate updating of the assortment over time in line with this growth.

In both cases, I also control for the time since last purchase. Simply repeating the main specifications with this new measure will generate selection that will bias the estimates. The platform is generally growing over time, so households who more recently went online are going to have higher assortment sizes, all else equal. However, their recent purchase or search also can reflect higher engagement and purchase likelihood overall, which could lead to spurious positive correlation between the size of the assortment and purchase probability.

Results from this alternative measure, which would allow for the possibility that consumers aren't fully aware of assortment changes, are consistent with a marginal negative effect on weekly ordering. This interpretation, however, does not square exactly with the prior finding that concurrent assortment changes negatively impact ordering. Because growth of the assortment is correlated across time within an area – i.e. high growth areas remain so throughout the sample – it's not possible to directly test the 'lagged perception' against current changes. I estimate a version of the specification with both measures. The effect of the concurrent assortment size on consumption remains significant and negative, while the alternative measure now has a noisy null effect. From these, I conclude that the marginal effect of assortment growth may have some spillovers over time, and the effect is robust to alternative consumer updating frequency.

3.4.3 Differential Effects across Restaurants

Based on the findings in Ershov (2018), I consider that new additions to the choice set may be meaningfully worse than incumbents. If new restaurants are worse, the average product quality could decline, which could partly explain the decline in consumption I have documented. This mechanism should be more muted in my context, as the introduction of low-quality new products does not, in principle, affect the consumption value of existing preferred products. Using observable quality measures, I found above that a high-quality new entrant is equally detrimental as a low-quality new entrant.[13]

I observe some measurable differences in attributes when comparing entrants (restaurants that enter during the sample) and incumbents. Table OA17 shows the distribution of product attributes among incumbents versus entrant restaurants. The prices charged by these restaurants differ slightly, but this occurs only in terms of delivery fee versus food costs. The total cost is very similar across the two groups. Older (i.e. incumbent) restaurants have many more reviews on average than new restaurants, though this is unsurprising as they have had longer to accumulate them. Incumbents are marginally higher rated on average than entrants. This may reflect selection on surviving restaurants: the incumbents that remain into the panel are ones that have not yet closed. New restaurant entrants also have lower sales, on average, than existing restaurants on the platform. This quality selection could contribute to quality dilution by new entrants, which in turn may contribute to the negative impact of assortment expansion on purchasing.

I conclude that the negative effect is not driven by higher prices or lower observed quality from new entrants, and that it is possible some of the effect is driven by unobserved restaurant quality. However, the marginal negative effect on overall orders in the repeated-purchase context cannot be fully explained by quality in a standard demand model. In particular, even if the average entrant is of lower (unobserved) quality, the negative effect on sales of *existing incumbent* restaurants is a violation of the independence of irrelevant alternatives (IIA).

3.5 Discussion

In this section, I showed how larger assortments marginally improve customer acquisition and reduce consumption among existing customers. This reduction occurs through increased interpurchase time, not smaller baskets. Descriptive evidence is consistent with the presence of consumer search frictions and incomplete information. Consumers who avoid search (by only using the

[13] Yelp rating may be a very noisy measure of quality. Even if that is the case, this information is displayed to consumers on this and many other platforms.

platform to repeat-consume) are unaffected by assortment growth. In contrast, consumers who frequently try new restaurants are most strongly negatively impacted by assortment growth. Search duration (the total time cost expended prior to purchase) is higher when choice sets are larger. This, in turn, increases the time between purchases (Table 5). Consumers experience higher search costs prior to a given purchase, which then reduces the speed at which they subsequently return to the platform. This effect is strongest following consumers' purchase from a novel-to-them restaurant.

These results are inconsistent with full information demand models, but the exact mechanism by which search costs are higher under larger assortments is unclear. One possibility, as in Kuksov and Villas-Boas (2010), is that consumers update their expectations about total search costs, though the per-product search costs are unaltered. This would also be consistent with consumers' expectations about match value changing with the size of the assortment. Another possibility is that per-product search costs are higher, since consumers have to sort through more products to acquire information about any particular restaurant.

I rule out several mechanisms through the research design and through robustness checks. First, this reaction is not through observable quality or price differences (Ershov, 2018) between entrants and existing restaurants. Second, since I have granular time and area-time fixed effects, I rule out that this is occurring because of platform-level promotional activity. Third, the search data is suggestive that this impacts the extensive margin of whether to engage in any on-platform search, rather than the intensive margin of how much to search at all. Across all users, conditional on ordering, larger assortments induce higher rates of repurchase/lower rates of experimentation. Given the structure of the user experience, this is consistent with users potentially searching in a less costly manner, by navigating from the home page which often presents recently ordered-from restaurants.

In the next section, I will build a structural model of consumer demand with two aims. First, I will distinguish between multiple ways in which the search process could be altered by changes to the assortment. Second, it will measure heterogeneous preferences, so that consumers' choice of particular restaurants can related to product characteristics. This will allow me to consider how removing any particular restaurant from the choice set will impact an individual consumer's purchase behavior.

While there are many benefits to the specific empirical context, the structure of the consumer purchasing decision (discrete choice) does prevent me from capturing the full benefit of variety. Grocery retail, where consumers typically purchase many goods that comprise a basket across categories, allows for a better measurement of the returns to consuming multiple goods. My interpretation of these results in broader contexts is that these negative effects of large assortments may be harder to detect, but still influence consumer behavior. The other limitation of this context is that I use only a narrow cohort of users, and there may be some adoption-time specific effects.

4 Limited Information Demand Model

Larger assortments reduce the purchase frequency of returning customers, but the prior section does not provide a clear remedy for platforms or retailers. I build a structural model of consumer information acquisition and demand which nests a test of the mechanism by which larger assortments reduce purchasing. Distinguishing this mechanism (along with estimating consumer preferences) is necessary in order for platforms and retailers to address the reduced purchasing among existing customers. I distinguish between assortments altering the cost of learning product information from altering consumers' expectations about product match. These mechanisms, while both directly addressable by reducing the assortment size, suggest different paths for how else platforms might improve retention. I use the model results to test several assortment reduction strategies. Offering smaller assortments to each consumer improves the expected revenue to the platform only if the reductions are targeted based on consumer tastes.

4.1 A Model of Discrete Choice under Incomplete Information

A consumer i faces a choice set of restaurants S_{it} on the platform at time t based on their location. After adoption, consumers choose to order from a single restaurant $j \in S_{it}$ (measured by choice dummy y_{ijt}) or the outside option (denoted $j = 0$) every period. Consumers have heterogeneous tastes over restaurant attributes. Their consumption utility from each option is $u_{ijt} = \delta_{ijt} + \zeta_{ijt}$. I assume consumers have incomplete information about each products consumption utility, though they know the contents of their choice set. Consumers know δ_{ijt} costlessly, but not ζ_{ijt}. ζ is mean-zero. Consumers can expend cost c_{it} (with inverse $\mu_{it} = 1/c_{it}$) to ascertain information about alternatives.

I assume consumers are rationally inattentive, as in Joo (2021), Matêjka and McKay (2015), Sims (2003), and Fosgerau et al. (2020). Before making a discrete product choice each period, consumers acquire some but not all information about products. The premise of Rational Inattenion (RI) models is that uninformed consumers optimally allocate their attention to better understand the attributes of their choice alternatives. I choose this framework for consumer information acquisition (as opposed to other models of consumer search) because of its flexibility and tractability in large choice sets. Attention is not binary - consumers may partially attend to multiple products, and attend to others not at all. The assumption driving the empirical predictions of the RI model is that consumers gather the information that will provide the most expected improvement in consumption utility. For example, if the consumer is trying to choose between two products, they will focus their attention on learning about product attributes that will allow them to distinguish which is most preferred.

Rationally inattentive consumers making a discrete choice proceed in three steps each period:

1. **Belief Formation**

   Consumers, given their purchase histories, the assortment size, and promotional activity, form subjective prior expectations about consumption utility for all restaurants.

2. **Information Acquisition**

   Consumers, given prior beliefs about each restaurant and the cost of acquiring information, select an (unobserved) information acquisition strategy. Consumers follow their information acquisition strategy and update their beliefs about restaurant consumption utility.

3. **Product Choice**

   Consumers select the highest expected utility alternative based on their posterior beliefs about consumption utility. Randomness in choice comes from uncertainty on the part of the consumer, rather than an unobserved restaurant-specific demand shock.

I assume that the cost of gathering information about products is proportional to how much the information reduces uncertainty in consumers' beliefs.[14] Appendix A1 details further assumptions and model notation. In the version of the model used in this paper, as in Joo (2021), product attributes may be fixed over time, and uncertainty arises because consumers' expectations are subjective. Let the unconditional probability of purchase be $\pi(\delta_{ijt})$[15], and $\mathbf{u}_{it}$ be the vector of consumption utilities. The probability that consumer i chooses restaurant j in period t is:

$$P_{ijt}(\mathbf{u}_{it}) = \frac{\pi(\delta_{ijt})exp(\mu_{it}u_{ijt})}{1 + \sum_{k \in S_{it}} \pi(\delta_{ikt})exp(\mu_{it}u_{ikt})} \tag{1}$$

The basic Rational Inattention framework for discrete choice does not necessarily encompass the effects of asssortment size document above. Adding a new product to the choice set, as shown in Joo (2021), increases the probability of purchase but may decrease consumer welfare. I need the model to be flexible beyond this. In particular, while adding products to the assortment may add

[14]In particular, I assume that the information cost proportional to the reduction in Shannon entropy between the unconditional purchase probabilities and the conditional choice probabilities after search.

[15]This unconditional purchase probability integrates over possible realizations of consumption utility given the consumer's prior beliefs about utility.

new characteristics, they also may alter consumers' unconditional probability of purchase ($\pi(\delta)$) via expectations about choice-specific match value. Additionally, larger assortments may alter the cost of accessing product information. In either case, this may alter whether consumers make a purchase and which products they choose to purchase. I assume that the assortment does not directly alter consumption utility of products.

These two channels (expectations and information costs) by which assortment size impacts consumers' information acquisition and product choice have distinct predictions for consumer behavior. If the cost of searching products increases with the size of the choice set, ceteris paribus, consumers will become less sensitive to the hidden portion of product utility. This could be consistent with consumers making higher-price or lower-quality selections, conditional on purchase. The impact of assortment size on consumers' expectations is less straightforward. If the restaurant equilibrium was modelled, we could recover how average match value might change with the size of the assortment. Absent that, however, assortment growth that alters expectations does not change how sensitive consumers are to the post-search characteristics, ceteris paribus.

These channels by which the contextual information about the size of the choice set impacts choice—prior beliefs about value, and information costs—reflect existing theoretical explanations for choice overload effects. Kamenica (2008) and Kuksov and Villas-Boas (2010) provide two accounts for how larger numbers of products reduce the probability of choice. In both cases, a larger number of products implies that the consumer will be worse off in expectation from consumption, either because the match value of the product is worse, or because the expected information costs outweigh the benefits from a better match value. The average search cost of each product could be altered by the total number of products due to more total search results from querying or due to higher information processing costs in the presence of more products. First, consumers have to sift through more search results, on average, to reach any product; search costs increase considerably with search result position (Ursu, 2018). Second, the cost of searching more intensively (paying more attention cost) may be altered by the amount of information displayed in total (Chandon et al., 2009; Gu, 2016).

