Essays on Modeling Choices in Experiential Goods Categories

by

Gwen Geun Hae Ahn

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Business Administration)
in the University of Michigan
2023

Doctoral Committee:

Professor Fred M. Feinberg, Chair
Associate Professor Elizabeth E. Bruch
Professor Peter S. Fader, University of Pennsylvania
Associate Professor Jun Li
Associate Professor Eric M. Schwartz

Gwen Geun Hae Ahn

geunhae@umich.edu

ORCID iD: 0000-0001-5707-7277

Dedication

This dissertation is dedicated to my family.

Acknowledgements

As I reflect on this milestone achievement of getting a PhD, I realize how this journey wouldn't have been possible without the unwavering support of a loving and caring community. And for that I am deeply humbled and overwhelmed with gratitude. If karma indeed plays a role in life, I can't help but feel that I've been fortunate to have done some things right in the past.

First and foremost, I am deeply indebted to my advisor Fred Feinberg for the endless support he showed me throughout the seven years of the program. Despite his overflowing inbox, he always made time for me and patiently sat down with me to help me navigate through this long journey. I learned from him not just how to think, model, and write, but how to be a better person. I still remember getting an email from Fred as I accepted the PhD admission offer from the Ross School of Business, where he said, "we'll all work to ensure you come to see it as among the best of your life," and it indeed was one of the best decisions of my life.

I want to thank Eric Schwartz for encouraging me to fearlessly delve deeper into the unknown and helping me formulate my thoughts more clearly. His positive energy has been a constant source of inspiration. I am also grateful to Peter Fader, for opening up a fascinating research opportunity and patiently guiding me through the project. I deeply appreciate my committee members, Elizabeth Bruch and Jun Li, for generously sharing valuable feedback on my dissertation.

I would be remiss if I didn't express my heartfelt gratitude to both the marketing group at Ross and our PhD community for their invaluable support throughout my journey. I feel truly lucky to have been part of Michigan community where I knew that I was supported and loved. I want to thank Puneet Manchanda for always sharing a good joke and wisdom, Anocha Aribarg, for helping me understand my paper even better, Justin Huang, Yesim Orhun, Jessica Fong, and S. Sriram for sharing thoughtful feedback and advice along the journey. Carolyn Yoon and Scott Rick were the most supportive PhD coordinators one could ever hope for, and for that I am grateful. I was fortunate to have developed lifelong friendship with my PhD friends, especially Tim Doering and Jangwon Choi, with whom I shared a lot of laughter and tears together. It meant so much to share the journey with friends championing each other's success. I am also thankful to Prashant Rajaram, Longxiu Tian, Rebecca Chae, Steve Shaw, and Dana Turjeman for their help and support.

Finally, none of this pursuit would have been possible without the support from my family. I literally couldn't have been who I am today without the love and support of my parents. It meant so much to have grandparents who were always proud of me and always supported my decision. I wish all of them could have seen me finish the program. I cannot thank my sister So Yoon enough for being my trusted advisor, mentor, and lifelong trailblazer. Talking to her always made me feel like I could do this. And lastly and certainly not the least, my husband Heewoo has always been my secure base whenever I was in doubt, supported me through good and bad, and made it possible to appreciate life more in Ann Arbor and feel at home. I am forever grateful for his presence in my life.

Table of Contents

List of Tables

List of Figures

Abstract

My dissertation explores how firms can use individual longitudinal choice data to increase usage and spending in experiential goods categories, specifically in two different live events domains: art performances and sports events. In the first essay, I study how consumers choose customizable bundles of art performances while balancing preferences for the constituent items and within-bundle variety. Known as the diversification effect, consumers prefer higher levels of variety when making choices for multiple consumption occasions. This suggests that picking the right level of variety is a crucial part of art performance bundle choices. While customizability adds flexibility to consumers' choices, due to the large number of possible bundles, it could increase consumers' cognitive costs. Based on the proposed model, I make individualized recommendations that reflect heterogeneous preferences for not only the performances but also for the level of variety.

To model consumers' choices of art performance bundles, it is necessary to 1) characterize the performances, 2) define and measure the variety of prospective bundles, and 3) devise an efficient way of estimating bundle choices in the prohibitively large combinatorial space of possible bundles. Using natural language processing, I extract latent dimensions of the art performances that can be used to characterize the performances. I use these latent attributes to construct variety metrics that capture consumers' perception of bundle-level variety. Additionally, I devise a novel Monte Carlo approach to integrate over the space of unobserved order in which the bundle was assembled to tame the curse-of-dimensionality in the estimation process. I find that including variety metrics substantially improves predictive performance of the model, allowing the performing arts organization to make better individualized recommendations.

In the second essay, I study the cross-channel structure of the National Football League (NFL) ticket markets and consumers' purchase channel choices with the goal of devising optimal dynamic pricing and inventory policies across different channels. Professional sports teams have widely adopted dynamic pricing policy, which resulted in significant revenue improvements. At the same time, the growth of legal secondary markets has contributed to the development of a complex market structure with multiple channels. Understanding the cross-channel structure and consumers' channel choice process allows teams to make more informed pricing and inventory decisions.

Partnering with an anonymous NFL team, I collected time-series data on the availability and pricing of tickets on primary and three major secondary channels and combine it with transaction data. I propose a three-part model to understand the supply and demand dynamics: sellers' supply decision, buyers' purchase decision, and channel choice to capture the evolution of the choice environment where ticket availability and prices vary over time and channels. I find that there exist significant price differences across channels even after controlling for seat quality, and that channel choices reveal differential price sensitivities, effects of time-until-game across channels, and strong past dependence. Importantly, the investigation into row-level supply decisions reveal potential cross-channel effect of sudden increase in ticket availability on the primary channel due to an unexpected buyback from the brokers, opening a window to investigate the causal effect of supply changes across the channels.

Chapter 1 Leveraging Variety Preferences to Customize Bundles for Experiential Goods: An Application to Performing Arts Subscriptions

1.1 Abstract

Firms offering experiential goods like performing arts and meal kit services often allow consumers to customize the "bundle" they purchase, allowing enhanced flexibility and variety. This, however, comes at the cost of greater cognitive demand, which firms can mitigate through suitable recommendations. To this end, we propose a model that captures the unobserved sequential process by which consumers assemble their bundles, balancing utility for individual component items with "holistic" bundle-level variety. The model is applied to create-your-own performing arts bundles; to enable ready extension to future seasons, we use natural language processing to extract latent performance attributes, using them to construct and test a multitude of variety metrics to capture bundle-level variety. The model accommodates heterogeneous preferences for variety and performance attributes via a hierarchical Bayes approach. To tame curse-of-dimensionality, we devise a permutation-based Monte Carlo approach to integrate over the unobserved order of item addition to bundles. Results suggest that including variety metrics substantially improves the model's predictions for out-of-sample bundle choice, relative to accounting only for component item characteristics, by 34.5%. The model has wide applicability

for firms offering experiential goods by allowing them to leverage user histories to make superior customized bundle recommendations.

1.2 Introduction

Firms offering "experiential" products – concerts, sporting events, speaker series, opera – thrive when consumers engage on an ongoing basis. Rather than simply attending single events that appeal to them, some consumers opt to purchase a subscription package or "bundle". Various marketing tactics encourage this behavior, from price promotion, exclusive events "for subscribers only," or early access to shows through ticket bundles before they are sold separately. Traditionally, bundles of experiential goods, such as city passes and movie packs, were designed by projecting overall demand for curated assortments based on their individual component demands. Recently, firms are increasingly offering consumers the option to customize their own bundles, and the range of experiential goods offered for bundling are widening. For example, meal kit services such as Blue Apron allow subscribers to pick which 2 to 4 meals to include in their delivery each week. Many other retailers offer similar customizable bundle options for a wide range of experiential products, including wines (e.g., Winc), online games (e.g., Fanatical), digital design tools (e.g., tools for Photoshop at True Grit), and cosmetics (e.g., Ipsy).

Customizable bundles allow users to freely choose their component items and users can receive price discounts and other benefits. Despite some firm-determined restrictions (e.g., buy five or more items), this gives users flexibility. However, this flexibility does not come without cost, as cognitive demand grows with the space of options. Choosing just five tickets from a modest set of 50 performances entails "considering" over 2 million possible bundles; doing so for bundles of five out of the 81 home games in a major league baseball season offers over 25 million possibilities. Populating customizable bundles can be especially cognitively demanding when the sets of offered items can change from one "season" to another, as they typically do with experiential goods. Both consumers and the analysts modeling their choices must contend with a staggeringly large space of bundling possibilities.

A consumer faced with hundreds of potential concerts, but only able to attend a handful, may find the search-and-optimization task so daunting as to put it off, buy fewer tickets, or avoid it entirely (Iyengar & Lepper, 2000; Kuksov & Villas-Boas, 2010). As such, firms that offer customizable bundles of experiential products need to understand how users customize their own bundles to mimic that process through dedicated models and, by leading individual users to superior option bundles, lower their time and cognitive costs. Further, improved recommendations can enhance user satisfaction for their selected bundles, lowering attrition, and reducing costs (e.g., for returns or replacements of items not liked by consumers).

To that end, we propose a model that captures consumers' bundle choice processes for experiential products. We focus on an important feature of bundle choice that differentiates it from individual item choices: "holistic" (i.e., bundle-level) characteristics, particularly the variety of the overall set (Bradlow & Rao, 2000; Evers et al., 2014; Read & Loewenstein, 1995; Shaddy & Fishbach, 2017; Simonson, 1990). This is especially pronounced for experiential goods, where benefits are intangible, and many consumers intentionally seek out novel or varied experiences. Further, the fact that many experiential product categories may not even have well-defined, readily available attributes adds another layer of complexity, which we address by leveraging textual descriptions of experiential goods, distinguishing this research from the extant literature on variety and bundle design. Specifically, we explore the question of how consumers balance their preference(s) for a certain type of experience against desire for the new and different.

We apply our model in a large-scale performing arts industry setting. Performing arts organizations, including both large organizations such as New York Philharmonic and Carnegie Hall as well as smaller local ones, commonly offer customizable subscription packages for which each patron can choose performances to attend over the season. The subscription package should include a minimum number of performances to qualify for bundle benefits such as price discounts, early seat access, easy exchange policies, and preferential parking. Compared to the organization-designed bundles that are mostly based on pre-established "genres" (e.g., five jazz performances), these create-your-own bundles offer far greater flexibility, making them a popular alternative. Consumers assembling custom bundles face a trade-off intrinsic to many purchase situations: balancing their preferences for certain types of performances with their preference for variety across the set of performances. For instance, one may love solo violin concerts, but would one choose a sequence of five of them and nothing else? Instead of the fifth solo violin concert, perhaps one may choose a symphony orchestra performance for something different, but not too different. Our aim is to explicitly model such trade-offs to guide organizations in designing bundles to recommend to their customer base, selecting from a constrained set of candidate experiential goods that may lack pre-established objective attributes.

To this end, we need to model users' utility for within-bundle variety, over and above their utilities for the bundles' component items. Yet, despite the importance of variety in consumer choices and a vast literature on the topic, it is not obvious how this trade-off can be formally modeled. For example, consider a stylized example of choosing bundles of different cuisines for an ordered series of future consumption occasions: user A chooses {French, Polish, Italian, Spanish, German} dishes, and user B {French, Italian, Austrian, Thai, Chinese}. The first, and in some sense primary, challenge here is to "measure" how much variety there is in these two bundles before even considering variety's impact. One obvious metric that can be used to characterize variety is the number of different cuisines. Yet both users A and B chose five dishes from five separate cuisines, making their choices appear equally "varied" in that sense alone. However, user A chose only Western cuisines, while user B's set included both Eastern and Western ones, so in some sense A's bundle is less varied than B's, although this presumes there is more intrinsic variation across continents than within them. Further, it is possible that, depending on the specific dishes, some sets across cuisines – for example, five chicken dishes – may be "closer" to one another than a set of ingredient-varied dishes within the same cuisine type. We stress the challenge of creating attributes to describe items to aid in measuring variety, and note it is particularly important in harder-to-describe experiential goods.

A related, yet distinct, challenge is how any chosen measure of variety should be incorporated in making future bundle recommendations. While the details of the chosen dishes would be helpful, it is unclear both which dimensions of the dishes should be considered and how to make recommendations for future bundles based on a limited number of observed bundle choices in the past. A further challenge is to consider, for example, dishes or cuisines that have not appeared before in anyone's choice set. In the context of meal prep services, concert series, and other bundle-based services, such issues are exacerbated when the set from which one must choose is changing from season to season or even occasion to occasion.

These and other challenges arise in applying the model in our empirical context of performing arts bundles. Specifically, an initial challenge is characterizing the items (performances) using detailed attributes that can capture their complex nature and yet are readily extendable to previously unobserved items, e.g., artists and troupes who have never appeared before at that venue. Arts performances being complex experiential products renders off-the-shelf, structured attributes insufficient to fully characterize them. A key implementation issue is that it is not at all obvious what sorts of attributes best characterize performing arts performances, let alone bundles of them. While one may like solo violin performances, the identity (nationality, ethnicity, or gender) of the violinist, musical era, and specific repertoire may also play pivotal roles. Furthermore, lineups change season-to-season, with new works or performers making their debuts. Compared to the complexity of the items themselves, the paucity and generality of genre labels like "dance" or "orchestra" renders them useful only in the broadest sense: it is possible to love some choreographers or performers far more than others. Even were one to have strong general enthusiasm for a particular genre, there still exists substantial within-genre heterogeneity that can mask important differences among component items. For example, the "theatre" genre spans anything from Shakespeare's Julius Caesar to contemporary experimental pieces like Young Jean Lee's Untitled Feminist Show. Because different "theatre" performances can offer radically different experiences (and thereby attract distinct audiences), we leverage the rich descriptions afforded by widely distributed brochures – which are consumer-facing and available online at the time of choice – to extract latent, presumably less-standard and more holistic, performance attributes, while critically allowing direct extension to a new set or season of performances. These detailed descriptions transcend genres in a patron-relevant manner: Shakespeare and Beethoven might appeal to those interested in "classics" or "the canon," while Untitled Feminist Show and the pianist Yuja Wang may attract those who support women or Asian artists.

A second challenge, and one that to our knowledge is novel in this setting, is in measuring within-bundle "variety". Users, unlike when they evaluate attributes of items already curated in a bundle, may evaluate prospective bundles in an evolving or sequential manner, as they consider various potential combinations. It is unclear which dimension, if any, of the attributes or items users (at least those who value or dislike varied experiences) consider when evaluating variety implications of adding items to such bundles, let alone how "variety" itself should be measured across multiple (latent) product attribute dimensions. Various measures of variety based on genres have been suggested in the literature (e.g., Datta et al., 2018; Orhun et al., 2016), but again, genres often fail to capture the key details of arts performances. Hence, we turn to the extracted attributes from the performance descriptions that capture textured details that could be relevant to construct variety metrics aimed at capturing the evolving within-bundle variety.

Critically, we construct a broad spectrum of variety metrics, empirically testing which aspects of a set's items best capture how consumers (heterogeneously) incorporate bundle variety into their preferences and downstream purchases. We find that topic representations and variety metrics based on the extracted topics substantially improve model fit compared with genre information alone, and that including "holistic" bundle metrics (e.g., average distance across all pairs of performances) substantially improves predictive performance in a hold-out season.

Prohibitively large sets of possible bundles pose another challenge – this one purely computational – in modeling choices of customizable bundles. The unit of choice for both bundle choice models and multi-item choice models are often sets of items (Bradlow & Rao, 2000; Chung & Rao, 2003; Farquhar & Rao, 1976; Rao et al., 2018 for a review). However, due to the small number of observations and large number of alternatives (bundles), it is infeasible to estimate either bundle or multi-item choice as-is, i.e., a pick-1-of-1000000+ multinomial model. Rather, we approach the bundle choice problem as a latent sequential process, circumventing the infeasible multinomial problem in a manner practicable across seasons and extendable to future ones.

One final, and related, challenge concerns estimation and prediction for extremely large choice sets, particularly those whose order is latent, given the latent sequential process modeling framework. Modeling the choice process as sequential has two critical empirical implications: that the choice set from which one chooses from will change (i.e., shrink) across choice occasions as chosen items are removed from the choice set; and that items that remain in the choice set will be evaluated both in terms of their characteristics and the variety implications with regards to the items already in the chosen bundle. Despite these implications, we do not observe which performance was added "first" in consumers' minds, and it is possible that this is opaque even to consumers themselves. For example, one might purchase a bundle including the pianists Yuja Wang and Benjamin Grosvenor, but the analyst does not know that the former was considered indispensable, conditionally lowering the utility of adding the second for consumers who prize variety. We suggest a practical method for navigating and stochastically integrating over these extremely high-dimensional choice and prediction scenarios in a fully Bayesian setting.

The rest of the paper is organized as follows. In section 1.3, we briefly review relevant literature and discuss our contribution, and describe our empirical setting in section 1.4. Then, in section 1.5, the modeling framework is introduced, and the results are discussed in section 1.6. Lastly, in section 1.7, we discuss future directions.

1.3 Prior research and present contribution

1.3.1 Literature on Variety

Variety, proverbially called "the spice of life," has been studied extensively in the academic marketing literature. It is a critical contextual factor that has been shown to affect consumption decisions in multiple contexts, including in consumers' "variety-seeking" choice behaviors and assortment choices both for retailers and consumers.

1.3.1.1 Consumers' variety-seeking behavior

Consumers seek variety to increase the overall "utility" of their consumption across a series of choice occasions (Kahn, 1995, 1998). Specifically, in hedonic consumption, consumers find the experience more enjoyable if they perceive the experience to be more varied (Galak et al., 2009; Kahn & Ratner, 2005; Ratner et al., 1999), which specifically highlights the importance of accounting for variety in the context of experiential goods, the purview of this study. Researchers have also proposed that one's need for stimulation and desire to overcome satiation (e.g., McAlister, 1982) as intrinsic motivations that drive consumers' variety-seeking behavior. More recently, variety has been proposed as a goal in and of itself (see Kahn & Rafieian (2022) for a review).

The literature also addresses the various factors that affect variety-seeking behavior, including – and especially pertinently for the present study – the nature and format of the choice scenario. When consumers make choices from an assortment (i.e., choosing multiple items now for later consumption), they tend to opt for greater variety than when making those same choices at the time of consumption. This "diversification effect" (Simonson, 1990) has been verified in multiple settings (Galak et al., 2011; Read et al., 2001; Read & Loewenstein, 1995), including the latent process of consideration set formation (Salisbury & Feinberg, 2012).

Researchers have proposed various models to estimate the effects of variety using observed choices. In one stream, variety-seeking is modeled as a state-dependent behavior, wherein previous brand choices affect subsequent ones (Chintagunta, 1998; Feinberg et al., 1992; Givon, 1984; Lattin & McAlister, 1985; McAlister, 1982). For example, McAlister (1982) proposed dynamic satiation – an attribute's marginal utility waning over choice occasions where it is consumed – as a source of observed switching behavior. On the other hand, Lattin & McAlister (1985) and Feinberg et al., (1992) suggested that previous choices can affect and alter preferences for subsequent ones, whereas Givon (1984) postulated that consumers derive utility both from the chosen item and from the act of switching itself.

A key feature of these models is that they rely on a panel data structure, i.e., brand switching patterns are fully observed. But we cannot readily adapt such models to settings where "switching" is not clearly defined. For many experiential products, particularly for "live" events like performances or sports games, the idea of "switching" is not quite relevant. One can only rarely re-experience anything similar – even the same band might perform different songs – or, more to the point of our eventual application, a subsequent season may have little or no overlap in performers or works with prior ones. In addition, the temporal sequence of choosing each item is required for most variety-seeking models, since they operate by comparing each item to those chosen or consumed before or after. But the order in which items are added to consumers' customizable bundles is often unobserved by the researcher (and may not even be clear to the decision-maker), as is the case with many bundles of experiential goods. The lack of an observed temporal choice sequence presents conceptual and computational challenges to applying extant model formulations to capture variety-seeking for customizable bundle choice and downstream optimization.

1.3.1.2 Assortment variety: perceptions and choices

"Assortment variety" plays an important role in consumers' choices (for a review, see Chernev, 2011), ranging from store selection to "whether, what, and how many" they will buy (Chiang, 1991; Gupta, 1988). Given its critical role, researchers have extensively studied how consumers form perceptions of variety, as well as how it affects consumers' choices.

Research has verified that perceptions of variety in an assortment is driven by more than simple summaries like the number of SKUs offered by retailers. Broniarczyk et al., (1998) find that, in addition to the number of SKUs, the availability of a consumer's favorite item and the amount of shelf space dedicated to the category affect how consumers perceive the variety of assortments. For example, if shelf space is held constant and favorite items are still available, reducing the number of items does not negatively affect perceptions of the assortment. Hoch et al., (1999) and Van Herpen & Pieters (2002) focus on the role of the multi-attribute structure and spatial organization in consumers' perception of assortment variety and find that information structure (i.e., how much the alternatives differ on multiple attributes) plays a significant role in variety perception. Kahn & Wansink (2004) show the downstream effects of assortment structure: organization and symmetry in the assortment affect perceptions of variety, which in turn affect consumption quantities. Additional assortment characteristics, such as the number of categories (Mogilner et al., 2008), graphical elements of an assortment (Townsend & Kahn, 2012) and physical alignment of the assortment (Deng et al., 2016) are documented to affect consumers' perceptions of assortment variety.

Another stream of literature explores the role of assortment variety on consumer choices. While consumers value assortment variety to be among their three most important factors in choosing a physical store (Arnold et al., 1983; Louviere & Gaeth, 1987), offering larger assortments with more alternatives is not always better. In their celebrated "jam study," Iyengar & Lepper (2000) find that consumers faced with too much choice may either delay or opt out entirely, a general phenomenon known as choice overload (see Chernev et al., 2015 for a review; Kuksov & Villas-Boas, 2010). In a recent work, Natan (2022) finds that, while expanding assortment increases acquisition of new customers, it lowers the frequency of usage among existing customers (specifically, for a food delivery app), suggesting that the relationship between assortment variety and sales can be nuanced and multidimensional.

Despite providing insights into the role of variety in "choice" and the nature of perceptions of variety, these studies focus on product categories with clearly defined sets of attributes, such as colors, sizes, and flavors, mostly from CPG retail categories. This paper departs from that tradition by focusing on experiential products. And while experiential products can and do have some such structured attributes – venue, seating, event timing, etc. – the truly defining aspects that can capture an experience's "qualia" and complexity are difficult to capture, yet they are arguably more important to understanding choice (Cooper-Martin, 1991). The literature leaves unclear how the information structure of sets of such complex items can and should be incorporated into a model of bundle choice. Moreover, because research on assortment variety centers around static assortments, relatively little is known about how introducing additional novel items will affect perceived variety for subsets that include them.

More subtly, and perhaps importantly for practical modeling purposes, perceptions of variety are typically elicited from consumers (Hoch et al., 1999; Kahn & Wansink, 2004; Van Herpen & Pieters, 2002), an impracticable methodology for customizable bundles with hundreds of thousands of new possibilities every season. Further, the proxies of variety perception are often summary metrics of assortment characteristics, such as number of categories, and actual SKUs. However, customizable bundles have relatively similar sizes, and summaries of performing arts bundles, such as number of "genres," are not generally sufficient to capture the complex nature of experiences, singly or as bundles. The present challenge includes not only identifying relevant item (performance) attributes important for choice, but also quantifying perceived variety along and across those attribute dimensions.

1.3.2 Literature on Bundling

Another stream of related research concerns "bundling". Bundling has received a good deal of attention given the widespread availability of various types of grouped products; both bundling strategies and welfare consequences have been studied, and prescriptive guidance for how marketers should go about designing and pricing potential bundles are also addressed (a full review is provided by Venkatesh & Mahajan, 2009).

Our study is closely related to the literature addressing the question of which products should be included in the bundle. An early example of modeling approach to this issue is the "balance modeling" proposed by Farquhar & Rao (1976). This approach explains variety-seeking as a process through which consumers try to obtain a balanced set of attributes, which could involve either homogeneity or heterogeneity within the attributes, depending on their nature. Bradlow & Rao (2000) proposed an extension of this balance model, estimated at the individual level using a hierarchical Bayesian approach. Later, Chung & Rao (2003) further developed it to accommodate bundles for products with different sets of attributes, such as a computer's monitor size and storage capacity.

While these models can be useful in understanding user choices for bundles, they require that the researcher observes many choices of bundles to estimate heterogeneous preferences for attribute levels and dispersion, conditions that rarely hold for custom bundle choices of experiential products. Further, the model measures users' attribute-specific variety-seeking (or variety-avoiding) behavior through preferences for the dispersion within each relevant attribute; although this can be readily applied when consumers evaluate and choose among curated bundles, this is not the case with customizable bundles, whose enormous combinatorial possibilities preclude a single consumer evaluating even a small fraction of them.

More recently, Kumar et al., (2020) propose use of co-purchasing and co-viewing histories to create product feature embeddings in order to suggest product bundles for large-scale retailers with copious assortments. These are then tested in a field experiment, demonstrating the feasibility and validity of the approach. However, their goal is to propose bundles at the market- rather than individual-level bundle, and therefore they sidestep measuring and accounting for individual preferences for variety.

1.4 Empirical Setting and Data

Our empirical application concerns customizable bundles for performing arts organizations, often termed along the lines of "create-your-own" subscriptions. Most performing arts organizations, including Carnegie Hall, New York Philharmonic, Detroit Opera House, offer customizable bundles alongside organization-curated ones. In exchange for pre-commitment to attend a minimum number of performances, consumers who subscribe to these programs can freely pick among eligible performances, often receiving benefits like fixed-percentage price discounts, early access seat choice, or simplified exchange policies. The organizations themselves benefit by lowering the risk of having many empty seats to sell at a last-minute discount, incurring future marketing costs, and a predictable cashflow early in the season, which helps plan expenditures.

We estimate our model on a rich dataset from a performing arts organization in the midwestern U.S. that offers create-your-own subscription programs. The program requires that consumers pick five or more performances in return for a 10% discount, as well as preferential seating selection and easy ticket exchange policies. The data span seven seasons, from 2011 to 2017, during which time the organization hosts an average of 54 performances a year, out of which roughly 3/4 are eligible for create-your-own subscriptions each season. The excluded performances are usually highly popular ones, some so much so that tickets are available only via lottery. The performances are classified into seven "genres" (chamber, choral, dance, orchestra, theatre, and others)[1], and are performed in several major venues within a five-mile radius. The data comprise two parts: transaction records and performance brochures with concisely informative textual descriptions of each performance.

1.4.1 Transaction data

The transaction data include 75,947 time-stamped records of bundle transactions over seven seasons, for both the curated and create-your-own bundle purchases, as well as add-ons and exchanges related to the subscription packages. Each transaction entry includes user ID, performance name and date, order date, purchase type, number of tickets, and price paid. The composition of create-your-own bundle purchases is reconstructed from order date and purchase type. Importantly, being a latent (and arguably mental) construct, the order in which the performances were "added" to create-your-own bundles is not observable, since the entire bundle order is entered into the system at once. We focus on create-your-own bundles with five or more unique performances (per the restriction by the organization) that were bought on a single day.

The dataset thereby consists of 1204 users who purchased create-your-own bundles at least once over the seven seasons. Each bundle is indexed by user-season pair, as some of these users bought the subscription package in multiple seasons, including some who purchased create-your-own bundles in all seven seasons. A total of 2142 create-your-own bundles were purchased, which included 288 unique performances. These varied in size from 5 to 10 performances, with an average of 5.9 performances per bundle. Among all users ever buying bundles, 61.3% of users (738 out of 1204) bought the bundle only once over the observation

---

[1] Genre is based on the performer, performance type, and content, are hand-labeled by staff at the organization, and each performance belongs to only one genre.

period of seven seasons.[2] We retain all users with at least two bundle purchases. The final sample consists of 466 users who bought 1404 create-your-own bundles with 8392 performances. We use data from seasons 2011 to 2016 as an estimation sample (466 users with 1175 bundles and a total of 6938 performances) and purchases in 2017 as a hold-out sample (229 users with 229 bundles and 1409 performances). Table 1.1 lists the summary statistics for the season characteristics.

Table 1.1 Summary of season characteristics

Season	Number of performances	Number of users	Average bundle size
2011	41	165	5.93
2012	36	174	5.61
2013	36	207	5.98
2014	38	185	5.85
2015	49	218	6.22
2016	43	226	5.98
2017	45	229	6.15

1.4.2 Performance data

Across 7 seasons, there are 536 unique performance showings. However, some have multiple (as many as 15) showings of the same performance, and these (the same performances by the same performer, with the same program, on several consecutive days) are collapsed, resulting in a total of 377 "unique" performances. Among those, 288 performances were eligible for create-your-own bundles.

[2] We removed users whose bundles ever included more than 10 performances (86 users) as these were likely firm coding errors in categorizing of bundle purchases, since bundles were largely exactly five performances (51.1% of all bundles, with over 90% including nine or fewer performances in their bundle.)

Data on performances included both structured and unstructured information: structured information included the date of the performance, genre, venue, average price of the tickets, as well as a binary weekend variable[3]; unstructured information included short descriptions (average = 163.2 words) of the performances from the brochures. These brochures were mailed to patrons from past seasons, and the same descriptions were available on the website, suggesting that the vast majority of users were likely exposed to these descriptions when considering their potential bundle purchases.

We use these descriptions to extract latent performance dimensions using Latent Dirichlet Allocation, "LDA" (Blei et al., 2003; Griffiths & Steyvers, 2004), to characterize the performances more richly than simple genre assignments alone.[4] We then construct variety metrics, which are summary measures of the topic distributions of the prospective subset under consideration, aimed at proxying its perceived variety. It is critical to note that the values of variety metrics are updated as additional items are considered and added to the set, which means that the variety of an existing set of performances can be affected more by some additions than others. We describe how we construct these measures in Model Development section. These variety metrics are later tested on how well they capture consumers' variety perceptions when evaluating potential bundles. Below, we briefly describe how we processed the text and discuss the LDA results.

---

[3] As we collapsed multiple showings of the same performances to a single "unique" one, the weekend variable was also collapsed by creating an "include weekend" dummy indicating whether the performance had at least one weekend showing.

[4] We chose LDA for its transparency, scalability, and extensive use in validated applications. There are many newer methods that build on and extend it, but exploration of alternatives did not show sizable substantive differences or unambiguous superiority in our application context.

1.4.3 Topic extraction

Among the unique features of performance descriptions as in our corpus of brochures is that names (of organizations, performers, venues, and authors) appear frequently in the text. For example, this is an excerpt from the description of American Ballet Theatre (ABT)'s Romeo and Juliet:

> *Kevin McKenzie, artistic director, Choreography by Kenneth MacMillan Music by Sergei Prokofiev ... Kenneth MacMillan's masterful interpretation of Shakespeare's enduring romantic tragedy has become one of ABT's signature productions. The story of Verona's tragic star-crossed lovers is woven throughout a dance tapestry rich in character nuance and sensuality, with Renaissance Italy providing a sumptuous and period-perfect background. Sergei Prokofiev's instantly recognizable music, performed live by the ...Opera Theatre Orchestra, underscores the lyric beauty and passion of this beloved ballet. ...*

As this short excerpt forcefully illustrates, names like "Kevin McKenzie," "Kenneth MacMillan," and "Sergei Prokofiev" appear frequently in the text. Moreover, newspapers and magazine names also appear frequently when previous reviews are quoted. We explored n-grams (a sequence n-words; for example, Sergei Prokofiev is a 2-gram and New York Times is a 3-gram) to prevent treating unique names such as Sergei Prokofiev separately as Sergei and Prokofiev. More than 1000 n-grams specific to the setting were identified. We concatenated n-grams that appeared in two or more documents, keeping 24.6% of n-grams under consideration. For example, Los Angeles Times appeared in 12 different documents, likely when their reviews were cited for the performances. Given that this 3-gram appeared more than twice, whenever Los

Angeles Times appeared in the document, it was treated as a unit rather than three separate terms, Los, Angeles, and Times.

We followed standard pre-processing steps, tokenizing and removing stop words from the document (see Berger et al., (2020) for detailed information on preprocessing text). We had 6501 unique lemmatized terms, from which we removed the bottom 5% in terms of tf-idf score (a metric reflecting a term's importance in the corpus). This leaves 6176 unique terms from 350 documents. We used standard LDA with priors suggested by Griffiths & Steyvers (2004) on the text data processed as described above. We chose to use 7-topic result, which is suggested as "optimal" number of topics following Cao et al., (2009). Further, the 7-topic model has the added benefit of having the same number of topics as the genres, making any comparisons – either overall or incrementally – from genres to textual information more defensibly apples-to-apples. We explored different number of topics; however, gains from using a larger number of topics was not substantial enough to justify the extra individual-level parameters, especially given the limited number of choices per person.

Figure 1.1 lists the top 10 words associated with each of the seven topics, which capture various dimensions of the performances, but are also highly correlated with genres. Going back to the example of ABT's Romeo and Juliet ("dance" genre), its topic loadings are shown in Table 1.2.

Table 1.2 Topic loadings for American Ballet Theatre's Romeo and Juliet

Title	Topic1	Topic2	Topic3	Topic4	Topic5	Topic6	Topic7
Romeo & Juliet	0.281	0.018	0.088	0.018	0.246	0.281	0.070

The performance loads highly on three topics – 1, 5 and 6 – which relate to "choral and symphony orchestra," "theatrical elements," and "dance and choreographic elements" of performances, respectively. Romeo and Juliet, being a Shakespearean work performed by a major ballet company, that the performance loads high on dance- and theatre-related topics makes intuitive sense. Intriguingly, having a live music component by an opera theatre orchestra appears to lead to a high loading on topic 1, which has both symphonic and choral elements. Importantly, the topic representations capture performance dimensions not easily captured by the broad genre alone, in large part because the loadings are not binary, but reflect degree of confluence, and are in fact a measured simplex variable, as opposed to binary judgment calls based on pre-established categories.

1.5 Model Development

We propose a model of consumers' (customized) bundle construction, based on a latent process of sequentially adding individual performances to an expanding subset. Unlike choices among available bundles pre-curated by the firm, customizing a bundle involves mentally adding in eligible individual items, that is, a (latent) sequential process rather than a one-shot decision. Moreover, this circumvents the "curse of dimensionality" of modeling bundle choice as a multinomial choice out of a prohibitively large number of potential bundles. For instance, in our current dataset, the average number (across seasons) of possible bundles of five performances exceeds 870,000. Further, because consumers often choose bundles that include more than the modal bundle size of five performances, the full set of possible bundles is potentially on the order of 2 (number of performances).

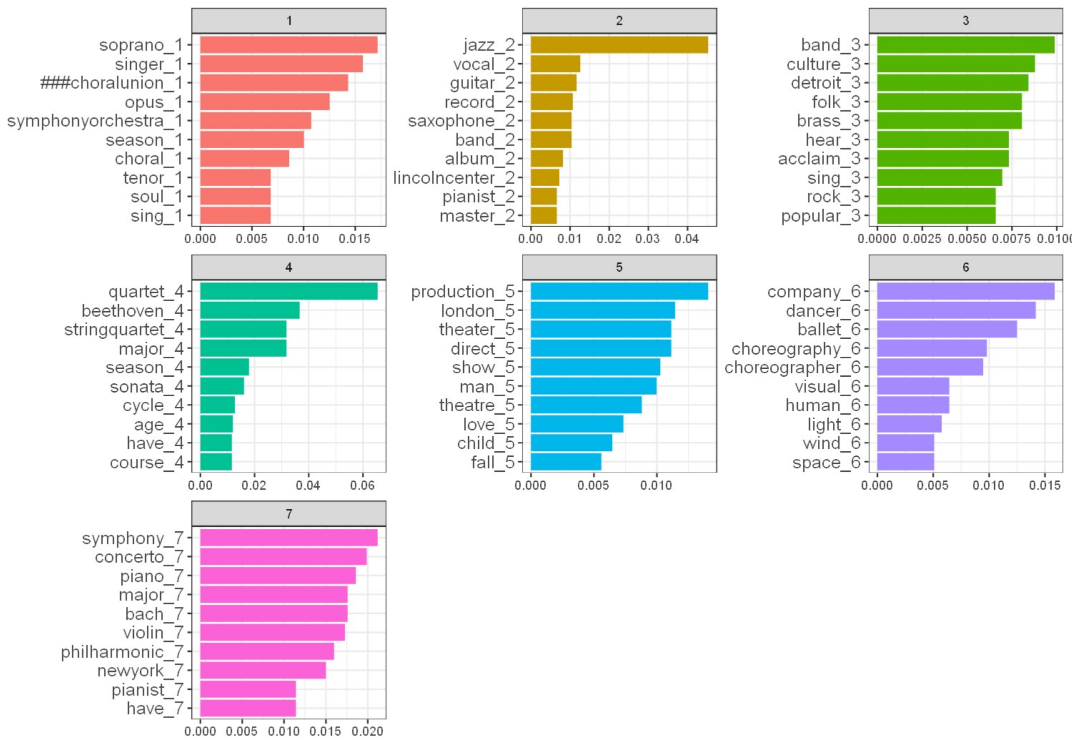

Figure 1.1 Top 10 words in each topic[5]

In addition to alleviating the dimensionality concern, modeling bundle construction as a unobserved sequential process allows us to uncover the "timing" (i.e., nth inclusion choice) at which consumers start considering the subset-level variety implications of additional performances. This approach contrasts with previous models of grouped choice (e.g., balance model; Bradlow & Rao, 2000; Chung & Rao, 2003; Farquhar & Rao, 1976), which treat variety as a static concept, such that consumers consider attribute dispersion of all the items in the bundle. In sequential customizable bundle choices, however, variety is dynamic, evolving as new items are (covertly, to the researcher) added to the provisional subset. Consumers may not take

[5] The name of the organization is masked as ### (in topic 1).

variety into account for all item inclusion decisions, especially for earlier ones, e.g., because "variety" is a fuzzy concept when there are only one or two items in the subset; it is also possible that variety becomes more (or less) important as more items are added to the set. We treat this as an empirical question. Our modeling approach allows us to empirically test when variety "kicks in" during the bundle assembly process, with a null model of "all items equally impact perceptions of variety" (one we will be able to reject for these data).

The model captures a process wherein consumers add performances to an evolving subset, balancing preferences for the component performances (based on their many attributes) and variety implications of adding a specific one to the extant subset (i.e., how its inclusion affects the variety of the intermediate bundle). We estimate consumers' heterogeneous preferences over performance attributes (genres and latent topic representations) and over variety metrics, which are constructed based on the latent topic representation of the performances to capture the perceptions of variety. For variety metrics, we construct various measures of variety and empirically test which one best appears to capture the marginal value of the underlying variety construct. We first describe the model itself and then explain how the different variety metrics are constructed.

1.5.1 Model

Consumer i in season s chooses an unordered bundle b_{is}, which includes $K_{is}(\geq 5)$ performances from L_s, the set of all performances in season s. We use ( ) to denote an unordered bundle, and {} an ordered bundle (of a specific permutation). Consumer i's (unordered) bundle in season s is defined as

$$b_{is} = \left(l_{is1}, l_{is2}, \cdots, l_{isK_{is}}\right),$$

where l_{isk} $(k = 1, \cdots, K_{is})$ denotes specific performances included in the bundle. We suppress season subscript s for simpler notation where possible.

Since we model bundle choice as a sequential process of adding performances to an evolving subset, we need to consider the order in which the performances are added, which we do not observe in our empirical setting (nor is it typically observed in such settings). Hence, we need to account for all possible permutations of the K_i elements in b_i. A priori, all permutations are assumed to be equally likely, and the unconditional probability of choosing (the unordered set) b_i is therefore the sum of the conditional probabilities of all permutations of its component performances.

Let R_{b_i} denote the set of all permutations of all performances included in the bundle b_i, such that

$$R_{b_i} = \{\{l_1, l_2, \cdots, l_{K_i}\}, \{l_1, \cdots, l_{K_i}, l_{K_i - 1}\}, \cdots, \{l_{K_i}, l_{K_i - 1}, \cdots, l_2, l_1\}\}$$

and refer to m^{th} permutation of K_i performances as $r^{(m)} = \{r_1, r_2, \cdots, r_{K_i}\}$, where r_1 is chosen first, r_2 second, and so on.

We specify the probability that user i chooses bundle b_i in a specific permutation order $r^{(m)} \in R_{b_i}$ as an exploded random coefficients logit model (Allison & Christakis, 1994). Then the probability that user i chooses b_i given a particular permutation $r^{(m)} \in R_{b_i}$ can be written as

$$P(b_i = (l_{i1}, l_{i2}, \cdots, l_{iK_i}) | \, r^{(m)}) = P_i(r_1) P_i(r_2 \mid r_1) \cdots P_i(r_{K_i} | r_1, r_2, r_3, , r_{K_i - 1})$$

Summing over all possible permutation orders, the unconditional probability that a user i chooses unordered bundle b_i is

$$P\left(b_i = (l_{i1}, l_{i2}, \cdots, l_{iK_i})\right) = \sum_{r^{(m)} \in R_{b_i}} P_i(r_1) \prod_{k=2}^{K_i} P_i(r_k \mid r_1, \cdots, r_{k-1})$$

where the choice probabilities of selecting the first performance is

$$P_i(r_k) = \frac{\exp(U_{ir_k})}{\sum_{l \in L_{s,k}} \exp(U_{il})} \text{ for} k = 1$$

and

$$P_i(r_k \mid r_1 \cdots, r_{k-1}) = \frac{\exp(U_{ir_k})}{\sum_{l \in L_{s,k}} \exp(U_{il})} \text{ for} k = 2, \cdots, K_i$$

where $L_{s,k}$ is the set of all eligible performances at k^{th} $(k = 2, \cdots, K_i)$ choice occasion, excluding performances chosen in previous choice occasions from L_s.

Utility specification

The conditional choice probabilities are driven by the user's utility for adding a given performance to the partially constructed, intermediate subset. The utility for adding a performance is driven by two different factors: preferences for the performance itself and preferences for variety of the resulting set with a new performance. First, the preferences for the performance enter the utility function in the form of the attributes of each performance considered in bundle construction. Performance-level attributes for a performance r_k populate vector X_{r_k}, which includes performance characteristics, e.g., genres, latent topic representations, average price of a ticket, weekend offering, and others. We further control for the venue of each performance (Z_{r_k}), which is a proxy for its relative expected popularity.

Second, and the focus of this research, the variety metrics V_{ir_k} – associated with the addition of performance r_k to the existing subset $\{r_1, \cdots, r_{k-1}\}$ – capture the interaction between performances that are already in the set and the performances that are currently considered as a k^{th} addition $(\forall l \in L_{s,k})$. We differentiate the implications of variety between earlier and later choices. Specifically, we allow for these variety metrics to only factor into later choices ($k \geq k^*$) for the bundle, once there are already "enough" chosen performances, while we allow them to not be relevant for earlier choices ($k < k^*$). We consider different values of k^* empirically ($k^* = 3,4$). Additionally, we allow for heterogeneous preferences for performance-level attributes X_{r_k} and variety metrics V_{ir_k}, but not for Z_{r_k}.Therefore, the utility for all performance choices is:

$$U_{ir_k} = X_{r_k}\beta_i + Z_{r_k}\gamma \text{ for } k = 1, \cdots, (k^* - 1)$$

$$U_{ir_k} = X_{r_k}\beta_i + Z_{r_k}\gamma + w_k V_{ir_k}\xi_i \text{ for } k = k^*, \cdots, K_i$$

$$\text{where } \beta_{ip} \sim N\left(\bar{\beta}_p, \sigma_{\beta_p}\right) \text{ and } \xi_{iq} \sim N(\bar{\xi}_q, \sigma_{\xi_q})$$

where w_k represents different weights on variety metrics for different choice occasions. This allows the model to capture whether users put more (or less) weight on variety at different choice occasions; for example, it is possible that consumers consider variety more important later in the (latent construction) sequence.

Estimation

Since we do not observe the order in which the performances are added to the bundle, we must compute the probabilities of all possible permutations and sum them. For instance, the probability of choosing a particular bundle of size $K = 5$ needs to be evaluated $5! = 120$ times, which is computationally costly (and even more so for some users who purchase more than five performances). The number of bundles to be evaluated increases factorially (i.e., $O((K/e)^K)$ in the bundle size, K performances. To avoid this heavy computational burden, we turn to Monte Carlo methods and select a random order of addition for each bundle (at user-season level) and fix the sequence of addition for estimation, essentially integrating out the unobserved order by stochastically sampling over possible orders. We show that this method allows us to faithfully recover the effects of attributes and variety of the bundle via simulation. Simulation results using both simulated and holdout season data and estimated coefficients appear in the appendix. We estimate the model with No U-Turn Sampling (NUTS) of HMC algorithm implemented in the Stan probabilistic programming language (Carpenter et al., 2017; Hoffman & Gelman, 2014) using 2000 draws, with the first 1000 draws used for burn-in.

1.5.2 Variety metric construction

Denoted V_{ir_k} above are variety metrics, which are constructed using the latent topic representations of the performances. With limited guidance from the extant literature on which dimensions of the attributes govern the perception of variety and its evolution, we consider a broad spectrum of variety metrics and empirically test which metric best captures the marginal value of variety to the set. We consider four sets of metrics summarizing pairwise Euclidean distances[6] of the topic representations. We refer to them as total, incremental, and delta metrics, along with one intermediate metric type, which we refer to as conditional. Further, four quantities "mean, minimum, maximum, and standard deviation" within each of those four approaches are computed. These sets are meant to capture different aspects of variation in the set variety as new performances are added to the set. Table 1.3 shows the list of variety metrics that we consider, and we explain the construction of each variety metric in detail below.

For ease of exposition, consider an example of a bundle of five performances shown in Table 1.4, which we assume are added to the set in the shown order. We explain the metrics assuming that the first three performances are currently already "in" the set: {Ballet Preljocaj (BP); Denis Matsuev, piano (DM); Yuja Wang (YW)} and the user is considering which fourth performance to add out of the remaining 38 eligible performances in the season (although

[6] We also constructed the same variety metrics using taxicab distance in earlier stages and found that the results were similar to those using Euclidean distances.

we only show 2 of the 38 here: Hamburg Symphony (HS) and Jazz at Lincoln Center Orchestra (JLCO)). Pairwise distances of these performances are listed in

Table 1.5 and

Table 1.6 summarizes the values of different variety metrics, continuing the illustration.