4.1.1 Impact of Assortment Expansion

Previous empirical implementations of rational inattention do not accommodate my primary empirical finding—that the probability of purchasing any product declines in the size of the choice set. Given the assumption on information costs, the logit-like choice probabilities suffer from the same problem as a standard multinomial logit. Holding price and other attributes fixed, the addition of an alternative weakly improves the probability of selecting an inside option. Unless the size of the assortment directly enters into either $\pi_{i,k}$, μ_i, or $u_{i,k}$, the size of the assortment can only increase the probability of purchase.

To address this, I assume that the cost of information and beliefs about quality are impacted by the size of the choice set. Allowing $\pi(\delta_{ijt}) = \pi(\delta_{ijt}, |S_{it}|)$ allows for the unconditional probability of choice to be altered by the size of the assortment. Letting $\mu_{it} = \mu_{it}(|S_{it}|)$ allows the size of the assortment to impact all alternatives at once, including incumbent and previously purchased ones.

4.1.2 Comparison to Full Information Discrete Choice Demand

The previous equation bears a close resemblance to the choice probabilities in a multinomial logit model of demand, but they differ in two important ways. First, this model does not have unobserved taste shocks ϵ; its randomness in choice comes from consumers' incomplete information. This is beneficial to my setting where the size of the choice set is very large, because it partially addresses the problem highlighted in Ackerberg and Rysman (2005). In this model, new products do not introduce a new, equally valued attribute that differentiates the new product. Additionally, given the size of the choice set, it is unreasonable to assume consumers have full information about hundreds of products. Moreover, the reduced form findings—that purchase frequency declines in the size of the choice set —cannot be rationalized by a full information model of consumer choice.

Second, the form of this model is equivalent in prediction of behavior to a particular parameterization of a Logit model of demand. Existing logit models (e.g., Draganska and Jain (2005); Ershov (2018)) include a congestion term that may account for this level shift, but they cannot account for the changes due to attention cost in sensitivity to consumption utility. My approach allows for a structural interpretation of this congestion term. As the cost of information rises, this model predicts consumers become less sensitive on average to price and quality information that requires attention or search. My approach provides a justification for the inclusion of the congestion term with a specific interpretation.

4.2 Implementation and Identification

I estimate this model on a subset of consumers[16] from the consumer panel. To identify separately the parameters of interest, I parameterize the model as follows:

$$u_{ijt} = 1 + x_{ijt}\beta_i$$

$$\mu_{it} \propto exp(w_{it}'\theta_i)$$

$$\pi(\delta_{ijt}) \propto exp(d_{ijt}'\gamma_i)$$

Fixing the intercept in utility u allows for the multiplicative identification with the inverse of the search cost μ. I normalize the outsize option to have $u_0 = 0$ and $d_0 = 0$. I allow for intercept level shifts through d and discrete x attributes (in this case, cuisine). I assume the cost of attention

[16] I restrict the panel to consumers who are never mobile in their choices. I also consider only consumers who purchase at most once per week, since this model takes discrete choice as a premise. Future work will examine robustness to these assumptions.

is positive. This parameterization yields choice probabilities:

$$P_{ijt} = \frac{exp(d'_{ijt}\gamma_i + exp(w'_{it}\theta_i)(1 + x'_{ijt}\beta_i))}{1 + \sum_{k \in S_{it}} exp(d'_{ikt}\gamma_i + exp(w'_{it}\theta_i)(1 + x'_{ikt}\beta_i))} \tag{2}$$

	Included Covariates
w_{it}	Assortment Size, Demographics (Income, Employment)
d_{ijt}	Assortment Size (separately for previously visited vs unvisited restaurants), # platform and restaurant visits, proxy for on-platform ads, # of Yelp reviews
x_{ijt}	Price (post-tax, delivery inclusive), Cuisine, Delivery Distance

I identify μ from consumer differential response over time to consumption utility ($x\beta$) relative to prior information shifters as a function of w - consumer demographics and the size of the choice set. Much of this variation is cross-sectional, similar to interacting consumer demographics with preference parameters. I identify the elements of γ which vary with the choice set size from differential changes to choice probabilities, conditional on changes to $\mu(1 + x'\beta)$. I identify β and the remainder of γ from cross-sectional and temporal changes to choice shares in response to x and d, holding the size of the choice set fixed. In the case of identifying price sensitivity, I rely mostly on across-restaurant variation in prices, though changes to municipal sales taxes allow me to use some variation over time in prices.

I take a simplified approach to allow for heterogeneity in consumer tastes. I use observable variation in consumer demographics, location, and average order frequency to segment consumers before estimation, and I assume consumers' tastes are uniform within a segment. This procedure is similar to Bonhomme and Manresa (2015) and Bonhomme et al. (2021), but there are no shared parameters across clusters. Within each cluster, I estimate the model using maximum likelihood. To address the uncertainty inherent in the clustering procedure and to allow for inference in counterfactuals, I sample a subset users from the panel 700 times to conduct inference. The results are detailed in Appendix A2.

I estimate three specifications that vary in terms of how the size of the choice set impacts consumers. I test a version of the model where the size of the choice set impacts δ alone, where it impacts μ alone, and where it impacts both. The size of the choice set does not significantly impact the information cost. The size of the choice set does significantly impact prior beliefs about utility for restaurants that the consumer has not tried in the past, though not for restaurants that they have consumed. Specification tests reject the version of the model where the size of the choice set enters the information cost alone, and the version of the model where $|S|$ effects both μ and δ. I explore counterfactual assortment strategies using the model where only consumer prior beliefs about products are a function of the choice set size.

Table 14: Elasticity Estimates

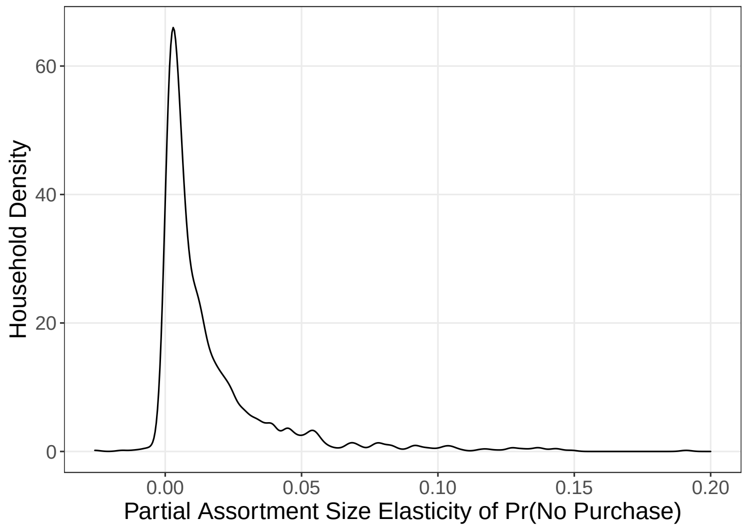
	Assortment Size on No-Purch	Own-Price	Restaurant Distance
Mean Household	0.02128	-2.703	-0.290
Median Household	0.00785	-2.531	-0.267
Variance of Means	0.00394	4.507	0.101

The main model effect of interest is how choice probabilities change with the addition of other alternatives. This partial elasticity is not equivalent to a full counterfactual of removing choices. Instead, it tells us how the probability of the no purchase changes as expectations and/or information costs adjust to the assortment size, holding the real size of the set and its contents fixed.[17] For most consumers, this elasticity is weakly positive. However, for many consumers this elasticity is very close to zero - there is heterogeneity in how users are negatively impacted by larger assortments. Figure 1 shows the distribution across consumers of this sensitivity. The average consumer has an partial-elasticity of 0.02, and the median consumer's partial-elasticity is 0.01. Consumers are sensitive to the assortment size, but the effect is very small. I report further summary statistics for assortment size, price, and restaurant distance elasticities in Table 14.

While the estimates suggest that consumers form expectations about their net returns to at-

[17] Specifically, I am interested in $\frac{\partial P_{i0t}}{\partial |S_{it}|} \frac{|S_{it}|}{P_{i0t}}$

Figure 1: Assortment Size Semi-Elasticity of No-Purchase



tention/search and consumption based on the size of this choice set, I don't explicitly model this connection. I include a linear term for the size of the choice set (interacted with restaurant experience), but future work should explore how to explicitly empirically model consumers' expectations of each product in a large assortment.

4.3 Counterfactual Assortment Reductions

Given the reduced form and structural evidence for a penalty to large choice sets, I conduct counterfactual restrictions of each individual's choice set. Unlike in most demand models (e.g. removing options in a logit demand system holding attributes fixed), the direction of these counterfactual results are not ex-ante clear. Whether reducing the size of the assortment benefits the probability of purchase depends two factors: how beneficial an option is in terms of consumption, and how large the penalty is relative to this benefit.

I use several targeting metrics to test assortment size reductions for each individual. I target which restaurants to remove based on the consumer choice model. I perform this targeting based on three model-driven metrics: expected revenue ($P_{ijt} * r_{ijt}$), choice probability (P_{ijt}), and consumption utility $u_{ijt} = 1 + x_{ijt}\beta_i$. The expected revenue metric reflects the platform's objective function, where a commission is earned on each sale. The expected revenue from an order varies based on the average basket size at the restaurant. I compare these targeting metrics to reducing the assortment size by removing alternatives at random.

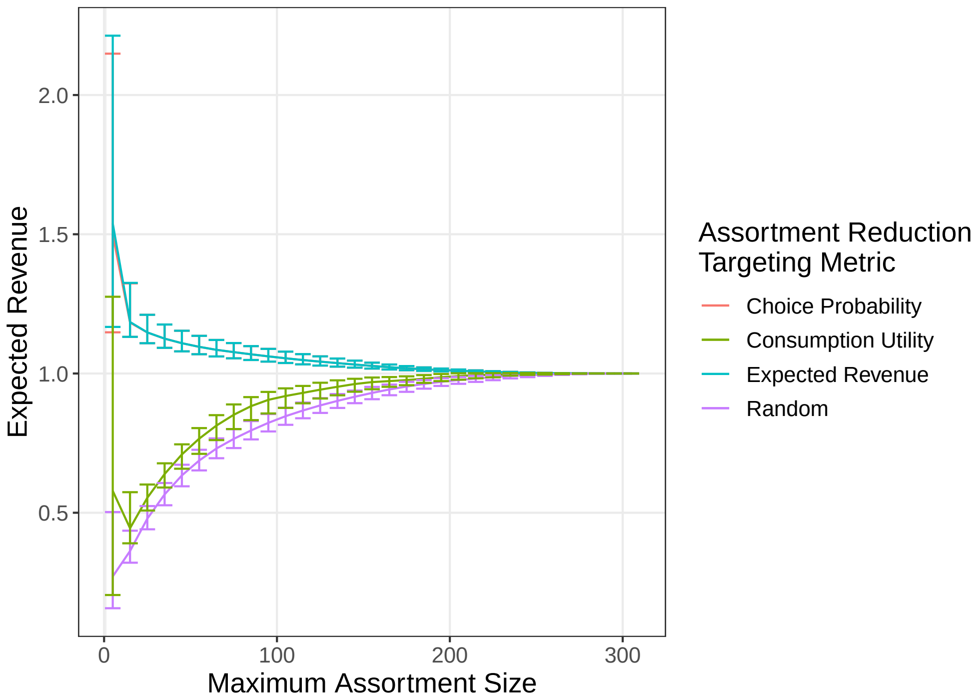
For each targeting metric, I simulate a series of assortment reductions for each individual. The counterfactual is evaluated at each assortment size as if all consumers cannot have more than the proposed number of restaurants, i.e. the assortment size is a cap. Consumers may still face smaller choice sets if their area has fewer restaurants supplied than the targeted size, but no restaurants are added.

In this model, removing choices alters the distribution of match values/consumption utilities and the consumer's prior choice probabilities. The first mechanic is captured by many demand models. If the removed restaurants were unlikely to be chosen, this may minimally impact their probability of any inside purchase. The second mechanic is what allows the model to improve the inside share with a smaller choice set. If the pre-attention choice probabilities decrease with the size of the choice set, consumers will be less likely to purchase any product.

Figure 2 plots the average normalized change in revenue by targeting metric for a sample date if all consumers face the same maximum assortment size. Reducing assortments improves the platform's revenue from existing consumers when the reductions are targeted using consumers' choice histories (quantity or revenue prediction targeting), but not when they are randomized or attribute targeted, shown in Table 15. This contrast highlights the interaction between the size of the assortment and its contents. Reducing the assortment can increase weekly revenues 55% if the reductions favor likely-to-be-chosen restaurants. However, this restriction is imposed uniformly on customers, which hides that the best-case assortment size could vary considerably across customers.

I contrast these platform-uniform restrictions with offering each individual consumer in the data

Figure 2: Counterfactual Revenues: Uniform Assortment Size Maximum



Note: This takes the average normalized revenue for the assortment size cap across subsampling iterations, error bars show 2.5 and 97.5 quantiles

Table 15: Uniform Assortment Size Revenues

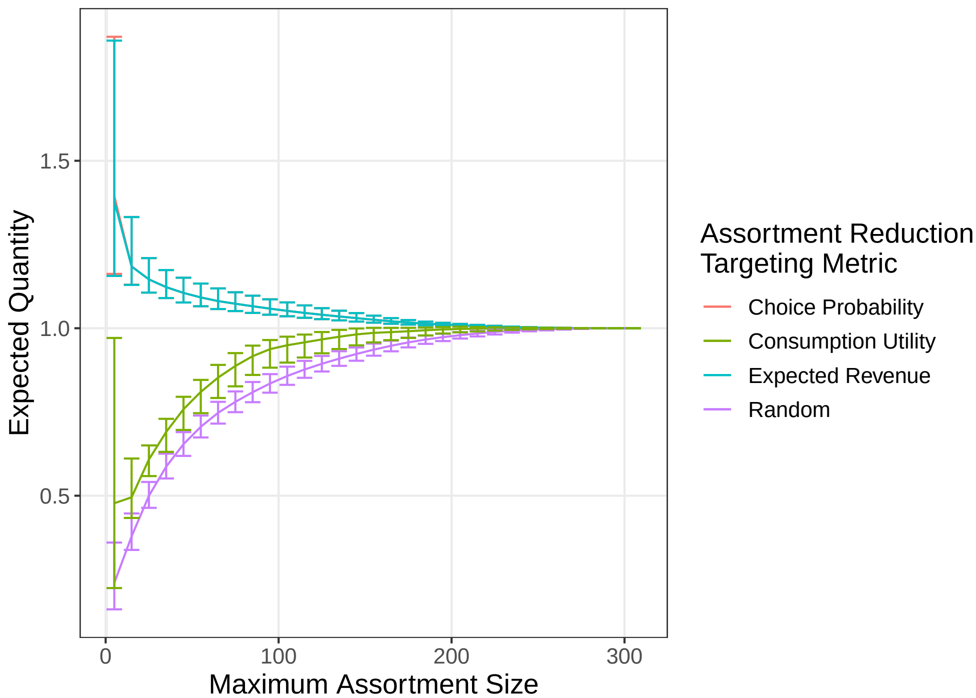
Targeting Metric	Mean Improvement	Median Improvement	2.5 Pctile	97.5 Pctile
Choice Probability	1.4996	1.2797	1.1479	2.1485
Random	1.0000	1.0000	1.0000	1.0000
Expected Revenue	1.5392	1.3335	1.1673	2.2131
Consumption Utility	1.0001	1.0003	0.9992	1.0005
	Results reported as a ratio with base of current assortment			

their own targeted maximum assortment size and contents. Targeted assortment sizes *and* contents improve weekly revenue by 55%. The lack of additional improvement occurs here because many targeted choice set sizes are the same as a uniform cap. The distribution of these assortment size maxima are plotted by targeting schema in Figure 4. Under this model, a minimal assortment size maximizes revenue or sales quantity for most consumers. However, if the firm is restricted to using only restaurant attributes to target choice set contents, using the individually targeted choice set size improves expected revenue significantly (versus in the uniform case, where it does not).

Table 16: Uniform Assortment Size Sales

Targeting Metric	Mean Improvement	Median Improvement	2.5 Pctile	97.5 Pctile
Choice Probability	1.3947	1.3213	1.1623	1.8701
Random	1.0000	1.0000	1.0000	1.0000
Expected Revenue	1.3797	1.2976	1.1561	1.8587
Consumption Utility	1.0001	1.0002	0.9994	1.0004
				Results reported as a ratio with base of current assortment

Figure 3: Counterfactual Sales: Uniform Assortment Size Maximum



Note: This takes the average normalized sales for the assortment size cap across subsampling iterations, error bars show 2.5 and 97.5 quantiles

Tables 16 and 18 show the expected increase in weekly order quantities, which in the both uniform and targeted assortment sizes increase 40% in the best case. I contrast this to revenue improvement (52 to 55%) to point out that most of the gain is getting consumers to order more often. However, the platform has an incentive to target the assortments towards higher-platform-margin restaurants when targeting is done by expected revenue. Targeting by choice probability (expected quantity) produces similar quantity improvements, but it does not capture as much platform revenue.

Figure 5 shows the distribution of the optimal assortment restriction to the individual's baseline assortment. Since my approach respects the supply of restaurants on the platform, I cannot rule out that some consumers would purchase more if supplied with larger assortments. For most consumers, this represents a considerable reduction in the realized choice - removing over 75% of the existing restaurants.

Figure 4: Distribution of Individual-Specific Assortment Size Caps

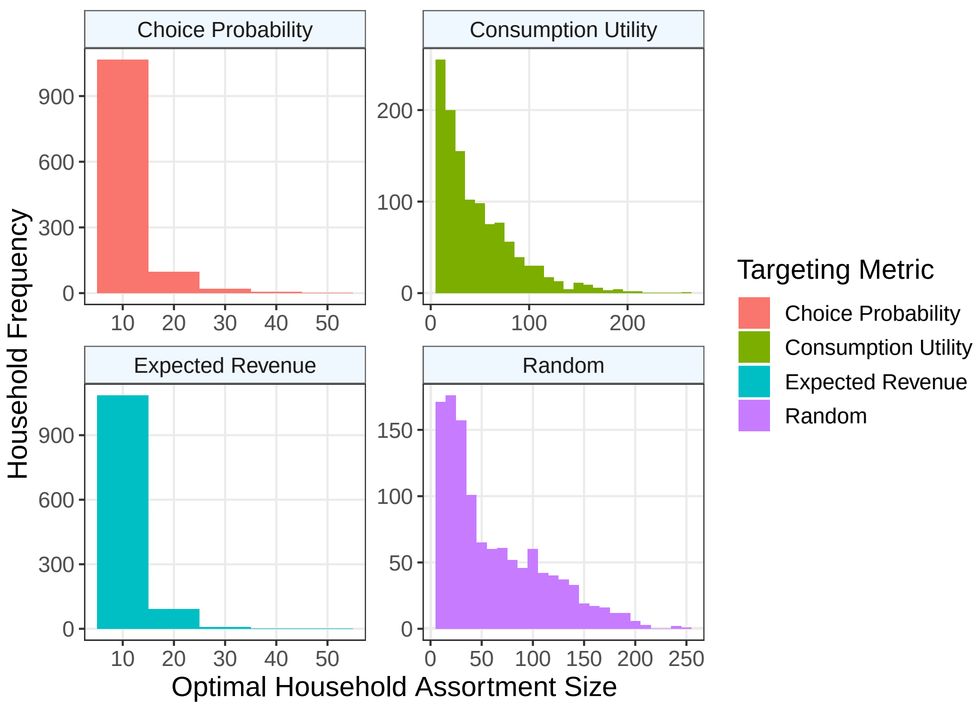


Table 17: Targeted Assortment Size Revenues

Targeting Metric	Mean Improvement	Median Improvement	2.5 Pctile	97.5 Pctile
Choice Probability	1.5235	1.3204	1.1621	2.2032
Random	1.0900	1.0571	1.0265	1.2642
Expected Revenue	1.5497	1.3553	1.1720	2.2417
Consumption Utility	1.3560	1.1738	1.0539	1.9947

Results reported as a ratio with base of current assortment

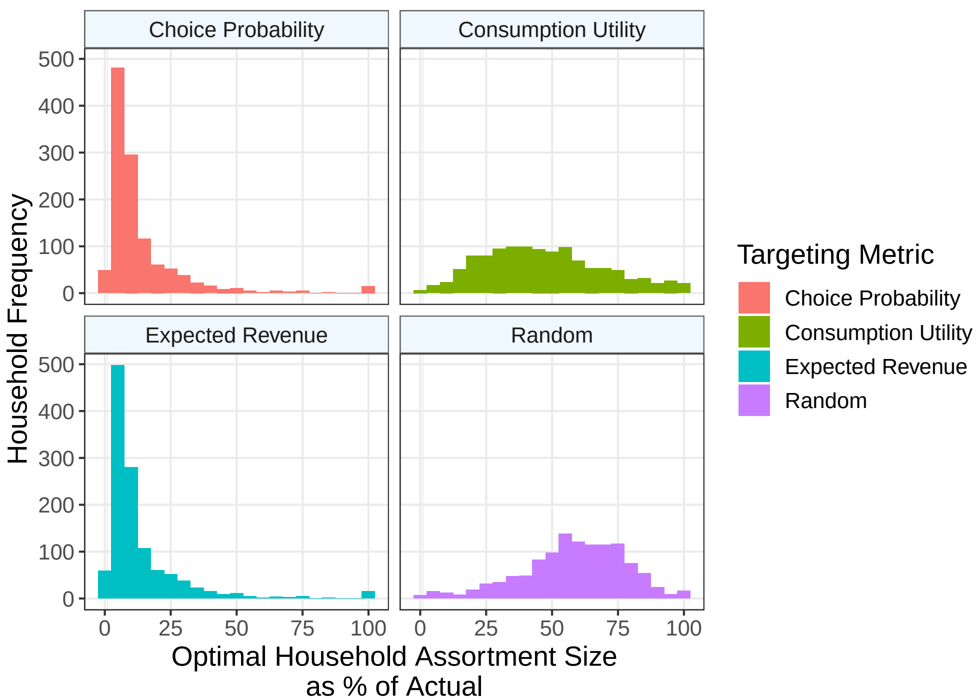
The optimal choice set sizes discussed above vary across consumer segments in terms of the flatness of the firms' objective function. For most consumers, removing alternatives improves purchase probability and revenue, but the shape is not as pronounced as in the aggregate total

Table 18: Targeted Assortment Size Sales

Targeting Metric	Mean Improvement	Median Improvement	2.5 Pctile	97.5 Pctile
Choice Probability	1.4039	1.3358	1.1647	1.8985
Random	1.0590	1.0492	1.0241	1.1390
Expected Revenue	1.3953	1.3234	1.1618	1.8892
Consumption Utility	1.2342	1.1651	1.0520	1.6706
	Results reported as a ratio with base of current assortment			

(Figure 4). In other words, the firm's objective function is fairly flat for some consumers, but not for others. Table 17 Future work will explore how well this out-of-sample prediction (large reductions in each individual's choice set) in experimental contexts.