Table 1.3 Overview of variety metrics

Variety Type		Variety Metrics
Total	×	Minimum
Delta		Mean
Incremental (Conditional)		Maximum
		Standard deviation

Table 1.4 Example bundle of five performances

Performance name	Topic1	Topic2	Topic3	Topic4	Topic5	Topic6	Topic7
Ballet Preljocaj (BP)	0.070	0.018	0.018	0.018	0.368	0.456	0.053
Denis Matsuev, piano (DM)	0.027	0.055	0.014	0.123	0.178	0.027	0.575
Yuja Wang, piano (YW)	0.121	0.010	0.202	0.121	0.020	0.040	0.485
Hamburg Symphony (HS)	0.189	0.057	0.057	0.019	0.189	0.415	0.075
Jazz at Lincoln Center Orchestra (JLCO)	0.009	0.780	0.037	0.083	0.009	0.046	0.037

Table 1.5 Distance matrix of the example bundle

	BP	DM	YW	HS	JLCO
BP	0.0	0.713	0.727	0.227	0.942
DM		0.0	0.282	0.663	0.920
YW			0.0	0.612	0.914
HS				0.0	0.854
JLCO					0.0

Table 1.6 Values of variety metrics from the example bundle

Performance	Variety Type	Min	Max	Mean	SD
HS conditional on (BP, DM, YW)	Conditional	0.282	0.727	0.574	0.253
	Total	0.227	0.727	0.537	0.223
	Delta	-0.055	0	-0.037	-0.029
	Incremental	0.227	0.663	0.501	0.238
JLCO conditional on (BP, DM, YW, HS)	Conditional	0.227	0.727	0.537	0.223
	Total	0.227	0.942	0.685	0.254
	Delta	0	0.215	0.148	0.031
	Incremental	0.854	0.942	0.908	0.037

1.5.2.1 Conditional metrics

Conditional metrics are a set of summary statistics of pairwise distances of only the performances that are already in the set (i.e., not including performances that the user is considering adding to the bundle). Summary statistics {min, mean, max, SD} of all pairwise distances are computed. In the example, given that there are three performances in the set (BP, DM and YW), the distances between pairs of performances are 0.713, 0.727 and 0.282, the summary values of these metrics are conditional min distance = 0.282, conditional max distance = 0.727, conditional mean distance = 0.574 and conditional SD = 0.253 for all performances in the consideration set.

As conditional metrics describe the variety level of the subset that is already chosen, the values of conditional metrics do not capture the variety implications of adding a new performance to the set. Conditional metrics are therefore not "attributes" of the candidate performances, and they do not directly enter utility function as variety metrics. Rather, these metrics are used to construct delta metrics, as explained below.

Whereas conditional metrics are aimed at capturing static variety of the already-constructed intermediate bundle, total, incremental, and delta metrics can capture the changes in variety of the set brought by the addition of candidate performance.

1.5.2.2 Total metrics

Total metrics capture the overall bundle-level variety of the candidate bundle that is being evaluated. They are constructed as summary measures of pairwise distances of both what is already in the set and what is currently being considered for addition. In the example, suppose the user is considering adding Hamburg Symphony as the fourth item to be included in the set. That makes {BP, DM, YW, HS} the temporary bundle that is being evaluated. With these four performances in the set, we get total min distance = 0.227, total max distance = 0.727, total mean distance = 0.537 and total SD = 0.223.

It is important to note that the value of total minimum and maximum may not necessarily change with the addition of the new performance, since the additional performance may not be in the minimum or maximum distance pair. In the case of adding Hamburg Symphony to the set, the maximum distance stays the same as distance between Ballet Preljocaj and Yuja Wang (0.727), whereas the minimum distance changes from 0.282 to 0.227. This suggests that if total metrics best capture the value of variety, the level of variety in the resulting set is likely the most salient aspect of variety in the bundle choice rather than the changes brought by the addition of new performance to it.

1.5.2.3 Delta metrics

Delta metrics are designed to capture the changes in the set-level variety rather than the level. They are defined as the difference between total and conditional metrics for all summary measures. In the example, delta metrics will capture the differences between the bundle of three performances {BP, DM, YW} and a bundle of four {BP, DM, YW, HS}. When Hamburg Symphony was added to the set, the minimum distance of the set changed from 0.282 to 0.227, so we define delta min distance = -0.055, whereas delta max distance = 0 since the maximum distance did not change. Further, delta mean distance = -0.037 (= 0.537 – 0.574) and delta SD = -0.029 (≈ 0.223 – 0.253). Overall, these measures would suggest that the addition of Hamburg Symphony to the set reduces the variety of the set, bringing the sets of performances closer to each other.

These delta metrics are more focused on the changes brought by adding a new performance, but they should still capture the resulting overall set-level variety. If consumers paid attention to changes in the sets' variety rather than their levels, delta metrics should best capture consumers' valuation of how much the newly added performance may contribute to the variety of the set.

1.5.2.4 Incremental metrics

Incremental metrics focus on the distances between the currently considered performance and each of the existing ones (i.e., distances among the performances already in the set are not considered). While these metrics do not fully capture the overall set variety implications, they do capture the relationship between the newly added performance and existing ones. In the example, incremental distances are 0.227, 0.663 and 0.612, giving incremental min distance = 0.227, incremental max distance = 0.663, incremental mean distance = 0.501, and incremental SD = 0.238.

If incremental metrics turn out to provide the best model performance (by representing the marginal value of variety of adding the performance to the set), it is consistent with consumers' focusing more on imminent changes brought by the addition of the new item rather than resulting variety of the overall new set.

It is noteworthy that delta and incremental metrics are similar in that they focus on changes. Hence, the dimensions of variety captured by incremental and delta metrics will overlap to a certain degree. In case of adding Hamburg Symphony to the set, for example, it brings the closest pair of performances even closer, which are captured by both delta and incremental metrics. The difference in the two measures is illustrated in the case of maximum distance: Hamburg Symphony does not broaden the spectrum of performances in the bundle because adding the performance to the set doesn't change the maximum distance of the set, which is better captured by delta metrics.

1.5.3 Model performance evaluation

We evaluate the performance of the model on holdout data, which include season 2017 bundle choices of users who had purchased at least one bundle in previous seasons (season 2011-16). For evaluation, we compute the choice probabilities of all possible bundles of five following (2) and compare the predictive performance across different specifications for the chosen bundles. Despite there being some consumers choosing to include more than five performances in their custom bundles, we restrict our purview to bundles of five performances for several reasons. First, roughly half of the bundles bought in the holdout period contain five performances. Second, as we are abstracting away from the number of performances, there is no formal way of addressing how many performances we should recommend to any one user, and not nearly enough variation in this regard to systematize the model to capture that aspect of our particular data. Third, because the prediction step requires evaluation of all possible permutations of a given sized bundle, allowing for a greater number of performances would substantially and prohibitively expand the space of combinations that need to be evaluated.

That said, our evaluation procedure is as follows:

Step 1. Variety metrics for all possible 5! permutations of ${}_{45}C_5 = 1{,}221{,}759$ combinations of five performances are computed.

Step 2. Using N draws from the joint posterior $P(\beta_i, \gamma, \xi_i | X, Z, W)$, the posterior probabilities of user i choosing a bundle b out of the set of all possible bundles B is computed as following:

$$P_i\left(b | \beta_i^{(n)}, \gamma^{(n)}, \xi_i^{(n)}\right) = \sum_{r \in R_b} P_i(r_1) \prod_{k=2}^{5} P_i(r_k | r_1, \cdots, r_{k-1}) \; \forall b \in B$$

Step 3. The probability of the chosen bundle b^i is compared across models.[7]

1.6 Model Results

With the extracted attributes (topics) from the performance descriptions and proposed variety metrics, we explore various specifications to narrow the search for variety metrics to be further tested with out-of-sample predictive performance. We first test whether extracted attributes allow us to better capture individuals' preferences over performances and then compare various combinations of variety metrics.

[7] There are users who have more than five performances in their sets, which means that there could be multiple bundles that are nested in the chosen bundle b_i. For example, if a user included seven performances in her bundle, there are 21 bundles of five performances that are nested within it. In that case, we use the maximum of the 21 possible probabilities and compare it across models.

Table 1.7 Model fit comparison

Type		Model	3rd on	4th on	Weighted (3rd on)
	(1)	Genre		-22853.2	
Baseline	(2)	Topic		-22477.3	
	(3)	Genre + Topic		-21800.5	
	(4)	Min + Max	-21719.4	-21755.7	-21712.5
Total	(5)	Mean	-21765.6	-21766.6	-21762.8
	(6)	SD	-21793.6	-21789.0	-21788.7
	(7)	Max + SD	-21792.6	-21783.7	-21785.1
	(8)	Min + Max	-21713.6	-21755.1	-21717.1
Delta	(9)	Mean	-21759.6	-21757.4	-21760.9
	(10)	SD	-21792.4	-21791.5	-21783.7
	(11)	Max + SD	-21791.2	-21784.5	-21793.3
	(12)	Min + Max	-21733.1	-21733.9	-21728.5
Incremental	(13)	Mean	-21731.9	-21743.8	-21749.0
	(14)	SD	-21792.4	-21791.5	-21790.0
	(15)	Max + SD	-21744.2	-21751.7	-21745.4

* All variety models are over genre + topic baseline

Table 1.7 contains model likelihoods evaluated at the median of the parameter estimates. We find that performances' topic information improves model fit – over genres alone (comparing (2) "baseline with topics" with (1) "baseline with genres") and in addition to genres (comparing (3) "baseline with genres and topics" with (1) "baseline with genres") – suggesting topics provide additional insight into consumer preferences rather than simply subsuming genre information. Rather, the additional attributes extracted from the text can represent contextually nuanced details of the performances that may transcend those captured by the genres. For example, going back to the example of ABT's Romeo and Juliet, its genre is dance, but the extracted attributes reveal that it has choreographic elements, a dramatic component (it is set to Shakespeare's play), and classical musical element. These additional elements need not be (and, in fact, are not) representative of all performances in dance genre, nor must this performance be emblematic of dance in its entirety, and this greater subtlety and flexibility may explain the additional improvement in model fit when both the genre and extracted attributes are appear as covariates.

Importantly, we explore various combinations of variety metrics. Specifically, we propose three categories of metrics for variety (total, delta and incremental) and four summary measures within each category (shown in 3). Based on preliminary exploration of the configuration space, we limit our search to combinations that include up to two variety metrics from the same category, testing the following four combinations within each (Table 1.8).

Table 1.8 Tested variety metric combinations

Variety Type		Variety Metrics
Total Delta Incremental	×	Mean SD Min + Max Max + SD

Because “variety” involves the span or extent of a set of items, these variety metrics meaningfully summarize the set only from the third choice on. For first choices, none of the metrics are defined, as there is no ‘pair’ to compute pairwise distances for. We do observe one pair of performances for which we can compute the distance for second choices – however, precisely because there is only one pair, these variety metrics are either not defined or provide redundant information. As there is no conditional distance defined for second choice, none of the delta metrics can be computed. Nor are total and incremental SD defined as there is only one pair in the set at the second-choice occasion. Even for total and incremental metrics that are technically defined, mean, minimum and maximum distances are identical to that unique pairwise distance. We therefore allow the variety metrics to enter either from the third or fourth choices, and we test which is superior.

Similarly, it is possible that individuals select items to provide variety earlier (vs. later) in assembling their choice sets; consequently, we also estimate a "weighted" version, for which different choices have differential relative impact, with "all weights equal" nested as a special case. [Note that, because we do not observe the order of items' entering the bundle – which may be latent even to the decision-maker – the model's likelihood integrates out this order and estimation can proceed by stochastically sampling over possible orders, as discussed in the previous section.]

1.6.1 Gains in out-of-sample prediction

Upon testing the models, although the differences are not stark, we find that across the different variety metric types, min + max and mean consistently show the largest improvement over the baseline model, which includes genres, topics, and control variables (venue and number of shows; weekend performance/not). That said, none of the variety metrics uniformly dominates the others across all specifications. While the incremental gains in terms of likelihood is not large, this should be understood with some context in mind; first, given that the variety metrics start kicking in at either the 3 rd or 4 th choices, only 66.5% (49.8%) of choices benefit from the addition of any variety metric. Further, given that we are including one or two extra variables (from which not everybody benefits, as we discuss again in the predictive performance of the model), the gains in model fit are moderate.

Given the extreme computational requirements of the hold-out posterior predictive performance checks, we focus on these two variety metrics (min + max and mean), allowing them to "kick in" at different timings from 3 rd-choice-into-set onward. We evaluated the probabilities of all possible bundles of five performances at the user level using 15 posterior draws. To evaluate predictive performance, we compare the predicted probability distribution of the chosen bundle across different models. At the individual level, we summarize the predicted probability distribution (of 15 probabilities per user, per bundle) using the median across draws and assess the gains from using variety metrics using the percentage gains of median probability under the variety model over the median probability under the baseline model. If a user had more than five performances in the chosen bundle, we focus on the bundle with the highest probability among the nested bundles of five. With 229 users, this leaves us with 229 observations on gains from using variety metrics.

Across users, to mitigate the impact of outliers and provide robust comparisons, we summarize the posterior predictive model performance with trimmed means of the percentage gains across users (1% and 5% from the top and bottom). Table 1.9 lists average predictive gains (across users) over the baseline model (Table 1.15 in the Appendix shows the same comparison using means of draws to summarize the predictive probability distribution at the individual level). Note that trimming decreases the magnitude of mean gains, as seen by comparing 1% vs. 5% values.

While the predictive gains vary somewhat across variety metric specifications, we find that allowing weights on variety preference to differ across choice occasions improves predictive performance for all but the incremental min + max model in the 1% trimmed result. Among those, the gains are especially noticeable in three specifications: for those with total mean, delta mean and delta min + max as variety metrics, allowing weights to vary across choice occasions leads to almost double the gains from specifications that assume the same weight for all choice occasion from the 3 rd / 4 th choices. These three specifications maintain their superior performance under both 1% and 5% trimmed results, showing robust, substantial improvement.

Table 1.9 Average predictive gains (in percentage) over baseline models across users (median across draws at the individual level)

	Users with improvement			Average gains 1% trimmed			Average gains 5% trimmed		
	3rd	4th	W	3rd	4th	W	3rd	4th	W
Total mean	98	109	122	17.78	14.85	**34.53**	7.73	7.99	**22.10**
Total Min + Max	116	119	112	19.75	20.42	21.86	13.99	14.58	10.09
Delta mean	102	105	**125**	15.93	14.33	28.20	8.03	6.68	**22.11**
Delta Min + Max	118	118	123	15.79	18.82	31.95	10.42	13.52	21.54
Incremental Mean	116	111	112	19.43	14.79	20.21	14.02	7.62	14.18
Incremental Min + Max	131	109	122	23.59	14.56	21.29	18.24	9.24	15.60

*W = Weighted, N = 229

We do not find a single specification that strictly dominates all others, in terms of magnitude of predictive gain or number of users for whom variety improves the predictive performance. It is not too surprising given that these variety metrics are formulated to capture the same underlying construct – how consumers might "conceptualize" variety across experiential items – and so some degree of correlation in their values is inevitable. Lack of uniform dominance, however, is not a shortcoming; instead, it robustly supports the importance of variety as a substantive component of modeling bundle choices.

Regardless of which variety metric is adopted, as the totality of 9's metrics clearly shows, accounting for within-bundle variety improves out-of-sample performance. While it might be tempting to believe that all users benefit from this additional information, this is not so: among users for whom the variety metric matters – in fact, only about half – it matters a lot. While we must stop short of any claims that "total mean" is the best metric in all situations, we adopt it in the subsequent sections, with differential weights for variety preferences depending on the timing of the choice within the sequence.

1.6.2 Model Estimates

Table 1.11 lists parameter estimates of average consumer preferences over performance and bundle characteristics in our focal specification. The baseline model (estimates shown in Table 1.10) includes six genres and six topics, with theatre and topic 7 (orchestra, symphony-related) as the baseline. Controlling for genres, consumers reveal significantly lower preferences for performances scoring high on topics 2, 4, and 6 (relative to reference level topic, topic 7), which capture jazz & band, chamber-music, and dance-related components, respectively. Comparing the parameter estimates between baseline and variety model, we find that the coefficients for the variables are remarkably similar across specifications (posterior distribution plot of the mean effects across these two specifications are shown in Figure 1.3 in the appendix), suggesting that preference for variety is likely coming from an independent variation in the dataset.

At the individual level (Table 1.12), roughly one quarter of the users (105/ 466 ≈ 22.5%) reveal a significant preference (95% posterior interval excludes zero) for or against any performance content attribute (i.e., topics and genres). Among those, 23 users (4.9%) reveal a strong preference for specific genres, while 88 users (18.9%) reveal both strong positive and negative preferences for topics. Within genres, the greatest number of users prefer dance, followed by orchestra, then choral performances. On the other hand, we observe polarized preferences for topics. For example, 51 users (10.9%) strongly dislike and 7 users (1.5%) strongly like topic 2, which is related to the jazz-musical element of the performances. We observe a similar pattern for topic 5, related to theatrical components; 16 users with strong liking of, and 3 with strong disliking of the topic.

Table 1.10 Parameter Estimates - baseline with genres and topics

Name	Median	[95% CI]		Name	Median	[95% CI]	
chamber	-0.104	-0.28	0.08	weekend	0.016	-0.049	0.083
choral	-0.089	-0.292	0.111	average price	0.284	0.256	0.313
dance	-0.051	-0.242	0.134	n showings	-0.069	-0.108	-0.031
jazz	0.13	-0.08	0.335				
orchestra	-0.089	-0.286	0.107	topic1	-0.046	-0.096	0.004
other	-0.058	-0.232	0.118	topic2	-0.199	-0.266	-0.133
				topic3	-0.014	-0.055	0.03
venue A	0.422	0.328	0.516	topic4	-0.182	-0.245	-0.123
venue B	0.346	0.23	0.466	topic5	-0.019	-0.072	0.037
venue C	0.145	0.019	0.273	topic6	-0.077	-0.141	-0.018
venue D	0.407	0.278	0.537				

Table 1.11 Parameter Estimates - weighted total mean distance

Name	Median	[95% CI]		Name	Median	[95% CI]	
chamber	-0.066	-0.248	0.118	weekend	0.016	-0.051	0.086
choral	-0.039	-0.242	0.157	average price	0.277	0.249	0.307
dance	-0.015	-0.206	0.167	n showings	-0.055	-0.094	-0.016
jazz	0.179	-0.024	0.379				
orchestra	-0.052	-0.249	0.14	topic1	-0.04	-0.089	0.007
other	-0.025	-0.195	0.149	topic2	-0.19	-0.253	-0.129
				topic3	-0.015	-0.057	0.027
venue A	0.427	0.327	0.526	topic4	-0.168	-0.226	-0.111
venue B	0.331	0.208	0.447	topic5	-0.013	-0.066	0.04
venue C	0.142	0.014	0.272	topic6	-0.082	-0.141	-0.026
venue D	0.424	0.287	0.558	**total mean dist**	**-0.212**	**-0.290**	**-0.133**

Table 1.12 Number of users with significant preferences (genres and topics)

Name	+	-	Name	+	-
chamber	1	0	topic1	0	0
choral	3	0	topic2	7	51
dance	14	0	topic3	7	0
jazz	1	0	topic4	0	6
orchestra	5	0	topic5	16	3
other	0	0	topic6	1	2

1.6.3 Variety and Covariate Effects

As shown in 11, consumers' average preference for variety (total mean distance) is negative and significant ($\hat{\beta}_{total\ mean\ dist} = -0.212; 95\%\ CI\ [-0.290, -0.133]$). This may appear counterintuitive at a first glance, as consumers are expected to prefer more variety when they make purchases in bundles compared to when making separate choices for immediate consumption (Read & Loewenstein, 1995; Simonson, 1990). However, this expectation must be interpreted as a comparison made in reference to one's own level of variety, i.e., if the user would have been making separate choices for immediate consumption. Since we are not making such within-individual counterfactual comparisons, our result does not suggest consumers seek less variety compared to some hypothetical alternative separate choice scheme.

Rather, the "negative variety preference" result can be understood in relation to the set of performances offered by the organization, in combination with how clearly consumers' preferences are defined around specific types of performances. The focal performing arts organization offers a wide range of performances across seven genres, varying from classical pieces by composers like Mozart and Beethoven to more novel and contemporary pieces, such as a modern dance performance accompanied by Azerbaijani opera music. Consumers choosing performances with similar themes are likely to result in a relatively concentrated set (e.g., romantic music from 19 th century) compared to the much larger variety of the available performances, a tendency that is likely further enhanced with experience as one's preference becomes more refined and clearly defined (Clarkson et al., 2013; Read & Loewenstein, 1995; Sela et al., 2019). As users tend to choose relatively similar performances (i.e., relatively smaller distances among the pairs of performances), their estimated preferences for variety are on average negative. The narrow range of the chosen values of the variety metric negatively affects the estimate of the coefficient on that variety metric, all else equal.

Additionally, across users, we find that the average price coefficient is positive and significant ($\hat{\beta}_{price} = 0.277; 95\%\ CI\ [0.249, 0.307]$).While it's important to note that this only serves as a "control" for preference (rather than a suggestion for ticket price optimization), one should bear in mind that bundle-purchasers elected to do so via early access, foregoing additional savings opportunities in the future, including night-of-show and student or senior discounts.

The coefficient for the number of showings per performance is significantly negative ($\hat{\beta}_{nshowings} = -0.055; 95\%\ CI\ [-0.094, -0.016]$). This may be due to the skewed distribution of showings across genres and performances, with 77.4% of all performances having a single showing only, and theatre performances accounting for most of the performances with three or more showings.

1.6.4 Individual-Level Comparisons

At the individual level, even fewer users – five, and all of them negative – have significant preferences for variety, but this is likely due to limited number of individual-level observations, which is further exacerbated by the fact that preferences for variety is informed by even fewer choices, excluding first two (three) choices. While none of the users have a strong, positive preference for variety, for illustrative purposes, we compare two users: one with strong preference against variety and another with largest median preference for variety. Table 1.13 compares the purchase histories of these two users: one with negative and one with positive variety preferences. User 147 (with negative variety preference) purchased larger bundles (7.1 performances per bundle) from fewer genres compared to User 212 (with relative positive variety preference) who has 5.4 performances per bundle. Moreover, despite having smaller bundle sizes, User 212's bundles have greater variety, expressed as sum of standard deviations across topics, compared to User 147, suggesting that the User 147's bundle contains a much more "concentrated" set of performances in the space of topics. Lastly, comparing the average of total mean distance values across choices, we see that the average total mean distance, another measure of variety, is also larger for User 212. This suggests that the model faithfully captures individual users' variety preferences using proposed metrics.