Figure 5: Distribution of Individual-Specific Assortment Size Reductions



The structure of this counterfactual is a partial-equilibrium concept; it does not consider how restaurants might react to restricting individual choice sets. However, since the platform is restricting the size of the potential consumer base by pruning low-purchase-likelihood consumers, this should not be too detrimental to restaurants. Moreover, since the restriction improves the probability of purchase, this could net-benefit many restaurants. I don't explore the compositional effect on restaurants further, since this estimation exercise is focused only on a small subset of consumers in one area. Future tests should consider how restaurants might react.

The counterfactuals are stylized versions of potential improvements that can be made by platforms. Platforms may instead consider heavy personalization as an alternative (Donnelley et al., 2021). The platform does not need to prevent their existing customers from ever accessing all the restaurants that serve their location. In practice, platforms may choose to allow consumers to find any restaurant in their set if they search for it directly. The results of these counterfactuals suggest that there is scope for targeting assortment reductions to improve purchase frequency.

5 Conclusion and Future Work

In this paper, I document that 'more is not better' for some consumers - particularly, those who already have participated in a category. Unlike previous work, I show this effect in large, real-world assortments over the consumer lifetime. In online restaurant delivery markets, larger product assortments - more restaurants - drive increased consumer adoption of the platform, but lower the rate at which existing consumers order. This effect cannot be rationalized by most demand models; I use a model of attention allocation where the size of the choice set impacts consumers' beliefs and attention costs to show how intermediaries can reduce the assortment for each individual in a targeted manner.

The online restaurant delivery market provides an ideal lab for isolating the effect of assortment size, since assortments observably vary frequently and across individuals. However, by construction, choice in this market is always discrete - consumers only order from one restaurant at a time. The discreteness of the choice allows me to identify choice frictions from larger assortments among repeat customers, but the net negative effect may not generalize to basket situations. For example, in grocery retailing or other markets where the norm is baskets containing multiple categories and multiple products within category, the benefit of variety may outweigh its costs. This could explain differences between the findings in this paper and in related work in grocery retailing (Borle et al., 2005). Still, there are many markets where discrete choice is relevant, and choice frictions may dominate. In infrequent, large-ticket categories (computers, cars), these choice frictions from variety may be hard to measure, but findings in this work can shed light on the potential drawbacks to more variety in these markets.

Future work on this topic falls into three groups. First, the findings from this paper can be tested by platforms. In my counterfactual exercises, I find that using imperfectly targeted measures to reduce the choice set can still produce gains. Platforms' internal recommendations model can be tested as ways to reduce assortments. Second, the results here can be extended to include menu-item level analysis if such data were available. Third, more detailed consumer search data can be leveraged to better understand the exact mechanism for these choice frictions in large assortments. Exploring this issue further will help determine where, when, and how assortment reductions should be implemented.

References

Ackerberg, Daniel and Marc Rysman, "Unobserved Product Differentiation in Discrete-Choice Models: Estimating Price Elasticities and Welfare Effects," *RAND Journal of Economics*, 2005, pp. 771–788.

Aguiar, Luis, Joel Waldfogel, and Sarah Waldfogel, "Playlisting Favorites: Measuring Platform Bias in the Music Industry," *International Journal of Industrial Organization*, 2021, p. 102765.

Berry, Steven, Alon Eizenberg, and Joel Waldfogel, "Optimal product variety in radio markets," *RAND Journal of Economics*, 2016, *47* (3), 463–497.

Berry, Steven T. and Joel Waldfogel, "Free Entry and Social Inefficiency in Radio Broadcasting," *The RAND Journal of Economics*, 1999, *30* (3), 397.

Boatwright, Peter and Joseph C. Nunes, "Reducing assortment: An attribute-based approach," *Journal of Marketing*, 2001.

Bonhomme, Stéphane and Elena Manresa, "Grouped Patterns of Heterogeneity in Panel Data," *Econometrica*, 2015, *83* (3), 1147–1184.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa, "Discretizing unobserved heterogeneity," 2021. Working Paper.

Borle, Sharad, Peter Boatwright, Joseph B. Kadane, Joseph C. Nunes, and Galit Shmueli, "The effect of product assortment changes on customer retention," *Marketing Science*, 2005.

Briesch, Richard A., Pradeep K. Chintagunta, and Edward J. Fox, "How does assortment affect grocery store choice?," *Journal of Marketing Research*, 2009.

Broniarczyk, Susan M., Wayne D. Hoyer, and Leigh McAlister, "Consumers' Perceptions of the Assortment Offered in a Grocery Category: The Impact of Item Reduction," *Journal of Marketing Research*, 1998.

Bronnenberg, Bart J., "The provision of convenience and variety by the market," *RAND Journal of Economics*, 2015.

Callaway, Brantly and Pedro HC Sant'Anna, "Difference-in-differences with multiple time periods," *Journal of Econometrics*, 2020.

Caplin, Andrew and Mark Dean, "Revealed preference, rational inattention, and costly information acquisition," *American Economic Review*, 2015.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 2020, *110* (9), 2964–96.

Chandon, Pierre, J. Wesley Hutchinson, Eric T Bradlow, and Scott H Young, "Does In-Store Marketing Work? Effects of the Number and Position of Shelf Facings on Brand Attention and Evaluation at the Point of Purchase," *Journal of Marketing*, 2009, *73* (6), 1–17.

Chernev, Alexander, Ulf Böckenholt, and Joseph Goodman, "Choice overload: A conceptual review and meta-analysis," 2012.

Chu, Junhong and Puneet Manchanda, "Quantifying Cross and Direct Network Effects in Online Consumer-to-Consumer Platforms," *Marketing Science*, 2016, *35* (6), 870–893.

Dixit, Avinash K. and Joseph E. Stiglitz, "Monopolistic competition and optimum product diversity," *The American Economic Review*, 1977, *67*, 297–308.

Donnelley, Robert, Ayush Kanodia, and Ilya Morozov, "The Long Tail Effect of Personalized Rankings," *Working Paper*, 2021.

Draganska, Michaela and Dipak C. Jain, "Product-line length as a competitive tool," *Journal of Economics and Management Strategy*, 2005.

Drèze, Xavier, Stephen J. Hoch, and Mary E. Purk, "Shelf management and space elasticity," *Journal of Retailing*, 1994.

Eisend, Martin, "Shelf space elasticity: A meta-analysis," *Journal of Retailing*, 2014, *90* (2), 168–181.

Ershov, Daniel, "Competing with Superstars in the Mobile App Market," *SSRN Electronic Journal*, 2018.

Farronato, Chiara, Jessica Fong, and Andrey Fradkin, "Dog Eat Dog: Measuring Network Effects Using a Digital Platform Merger," 2020. NBER Working Paper.

Fosgerau, Mogens, Emerson Melo, Andre De Palma, and Matthew Shum, "Discrete choice and rational inattention: A general equivalence result," *International Economic Review*, 2020, *61* (4), 1569–1589.

Gu, Naiqing, "Consumer Online search with Partially Revealed Information." PhD dissertation, University of Chicago 2016.

Halaburda, Hanna, Mikołaj Jan Piskorski, and Pinar N. Yildirim, "Competing by restricting choice: The case of matching platforms," *Management Science*, 2018.

Hotelling, Harold, "Stability in Competition," *The Economic Journal*, 1929.

Illanes, Gastón and Sarah Moshary, "Market Structure and Product Assortment: Evidence from a Natural Experiment in Liquor Licensure," Working Paper 27016, National Bureau of Economic Research April 2020.

Iyengar, Sheena S. and Mark R. Lepper, "When choice is demotivating: Can one desire too much of a good thing?," *Journal of Personality and Social Psychology*, 2000.

Joo, Joonhwi, "Rational Inattention as an Empirical Framework: Application to the Welfare Effects of New Product Introduction," 2021. Working Paper.

Kahn, Barbara E. and Brian Wansink, "The Influence of Assortment Structure on Perceived Variety and Consumption Quantities," *Journal of Consumer Research*, 2004.

Kamenica, Emir, "Contextual inference in markets: On the informational content of product lines," *American Economic Review*, 2008.

Kim, Yewon, "Consumer Retention under Imperfect Information," 2021. Working Paper.

Kuksov, Dmitri and J. Miguel Villas-Boas, "When More Alternatives Lead to Less Choice," *Marketing Science*, 2010.

Lancaster, Kelvin, "Socially optimal product differentiation," *American Economic Review*, 1975.

_ , "The Economics of Product Variety: A Survey," *Marketing Science*, 1990.

Li, Jun and Serguei Netessine, "Higher market thickness reduces matching rate in online platforms: Evidence from a quasiexperiment," *Management Science*, 2020, *66* (1), 271–289.

Lin, Xiliang, "Disaggregate network effects on two-sided platforms." PhD dissertation, The University of Chicago 2017.

Matêjka, Filip and Alisdair McKay, "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *American Economic Review*, 2015.

Quan, Thomas W. and Kevin R. Williams, "Product variety, across-market demand heterogeneity, and the value of online retail," *RAND Journal of Economics*, 2018.

Reshef, Oren, "Smaller Slices of a Growing Pie: The Effects of Entry in Platform Markets," 2020. Working Paper.

Romano, Joseph P, Azeem M Shaikh et al., "On the uniform asymptotic validity of subsampling and the bootstrap," *The Annals of Statistics*, 2012, *40* (6), 2798–2822.