Lastly, Table 1.14 lists estimates of weights for variety for p^{th} choices within different bundle sizes (n) (i.e., $w_{n,p}$ is the weight for the pth performance in a bundle of size n). We do not estimate separate weights for 6 to 10 th choices, given the limited number of observations. For example, for bundle size of seven performances (third row), the table shows separate weights for 3 rd, 4 th, 5 th choices ($w_{7,3} = 0.430, w_{7,4} = 0.562, w_{7,5} = 0.710$) and the shared weight for 6 th and 7 th choices ($w_{7,6} = w_{7,7} = 1.639$). The estimated weights suggest a stronger role for variety (whether positive or negative) later in the bundle creation process. This pattern could be driven due to the salience of variety as more items are added to the subset; specifically, as more items are added to the subset, the mean distances among performance pairs likely will increase, which could in turn trigger consumers to place more emphasis on variety implications.

Table 1.13 Comparison of user histories with positive and negative variety preferences

	ID	Season	Bundle Size (A)	Number of genres (B)	Mean performances per genre (A/B)	SD across topics	Average of total mean dist
	212	2012	5	3	1.67	1.12	0.71
	212	2013	5	5	1	1.21	0.70
+*	212	2014	5	5	1	1.56	0.96
	212	2015	5	3	1.67	1.12	0.71
	212	2017	7	6	1.17	1.19	0.64
	147	2011	5	3	1.67	0.86	0.58
	147	2013	7	5	1.40	0.96	0.56
	147	2014	5	2	2.50	0.86	0.61
-	147	2015	10	3	3.33	0.67	0.37
	147	2016	8	2	4	0.70	0.32
	147	2017	8	3	2.67	0.73	0.44

* Note that the user's variety coefficient is statistically not significant

Table 1.14 Median parameter estimates of weights in total mean distance model

Bundle size	3rd	4th	5th	6th+
5	0.753	0.991	1.250	
6	0.510	0.669	0.843	1.950
7	0.430	0.562	0.710	1.639
8	0.388	0.511	0.643	1.481
9	0.363	0.477	0.602	1.386
10	0.346	0.455	0.575	1.322

1.6.5 Posterior Predictive Performance

We made hold-out season predictions for 229 users (49.1% of 466 users for estimation) who made a bundle purchase in season 2017. On average, they had 2.44 bundle purchases prior to the hold-out season, with 36.7% of users purchasing only once in the past. There were 45 performances in season 2017, slightly higher than the 40.5 average in prior seasons. 45 performances implies that there are ${}_{45}C_5 = 1{,}221{,}759$ possible bundles of five, each with 120 (=5!) possible latent orders.

Across all users, including the variety metric in the model increased the mean predictive probability by 42.7% over baseline model prediction and 205 times better than purely random chance (1/1,221,759). Given that the purpose of the model is to capture individual-level preferences over performances and variety, we focus on within-individual comparisons between the baseline model with topics and genres, and the model that includes total mean distance as the variety metric. For robustness, we report results of trimmed mean in 9 (1% and 5% from top and bottom) and use 1% trimmed as main result while showing 5% trimmed results in parentheses.

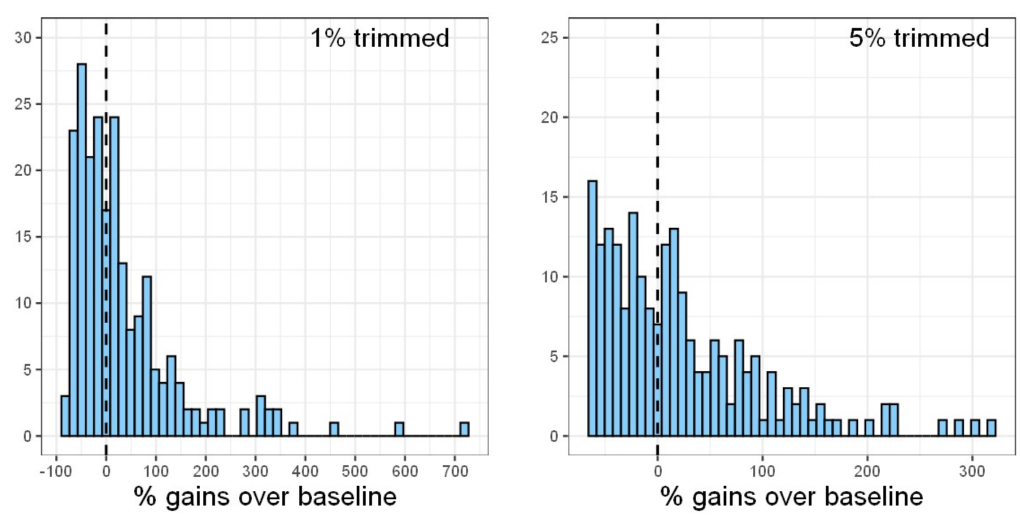

Figure 1.2 Distribution of percentage gains in predictive probability using variety metric (total mean distance) over baseline

Figure 1.2 shows the distribution of within-user percentage gains over baseline performance (1% trimmed on the left, 5% on the right). These offer a “scale-free” summary of the model's performance: 122 (53.3% of 229) users show increase in their predictive probability. Further, across all users, the (1% trimmed) average gain is 34.5% (22.1% for 5% trimmed). Zeroing in on users whose performance improved by use of variety metric, users gain on average a hefty 99.5% (73.0%), confirming that, among users who benefit from the variety metric, the improvement in predictive performance is substantial, reflecting a doubling in performance. This gain is notably higher for those with significant variety preferences; of the five users who have significant variety preferences, three users made a bundle purchase in 2017. The average gains for these three users is 219.1% (with one of the users in top 5%; removing that user, the average gain for the two users is 152.97%), compared to 32.0% for all other users whose variety preference is not significantly different from zero (again, trimming the top and bottom 1%; 20.8% with 5% trimmed). The contrast is stark, although the number of users with strong (dis)taste for variety is limited.

Lastly, we further investigate the composition of specific bundles the two models – baseline and variety – recommend, out of all possible bundles of five nested in the actual chosen bundle. 117 out of 229 users who made a bundle purchase in 2017 had more than five performances in the bundle. The variety and baseline models assign the highest probability to different bundles of five (i.e., the composition of the recommended bundle is different) in 64 out of the 117 bundles with more than five performances. The average probability gains are higher among users for whom the recommended bundle differs under variety versus baseline model prediction (43.3% vs. 34.2% for baseline after 1% trimming; 28.9% vs. 22.5% under 5% trimming). The gains for bundles of size five are the smallest, with 29.8% (17.8% under 5% trimming) gains over baseline, as there is no room for change in composition. This suggests two findings: first, using variety metrics can result in compositional changes in predicted bundles, allowing substantial gains in prediction. Secondly, the gains in predictive probability, even for the same recommended bundle, suggest that variety provides additional insight into preferences for variety itself, over and beyond the preferences for performances.

1.7 Conclusion

Firms offering their customers various rosters of potential experiences face a perpetual challenge in making recommendations of novel combinations for the future. To help consumers navigate these possibilities on their own, firms typically curate "bundles" of options that seem to go well together. Yet such bundles, by their market-level nature, cannot take into account individuals' preferences or their idiosyncratic desire for variety. Even when firms try to circumvent this problem by allowing consumers to customize their own bundles, many demur, owing to the sheer scope of possibilities.

In this paper, we meet this challenge by proposing a comprehensive model for customizable bundles of experiential products, one for which users balance utility for the individual component characteristics and the "holistic" variety of the bundle. Applying it to create-your-own performing arts subscription program from a midwestern performing arts organization, we find that accounting for variety via such holistic bundle-level metrics can improve hold-out predictive performance by up to 34.5%. Importantly, while the variety metric does not improve prediction for everybody, users who benefit do so substantially.

Our model addresses four challenges in both practice and the extant literature that arise in modeling choices of customizable bundles in the experiential products domain. First, by characterizing performances using attributes extracted from brochure descriptions, we can capture textured performance details that alleviate the coarse granularity of "genre" typologies; while genres often leverage expert opinion to encapsulate the complexity of performances, their intrinsic conceptual broadness is such that heterogeneity within a performance genre can sometimes be even larger than that across them. The additional detail and flexibility of extracted attribute information improves model fit significantly and offer the additional benefit of being extendable to future seasons containing previously unobserved performances and new artists, defining features of performing arts over seasons.

Second, we construct and test variety metrics, which are attribute dimensions of the bundles that can capture consumers' perceptions of bundle-level variety. Relatively little is known about which attribute dimensions consumers appear to attend to or how perceptions of variety evolve with the addition of items to a bundle. While we do not find a single metric that outperforms all others, we observe substantial improvement in predictive performance across various variety metric specifications, supporting the importance of accounting for variety preferences; empirically, we find that total mean distance of the bundle – computed as the mean of all pairwise distances between performances in the bundle – to offer the strongest performance for our dataset. More importantly from the perspective of application and generalizability, using these variety metrics substantially improved the model's predictive performance.

Third, we circumvent exploring the prohibitively large space of possible bundles by modeling bundle choice as a latent sequential addition process, as opposed to a multinomial one across potentially millions of bundle options. Doing so reduces the problem to a sequence of pick-one-of-$(n-k)$ problems, conditional on a specific order of addition, which unfortunately is not observable by the analyst. To overcome this, our fourth challenge, we rely on a novel Monte

Carlo approach by randomly drawing an order of addition for each customer, in effect stochastically integrating out the latent order while providing estimation stability within each individual bundle and demonstrate the validity of the approach via a series of parametric recovery exercises.

While our work proposes a novel methodology to leverage observed bundle choices to estimate individual users' preferences over performances and variety, as well as to make recommendations, there are inevitable limitations that open up avenues for further research. Notably, while relatively brief customer history is a common phenomenon in bundle choices, the annual cycle of seasonal subscriptions renders these individual histories even shorter, making it challenging to extract individual estimates with high precision, even via hierarchical Bayesian techniques. This problem is even more severe for variety preference parameters, the focal operative element of the entire approach, as we do not use all choice occasions to infer users' preferences over variety. Other settings in which the model can be applied often have shorter cycles, e.g., meal kit choices, which takes place either monthly or weekly, substantially alleviating such data limitations. Relatedly, our proposed approach, while accommodating heterogeneous preferences for variety, nonetheless does so in a stationary manner: evolution in one's preference for variety could potentially be modeled explicitly with longer user purchase histories.

Further, as the current work focuses on making recommendations to retain users who have previously made bundle purchases, our ability to extend recommendations for new users or those who only have purchased individual tickets is limited, either at population or segment levels. A data fusion approach (Feit & Bradlow, 2021), aligning users who have bought bundles before with those who have only made individual ticket purchases, would allow researchers to have a better understanding of variety preferences of the latter group. Combined with a field experiment aimed at understanding how the broad population responds to recommendation policies would directly aid the organization in refining their bundling strategies.

Lastly, the ultimate test of any theory or modeling framework is its empirical evaluation in the field. Despite the model's excellent "out-of-sample" performance in predicting results for a hold-out season, the true proof-of-concept would be a field experiment where the model's predictions are used to generate bundle suggestions for individual consumers. This would not only allow an evaluation of uptake (i.e., how likely suggested bundles are chosen), but other spillover effects like willingness to purchase tickets in future seasons and perhaps even patronage in the form of donations. While we hope to be able to report on such an experiment in the future, the model at present can be used "out of the box" by organizations that offer customizable bundles to improve their implementation of and eventual success with that critical customer-facing feature.

The model can also be applied in settings where consumers make multiple bundled purchases of similar products or experiences for future consumption, ranging from "intangibles" like performances and sporting events, to consumables like snacks, meal kits, and wines, through durables like books. Because customization increases the cognitive and attentional resources required of consumers, it can discourage them from committing to a full-sized bundle, or even purchasing anything at all. However, with more relevant, customized recommendations that incorporate individuals' component item preferences along with preferences for variety, firms can lower the entry barrier to purchasing customized bundles for consumers and reduce costs involved with return and replacement.

1.8 Appendix

1.8.1 Model results

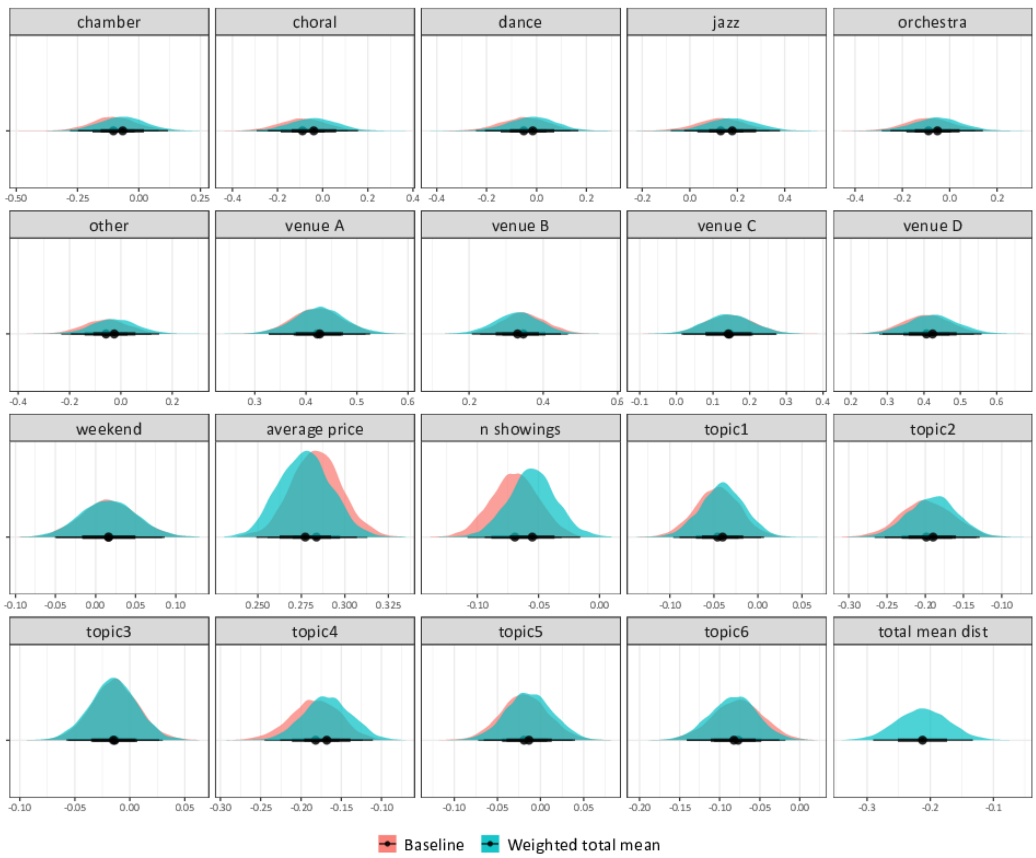

Figure 1.3 Posterior distribution of the mean effects: baseline and weighted total mean

Table 1.15 Average predictive gains (in percentage) over baseline models across users (mean across draws at the individual level)

	Users with improvement			Average gains 1% trimmed			Average gains 5% trimmed		
	3rd	4th	W	3rd	4th	W	3rd	4th	W
Total mean	101	114	130	16.3	10.41	27.63	6.43	6.51	21.99
Total Min + Max	121	117	117	22.31	19.93	24.2	14.64	13.66	17.53
Delta mean	106	104	120	15.12	7.59	21.03	9.42	3.83	16
Delta Min + Max	114	117	115	19.38	19.85	24.58	11.73	14.82	17.06
Incremental mean	114	120	112	17.01	17.08	13.69	13.48	11.7	9.94
Incremental Min + Max	119	116	116	19.73	12.59	25.05	13.73	9.23	18.26

* "W" = Weighted, N=229

1.8.2 Simulation studies

Here we discuss simulation results verifying parameter recovery using the Monte Carlo approach used for scalable estimation of the proposed model of custom bundle choice. Specifically, the approach entails randomly selecting and fixing one order of bundle addition, separately for each user, and thereby estimate the probability of choosing the bundle. That is, rather than summing over all probabilities arising from all possible unobserved orders (120 for bundles of five, but in general potentially far more), we "stochastically select" a separate order for each user, and thereby marginalize over them.

For the first simulation, we generated a relatively simple and small dataset, with 10 performances from which 400 users each chose five. The attributes included three generated topics and total mean distance computed for all possible bundles of five (${}_{10}C_5 = 252$). The specific steps involved are as follows.

Step 1. Attribute generation: for ten performances, three topics were generated using a Dirichlet distribution: $(x_1, x_2, x_3) \sim Dirichlet(1, 1, 1)$, where $\{x_1, x_2, x_3\}$ are topic loadings for each performance.

Step 2. The values of total mean distance (variety metric) for all possible (unordered) bundles of five performances were computed. Both topics and variety metrics are standardized.

Step 3. Parameters were set to have moderately differing means and variances:

$$\beta_{1i} \sim N(-0.041, 0.23)$$

$$\beta_{2i} \sim N(-0.14, 0.42)$$

$$\beta_{3i} \sim N(0.18, 0.34)$$

where β_{1i} is user i's coefficient for topic 1, β_{2i} for topic 2, and β_{3i} for total mean distance, and $i = 1, \cdots, 400$.

Step 4. Individual users' probabilities of choosing all possible (unordered) bundles were computed using

$$P_i(b|\beta_i) = \sum_{r \in R_b} P_i(r_1) \prod_{k=2}^{5} P_i(r_k | r_1, \cdots, r_{k-1}) \; \forall b \in B$$

where b is any possible unordered bundle of five performances in B, which includes 252 bundles. r refers to specific order of addition to the bundle, such that r_1 is added first to the set, r_2 second, and so on. R_b is the set of all permutations of performances in bundle b, hence $|R_b| = 5! = 120$.

Step 5. Each user chose one of the (unordered) bundles in a multinomial fashion, with computed probabilities as in Step 4.

Step 6. For each user, a random permutation order ($r_i \in R_{b_i}$) was drawn and estimated.

Step 7. Step 6 was repeated 10 times.

Estimates from the simulation appear in 5 (means) and 6 (heterogeneity), where the solid black lines are true values of the parameters. All six parameters are recovered well.

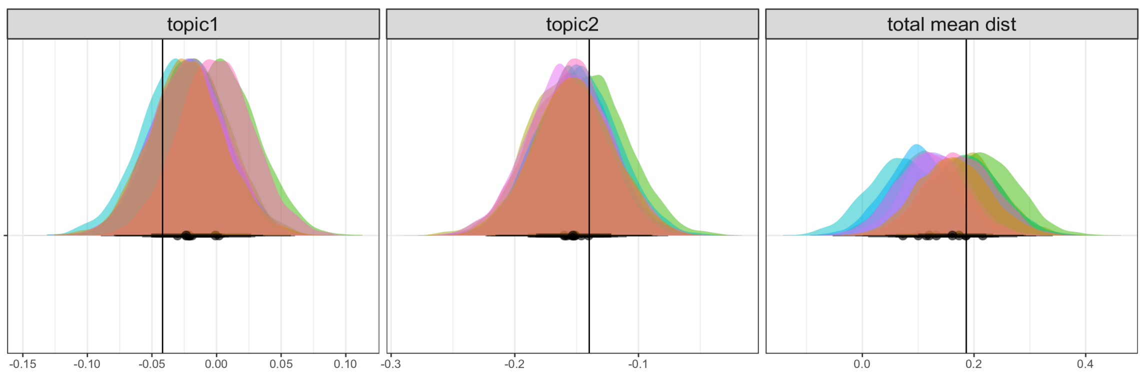

Figure 1.4 Recovery of mean effects: posterior distribution of the mean effects of topics 1, 2 and total mean distance

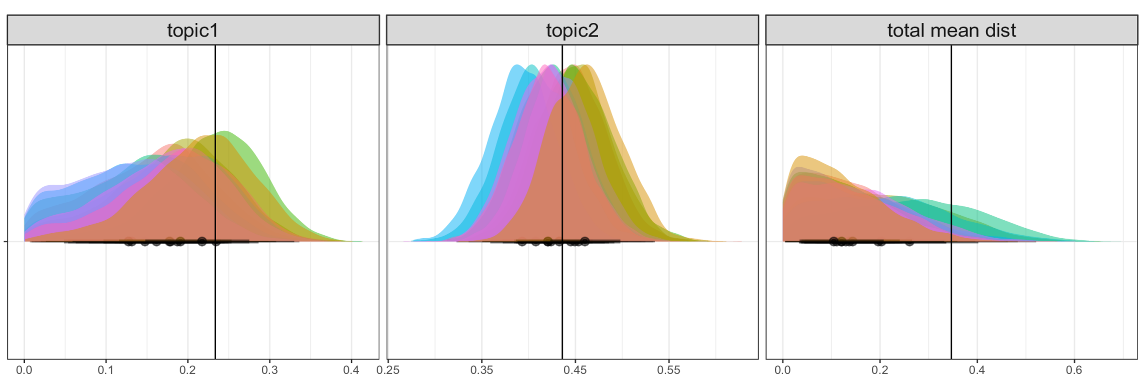

Figure 1.5 15 Recovery of heterogeneity parameters: posterior distribution of standard deviation coefficients for topics 1, 2, and total mean distance

For the second, data-based simulation, we used holdout season (season 2017) data with 45 performances and 229 users who made a bundle choice in the season. For the values of the true parameters, we picked one random draw from the posterior distribution of the coefficients in the model that used total mean distance as the variety metric, with variety kicking in from the third choice on.

Step 1. Using one draw s from the posterior distribution $(\beta_i^{(s)}, \gamma^{(s)}, \xi_i^{(s)}) \sim P(\beta_i, \gamma, \xi_i | X, Z, W)$ as the true parameter, we computed the probabilities of all possible bundles of five (1,221,759 bundles) by summing over all possible $5! = 120$ orders of addition following the equation below:

$$P_i\left(b | \beta_i^{(s)}, \gamma^{(s)}, \xi_i^{(s)}\right) = \sum_{r \in R_b} P_i(r_1) \prod_{k=2}^{5} P_i(r_k | r_1, \cdots, r_{k-1}) \; \forall b \in B$$

where R_b is the set of all permutations of unordered bundle b, and B is the set of all possible bundles of five among the 45 performances in season 2017 ($|B| = 1{,}221{,}759$).

Step 2. Based on the computed probabilities from Step 1, each of 229 users picked 1 out of 1,221,759 possible (unordered) bundles.

Step 3. For each user, a random order of addition for the five performances included in the bundle chosen in Step 2 were assigned.

Step 4. Parameters were estimated using the one random order of addition for each user.

Step 5. Steps 3 and 4 were repeated 50 times.

Given the large number of parameters (20 mean effects), we illustrate parameter recovery using plots of the posterior distribution of each coefficient (6, 7, 8; again, the solid black line indicates the true parameter value). While we find that all the parameters are well recovered, the variation of the posterior distribution around the true value is the largest for total mean distance. However, this is expected for several reasons: first, in the simulation, we allow users to have five choices per person, and using three of those choices to inform individual users' preferences for the parameters serves to limit the amount of information available. Secondly, variety metric is the only variable that changes values as the orders in which the items are added to the set, adding to the variation of distribution around the true parameter.

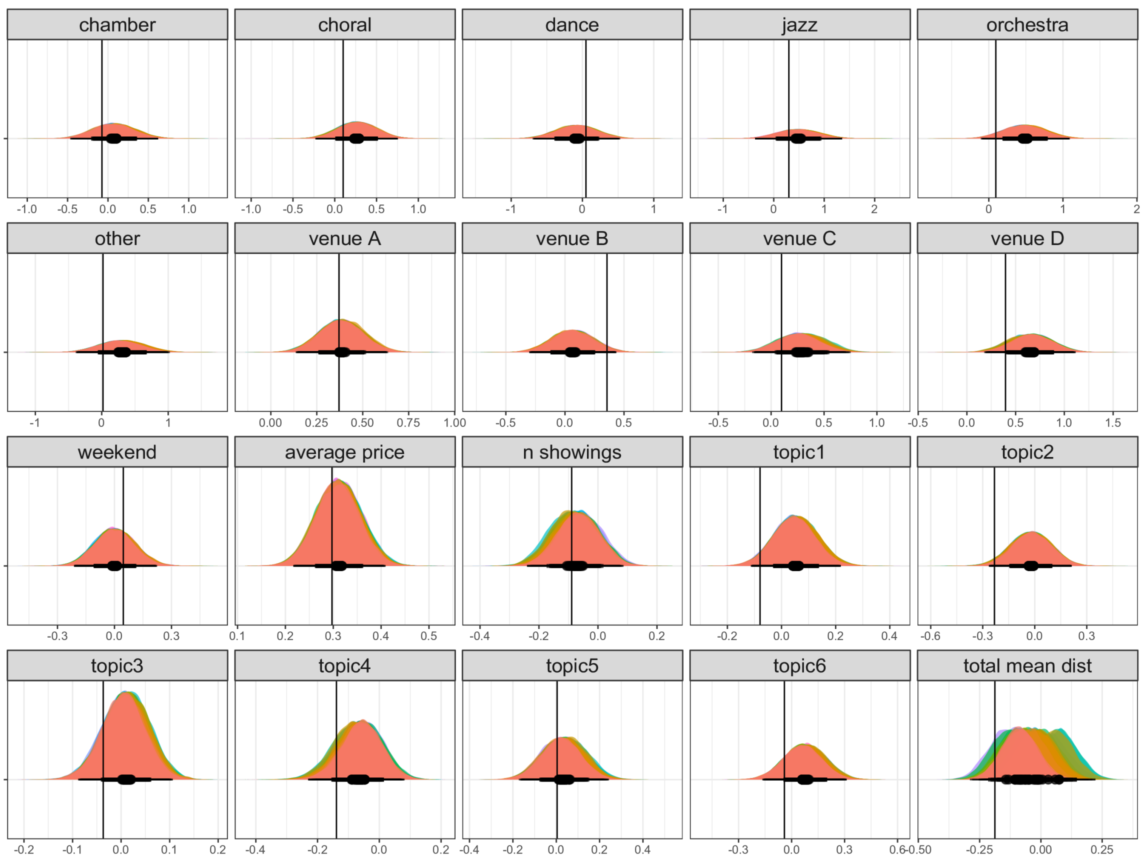

Figure 1.6 Posterior distribution of the mean effects

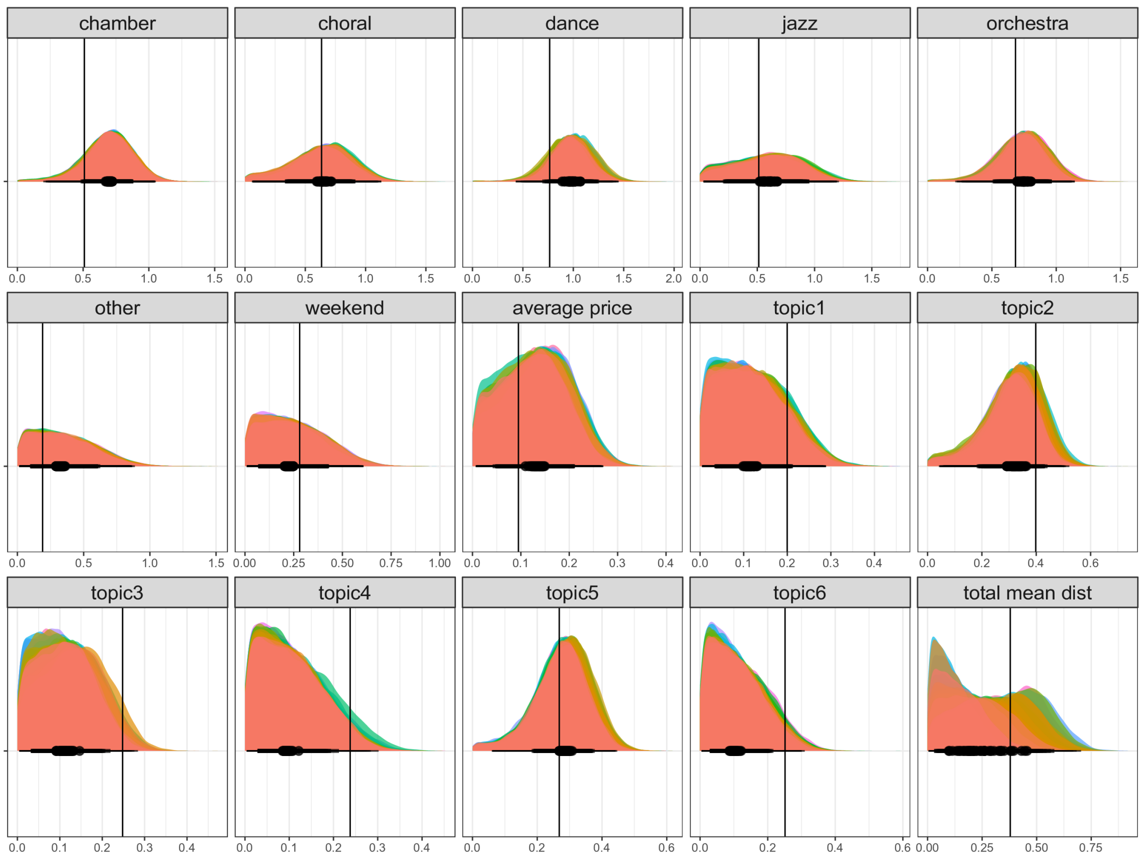

Figure 1.7 Posterior distribution of heterogeneity across users

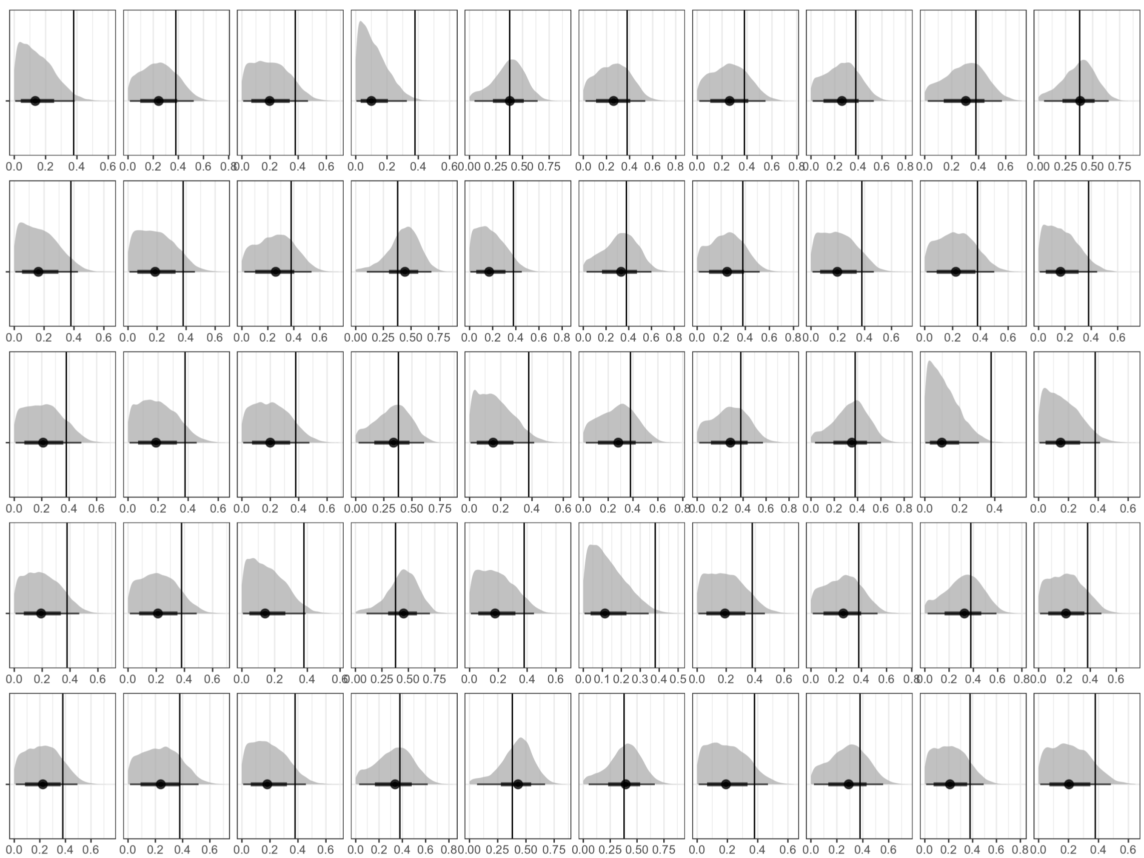

Figure 1.8 Posterior distribution of heterogeneity across users (for total mean distance)

Chapter 2 Cross-channel Price and Inventory Optimization for Live Events: An Application to NFL Ticket Purchases

2.1 Introduction

Dynamic pricing is a widely adopted practice across many industries, most prominently in hotels and airlines, and more recently in professional sports. In 2009, the San Francisco Giants first adopted dynamic pricing among professional sports franchises, and the team reported a 7% increase in revenue in that season ("Forty under 40: Barry Kahn," 2011). Since then, other professional teams followed suit, and dynamic pricing has become mainstream for professional sports franchises, moving beyond Major League Baseball (MLB), to include the National Football League (NFL) and National Hockey League (NHL). Another recent development in the sports ticket industry is the growth of legal digital secondary markets such as StubHub and SeatGeek. While teams initially resisted such changes, they eventually endorsed resale, sometimes partnering up with major secondary markets (e.g., NFL partnered up with StubHub as authorized ticket resale marketplace in 2017) (Courty & Davey, 2020).

Under these developments, primary and many secondary channels compete in the sports tickets market. The teams sell tickets through their primary channel, often in partnership with large online transaction platforms, the most prominent of which is TicketMaster. On secondary channels, resellers – including both ticket brokers and individuals – trade their tickets. While the two types of channels differ mainly in terms of seller identity, sellers in both types of channels engage in some form of dynamic pricing (Drayer et al., 2012; Sweeting, 2012). Given their sizable potential revenue implications, optimal dynamic pricing has received a great deal of attention. For example, Xu et al. (2019) propose optimal pricing policies for an MLB franchise based on a demand model for single game tickets and suggest that daily price re-optimization can increase revenue by 17.2%. Individual resellers and brokers can also easily engage in dynamic pricing through online platforms.

Despite the attention paid to dynamic pricing policies and channel effects, relatively less attention has been given to understanding how consumers choose among many possible channels. Given that searching across different channels has become easier, franchises could benefit from having not only an accurate estimate of future demand but also an understanding of supply across multiple channels. That is, as Zhu (2014) points out, franchises could improve on their revenue by incorporating competition from secondary channels. Further, given the numerous channels that compete in the sports tickets market, teams would benefit from understanding consumers' channel choice processes. Armed with this additional information, teams can refine their inventory and pricing policies through their primary channel, and as the team further considers entry into secondary markets, the potential to employ optimal inventory and pricing strategies there as well.

In this project, our research objective is to understand 1) the market dynamics – both pricing and inventory-wise – of sports event tickets across multiple channels, 2) ticket buyers' dynamic purchase decisions given market conditions (price and availability), and 3) consumers' channel choices. Based on this understanding, we aim to devise optimal distribution and pricing policies for the team – mainly for the primary channel, but also for secondary ones. Specifically, we study this issue from an (non-disclosed) NFL franchise's perspective.

NFL teams have a league-wide primary ticket exchange partner, along with multiple secondary channels, including major platforms such as StubHub, Vivid, and SeatGeek. The team collects rich data on transactions, events (e.g., team and opponent performance), and summaries of price histories across channels. Further, the team utilizes stadium characteristics to capture seat attractiveness for their dynamic pricing. However, the team does not have information on the details of inventory and prices on secondary markets (except for which tickets were bought, which are recorded through transaction logs), despite having a partnership with larger brokers who operate on multiple secondary channels.

To fill this gap, we collected a detailed supplemental dataset on NFL ticket market dynamics. This comprises listing information from three major platforms including the primary channel and three secondary channels[8]. Listing information includes ticket location (section / row), how many seats are offered, and at what prices. We fuse this data with the detailed transaction data, allowing us to roughly track which listings were sold, and which others remain unsold. Further, listings data across multiple channels allow us to model temporal evolution of prices and availabilities, all while taking the spatial relationship of the listed seats into account.

In the next section, we present a review of the relevant literature, followed by an overview of the datasets. Then we discuss the research problem in detail. We then present preliminary results on ticket listing prices, consumers' channel choices considering the market conditions, and dynamic ticket availability. We discuss extensions of the current models.

2.2 Literature Review

Our research focuses on the dynamic pricing of sports event tickets and related channel and seat choices. In this literature review, we discuss extant research on event pricing for both primary and secondary markets, as well as seat valuation and choices. Four major pricing

[8] The term "platforms" refers to particular ticket marketplace websites. One of the three ticket marketplace platforms operates both as a primary and secondary channel, resulting in a total of four channels across three platforms.

strategies are discussed in the literature: uniform pricing, tiered pricing, variable ticket pricing (VTP), and dynamic ticket pricing (DTP). Tiered pricing refers to pricing tickets based on location, compared to uniform pricing where all seats are priced the same. Variable ticket pricing takes a step further from tiered pricing by reflecting predicted demand for the event. While variable ticket pricing and dynamic ticket pricing share the idea that tickets should be priced based on demand, variable ticket pricing sets the price in advance, whereas dynamic pricing reflects fluctuating demand over time as the event approaches.

Leslie (2004) introduces a structural model of demand for a Broadway play and showed that tiered pricing can improve the firm's revenue by 5% relative to the uniform pricing policy, without a substantial impact on aggregate consumer welfare. Courty & Pagliero (2012) find a similar effect size of tiered pricing in concert settings. In the domain of sports event tickets, Rascher et al. (2007) show that using variable ticket pricing would have resulted in up to a 2.8% increase compared to fixed pricing, and as much as an extra $1.4M in revenue. Xu et al. (2019) model primary demand for single-game tickets for an MLB franchise and design dynamic pricing policies for an anonymous MLB franchise. They find that higher home team performance leads to lower price sensitivity, but section choices are not significantly affected by team performance. Further, they find that demand generally increases as the game day approaches, especially in the last two weeks prior to the game. Their results imply that the franchise can gain up to a 17.2% revenue improvement with flexible daily price re-optimization. Arslan et al. (2022) model college football ticket purchases across distinct segments with heterogeneous customers with multiple sales channels (subscription vs. single-game tickets), finding significant within-segment heterogeneity in price sensitivity and distance to the field. They suggest that price optimization based on their findings could provide a revenue increase of as much as 7.6%.

On the other hand, Tereyağoğlu et al. (2017) model the ticket sales of a non-profit performance organization and evaluates the impact of a decreasing monotone discounting (DMD) pricing policy under two types of customers – subscribers and occasional customers – using a competing hazards framework. They find that purchase patterns differ by different types of customers, including the timing of ticket purchase and the degree to which they value the discounts, as well as by the price tier of the tickets. Further, they find that commitment to the DMD policy can increase revenue per concert by as much as 6.7%.

Expanding the perspectives to secondary markets, Sweeting (2012) studies dynamic pricing behavior of secondary sellers for MLB tickets using MLB transaction data from two large secondary channels. He shows that, consistent with theoretical models of dynamic pricing, sellers decrease their ticket prices as the game nears because the opportunity cost of holding tickets decreases over time. Further, he finds that buyers are not strategic in the secondary market, and estimates that dynamic pricing can increase secondary market seller's payoff by 16% over fixed pricing. Zhu (2014) uses both primary and secondary ticket data from an MLB franchise and proposes optimal dynamic pricing. He studies the revenue implications of optimal pricing for the franchise when competition from secondary market exists under two conditions: when all consumers are strategic vs. not strategic. He finds that the franchise revenue could increase by up to 6.9% under nonstrategic consumers (up to 3.7% under strategic consumers) by dynamically pricing tickets as the game approaches.

Whereas previously mentioned research studies pricing strategies and suggest prescriptive optimal pricing policies, Shapiro & Drayer (2014) study factors that affect the dynamic ticket prices on the primary market and secondary ticket prices on StubHub for the San Francisco Giants, which was the first team to adopt dynamic ticket pricing in 2009. They find that team and individual performance variables, time-related factors (e.g., weekend game, days until game) and ticket-related factors, especially the seat locations, play significant roles in explaining the dynamic prices on primary channel, as well as the secondary ticket prices. Paul & Weinbach (2013) echo the findings of Shapiro & Drayer (2014), and identify variables such as weekday, performance, opponent, and promotions as factors that influence dynamic ticket prices across four MLB franchises. Further, Courty & Davey (2020) use panel data over 20 years and evaluate the impact of different pricing policies adopted by MLB franchises on their revenue and team value. They find that variable pricing increases revenue and team value by 4.2% and 9.5%, whereas introducing dynamic pricing and sponsored secondary markets (franchises partnering with secondary markets) had no significant effect on revenue or team value.

Another stream of research focuses on the impact of secondary markets for primary market sellers. Geng et al. (2007) study the impact of resale on primary seller's profit and proposes that resale does not necessarily hurt their profits. They propose a two-period model of ticket resale, and find that partial resale, where one can only resell his/her tickets in advance but not on the spot, can be profitable under limited capacity, small enough high-valuation buyers, and not too large number of early arrivers. Cui et al. (2014) study the impact of ticket resale on event organizers' revenues, using a model that varies resale transaction costs for consumers and speculators. They investigate scenarios under which resale market may be beneficial for the event organizers; for example, under fixed pricing policy, lower resale transaction costs benefit the firm. Lewis et al. (2019) empirically evaluate the value of secondary market options for the franchise and for the season ticket buyers using a structural model of supply and demand on the secondary market. They find that secondary markets are beneficial for both the team and the individual ticket holders. The option increases the season ticket revenues by $1 million per season for the team, and average monetary value of secondary market option for individual ticket holders is $138 per season ticket.

Lastly, our research incorporates users' seat choice decisions, and thus relates to stream of research on seat valuation and seat preferences. Veeraraghavan & Vaidyanathan (2012) study the value of seats in the stadium based on the location of the seat as well as consumer characteristics. Using survey data on the post-consumption experience, they propose the Seat Value Index (SVI) as a measure of seat value. They find that seat location attributes account for a major portion of the SVI, showing the importance of the seat location in consumers' seat valuation. Blanchard et al. (2020) study locational choices in event venues and captures heterogeneity across users' preferences to different aspects of seat choices, including proximity to other people and to focal elements of the event (e.g., screen or stage) using experimental data on seat location choices. They then show that the model estimates can help improve expected occupancy of the event by altering seat availability.

2.3 Data Description

We use two datasets on main market dynamics and three datasets that cover institutional details and additional user behavior. These datasets allow us to paint a highly detailed picture of the NFL ticket market. The first dataset provides detailed information on individual-level ticket transactions across different platforms. The second dataset includes snapshots of ticket availability and pricing on primary channel and three secondary channels. Additional datasets provide details of the events and the seats, along with partial observation of user browsing behavior.

2.3.1 Transaction data

The transaction dataset provides information on individual-level transactions over three seasons (seasons 2018/2019/2021). The team sells season tickets in advance (~30,000 season tickets are sold each season), followed by single-game ticket sales, where seats for individual games can be bought separately. This dataset only includes single-game ticket sales.

Single-game tickets can be bought either through the official partner (i.e., through the primary channel) or from resellers on secondary channels. Resellers include individuals who have tickets (e.g., season ticket buyers who want to sell part of their season ticket bundle) and professional brokers who advance purchase large number of tickets and sell them on multiple secondary markets. While the identity of resellers is not observed, we observe single-game ticket transactions across 15 channels.

The unit of observation is seat-level transactions; for each transaction, we observe the seat location (section, row, and seat number), paid ticket price, order date, and the channel on which the transaction took place. Over three seasons, we observe a total of 540,036 ticket transactions by 157,279 individual ticket buyers, after dropping users with $0 transactions and post-game transactions, and 28 users who had more than 100 tickets during the observation period. These 28 users made 13,432 ticket purchases in total, which suggests that these users could potentially be brokers rather than individuals purchasing tickets for their own consumption.

Of the 15 channels that are tracked, one is the primary channel, where buyers can purchase tickets through the team's official partner, and the remaining 14 channels are secondary channels that vary in their sizes. Note that we use the term platform and channels interchangeably in most cases except for Platform A, which serves both as a primary and secondary channel. We treat Platform A – primary and Platform A – resale as separate channels on the same platform. All other channels correspond to respective platforms and need not be distinguished between channels and platforms.

Primary channel accounts for 23.7% of the transactions across the three seasons that we observe, and its market share is growing over time. For secondary channels, there are only four channels with market share over 5%. Table 2.1 shows summary statistics of transactions on different platforms by season. Noticeably, Platform D, 4th largest in overall market share, had a big dip in its market share in 2019 due to contractual issues with the team.

While we observe transaction data over three seasons, only the latter part of the last season overlaps with the snapshot data (described in detail below) that provides information on the supply-side market conditions (i.e., availability and pricing). Restricting to the overlapping period (plus the last day of the season), a total of 26,253 ticket transactions were made over 54 days by 9,088 users, out of which 7,598 users (83.6%) had previous ticket transactions in the past. During this period, 60.66% of the transactions took place on Platform A, followed by Platform B (18.5%), Platform D (7.16%), and Platform C (5.8%).

2.3.1.1 Channel usage

We compute channel usage patterns using transaction data. Since observations are made at the seat level, we define a *transaction* as a set of ticket purchases that takes place on the same date, in the same section and row, on the same channel, for the same game. Of 15.12% ($N = 23{,}773$) of all users with at least two transactions, 49.54% of users ($N = 11{,}776$) use two channels (among two-channel users, $mean_{npurchases} = 2.49$), followed by 46.73% of users ($N = 11{,}109$) using only one channel (among single channel users, $mean_{npurchases} = 2.26$). That users tend to use a limited number of channels suggests that it is likely that previous channel usages significantly affect subsequent channel choices and that users may not browse all available channels.

2.3.2 Market snapshots

The dataset was collected from three major ticket transaction platforms. These platforms represent 86.87% of the market for the focal team across three seasons. The dataset covers a period of 54 days in the latter part of 2021 season (from November 2021 – January 2022). During this period, four home games took place and up to 16 snapshots were collected on each channel. The number of snapshots available for each game varies across channels based on the game date and the interval at which the market snapshots were taken. The intervals between snapshots range from 1 to 14 days. Of the four games that happened during the observation period, we focus on the last two games that have five or more snapshots on each platform.