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

Sims, Christopher A, "Implications of rational inattention," *Journal of monetary Economics*, 2003, *50* (3), 665–690.

Sloot, Laurens M, Dennis Fok, and Peter C Verhoef, "The short-and long-term impact of an assortment reduction on category sales," *Journal of Marketing Research*, 2006, *43* (4), 536–548.

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 2020.

Ursu, Raluca M, "The power of rankings: Quantifying the effect of rankings on online consumer search and purchase decisions," *Marketing Science*, 2018, *37* (4), 530–552.

A1 Subjective Prior Rational Inattention Discrete Choice Framework of Joo (2021)

This appendix provides an overview of the assumptions required and derivation of choice probabilities in the Subjective Prior Rational Inattention Discrete Choice framework as developed in Joo (2021). I refer the reader to Joo (2021) for the complete model derivation. Model notation will follow Joo (2021) closely.

Each period, consumers are endowed with an information structure, where they perceive all alternatives as identical, and a known information cost function. The consumption utilities (u_{ijt}) are deterministic and fixed. This is in contrast to Matêjka and McKay (2015), where consumption utilities are stochastic, and the attention process uncovers signals about the current realizations.

In the first stage, consumers form subjective prior beliefs over the consumption utilities of each restaurant. Similar to Joo (2021), I allow these priors to be based on the size of the choice set, the consumers' purchase history, and promotional activity. In the second stage, consumer optimally selects which restaurant information to attend to, given the known cost of information and subjective prior beliefs. Finally, the consumer acquires this information, updates their expectations about restaurant consumption utilities in a Bayesian manner, and chooses the product given their posterior beliefs.

Let S_{it} denote the set of restaurants available to consumer i in week t on the platform, and 0 denote the outside option. Restaurant-specific consumption utilities $\mathbf{u_{it}}$ have degenerate support (with distribution Q^0)- the true consumption utility of each restaurant during that period is fixed. Consumers, however, are unaware of this exact fixed utility, and instead have prior beliefs over their perceived consumption utility from each product. Denote these perceived utility expectations as $\mathbf{v_{it}}$. The shifters of these perceptions are the length $|S_{it}|$ vector $\mathbf{D_{it}}$, such as advertising and choice history. The subjective prior beliefs held by consumers take a distribution $Q_i(\mathbf{v_{it}}) = Q(\mathbf{v_{it}}||\mathbf{S_{it}}|, \mathbf{D_{it}})$.

After the consumer acquires an optimal amount of product information (which is never complete - as I will discuss below, information costs are increasing in precision), they update their beliefs about restaurant specific consumption utilities. From the consumers' perspective, they have learned about "realized" $\mathbf{v_{it}}$ - subjective consumption utilities - when they are updating these beliefs. However, the "realization" of $\mathbf{v_{it}}$ is just $\mathbf{u_{it}}$ from the researchers perspective. So consumers' post-search conditional choice probabilities are conditioned on true restaurant consmption utilities: $Pr(i \text{ chooses } j \text{ in } t|\mathbf{v_{it}}) = Pr(i \text{ chooses } j \text{ in } t|\mathbf{u_{it}})$. Given these conditional choice probabilities, I can define the unconditonal choice probabilities (namely, unconditional on realizations of search) over the consumer's subjective prior distribution:

$$\pi_{ijt} = \int Pr(i \text{ chooses } j \text{ in } t|\mathbf{v_{it}})Q_i(d\mathbf{v_{it}})$$

Given these assumptions on consumers' information and information acquisition, I will define the information cost function. In particular, assume that the cost of information is proportional to the reduction in entropy between the conditional choice probabilities (post-search) and the unconditional choice probabilities. In particular, assume the cost of information for a particular set of posterior belief-based conditional choice probabilities is the reduction in entropy from $\pi_{\mathbf{it}}$ to $Pr(i \text{ chooses } j \text{ in } t|\cdot)_{j\in S_{it}}$. This function is convex: the more certain the posterior belief, the more expensive its associated information cost.[18]

Under these assumptions, from the researcher's perspective, the conditional choice probabilities (post-search) are of the form

$$Pr(i \text{ chooses } j \text{ in } t|\mathbf{u_{it}}) = \frac{exp(log(\pi_j) + \mu_{it}u_{ijt})}{1 + \sum_{k\in S_{it}} exp(log(\pi_k) + \mu_{it}u_{ikt})}$$

I refer the interested reader to detailed discussions and derivations in Joo (2021).

[18] As noted in Joo (2021), under the subjective prior RI model, the interpretation of information costs differs slightly from the rest of the RI theory literature. In particular, it should be thought of as"the cost of consumers' choice adjustments associated with changing the choice probabilities from unconditional choice probabilities... to conditional choice probabilities" (Joo (2021), page 17).

A2 Estimation Details and Results

To do inference jointly over the household grouping and group-level maximum likelihood estimation, I construct confidence intervals based on the variation across subsamples (Romano et al., 2012). For each iteration of the subsampling procedure, I drop twenty households from the panel, cluster the households using K-means clustering into 10 groups, and estimate the model using maximum likelihood within the group. This approach will not allow for cluster-level inference, but I will report variation across individuals (who are clustered differently across iterations) and across the entire sample. Clustering is based on a vector of census tract matched demographic variables (income, education, race, and employment variables) and the average order frequency of the user.

A2.1 Selected Structural Estimation Results

I reported the distribution of elasticity estimates for 3 measures of interest in Table 14: the effect of assortment size on no-purchase, own-price elasticity, and the distance elasticity of restaurant-consumer matches. I also report summary statistics for all parameters (excluding time effects) in Table A1.

Figure 6: Effect of Assortment Size on Consumer's Subjective Beliefs

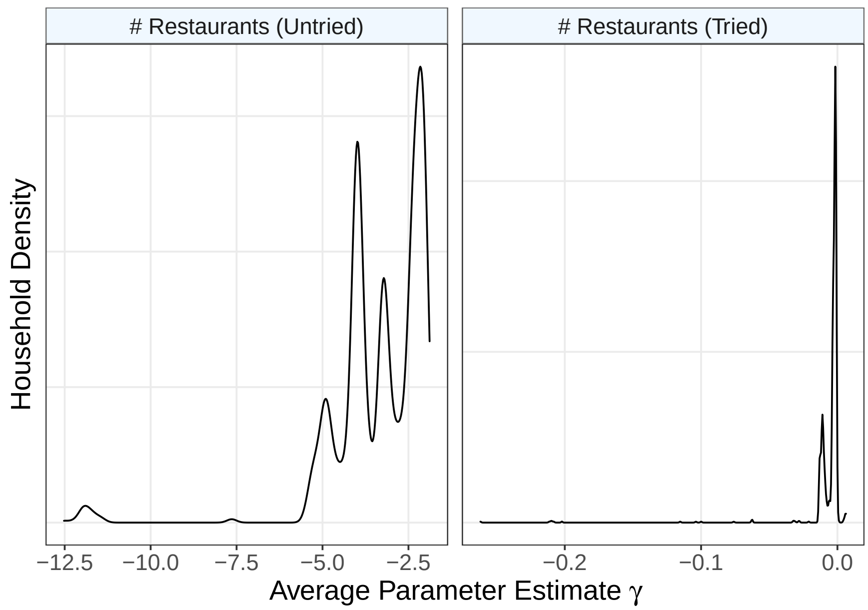


Figure 7: Own-Price Elasticity across Households

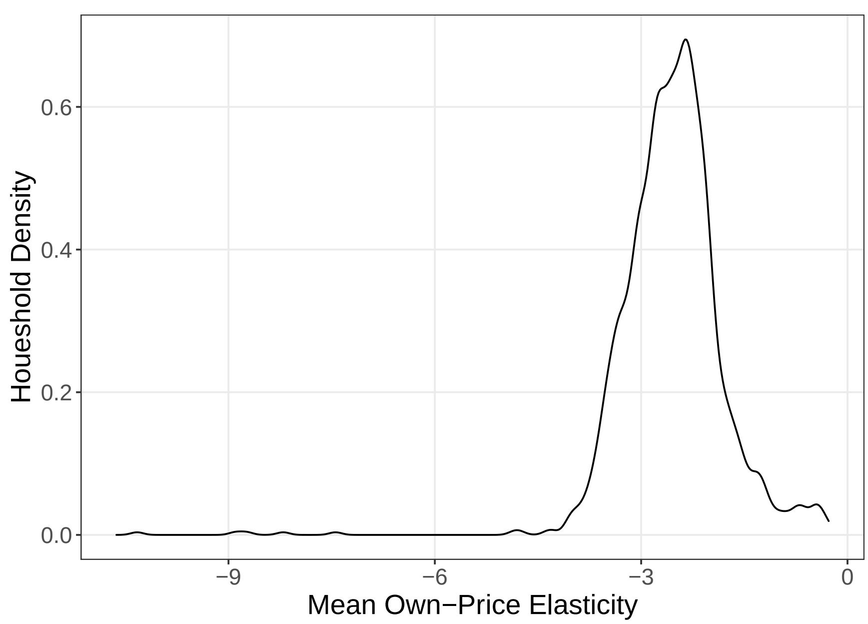
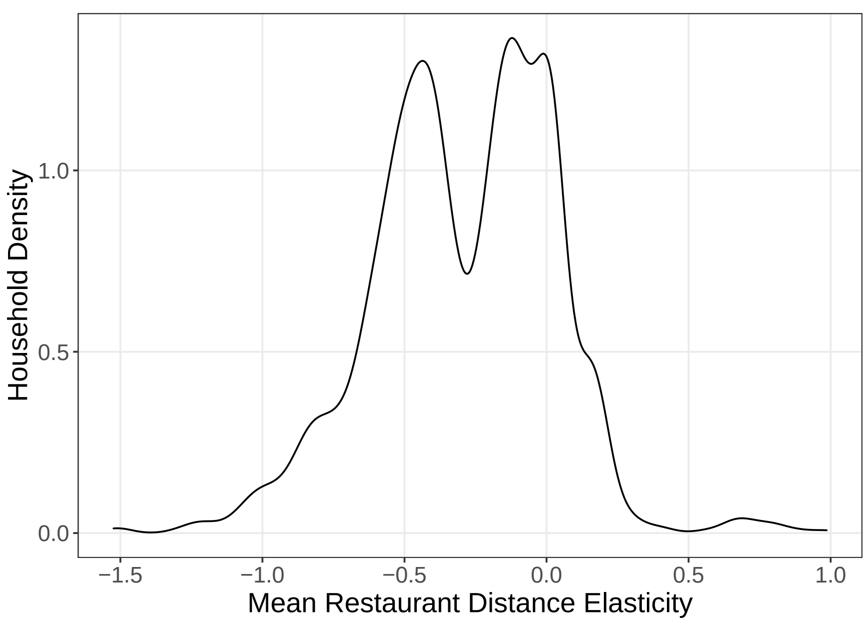


Table A1: Parameter Estimates

Variable	Mean	Median	Var	Q(2.5)	Q(97.5)	Parameter
American	-45.841	-1.039	39339787.0	-3.673	0.117	β
Asian	5.329	-0.408	531214.9	-14.241	1.422	β
Breakfast	14.799	-0.304	4247365.6	-2.795	1.983	β
Chinese	-15.453	-0.591	3544416.2	-3.539	0.646	β
Indian	0.430	0.215	1808034.8	-3.757	3.565	β
Italian	8.325	-0.902	36189740.5	-3.318	0.080	β
Japanese	-7.606	-0.521	479817.9	-4.384	3.115	β
Juice Bars Smoothies	0.874	0.497	22217.6	-33.990	48.497	β
Mediterranean	-7.065	-0.364	4363193.0	-9.683	3.227	β
Mexican	-3.940	-0.884	264884.4	-19.573	4.113	β
Noodles	-2.952	-0.671	379430.9	-3.855	0.987	β
Pizza	-5.287	-1.229	17195.9	-8.210	-0.037	β
Price (USD)	2.378	-0.091	124840.8	-0.176	-0.002	β
Rest Distance (km)	-18.611	-0.086	12485867.8	-0.523	0.236	β
Salads	2.795	0.012	20311.7	-11.794	19.147	β
Sandwiches	-41.684	-8.733	8553.9	-365.599	112.928	β
Prev Rest Orders	0.078	0.075	0.4	0.036	0.112	γ
Prev Orders	0.263	0.133	40.9	0.025	0.371	γ
Platform Ad Proxy	-0.053	-0.038	0.6	-0.127	-0.004	γ
Yelp Review Count	-0.000	-0.000	0.0	-0.000	-0.000	γ
Size New	-9.757	-1.623	1942107.7	-15.619	-0.349	γ
Size Return	-0.006	-0.002	0.0	-0.015	0.000	γ
Ordered Within 2 Mos	-29.562	-2.895	4832285.2	-55.056	1.518	μ
Unemployment Rate	-0.412	-0.013	8.2	-3.805	0.030	μ

Figure 8: Restaurant Distance Elasticity across Households



OA3 Data Summary

Table OA2: Summary Statistics: Consumer Panel

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)
Orders	19.84	31.76	2	3	8	22
Restaurants Tried	7.14	8.63	1	2	4	9
Spending	519.38	885.21	14.70	77.90	193.87	553.30

Figure 9: Census Tract Panel Choice Set Sizes

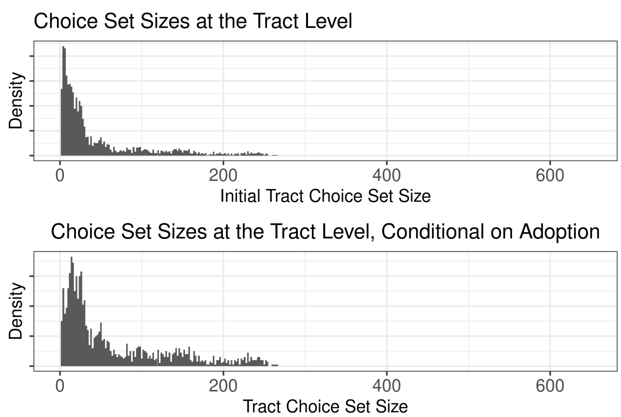


Figure 10: Individual Panel Choice Set Sizes at Start

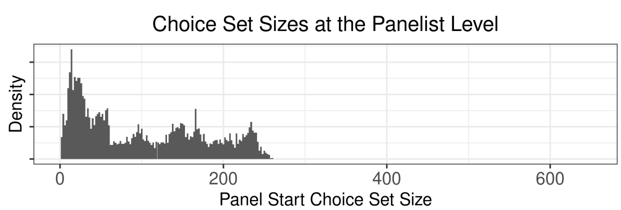
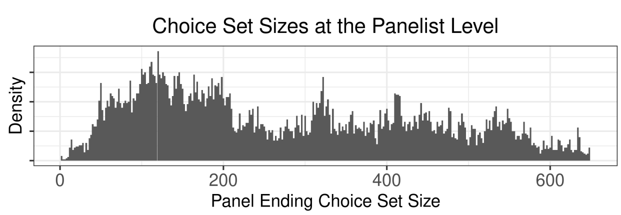


Figure 11: Individual Panel Choice Set Sizes at End



OA4 ONLINE APPENDIX Within- and Across-Neighborhood Variation

One main premise of the generalized difference-in-differences identification strategy used here is that consumers who live in the same neighborhood are more comparable than those who live in different neighborhoods. I will show that neighborhoods (measured by ZCTA in this paper), explain a considerable share of variation in consumer demographics.

To test whether census tracts are more similar within or across neighborhoods, I conduct a descriptive exercise by regressing demographic variables on a vector of ZCTA fixed effects. I report in Table OA3 the R-squared and F-statistics from these regressions. ZCTA fixed effects explain a significant and large share of the demographic variation in census tracts.

Table OA3: Variation in Demographics Within ZCTA

	R-Squared	F-Statistic
Pop2015	0.364	4.055
Hh Med Inc	0.672	14.545
Hh Mean Inc	0.691	15.848
Hh 100 To 150	0.501	7.131
Hh 150 To 200	0.562	9.114
Hh Over 200	0.673	14.600
Perc Hs	0.764	22.897
Perc Bach	0.805	29.324
Perc Hs 18to24	0.306	3.124
Perc Bach 18to24	0.532	8.071
Perc Hs Over25	0.649	13.134
Perc Bach Over25	0.771	23.865
Perc Grad Over25	0.769	23.543
Participation Perc	0.449	5.770
Unemployment Rate	0.428	5.314
Pop Black	0.683	15.257
Pop Asian	0.688	15.637
Pop Hispanic	0.643	12.774
Perc Black	0.785	25.820
Perc Asian	0.767	23.355
Perc Hispanic	0.792	26.939

OA5 ONLINE APPENDIX Two Way Fixed Effects Robustness

Recent work has highlighted the possibility of estimation failures in identifying effects for staggered adoption difference-in-differences designs (Callaway and Sant'Anna, 2020; De Chaisemartin and d'Haultfoeuille, 2020; Sun and Abraham, 2020). In particular, guaranteeing estimation of the desired static causal treatment parameter using two-way fixed effects estimation relies on further assumptions around treatment homogeneity, no anticipation of treatment, and treatment dynamics. While numerous estimators have been proposed to address these concerns, none so far can handle the large number of units and the continuous treatment (assortment size) I use in this project. I propose three broad sets of solutions to test the robustness of my estimated effects.

1. Where feasible, test for unobserved heterogeneity in treatment effects (De Chaisemartin and d'Haultfoeuille, 2020)

2. Construct difference-in-difference estimates for each entry experiment without staggered timing

3. Saturate model with observed heterogeneity

In the case of (1), I test the robustness of the adoption specifications (e.g. Table OA4) to unobserved treatment heterogeneity and negative weighting. I find that it is possible that the positive effect of assortment size on adoption is actually negative under sufficient treatment heterogeneity, or that the true effect of assortment size is on average 0 but has positive variance.

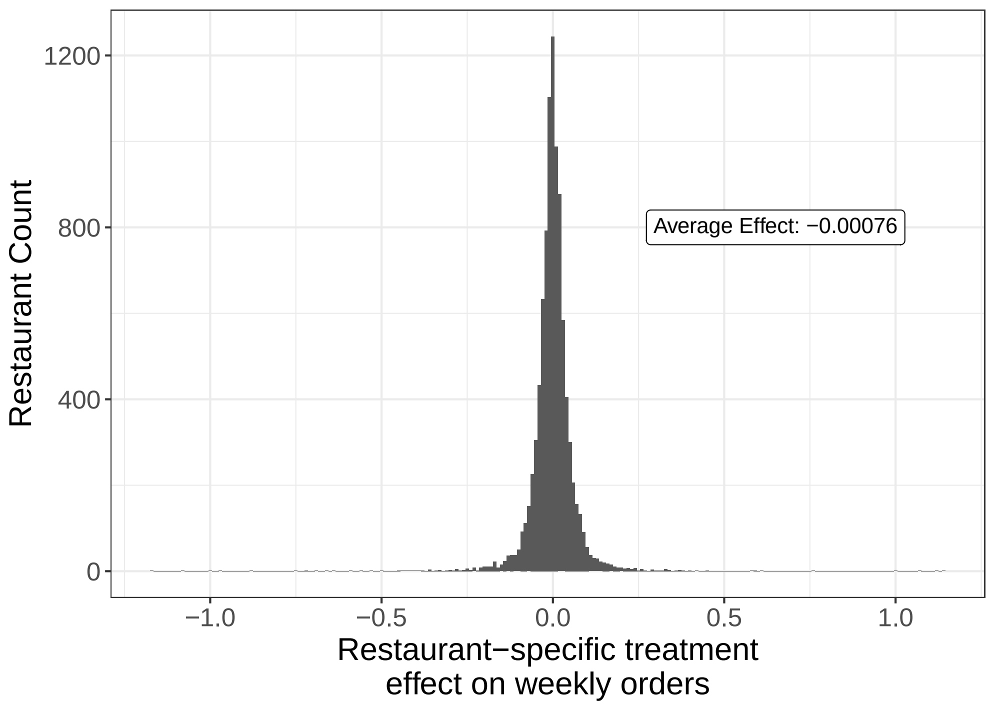
The size of my data prohibits me from considering the prior procedure for the effect of additional variety on returning customers. Instead, I will consider (2) and (3) as alternatives. (2) decomposes the continuous treatment of assortment size into its binary parts: specific restaurant entry and exit. I will estimate the marginal effect of each restaurant's entry on returning consumers using only two periods. The entry timing of each restaurant is uniform, so this design will avoid using any staggered treatment timing for each restaurant-specific event study design. However, by construction, I will estimate simple two-period difference-in-difference estimates for the contemporaneous effect of single restaurant entry onto the platform. In particular, I estimate:

$$y_{it} = \alpha_i + \alpha_{z(i)t} + \beta_j D_{ijt} + \epsilon_{it}$$

For restaurant entries j (with treatment dummy D_{ijt}), using only the immediate pre- and post-entry weeks for estimation and neighborhood-week controls $\alpha_{z(i)t}$. I expect these estimates to be very noisy. If the true average effect is similar to what I report in the body of the paper, these designs are underpowered (in particular since the treated group is often only several hundred observations. I show the mean and distribution of these effects in Figures 12 and 13 for outcome measures weekly orders and weekly spending.

The average effect across restaurants is consistent with the homogeneous effect estimated in the main specifications - restaurant entry reduces the probability of purchase and the average level of spending. However, because the effect varies so much across restaurants, I can't rule out that these effects are consistent with a zero-mean process.

Figure 12: Restaurant-specific Effect on Orders



Finally, while I can't directly address the possibility of unobserved treatment heterogeneity, the average effect of assortment size on consumption is mostly consistent (i.e. negative) across many observable sources of heterogeneity. For example, the marginal effect is consistently negative across assortment size, across consumers with different choice histories (excepting consumers who never vary their consumption), and across restaurants of different qualities.