Among the three platforms, Platform A serves both as a primary channel as well as a secondary channel, and the other two (Platforms B and C) are secondary channels. Each snapshot captures ticket availability and prices for the remaining games on a certain day on a specific platform. Across all four channels, for the remaining games, we observe how many consecutive seats are offered at which location, and the listed price for each. From this information, we construct listings information, defined as a set of consecutive seats offered by a seller in a specific row within a section for a specific game. All tickets within a listing are priced the same, such that if a seller decides to change the price of her listing, the price of all the seats within the listings will change. While a listing can include one or more tickets, not all tickets in a listing need to be sold together; they could be sold in smaller batches. In most cases, however, sellers do not allow selling $(N-1)$ tickets out of the N tickets that they offer, as a single remaining ticket could be harder to sell.

2.3.2.1 Platform A

Platform A is the most popular of the three platforms (61.85% of all transactions take place on Platform A), and the only one that offers both primary and resale tickets. On the website, the two types of tickets are offered side-by-side, so that buyers can easily browse across ticket types, and buyers can easily distinguish the two types of tickets. Platform A provides seat-level ticket information. Snapshots also include which seats are available for a specific game on a specific day. This allows the most direct comparison of the availabilities and prices across time to me made, both for users and for researchers.

While primary and secondary tickets are on the same page, we treat the primary and secondary as separate channels, as the sellers, as well as inventory (availability) and price trends differ across the two types of channels. However, it is natural to assume that the buyer will consider buying from either channel when they are browsing or purchasing tickets on this platform.

For both channels on the platform, we observe 6 and 13 snapshots for the two focal games, respectively, after excluding one snapshot from each game. The exclusion was due to an anomalous fluctuation where a significant drop in the number of observations was followed by a significant increase in the following period, reaching a level that is close to the snapshot taken two periods before. The decision was based on consultation with the focal NFL team, which suggested that this is likely due to an error during the data collection process.

The two games are 21 days apart and observed during roughly the same calendar period. Figure 2.1 shows snapshot availability across games and channels. Game X, which took place earlier in the season, was observed as far out as 33 days before the game, and as close as 2 days before the game on Platform A, and Game Y was observed 13 times, from 40 days to 1 day before the game.

Table 2.1 Summary statistics of market share and price distribution across channels

Channel	Season	Share	N users	Mean price	SD price
Platform A - primary	2018	19.858	14805	116.948	52.655
Platform A - primary	2019	24.051	18025	120.358	55.632
Platform A - primary	2021	28.504	15080	118.045	51.655
Platform A - resale	2018	42.973	28927	124.257	81.474
Platform A - resale	2019	35.662	22937	116.227	75.551
Platform A - resale	2021	35.024	16614	106.686	73.643
Platform B	2018	16.904	9848	94.08	58.305
Platform B	2019	25.285	16236	117.167	76.64
Platform B	2021	16.514	7585	107.861	72.836
Platform C	2018	3.807	2483	89.531	50.04
Platform C	2019	5.489	3758	115.707	67.497
Platform C	2021	6.319	3077	97.174	59.005
Platform D	2018	11.371	6185	99.76	54.612
Platform D	2019	0.577	330	151.283	77.157
Platform D	2021	6.838	2741	103.516	66.009
Platform E	2018	1.651	883	95.607	51.296
Platform E	2019	1.218	670	93	50.421
Platform E	2021	1.803	687	90.196	52.489
Platform F	2018	2.071	1181	107.848	65.14
Platform F	2019	2.043	1159	110.931	63.758
Platform F	2021	1.267	568	104.062	62.557
Platform G	2018	0.59	262	99.362	51.317
Platform G	2019	1.097	530	110.94	61.53
Platform G	2021	0.483	191	103.125	57.849
Platform H	2018	0.307	89	111.576	55.03
Platform H	2019	0.77	366	109.233	46.962
Platform H	2021	0.277	74	96.978	45.906
Platform I	2018	0.373	213	73.204	49.178
Platform I	2019	0.456	294	77.328	47.242
Platform J	2018	0.021	14	191.548	79.95
Platform K	2018	0.015	7	95.828	44.551
Platform K	2019	0.613	374	84.211	46.985
Platform L	2018	0.059	4	88.871	56.43
Platform L	2019	0.016	2	64.954	51.193
Platform M	2019	2.715	1739	81.268	43.941
Platform M	2021	2.973	1281	77.595	41.015
Platform N	2019	0.007	4	107.212	26.942

Across the two games, the two channels differ substantially in how the price and inventory trajectory evolves over time. Figure 2.2 shows price trend across platforms. Across the two games, per-ticket price is consistently higher for Game Y on both channels ($mean_X = \$108$ vs. $mean_Y = \$156, p < 0.001$) and tickets are generally priced higher on resale channel ($mean_{primary} = \$128$ vs. $mean_{secondary} = \$145, p < 0.001$). Also, whereas the price trend is relatively flat for both games on the primary channel, there is more fluctuation on resale channel, with a sharp drop in price before the game. In terms of inventory (Figure 2.3), there are slightly more available primary tickets than secondary tickets across time ($mean_{primary} = 5025$ vs. $mean_{secondary} = 2809, p < 0.001$). The primary channel shows a sharp increase in inventory, whereas the resale channel generally has a downward trend.

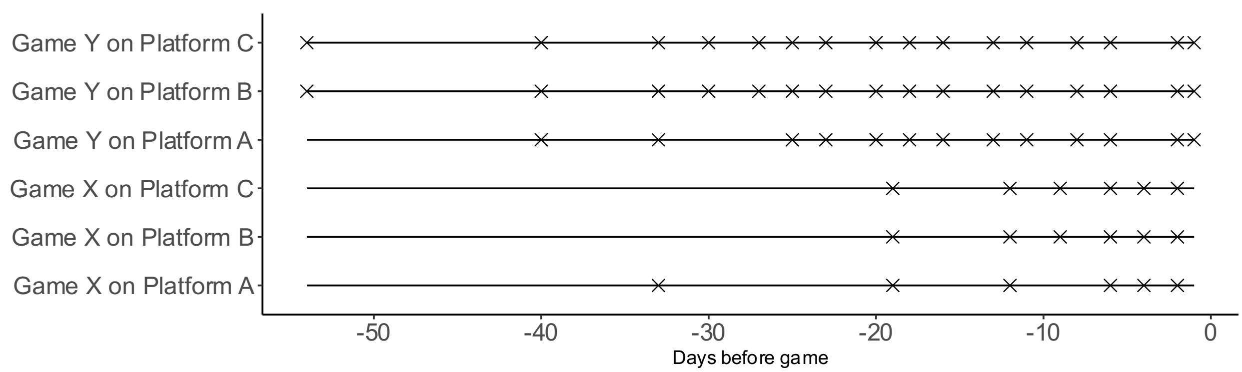

Figure 2.1 Market snapshot schedule

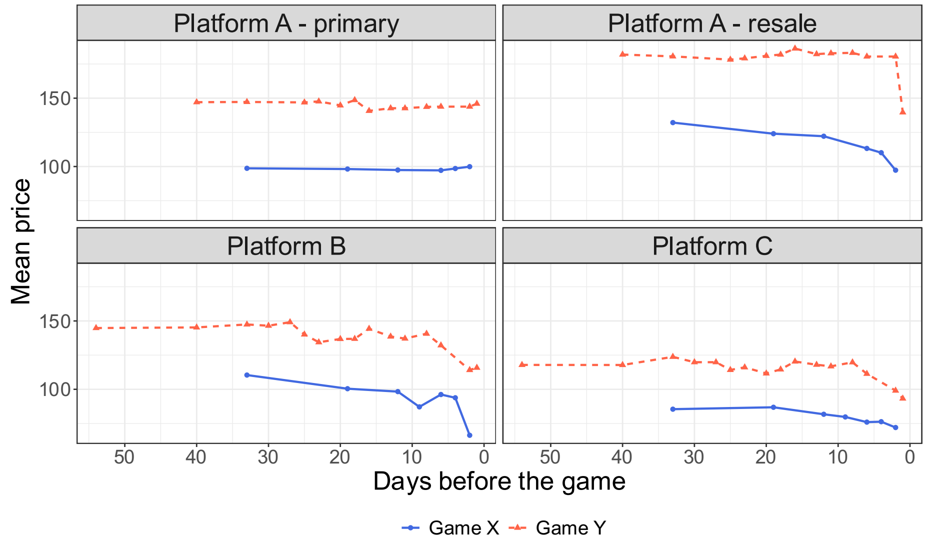

Figure 2.2 Listing price trend across channels

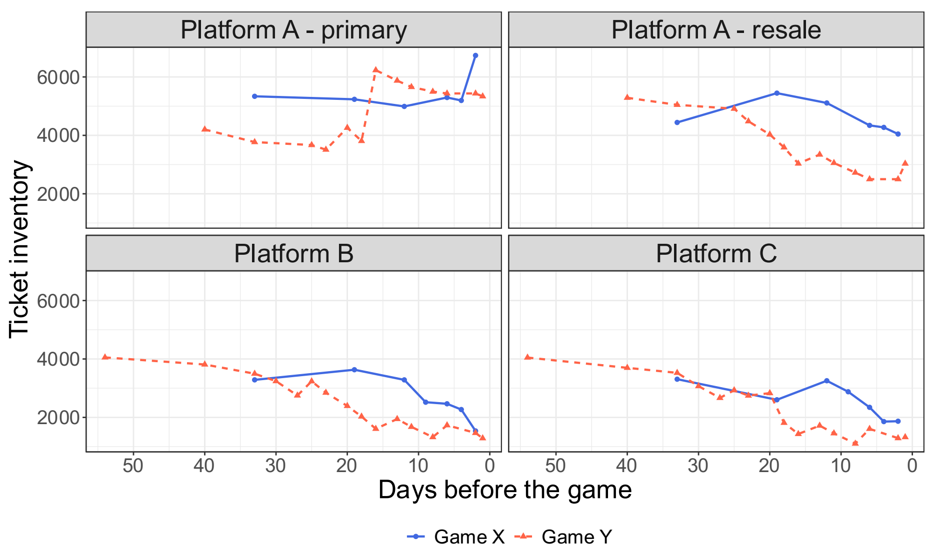

Figure 2.3 Inventory trend across channels

2.3.2.2 Platform B

Platform B is a resale-only platform that accounted for 19.9% of all single ticket transactions over the past three seasons. Resellers, including both brokers and individual resellers, post their tickets at their desired prices, and they can choose to remove their listings.

The platform provides information on the listing location (section, row, number of consecutive seats, and seat ranges) to buyers. However, buyers do not have full information on where their seats will be until they check out, unless they purchase the entire listing. Unfortunately, seat ranges (specific seat numbers) are not included in the snapshot, and we only observe listing-level details such as the aggregate location (section and row), price, minimum, and maximum number of tickets available for each listing.

On Platform B, we observe instances of tickets priced at $9999. While we cannot fully rule out the possibility that these are pricing strategies of individual sellers trying to prevent the seat from being sold, NFL team experts suggested that these are systematic posting errors that arise from the posting process. Therefore, we dropped all $9999 observations, which ranged from 0 to 120 offerings per snapshot (in the case of 120 offerings, they translated to 37 listings, which comprised under 5% of listings available for that snapshot).

For the focal games X and Y, we observe 7 and 16 snapshots, respectively. Game X is observed as early as 33 days before the game, and as close as 2 days before the game, whereas Game Y is observed over 53 days, from 54 days before down to 1 day before the game. Price and inventory patterns on Platform B are similar to those of Platform A – resale. Ticket prices are stable until they get closer to the game, and then they decrease, with sharper decreases right before the game. Further, per-ticket prices are generally higher for Game Y than for Game X ($mean_X = \$96.1$ vs. $mean_Y = \$140, p < 0.001$). Inventory, as is also true for Platform A – resale, generally decreases over time with no significant differences in the number of tickets available across the two games.

2.3.2.3 Platform C

Platform C is another major resale platform that allows brokers and individual sellers to post their tickets, and 5.09% of all single-game transactions across three seasons were made on this channel. The platform only offers aggregate location of the tickets (section/row) and how many consecutive seats are available, rather than specific seat locations within section/row.

Again, for the focal games X and Y, we observe 7 and 16 snapshots, respectively, on the same days on which snapshots on Platform B were collected. The general price patterns follow other resale channels; prices are higher for Game Y than X ($mean_X = \$80.2$ vs. $mean_Y = \$116, p < 0.001$). There is no significant difference in the number of available tickets for the two games. Summary statistics of price and inventory are shown in Table 2.2.

Table 2.2 Summary statistics of price and inventory across channels

Channel	Game	Mean (price)	SD(price)	Mean(inventory)	SD(inventory)
Platform A - primary	Game X	98.375	43.62	5464.833	634.243
Platform A - primary	Game Y	144.825	65.308	4821.769	963.819
Platform A - resale	Game X	116.418	64.755	4609.667	543.987
Platform A - resale	Game Y	178.276	101.202	3656.385	990.232
Platform B	Game X	96.131	55.592	2713.286	727.971
Platform B	Game Y	139.702	84.462	2430.938	925.328
Platform C	Game X	80.24	48.513	2587.857	599.72
Platform C	Game Y	115.628	72.319	2328.625	966.97

2.3.3 Other datasets

2.3.3.1 Event data

The demand for the event depends on various factors – including who the opponent team is, how the teams are performing in the season, what day the game is, etc. The event dataset provides such relevant event information: event date and time, weekday of the event, the focal team's, and opponent's cumulative win percentage over the past three seasons, how much the team spent on promotions, etc. One important variable is the event score, which is used by the team to understand and dynamically set the prices. It is a score based on the demand estimate for the game, taking various factors (e.g., seasonality, weather, opponent, etc.) into account, with an average score of 100, for an 'average' game in season 2019. This variable could serve as a benchmark against which we can compare our models.

2.3.3.2 Seat location data

The home field has a capacity of 65,000 and is divided into sections, rows, and seats. The seat location dataset provides information on the structure and quality of the seats, including the capacity of section-row, what amenities are available around the seat location, and the corresponding price-location ID. A total of 123 sections are available in the stadium (excluding the suites), with one side of the stadium being the home team side, and the other side the opponent team side. The stadium is also divided into price-locations, which reflect the price tiers of the seat locations for season ticket members. Price-locations and sections are not nested; price-locations can include one or multiple sections, and each section can include one or more price-locations. Figure 2.4 shows the layout of the stadium with sections and price-locations. Each line indicates a section, and the color of the line reflects the price-location. For example, in a fifty-yard line section close to the field, rows closer to the field are priced differently from those in the back (shown in blue in the front and green in the back). Thus, there are two price-locations within the fifty-yard line section. At the same time, there are two 50-yard line sections on both sides of the field – and the prices of these symmetric sections are the same, putting them in the same price-locations. Both sections are assigned the same two price-location IDs.

Additionally, we constructed row-level coordinates of seats based on the map of the stadium, allowing us to use the spatial structure of the seats to understand the ticket price distributions.

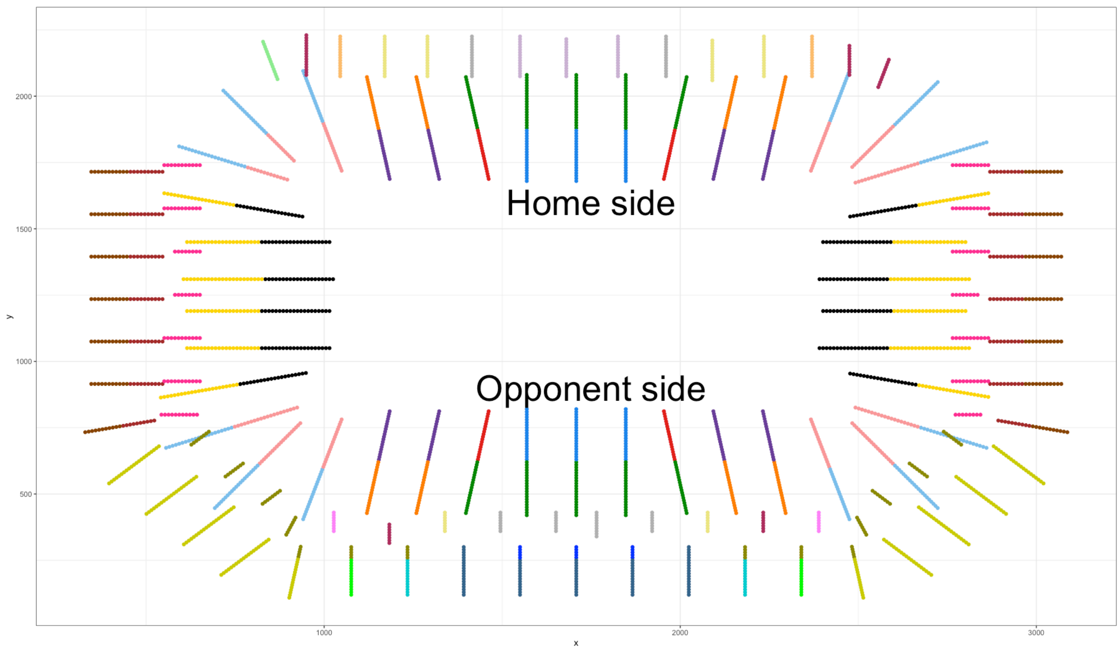

Figure 2.4 Section and price-location layout

Each dot represents the center of each row within a section, consecutive dots indicate a section, and the different colors represent different price tiers as indicated by the price-location ID. For example, there are three 50-yard line sections on each side, colored with green and blue; the rows colored in blue are higher tier seats compared to green-colored rows. Some sections show overlap with other sections; this is because some of the club-level seats overlap with rows in the back in the lower level on a 2D map.

2.3.3.3 Browsing data

We do not observe ticket buyers' full browsing history. Browsing dataset provides partial observation on browsing behavior for a subset of users. For a small subset of users who received an email from the team and clicked through, their browsing behaviors on Platform A are tracked and recorded, for seasons 2019 and 2022. The dataset includes 6,122 users' 11,918 search records; each observation includes when the search was created for which event. While this dataset provides only partial information about consumers' search behavior – that the user considered Platform A as part of their search process – it allows us to identify users who would have considered the two channels on Platform A during their channel choices, providing an extra piece of information to infer the latent channel consideration decisions.

2.4 Model Development

We break down the problem into three parts that are closely related to each other. The first part of the problem is modeling the market conditions – i.e., availability and pricing. Sellers decide whether to sell their tickets considering the expected demand and supply around the focal location of the tickets. For example, if there is limited supply with larger demand for tickets where the seller holds tickets, sellers could charge higher prices, which in turn could lead ticket holders to list their tickets for sale. On the other hand, when faced with larger competition for limited demand from surrounding listings, sellers could either decide to lower the prices or not to offer their tickets. Further, the temporal dimension should play an important role in deciding the availability and pricing decisions, as the opportunity cost of holding the tickets decreases over time (Sweeting, 2012). This part of the model captures the supply side market dynamics and governs what is available at a certain time.

The second and third parts of the model focus on dynamic demand and channel choices. Conditional on which game to attend and how many tickets they want to purchase, buyers choose where in the stadium they would like to sit. Based on where they choose to sit, buyers then make a channel choice. Since we do not observe their search process, we model their consideration sets (of channels) probabilistically, from which their final channel choice is made. The channel decisions are likely to be driven by channel-specific factors including ticket availability and prices as well as previous experience with different channels. We provide a detailed overview of model parts and how different parts relate to each other in the following.

2.4.1 Dynamic availability and pricing

There are 65,000 seats in the home stadium, out of which about half (~30,000) are sold as season tickets. Prior to season, part of the single-game tickets is sold to partnered resale brokers, who later post them on various resale platforms. The franchise releases the remaining single-game tickets on primary channel. Any ticket that has been sold through primary channel can be resold on secondary channels at any time, and the prices can be adjusted over time, allowing the resellers to easily engage in dynamic pricing. There are broadly three types of sellers: the team (through their primary channel partner), brokers (either partners of the team or smaller brokers), and individuals who have extra tickets, either from previous single ticket purchases or season tickets.

Resellers' decision to post their resale ticket at any time t depends on various temporal and spatial factors, as well as game-specific factors. For example, team performances could increase the demand for the ticket, which could induce resellers to make their tickets available on the market. Additionally, availability and pricing of tickets surrounding the focal seat location could affect resellers' decision to make their tickets available on the market. It is possible that certain sections have a large number of tickets available, which could induce competition among the listings and possibly drive the price down. In such cases, resellers may choose to wait until the excess supply is reduced, although such decisions would depend on the location of listings. For example, if the focal listing is in sections that usually don't sell out, they may not choose to wait. Lastly, these decisions to hold back or to release available tickets also depends on how much time is remaining until the game day. Since the resale tickets are perishable goods, resellers, especially if they do not intend to use the tickets themselves, will likely decide to post their inventory as the day of event gets closer.

Let Y_{srgt} indicate the decision of a seller with a ticket in section s, row r for game g will make the ticket available for resale at time t. At every period t, the seller decides whether to make the ticket available based on the demand for the ticket and the supply conditions of what listings are available near the seller's seat at what price. Let utility for posting the tickets be denoted U_{srgt}. It can be written as:

$$U_{srgt} = s_a(t) + s_a(r) + \psi_s + \beta_a X_{gt} + \delta_a W_g + \gamma_a Z_{srgt} + \xi_a A_{srgt-1} + \varepsilon_{a,srgt}$$

where $Y_{srgt} = 1$ if $U_{srgt} > 0$. The utility depends on $s_a(t)$, a flexible function over time, which allows us to flexibly capture the time trend that reflects the decreasing opportunity cost of holding on to the tickets (Sweeting, 2012), $s_a(r)$ and ψ_s, a smooth function over the effects of row and section fixed effects to capture the effect of seat location, X_{gt}, time-varying effects that affects demand, such as the focal team and opponent's recent performance or popularity, W_g, game-specific effects that can affect the demand, such as weekend or holiday games, Z_{srgt}, the availability and prices around the focal tickets that are being considered for resale and $A_{srg(t-1)}$ whether seat was available at time period $(t-1)$. An important empirical consideration is what consists as 'nearby' seats, and how we could use detailed locational information to regularize the effects of seats across the space.

Conditional on a seat being made available based on a variety of market factors, resellers set their prices for tickets in section s and row r (p_{sr}). Seat price would presumably be affected by the quality of the seat, as determined by the location. Over and above the location effects, the prices would be determined as a function of what is available around the tickets that are being offered and how much the prices are, as well as how much time is left until the game. Further, game-related factors, such as the recent performances of the focal team, or the opponent would influence the pricing decisions. The pricing model then can be written as:

$$\log(p_{srgt}) = s_p(t) + s_p(r) + \psi_p + \beta_p X_{gt} + \delta_p W_g + \gamma_p Z_{srgt} + \varepsilon_{p,srgt}$$

where $\begin{pmatrix}\varepsilon_a \\ \varepsilon_p\end{pmatrix} \sim N\left(\begin{bmatrix}0 \\ 0\end{bmatrix}, \begin{bmatrix}1 & \rho \cdot \sigma_p \\ \rho \cdot \sigma_p & \sigma_p^2\end{bmatrix}\right)$, following Heckman (1979).

2.4.2 Purchase decisions and channel choices

In modeling the purchase decisions and channel choices, we build on Xu et al. (2019) model that decomposes the demand into 1) game demand, 2) ticket quantity, and 3) seat section choice. On top of these three components, we add the layer of 4) channel choice at the last stage.