Additionally, I allow for neighborhood-specific effects of assortment size. This may capture the relative pain of choosing online relative to growth of offline choices. Similar to the restaurant-specific effects shown above, the effect at the neighborhood (ZCTA) level is similar on average to the main effects, but it varies widely across neighborhoods. Figure 14 plots the distribution of effects. This variability is not related on average to the relative number of panelists in the area (and thus sample size) or to the average total restaurant online availability in the area.

Figure 13: Restaurant-specific Effect on Spending

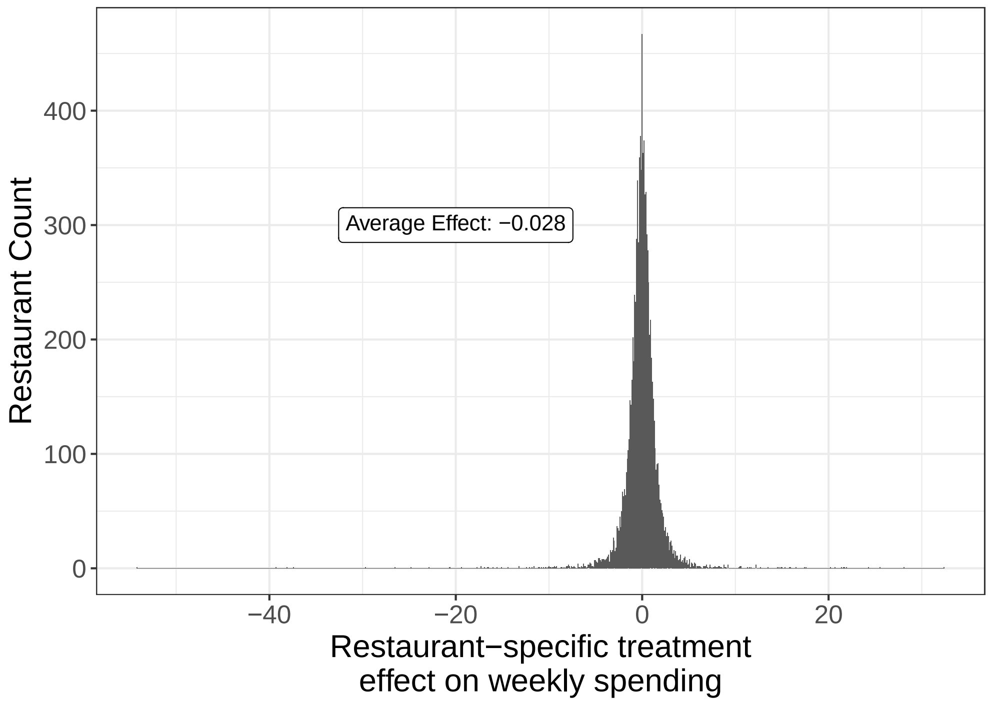
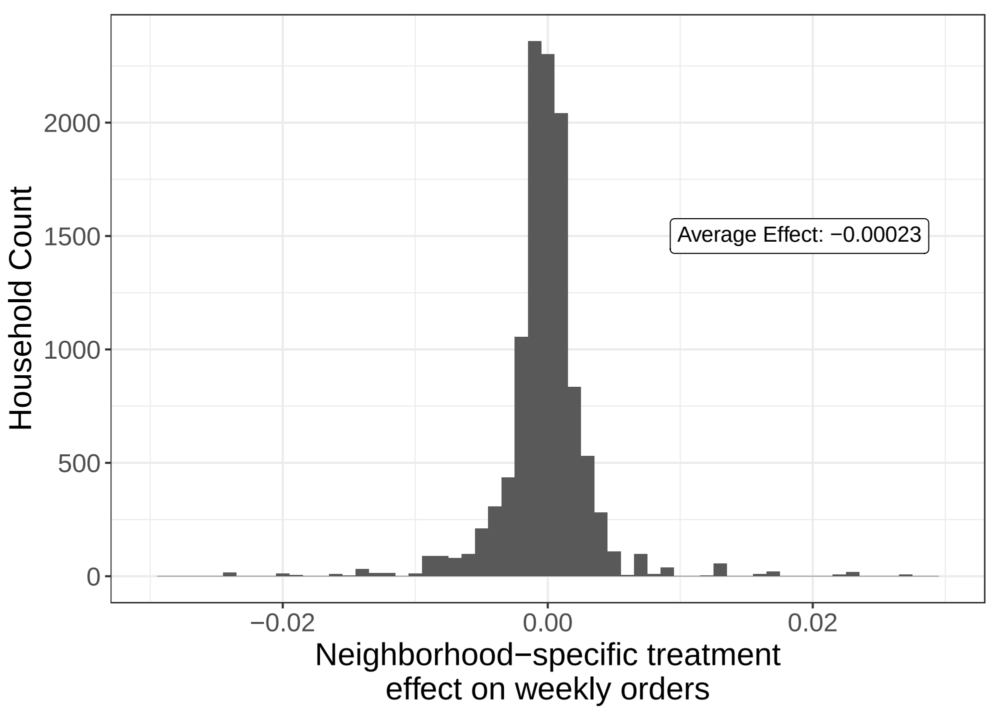


Figure 14: Neighborhood-specific Effect on Orders



I conclude from these robustness checks that I can rule out large sign reversals from my estimation techniques. However, the average negative impact of assortment size on purchase frequency may be misleading - the two way fixed effects estimation technique may obscure a true effect that is very close to zero.

OA6 Additional Tables

OA6.1 Additional Results

Table OA4: Impact on Adoption Rate at Census Tract Level

	Dependent variable:	
	Adoption Rate	
	(1)	(2)
Restaurant Count	0.0002***	0.0002
	(0.00002)	(0.0001)
Count Elasticity	0.4091	0.3182
ZCTA-Week FE?	N	Y
Observations	204,798	204,798
R^2	0.718	0.776
Adjusted R^2	0.713	0.740
Note:		*p<0.05; **p<0.01; ***p<0.001
		All Specs include Week and Tract FEs
		Standard Errors are clustered at tract level

To control for selection into search, I consider controlling for the residuals from a first stage linear probability model for whether the consumer engages in any search during the week. As shown above, controlling for selection in this manner does not alter the qualitative results for conditional search behavior.

Table OA5: Impact on Adoption Rate at Census Tract Level With Controls

	Dependent variable:		
	Adoption Rate		
	(1)	(2)	(3)
Restaurant Count	0.0002***	0.0003***	0.0002***
	(0.00002)	(0.00002)	(0.00001)
Cuisine Count		−0.0004*	
		(0.0002)	
Cuisine Entropy			−0.0003
			(0.001)
Count Elasticity	0.4091	0.4286	0.4081
Observations	204,798	204,798	204,798
R^2	0.718	0.718	0.718
Adjusted R^2	0.713	0.713	0.713
Note:	*p<0.05; **p<0.01; ***p<0.001		
	All Specs include Week and Tract FEs		
	Standard Errors are clustered at tract level		

Table OA6: Impact on Churn Rate at Census Tract Level

	Dependent variable:	
	Churn Rate	
	(1)	(2)
Restaurant Count	0.020**	−0.023
	(0.006)	(0.028)
Promotion Use Share	27.951***	29.356***
	(1.533)	(1.552)
Count Elasticity	0.03363	-0.03889
ZCTA-Week FE?	N	Y
Observations	104,218	104,218
R^2	0.118	0.331
Adjusted R^2	0.091	0.114
Note:	*p<0.05; **p<0.01; ***p<0.001	
	All Specs include Week and Tract FEs	
	Standard Errors are clustered at tract level	

Table OA7: Impact on Churn Rate at Census Tract Level With Controls

	Dependent variable:		
	Churn Rate		
	(1)	(2)	(3)
Restaurant Count	0.020**	0.027***	0.018**
	(0.006)	(0.007)	(0.006)
Cuisine Count		−0.270*	
		(0.121)	
Cuisine Entropy			−2.682*
			(1.341)
Promotion Use Share	27.951***	27.933***	27.912***
	(1.533)	(1.533)	(1.533)
Count Elasticity	0.03363	0.04492	0.02979
Observations	104,218	104,218	104,218
R^2	0.118	0.118	0.118
Adjusted R^2	0.091	0.091	0.091
Note:			*p<0.05; **p<0.01; ***p<0.001
			All Specs include Week and Tract FEs
			Standard Errors are clustered at tract level

Table OA8: Effect of Assortment Size on Weekly Spending by User Consumption Variety

	Dependent variable:	
	Weekly Spending	
	(1)	(2)
Rest Ct : High Variety User	−0.012***	−0.010**
	(0.002)	(0.003)
Rest Ct : Low Variety User	0.007***	0.009**
	(0.001)	(0.003)
Rest Ct : Medium Variety User	−0.002*	0.001
	(0.001)	(0.003)
ZCTA-Week FE?	N	Y
Observations	2,058,406	2,058,406
R^2	0.194	0.222
Adjusted R^2	0.190	0.191
Note:		*p<0.05; **p<0.01; ***p<0.001
		All specs include Individual and Week FEs
		Standard Errors clustered at individual level
		High and Low Variety Defined as Top and Bottom Quartile of Restaurants Ordered

Table OA9: Varying Marginal Effect of Assortment Size on Weekly Ordering

	Dependent variable:	
	Weekly Orders	
	(1)	(2)
1 to 140 Restaurants	0.00000	−0.0001
	(0.0001)	(0.0001)
140 to 278 Restaurants	−0.00004	−0.0001*
	(0.00005)	(0.0001)
278 to 417 Restaurants	−0.0001	−0.0001*
	(0.00004)	(0.0001)
417 to 555 Restaurants	−0.0001	−0.0001*
	(0.00003)	(0.0001)
555 to 700 Restaurants	−0.0001	−0.0001*
	(0.00004)	(0.0001)
ZCTA-Week FE?	N	Y
Observations	2,058,406	2,058,406
R^2	0.211	0.238
Adjusted R^2	0.207	0.212
Residual Std. Error	0.349 (df = 2046932)	0.348 (df = 1988201)
Note:		*p<0.05; **p<0.01; ***p<0.001
		All specs include Individual FEs
		Standard Errors clustered by individual

Table OA10: Impact of Relevant Restaurant Entry on Returning Customers

	Dependent variable:			
	Weekly Spending			
	(1)	(2)	(3)	(4)
Ever Consumed Cuisine Restaurants	−0.015***	−0.020***	−0.015***	−0.012**
	(0.002)	(0.003)	(0.002)	(0.004)
Never Consumed Cuisine Restaurants			0.004***	0.008*
			(0.001)	(0.003)
ZCTA-Week FE?	N	Y	N	Y
Observations	2,058,406	2,058,406	2,058,406	2,058,406
R^2	0.193	0.221	0.194	0.221
Adjusted R^2	0.189	0.190	0.189	0.190
Note:				*p<0.05; **p<0.01; ***p<0.001
				All specs include Individual and Week FEs
				Standared Errors clustered by individual

Table OA11: Effect of Assortment Size on Search Duration and Purchase given Any Search

	Dependent variable:							
	Weekly Search Duration Conditional on Search				Weekly Orders Conditional on Search			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Restaurant Count	-0.011^{*}	-0.006	0.020	0.008	0.0001	0.0001	-0.0005	-0.001
	(0.005)	(0.004)	(0.021)	(0.018)	(0.0002)	(0.0002)	(0.001)	(0.001)
ZCTA-Week FE?	N	N	Y	Y	N	N	Y	Y
Selection Controls?	N	Y	N	Y	N	Y	N	Y
Observations	135,414	135,414	135,414	135,414	135,414	135,414	135,414	135,414
R^2	0.261	0.322	0.488	0.530	0.297	0.298	0.521	0.522
Adjusted R^2	0.223	0.287	0.241	0.303	0.261	0.262	0.289	0.290

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include Week and Individual FEs
Omits search selection controls
Standard Errors clustered by individual

Table OA12: Effect of Assortment Size on Weekly Search Sessions

	Dependent variable:		
	Weekly Sessions		
	(1)	(2)	(3)
Restaurant Count	-0.0003^{***}	-0.0003^{***}	-0.0002^{***}
	(0.0001)	(0.0001)	(0.0001)
Cuisine Count		0.002^{*}	
		(0.001)	
Cuisine Entropy			0.038^{*}
			(0.015)
Count Elasticity	-0.3094	-0.3249	-0.2783
Observations	1,436,956	1,436,956	1,436,956
R^2	0.222	0.222	0.222
Adjusted R^2	0.217	0.217	0.217

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include Week and Individual FEs
Standard Errors clustered by individual

Table OA13: Effect of Assortment Size Conditional on Weekly Search

	Dependent variable:					
	Weekly Orders Given Search Sessions			Weekly Spend Given Search Sessions		
	(1)	(2)	(3)	(4)	(5)	(6)
Restaurant Count	0.0001	0.0001	0.0002	0.008	0.008	0.009
	(0.0002)	(0.0002)	(0.0002)	(0.006)	(0.006)	(0.006)
Cuisine Count		0.006			0.040	
		(0.004)			(0.133)	
Cuisine Entropy			0.023			1.067
			(0.053)			(1.585)
Residual from Search Indicator	0.090***	0.090***	0.090***	2.524***	2.524***	2.522***
	(0.012)	(0.012)	(0.012)	(0.373)	(0.373)	(0.373)
Count Elasticity	0.02409	0.01584	0.02665	0.05206	0.05005	0.05646
Observations	135,414	135,414	135,414	135,414	135,414	135,414
R^2	0.298	0.298	0.298	0.315	0.315	0.315
Adjusted R^2	0.262	0.262	0.262	0.280	0.280	0.280

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include Week and Individual FEs
Standard Errors clustered by individual

OA6.2 Robustness Checks

Table OA14: Event Study: Impact on Adoption Rate at Census Tract Level

	Dependent variable:			
	Change in Adoption Rate			
	(1)	(2)	(3)	(4)
Change in Restaurant Count	0.00004	0.001	−0.0001	0.001
	(0.0002)	(0.001)	(0.0003)	(0.001)
Change in Cuisine Count			0.001	−0.0003
			(0.001)	(0.002)
ZCTA-Week FE?	N	Y	N	Y
Observations	9,309	9,309	9,309	9,309
R^2	0.001	0.157	0.001	0.157
Adjusted R^2	0.0002	0.037	0.0002	0.037

Note: *p<0.05; **p<0.01; ***p<0.001
All Specifications have Week Controls
Standard Errors are clustered at tract level

Table OA15: Event Study: Impact on Churn Rate at Census Tract Level

	Dependent variable:			
	Change in Churn Rate			
	(1)	(2)	(3)	(4)
Change in Restaurant Count	−0.002	2.171*	0.371	2.197*
	(0.283)	(0.961)	(0.320)	(0.948)
Change in Cuisine Count			−3.249	2.651
			(1.864)	(3.984)
Change in Promo Usage	13.567*	11.547	13.451*	11.601
	(6.061)	(7.140)	(6.066)	(7.143)
ZCTA-Week FE?	N	Y	N	Y
Observations	3,270	3,270	3,270	3,270
R^2	0.002	0.319	0.003	0.319
Adjusted R^2	0.001	0.068	0.002	0.068
Note:	*p<0.05; **p<0.01; ***p<0.001			
	All Specs include Week FEs			
	Standard Errors are clustered at tract level			

Table OA16: Effect of Assortment Expansion due to Merger on Weekly Orders

	Dependent variable:					
	Difference in Weekly Orders					
	(1)	(2)	(3)	(4)	(5)	(6)
Change in Rest. Ct.	−0.0001	0.001	0.0001	0.001	−0.0001	0.002
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.002)
Change in Var Ct			−0.002	−0.002		
			(0.003)	(0.007)		
Change in Cuis Entropy					0.002	0.151
					(0.036)	(0.085)
ZCTA-Week FE?	N	Y	N	Y	N	Y
Count Elasticity	-0.1828	2.3467	0.1453	2.3467	-0.1943	3.1670
Observations	56,483	56,483	56,483	56,483	56,483	56,483
R^2	0.0004	0.027	0.0004	0.027	0.0004	0.027
Adjusted R^2	0.0003	−0.005	0.0003	−0.005	0.0003	−0.005
Note:	*p<0.05; **p<0.01; ***p<0.001					
	All Specifications have Week Controls					
	Standard Errors Clustered at Individual Level					

Table OA17: Comparison of New and Incumbent Restaurants

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)
Panel A: Incumbents						
Price	29.36	8.34	7.99	24.20	27.85	33.21
Delivery Fee	2.41	1.89	0.00	0.74	2.28	3.89
Yelp Rating	3.63	0.43	1.00	3.50	3.50	4.00
Yelp Review Count	776.35	748.32	1	184	403	1,340
Yelp Price Tier	1.63	0.51	1.00	1.00	2.00	2.00
Sales Quantile	0.66	0.32	0.00	0.49	0.78	0.91
Panel B: Entrants						
Price	26.88	9.57	8.28	20.29	25.35	31.55
Delivery Fee	3.77	1.26	0.00	3.36	3.99	4.47
Yelp Rating	3.52	0.69	1.00	3.50	3.50	4.00
Yelp Review Count	332.63	375.70	0	72	208	467
Yelp Price Tier	1.54	0.57	1.00	1.00	2.00	2.00
Sales Quantile	0.24	0.29	0	0	0	0.5

OA6.2.1 Monthly Results

Table OA18: Effect of Assortment Size on Monthly Orders

	Dependent variable:		
	Monthly Orders		
	(1)	(2)	(3)
Restaurant Count	−0.0004***	−0.0004***	−0.0003**
	(0.0001)	(0.0001)	(0.0001)
Cuisine Count		0.005*	
		(0.002)	
Cuisine Entropy			0.062*
			(0.030)
Count Elasticity	-0.1946	-0.2051	-0.1714
Observations	463,369	463,369	463,369
R^2	0.382	0.382	0.382
Adjusted R^2	0.366	0.366	0.366

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include Month and Individual FEs
Standard Errors clustered by individual

Table OA19: Local Effect of Assortment Size on Monthly Orders

	Dependent variable:		
	Monthly Orders		
	(1)	(2)	(3)
Restaurant Count	−0.0002	−0.0003	−0.0002
	(0.0005)	(0.0005)	(0.0005)
Cuisine Count		0.007	
		(0.004)	
Cuisine Entropy			−0.013
			(0.070)
Count Elasticity	-0.08581	-0.15066	-0.08697
Observations	463,369	463,369	463,369
R^2	0.401	0.401	0.401
Adjusted R^2	0.366	0.366	0.366

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include ZCTA-Month and Individual FEs
Standard Errors clustered by individual

Table OA20: Effect of Assortment Size on Monthly Search Sessions

	Dependent variable:		
	Monthly Sessions		
	(1)	(2)	(3)
Restaurant Count	-0.001^{***}	-0.001^{***}	-0.001^{***}
	(0.0002)	(0.0002)	(0.0002)
Cuisine Count		0.008	
		(0.004)	
Cuisine Entropy			0.134^{*}
			(0.066)
Count Elasticity	-0.2866	-0.2964	-0.2607
Observations	323,996	323,996	323,996
R^2	0.383	0.383	0.383
Adjusted R^2	0.367	0.367	0.367

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include Month and Individual FEs
Standard Errors clustered by individual

Table OA21: Effect of Assortment Size Conditional on Monthly Search

	Dependent variable:					
	Monthly Orders Given Search Sessions			Monthly Spend Given Search Sessions		
	(1)	(2)	(3)	(4)	(5)	(6)
Restaurant Count	0.0001	0.00002	0.0001	0.008	0.008	0.008
	(0.001)	(0.001)	(0.001)	(0.016)	(0.017)	(0.017)
Cuisine Count		0.010			0.143	
		(0.011)			(0.334)	
Cuisine Entropy			0.004			−0.591
			(0.151)			(4.321)
Count Elasticity	0.006897	0.001494	0.007166	0.029013	0.026089	0.027615
Observations	66,500	66,500	66,500	66,500	66,500	66,500
R^2	0.417	0.417	0.417	0.421	0.421	0.421
Adjusted R^2	0.358	0.358	0.358	0.362	0.362	0.362

Note: *p<0.05; **p<0.01; ***p<0.001
All specs include Month and Individual FEs
Standard Errors clustered by individual

OA6.3 Selection into Search Data

Only about three quarters of the consumer panel has any match in the search data. I restrict this match further to ensure that the searches accompanying purchase are present at least half of the time (they are otherwise inferred). The subset of users which have search data is not identical to those without - the summary statistics are presented in Table OA22. The means of these summary statistics can reject the null hypothesis that they are the same, but inspection of their magnitudes suggests that the samples are fairly similar.

Summary statistics of the search data are shown here. Conditional on having interaction with the platform, a user need not enter a search query to end up with a purchase (e.g. navigating from links presented on the home page such as previous purchases). The median user, however, does search via query. Desktop and mobile split total usage evenly in this sample. The typical session duration is quite short - only a few minutes are spent searching.

Table OA22: Comparison of Sample with and without Search

Statistic	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)
Panel A: Search Data						
Orders	18.40	28.24	2	3	8	21
Restaurants Tried	6.91	8.10	1	2	4	9
Spending	487.87	794.61	16.40	81.10	203.86	549.63
Initial Assortment Size	107.09	79.00	1	31	95.5	172
Panel B: No Search Data						
Orders	17.35	27.59	2	3	7	18
Restaurants Tried	6.60	7.98	1	2	4	8
Spending	434.91	744.19	14.70	70.02	159.95	457.88
Initial Assortment Size	97.17	77.57	1	26	68	162

Table OA23: Search Data Summary Statistics

Statistic	Mean	St. Dev.	Median
Weekly Sessions	0.329	1.362	0
Session Duration (minutes)	4.081	4.229	3.000
Search Queries	1.239	0.444	1.000
Desktop Share	0.530	0.451	0.500