We assume the following order in consumers purchase decisions and channel choices:

Step 1. Buyer i arrive at time t, knowing which game (g) to attend and how many tickets (q_{igt}) to purchase – these users have decided to go through with the purchase, based on their evaluation on the current ticket prices, team performance, and overall availability.

Step 2. They choose seat tier s based on overall market availability and prices at time t.

Step 3. Based on their location choice s, they choose channel c based on the current prices, availabilities, and their previous experience with channels (X_{ic}).

In step 1, let N_{gt} be the total number of buyers who purchase tickets on day t for game g. The demand will be driven by game-specific factors and game-time-specific factors. Game-specific factors (W_g) would include the popularity of the opponent and day and time of the game. Game-time-specific factors (X_{gt}) include the performance of the home team and opponent during the season, time until game, and current availabilities and prices. Previous research found that the demand generally increases as the game date gets closer (e.g., Shapiro & Drayer, 2014; Xu et al., 2019), and we find that this is in line with the empirical pattern that we see in the data. We expect higher price and availability will have a negative effect on (realized) demand, as more availability could imply that the game is not as attractive or that buyers could delay their purchases. The effect of availability might also depend on time until game as buyers may be less likely to delay their purchases as the game approaches.

$$N_{gt} \sim Negative\ Binomial\left(\mu_{gt}^{D}, \kappa^{D}\right), \text{ where}$$

$$\log\left(\mu_{gt}^{D}\right) = s_{D}(t) + \beta^{D} X_{gt} + \gamma^{D} W_{g}$$

A second component of the demand is the number of tickets for user i, game g at time t, q_{igt}. Again, we can use negative binomial regression to model the quantity needed using game-specific factors (X_g) and game-time-specific factors (W_{gt}):

$$(q_{it} - 1) \sim Negative\ Binomial\ \left(\mu_{t}^{Q}, \kappa^{Q}\right), \text{ where}$$

$$\log\left(\mu_{t}^{Q}\right) = s_{Q}(t) + \beta^{Q} X_{gt} + \gamma^{Q} W_{g}$$

Upon arrival (conditional on game choice and quantity choice), buyers choose where they would like to sit in the stadium. This decision is based on the current market availability and prices across channels; specifically, factors such as availability, price, and number of listings could affect their choice of seat location. We could use a multinomial logit model to capture the seat tier choices. Let the utility of chosen seat tier s be U_{sgt}:

$$U_{sgt} = \beta^{s} X_{sgt} + \xi^{s} Z_{s}$$

X_{sgt} includes variables such as occupancy, number of listings available at that time in the section, and the average price of tickets, whereas Z_s includes section-specific variables including the level of the section (lower, club, or upper) and whether it is close to various types of amenities. The probability of choosing seat tier s is:

$$P(s|g,t) = \frac{\exp(U_{sgt})}{\sum_{l} \exp(U_{lgt})}$$

In the last stage, buyers make a channel choice based on the current availabilities and prices. However, note that they might not consider all available channels at this stage. Some of them may have a strong preference for the channel that they've used before, whereas others may be highly price sensitive, in which case they might decide to explore all available channels. We take the probabilistic consideration set approach where the probability of choosing channel j is the sum of probability of choosing channel j out of set of considered channels C:

$$P(j) = P(j|C)P(C)$$

Conditional on the consideration set C, the probability of choosing channel j is modeled as a multinomial logit:

$$P(j|C) = \frac{\exp(U_{ijsgt})}{\sum_{c \in C} \exp(U_{icsgt})}, \text{ where}$$

$$U_{icsgt} = \delta_c + \beta^c X_{csgt} + \gamma^c Z_{ic}$$

δ_c captures channel-specific intercept, X_{csgt} includes channel-section-time specific variables that include prices and availabilities in the seat location s, and Z_{ic} include user- and channel-specific prior experiences.

We make three assumptions about channel consideration sets: 1) users always consider channels that they have used or browsed before and 2) users who use either channel on platform A always consider both channels (given the structure of the platform that displays primary and secondary options side by side, this assumption is not very restrictive). Second assumption allows us to reduce the number of possible consideration sets from $|C| = 15 = 2^4 - 1$ to $|C| = 7 = 2^3 - 1$, alleviating concerns for computational burden to estimate P_C.

2.5 Preliminary results

In this section, we present preliminary results on models of listing prices, channel choices and dynamic ticket availability. The results on listing prices shed light on how listing prices differ across channels, over time, and space. We find evidence that listing prices differ significantly across channels, even after controlling for the seat quality (based on their location) and for the temporal effects. Furthermore, we find that primary channels and secondary channels differ substantially in how the prices evolve over time. Whereas the ticket prices increase slightly over time on the primary channel, resellers sharply decrease their resale prices as the game approaches on secondary channels. Our findings also indicate that seat quality, proxied by seat location, is the most important predictor of listing prices, over and above channel and temporal effects. However, recent home team performance (as measured by recent win/loss) does not significantly predict listing prices (Xu et al., 2019). These factors - channel, temporal, and spatial - are likely key drivers of buyer's decisions on whether, when, and from where they buy tickets.

The channel choice model reveals important differences across channels and drivers of channel choices. For example, price sensitivities and planning horizons (i.e., how far in advance users choose to purchase tickets) differ across channels. Regarding the drivers of channel choice, we find that previous experience with a channel has a strong effect on subsequent channel choices, and that customers tend to prefer channels with more listings (i.e., higher availability) in the area where they are looking to purchase tickets.

Lastly, the dynamic availability model captures how supply conditions in the vicinity of focal seats impact sellers' decisions to make their tickets available. We find that past supply conditions (prior availability and prices) significantly affect subsequent availability; for example, more tickets were available when neighborhood prices were high in the past period. Furthermore, we find that the effect of time until game has opposing effects on primary and secondary channels. Closer to the game date, availability increased on primary channel while it decreased on secondary channels. Ticket buyback on primary channel, which took place about two weeks prior to Game Y, led to a sudden influx of tickets on primary channel, increasing availability on the channel as it got closer to the game. The ticket buyback seems to not only have affected supply on the primary channel, but also on other secondary channels.

2.5.1 Listing prices

We model listing prices across the four channels using covariates known to affect ticket prices. We use generalized additive model (GAM, Hastie & Tibshirani, 1987) to flexibly capture the effect of time and space on listing prices. The results reveal interesting differences across channels and spatial locations, as well as the effect of recent team performance on the listing prices. We focus on the last game, with as many as 16 market snapshots for some channels (Game Y). We estimate the model using all availability snapshots except the last, and then make price predictions on the withheld snapshot. The unit of observation is listings for which we observe the channel, seat location, and the number of tickets. We also have information on how many days before the game the listing was available, and recent team performance. We also construct variables to capture availability and prices in the neighborhood, which we define by section × price-location ID. We measure availability using the number of listings on the same day in the same neighborhood, and prices using their average prices. We remove the focal listing price when computing the average prices in the neighborhood, such that if there are N listings, we use $(N-1)$ listing prices to construct the average prices after removing the focal observation. However, for about 5.5% of all listings that are the only listing available in the location, we broaden the range of location and use the average price of listings in the same price-location, as it reflects the ticket price tiers.

2.5.1.1 Model results

Table 2.3 shows incremental value of different types of covariates in predicting the listing prices. Comparing Models 2 with Models 3 and 4, we see that seat location plays an extremely important role in explaining the listing prices (adjusted R^2 jumps from less than 10% to close to 80%.) In Model 3, we use fixed effects for sections, along with a smooth GAM function for the row to capture possible nonlinearities in prices across different rows. For example, while we expect that rows closer to the field will be more expensive, people may think that the differences in seat quality are larger in the front rows than in the middle of the section. It may be that the difference in perceived seat quality may be much larger between rows 1 and 2 than between rows 21 and 22. On the other hand, in Model 4, we use smooth functions over section-row-specific coordinates of the stadium to capture the effects of location on seat prices. While the two types of location variables (section and row vs. coordinates of each row) convey similar information, we find that the former have slightly higher explanatory power. This is somewhat unexpected given that the coordinates allow spatial regularization across sections that are close to each other, whereas section fixed effects do not capture such spatial relationships. We use both types of location variables in our chosen specification.

To proxy for the recent team performance, we use home team's most recent win/lose record prior to the snapshot date. We find little difference in the predictive results when recent team performance is included. We also tested the outcome scores of recent games to represent recent performances, but the results did not change. It may be the case that the game outcome was as expected in this relatively short time frame, which could explain why it did not affect listing prices. Or, it could be explained along the lines of the findings of Xu et al., (2019) which suggests that better team performances do not directly affect buyers' inherent valuation of the product.

Another set of variables provides information on other available listings nearby, including average prices and the number of listings. These variables help identify how tickets may be priced in reference to what else is available and at what price. While the impact of these variables is relatively small, incorporating them into the analysis provides some incremental value.

Table 2.4 displays estimates for Model 12, which shows the best holdout predictive performance (the results are similar to Models 8-11). In terms of channel effects, we observe that there are significant price differences across channels, even after controlling for various location effects to capture the quality of the seat and channel-specific temporal effects to account for price changes over time. Specifically, we find that Platform A – resale is more expensive and both Platform B and C are less expensive compared to the primary channel. Given the relative ease with which buyers can browse for similar seats across different channels, such significant price differences are surprising. Such persistent price differences (which we observe across specifications) suggests that substantial differences exist across channels, possibly including the users thereof. Furthermore, in terms of availability-related variables, we observe that, as expected, the average price in a similar location positively predicts the focal seat price.

Table 2.3 Model comparison for listing prices

Covariates	Model 1	Model2	Model3	Model4	Model5	Model6
Time	s(.)	s(.) by channel	s(.) by channel	s(.) by channel	s(.) by channel	s(.) by channel
Channel	FE	FE	FE	FE	FE	FE
Location - Section			FE		FE	FE
Location - Row			s(.)		s(.)	s(.)
Location - Coordinates				s(.) on x, y	s(.) on x, y	
Recent win						yes
Nearby listings						
adj. R^2	0.080	0.082	0.844	0.783	0.851	0.844
CCV	0.248	0.247	0.042	0.058	0.040	0.042
RMSE on holdout	0.431	0.432	0.217	0.249	0.216	0.217
Covariates	**Model7**	**Models**	**Model9**	**Model 10**	**Model 11**	**Model 12**
Time	s(.) by channel	s(.) by channel	s(.) by channel	s(.) by channel	s(.) by channel	s(.) by channel
Channel	FE	FE	FE	FE	FE	FE
Location - Section	FE	FE	FE	FE	FE	FE
Location - Row	s(.)	s(.)	s(.)	s(.)	s(.)	s(.)
Location - Coordinates	s(.) on x, y	s(.) on x, y	s(.) on x, y	s(.) on x, y	s(.) on x, y	s(.) on x, y
Recent win	yes	yes	yes	yes	yes	no
Nearby listings		mean ticket price N listings 1(unique listing)	mean ticket price N listings 1(unique listing)	mean ticket price N listings 1(unique listing)	mean ticket price	mean ticket price
adj. R^2	0.851	0.855	0.855	0.855	0.855	0.855
GCV	0,040	0.040	0.040	0.040	0.040	0.040
RMSE on holdout	0.216	0.201	0.200	0.200	0.200	0.200

* s(.) refers to a smooth function

Table 2.4 Estimated model parameters for listing prices

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	7.3826	0.2474	29.8418	<0.0001
PlatformA-resale	0.0165	0.0038	4.4118	<0.0001
PlatformB	-0.0280	0.0034	-8.1671	<0.0001
PlatformC	-0.1414	0.0035	-39.9718	<0.0001
Average ticket price in the neighborhood	0.0827	0.003	27.5778	<0.0001
Ssection FE		YES		
B. smooth terms	**edf**	**Ref.df**	**F-value**	**p-value**
s(days until game) - primary	1.0001	1.0001	4.4626	0.0346
s(days until game) - platform A - resale	3.3324	4.1025	4.4539	0.0013
s(days until game) - Platform B	6.4131	7.4347	20.8613	<0.0001
s(days until game) - Platform C	8.0464	8.7187	23.7523	<0.0001
s(row)	8.9432	8.9987	178.4422	<0.0001
s(x)	8.9120	8.9970	118.7741	<0.0001
s(y)	8.9005	8.9965	52.8862	<0.0001

Figure 2.5 shows estimated smooth functions for channel-specific temporal trends. First, we see that the primary and secondary channels show radically different temporal trends (top-left panel vs. others); on primary channel, prices linearly increase, although the magnitude is not large. However, on secondary channels, we see that after earlier fluctuations, prices decrease sharply as the game approaches – starting around 10 days before the game. The contrast is in line with the conflicting patterns suggested by Courty (2003) and Drayer & Shapiro (2009); Courty (2003) suggests that "diehard fans" are price sensitive and therefore purchases tickets earlier, whereas "business professionals" who choose to make decisions at the last minute are less price sensitive, which is in line with the primary channel price trends. Drayer & Shapiro (2009) found that secondary market prices decreased as it got closer to the game. Secondly, as we separately control for the temporal trend across channels, we are accounting for the last-minute deep discounts of tickets on the secondary markets. This in turn suggests that the channel-specific

price differences that we observe in the results are unlikely to be driven by price differences in a specific time window.

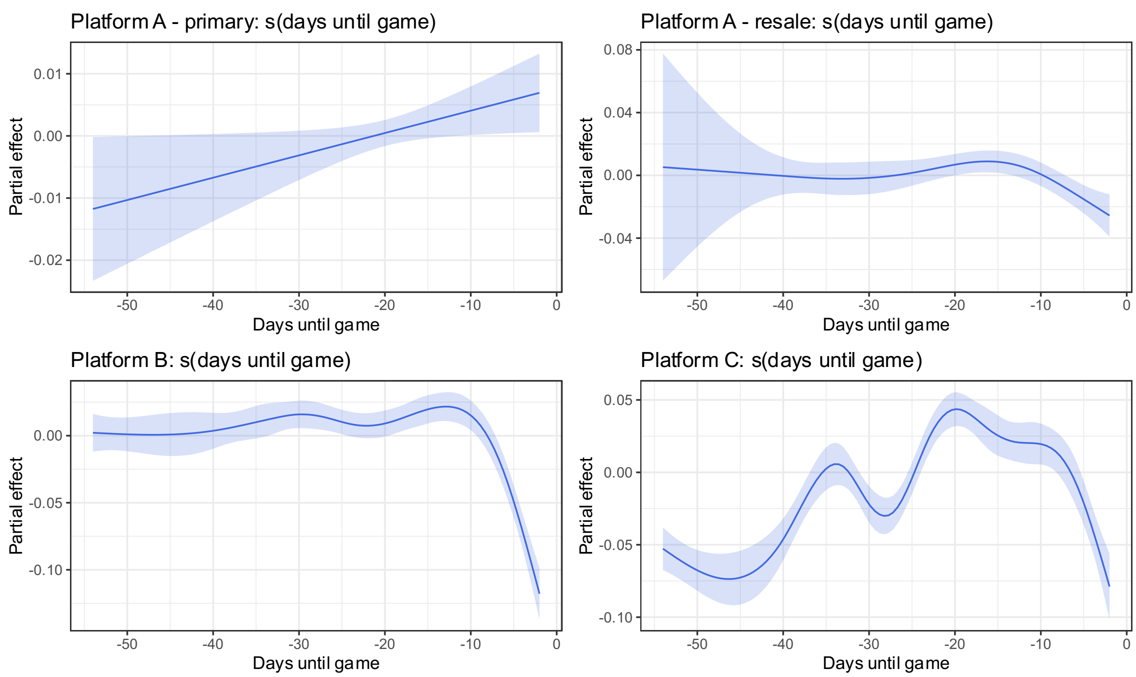

Figure 2.5 Partial effect of days until game on listing prices from listing price GAM

Figure 2.6 shows estimated smooth functions for the spatial effects of seats on listing prices. Left most plot shows the effect of rows, the middle plot depicts the effect of x-coordinates (length of the stadium), and the right one shows the effect of y-coordinates (different sides of the field). In terms of row effects, as expected, we see a generally decreasing trend as the seat moves toward the back of the section. Specifically, ticket prices decrease sharply in the first few rows and last few rows, whereas the middle part of the sections is relatively flat. This is well in line with the principle of diminishing sensitivity, where people are more sensitive to changes around the boundaries and less so away from the boundaries (Tversky & Kahneman, 1992).

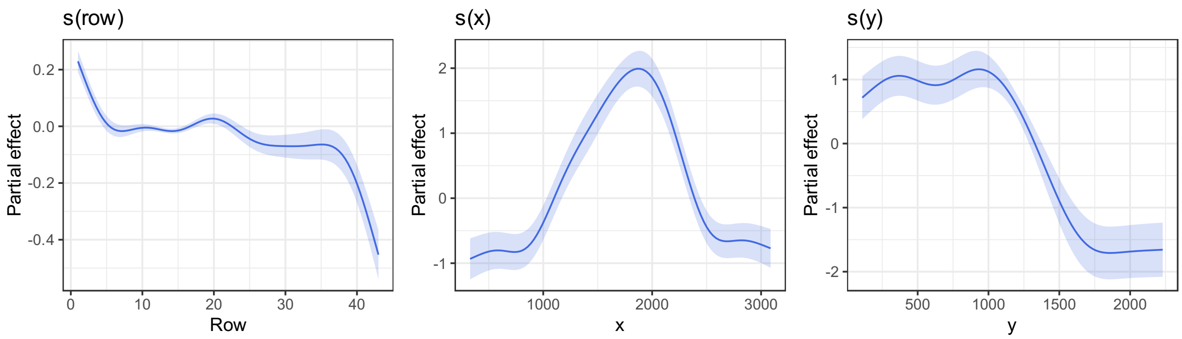

Figure 2.6 Partial effect of spatial location on listing prices from listing price GAM

In terms of the effect of the Cartesian coordinates, we see that prices are higher toward the middle of the stadium along the x-axis, which is consistent with our expectation that seats closer to 50-yard line will be more popular and more expensive. On the other hand, the y-axis captures the vertical dimension of the stadium; the lower values of y represent sections on the opponent side of the stadium and higher values represent the home team side. Mid-values of y represent endzone sections, which are not the best seats in the stadium due to the one-sided view of the entire stadium. That the home team side is priced relatively lower compared to the opponent side is somewhat surprising. One possible explanation could be that this pattern is unique to this game, as the opponent is a highly popular team whose geographic base is not too far from the home team's city.

2.5.1.2 Prediction

We made predictions on a holdout period using the last snapshot, which was taken the day before the game. The range of listing prices was \$9 to \$750 (Figure 2.7), with an average ticket price of \$131 ($SD_{price} = 66.2$). The average absolute difference between predicted and actual listing prices was \$16.5 ($SD_{diff} = 30.88$). Over 80% of listing prices were predicted within 20% of their actual value, and over 60% were predicted within 10%, indicating a generally good predictive performance.

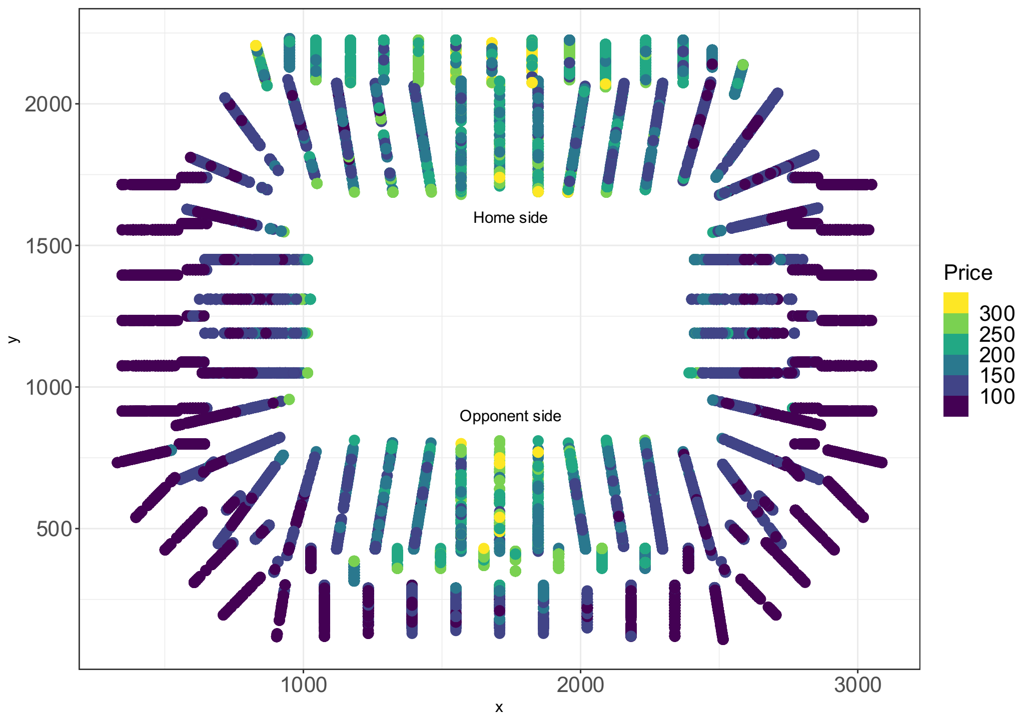

Figure 2.7 Listing prices in the stadium (5% trimmed; 2.5% from top and bottom)

From Figure 2.7, we observe that the highest-priced tickets are in the lower-level and club-level seats around 50-yard line, while the sides of the stadium and the endzones have relatively lower ticket prices. When we compare this information with Figure 2.8, which shows prediction errors by channel, we see that the more expensive regions in the stadium have higher levels of prediction error. We see that Platform A – resale shows the most frequent case of overprediction in the lower- and club-level sections around 50-yard line. This could be because the prediction was made for the day before the game, when there could be large price changes.

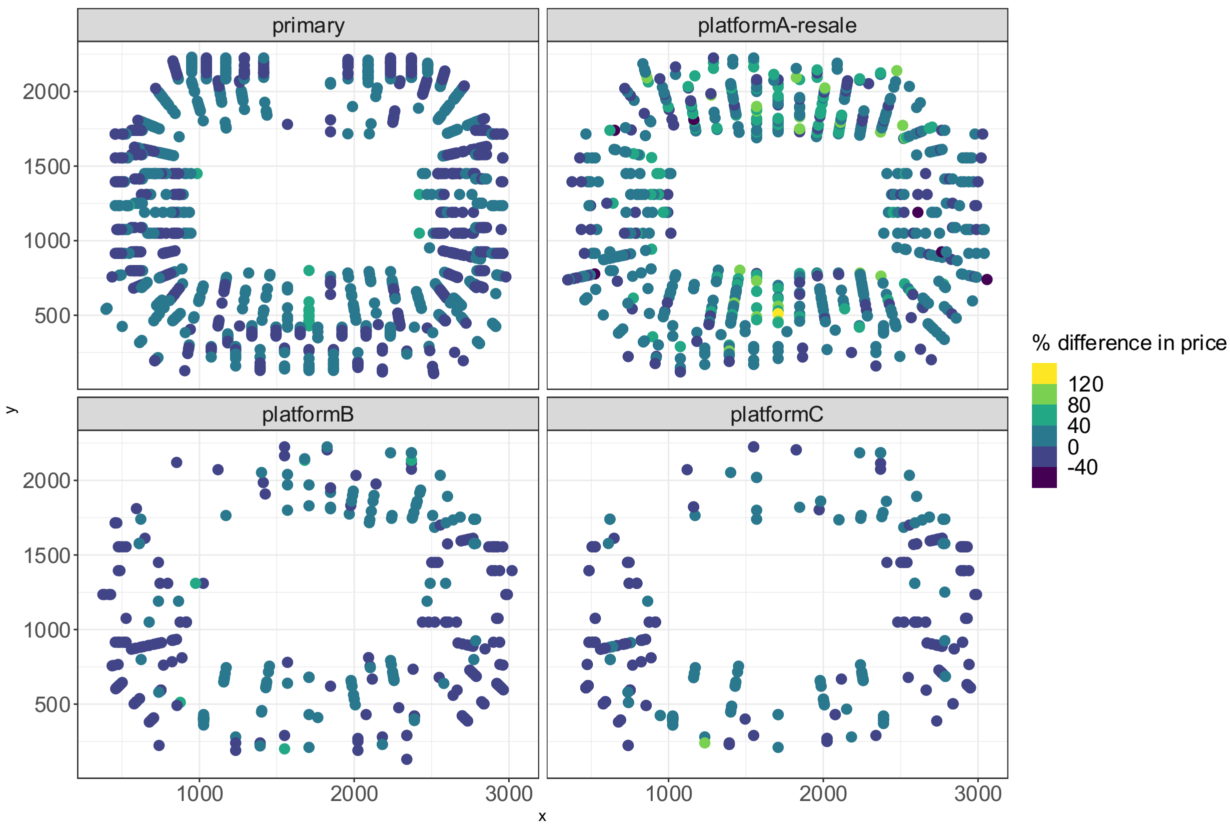

Figure 2.8 Prediction error across channels (5% trimmed on listing prices; 2.5% from top and bottom)

We compare the price fluctuations of holdout listings that had a matching snapshot[9] from the day before. Figure 2.9 shows how price changes vary across channels: we see that prices on Platform A – resale channel dropped the most, especially in the expensive areas of the stadium (higher values mean larger price drops). This could be because the ticket prices were highest on this channel, and there was not much time remaining to sell the tickets. This last-minute price drop on this channel could have led to higher prediction errors. We estimated the same prediction

[9] The listings are matched at the section and row level, and it is not guaranteed that the matched observations are the same listings as we do not observe the seat number. However, given the large overlap of the listings (82%) on the two consecutive days, it is likely that the matched ones are indeed the same listings. Even if this is not the case, assuming that seat values are going to be similar at the section/row level, the other matched listing could represent what the listing price would have been had it been available on the previous day.

model leaving two last snapshots as holdout, and we find that the prediction errors are larger on the last day (Figure 2.10).

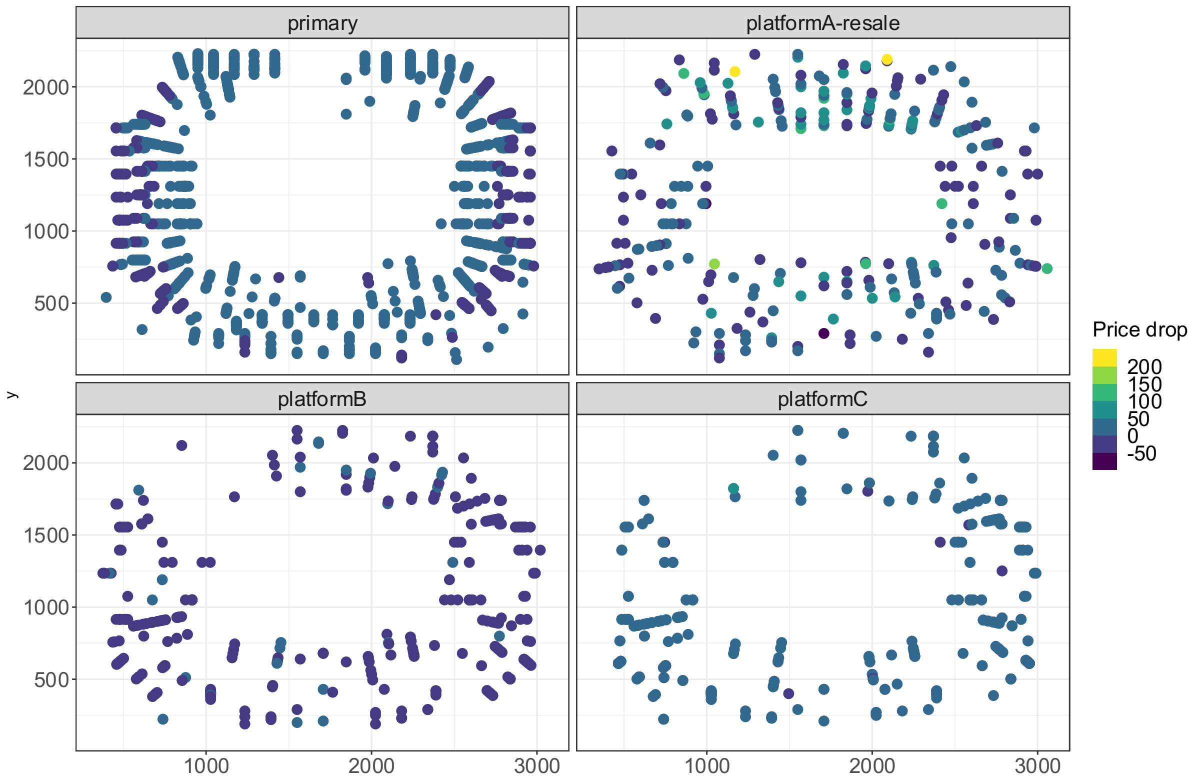

Figure 2.9 Observed price changes from the previous day

The results of price model reveal substantial differences across channels as well as the importance of accounting for temporal and spatial structure in this dynamic market. At this stage, we have not yet incorporated the “listing” decision in our model, discussed in the Model Development section. Currently, the proposed model assumes that the listing decisions are shared across channels (i.e., sellers post across channels). However, given that there are quite some differences in the available listings (e.g., see Figure 2.9 that shows the availabilities across

channels on the same day) and the significant price differences that persist, it may be more appropriate to incorporate channel decision in the dynamic availability and pricing models.

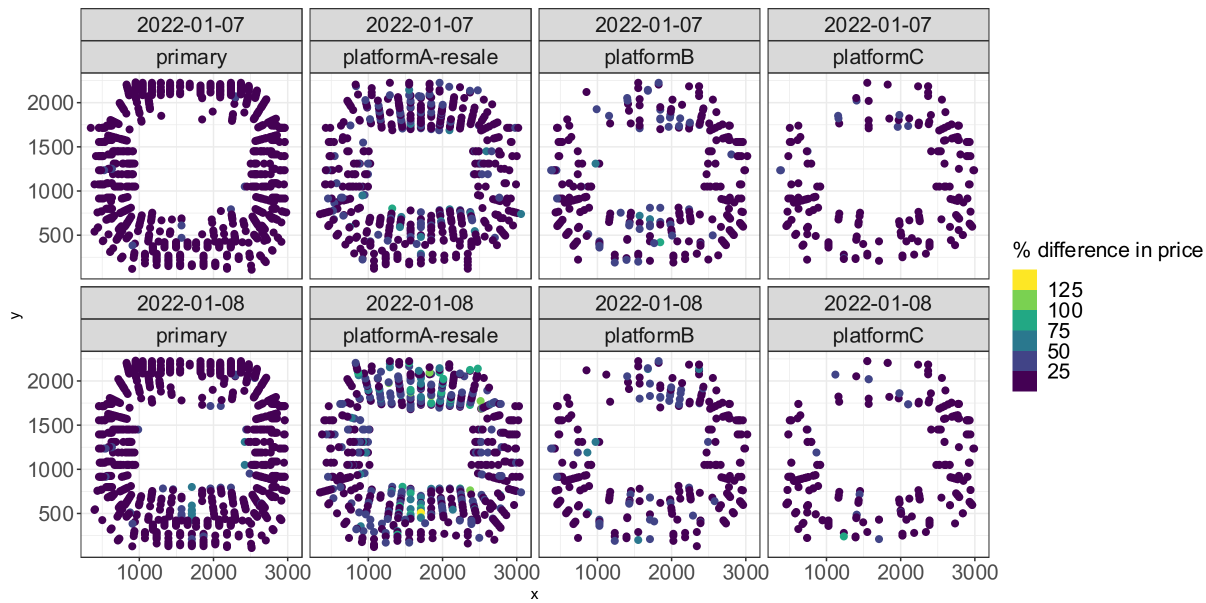

Figure 2.10 Prediction error across channels with a two-day holdout period

The results of price model reveal substantial differences across channels as well as the importance of accounting for temporal and spatial structure in this dynamic market. At this stage, we have not yet incorporated the "listing" decision in our model, discussed in the Model Development section. Currently, the proposed model assumes that the listing decisions are shared across channels (i.e., sellers post across channels). However, given that there are quite some differences in the available listings (e.g., see Figure 2.9 that shows the availabilities across channels on the same day) and the significant price differences that persist, it may be more appropriate to incorporate channel decision in the dynamic availability and pricing models.

Moreover, we are currently modeling the prices as a function of other tickets that are available at the same time in the same area. Although this approach allows us to capture how nearby availabilities affect pricing decisions, sellers cannot observe the exact availability of tickets. Therefore, for a more accurate predictive model, it would be more appropriate to model this as a function of recent availabilities and demand, as sellers can monitor changes in the listings, such as when a new listing becomes available or previously available listing is taken off the market (likely due to being sold).

Lastly, we could incorporate additional spatial structure into the model to capture the subtleties involved in location choices. For example, it is possible that the effect of row varies across section level (lower, club, upper); being closer to the field could be much more salient if one sits in the lower level compared to the club or upper level. Accounting for such structural aspects of seat choices will allow us to accurately capture the listing price patterns, but also how users make substitutions across different seat locations.

2.5.2 Channel choices

We model how consumers choose channels from which to purchase their tickets given the market conditions. Given the varying coverage and granularity of the datasets, we limit our attention to purchases for the two games that were 1) made on or after the date of the first snapshot, and 2) made on one of the four channels for which we observe the availability and prices. 70.8% of all transactions that took place during this window were for the two focal games of interest, and 63% of all single-game transactions for the two games were retained ($N = 5{,}938$). Further, we assume that users consider all four channels.

Whereas transaction data is available at the daily level, market snapshots are available on average every 3.5 days with irregular intervals (see Figure 2.1). We use the most recently available market availability and prices from the snapshot data (at the section / row level) to predict channel choices. More than 75% of the orders had a matching section-row level availability and pricing data are available within 3 days of the order date, and 94.8% of the orders had the data within 10 days of the order date. However, it is important to note that not all channels have listings available in the particular section and row of the stadium where the tickets were bought, in which case we create an indicator variable for "no matching listing". There are occasions where there are no matching snapshots at all; if a listing is made available on a channel and is sold in between snapshots, we cannot track the details of the listing. We drop additional 179 ticket-level transactions without any matched listings, leaving 5,759 transactions to estimate the model.

We use order-related information, availability and pricing information and previous usage information to explain channel choices. In terms of order-related variables, we use the size of the transaction (number of tickets), how many days before the game the order was placed, and which game the transaction was for. Availability and pricing information include number of listings, indicator for whether there was no seat available within the row, and average prices. It is important that we reflect that there might not have been any listing available in that location on some of the channels, so we include no listings variable to capture the impact of recent "stockout" (i.e., no listing available). Lastly, we use previous usage variables to capture loyalty, or stickiness of channel usage on following channel choices. Using transaction data prior to the snapshot observation period, the variables indicate whether a user has any previous history of using the channels.

2.5.2.1 Model results

Table 2.5 shows the results for multinomial logit model for channel choices for the two games, with Platform A – primary channel as the baseline. The results reveal interesting patterns in channel choices. Notably, Platform A – resale is more likely to be chosen than the primary channel, despite sharing the same platform. On the other hand, Platform B, and C are significantly less likely to be chosen. As expected, the effect of recent “stockouts” is significant and large, and having more listings in a row (compared to other channels) increases the probability of choosing a channel.

In terms of order size, when consumers purchase more tickets in one transaction, they are more likely to use Platform A – resale or Platform B compared to the primary channel. This may have to do with the fact that the proportion of 1-ticket transactions is more than double those of the other three channels (21.3% in the primary channel vs. 7.5% on average across the other three channels). This difference in transaction size could be due to the policies on how the listings can be broken into smaller offerings. While most resellers prefer not to sell $(N-1)$ tickets from their N-ticket listings, primary channels do not have such restrictions, allowing users to pick the tickets freely among any available seats in the listing. Given that the most common listing size is two in all four channels, single ticket buyers may be more prone to choose primary channels that provide more flexibility in seat choices.

Also, primary channel users exhibit different timing preferences and price sensitivity compared to the three secondary channels that we observed. Buyers tend to purchase significantly earlier on primary channels ($mean_{primary} = 16.47, mean_{secondary} = 12.6, p < 0.001$), a pattern consistently observed across seasons and games. A noticeable pattern that we see regarding price sensitivities is that the price coefficient is positive on the primary channel. This pattern remains robust across specifications, even when we control for the “quality” of the seats using `row_score`, a proxy of seat quality that is used by the team for dynamic pricing. This suggests that this may not simply be because buyers using primary channel purchase better quality seats. Further, this is in contrast with pricing patterns across the channels; Figure 2.2

shows that listing prices on Platform A – resale are generally more expensive than the primary channel. This could be due to complex interaction between how users choose the channels given the availability and prices (market conditions). For example, because it is easy to compare primary and resale prices on Platform A, it is possible that the primary tickets appear relatively cheaper than the higher price of the resale tickets, especially in the more expensive areas of the stadium.

Lastly, we find that previous usage of the channels positively affects following channel choices, as expected. Interestingly, Platform C, which is the least popular channel of the four channels that we observe, has the highest stickiness, followed by Platform B. Compared to these, both channels on Platform A show a smaller effect of previous usage on future channel choices. This may have to do with the fact that the two channels, despite being separate channels, share the same platform, and users may alternate between the two types of channels.

To summarize, the channel choice model highlights how channels differ in important ways. First, users of each channel differ in their price sensitivity, their planning horizon (i.e., how much in advance they decide to purchase the tickets), and their channel loyalty. Secondly, the ‘contents’ of the order, such as for which game the user is looking to buy tickets and number of tickets, also seem to guide channel choices. These factors need to be reflected in models of ticket demand and channel choices. Lastly, as illustrated in the case of positive price coefficient for the primary channel, there could be interactions between the market conditions of different channels.

Table 2.5 Multinomial channel choice for ticket purchases

	Dependent variable
	y
(Intercept):Platform A-resale	0.552*** (0.074)
(Intercept):Platform B	-0.147* (0.086)
(Intercept):Platform C	-0.925***(0.110)
number of listings	0.138***(0.039)
1(no listings)	-1.934***(0.105)
number of tickets sold:Platform A - resale	0.188*(0.046)
number of tickets sold:Platform B	0.133** (0.052)
number of tickets sold:Platform C	0.030 (0.066)
order-number of days before game:Platform A -resale	−0.453*** (0.049)
order-number of days before game:Platform B	−0.691*** (0.060)
order-number of days before game:Platform C	−0.632*** (0.076)
event - Game Y: Platform A - resale	0.624*** (0.102)
event - Game Y: Platform B	−0.078 (0.122)
event - Game Y: Platform C	−0.096 (0.158)
average price per ticket: Platform A - primary	0.308*** (0.058)
average price per ticket: Platform A - resale	−0.387*** (0.045)
average price per ticket: Platform B	−0.044 (0.047)
average price per ticket: Platform C	0.038 (0.069)
1(previous usage:Platform A - primary)	0.387*** (0.133)
1(previous usage:Platform A - resale)	0.729*** (0.101)
1(previous usage:Platform B)	1.944*** (0.132)
1(previous usage:Platform C)	2.520*** (0.254)
Observations	5,340
R2	0.285
Log Likelihood	−4,686.197
LR Test	3,730.883*** (df = 22)
Note	*p<0.1; **p<0.05; ***p<0.01

Another major assumption is that users will always consider all possible channels and their respective market conditions. Realistically, a large fraction of users will not visit all available channels – possibly not even the four channels due to the search costs involved in the process. Further, due to search cost heterogeneity (De Los Santos et al., 2012; Nishida & Remer, 2018), users will likely to differ not only in what channels they consider, but how many they are willing to consider. Given that such assumption is strong and likely violated in channel choice process, we propose using probabilistic consideration set approach (e.g., Andrews & Srinivasan, 1995) to address this limitation.

Additional assumption that we make in this estimation is that users consider only the price distribution of what's within the section – row, which is a relatively narrow area that includes up to 36 seats. Users are likely to consider surrounding areas, either within the same section or the same price-location as well as seats in surrounding sections, making such granularity an unrealistic assumption. Previous works on pricing and valuing MLB tickets used a much coarser level of seat aggregation (e.g., Xu et al., (2019) uses 14 sections for the whole stadium, Lewis et al., (2019) uses 6 tiers, Zhu (2014) uses 7 areas) in contrast to the current model of channel choice for a specific row. Using detailed information at the row level to model channel choices may amplify the effect of stockouts. We observe more occasions of stockouts across different channels when using row-level information than when using larger levels of aggregation. There are 25 unique price-locations within the stadium (determined based on the price tiers of the tickets) and 186 areas (defined by sections × price-locations), either of which may be a more realistic granularity. Exploring and choosing an appropriate level of aggregation would be important going forward.

2.5.3 Dynamic ticket availability

Using the snapshot data, we study dynamic seat availabilities across channels and games while accounting for the temporal and spatial dependence. Specifically, we focus on how recent availability and prices in the vicinity of focal seats affect the likelihood of these seats being made available for sale in the following periods, capturing the dynamic evolution of seat availability across space and time. The unit of observation for the dynamic ticket availability model is the number of tickets available at the row level at each snapshot period.

The fraction of rows without any available tickets ranges from 61.9 to 87.2% across channels and games. On primary channel, teams make supply decisions for the remaining tickets after season ticket and broker sales, which limits the number of tickets that can be made available for sale. Further, only a subset of ticket holders will choose to list their tickets for sale on secondary channels, contributing to the large fraction of rows with no availability. Additionally, the fact we are observing only the latter part of the season contributes to the large fraction of rows with no availability. Fraction of rows with zero availability, as can be expected, is higher on smaller platforms, platforms B and C. Figure X shows the histogram of number of seats available at the row level. In this section, we model the availabilities separately for Platform A – primary, Platform A – resale, and collapse Platforms B and C.

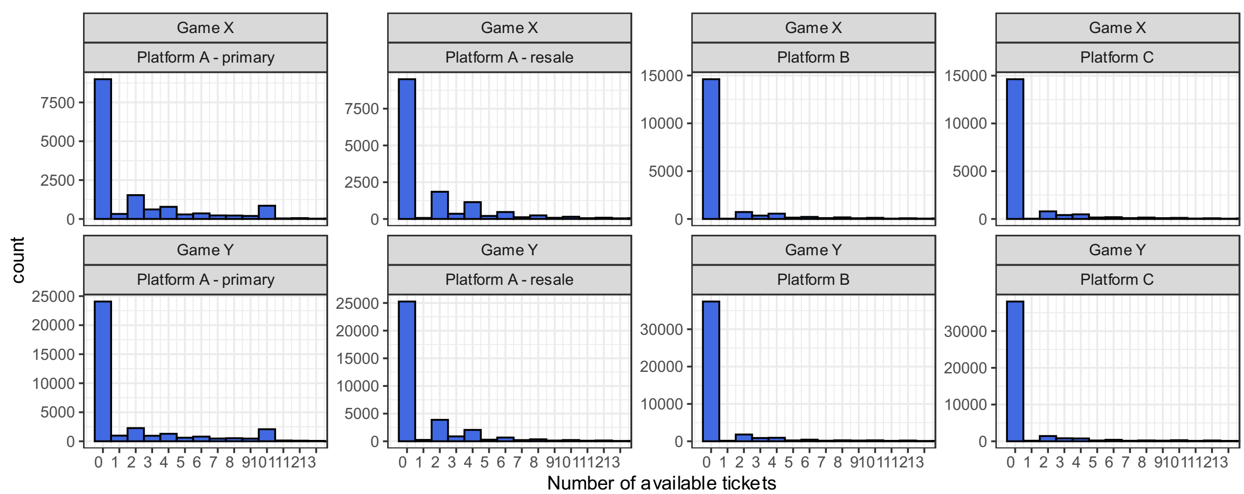

Figure 2.11 Histogram of number of available tickets (top 1% trimmed)

Given the large share of zeros, we use zero-inflated count models combined with generalized additive models (GAM) to capture the number of tickets available in a row. A key to capturing the dynamics of ticket supply, where sellers (including the team on primary channel) decide to put the tickets up on and off the market as they see fit, is leveraging the temporal and spatial information for the focal rows to understand how the supply in the neighborhood affects the ticket supply.

To do so, we construct variables to capture recent time- and location-specific supply conditions of the market, specifically the past-period availability and prices in the neighborhood of the focal row. We define sections × price-locations as the scope of neighborhood that affects supply decisions. We measure past availability ($availability_{t-1}$) and price ($price_{t-1}$) in the neighborhood as the number of available seats and average price of available tickets at time ($t - 1$). In constructing these variables, we remove past-period availability information from the focal rows to avoid the possibility of using the availability of the same seats (which might remain unsold from the past period) to predict the current availability. Due to the large fraction of rows with no available tickets on the platforms, there are occasions where nothing was available in the neighborhood in the previous period. For such rows, we use average transaction prices in the area in the past seasons (seasons 2018 and 2019) as past-period prices in the neighborhood and control for the effect of there being nothing available in the neighborhood area ($1[no\ availability_{t-1}]$). The proportion of observations without any previous availability in the neighborhood ranges from 2.1% to 29.7% across games and channels. Again, channels on Platform A tend to have lower shares of rows without any available tickets in the neighborhood in the past period whereas Platforms B and C have a higher share of such observations, consistent with the inventory patterns shown in Table 2.2.

For model estimation, we mean-centered availability, prices, and capacity for each game and combined the datasets for the two games within channels. For Platforms B and C, we combined the two datasets after mean-centering and combining across games. We used the first snapshot period as initialization period to construct the past availability and pricing information and left the last period out for predictive performance check.

2.5.3.1 Model results

We first discuss the results of zero-inflated negative binomial models. Table 2.6 shows the results of zero-inflated negative binomial models for Platform A – primary, Platform A – resale, and Platform B/C. We find several common dynamics across platforms. First, past period availability in the neighborhood has a positive effect on the number of available tickets in the following period and a negative effect on zero-inflation across all platforms (i.e., lower probability of excess zeros). That is, more seats are likely to be available in the areas where there were many seats available in the past period. This effect could be driven by area-specific levels of supplies that are not fully captured by section fixed effects. While section fixed effects can absorb some area-specific supply conditions, neighborhoods as defined by sections × price-locations are often much narrower than the sections, which could explain why there is residual area-specific effects. For example, the ticket supply around 50-yard line towards the front is scarcer than in other, less desirable locations, including seats in the rows in the back of the section on the 50-yard line. As such, the past availability in the past period could predict lower availability in the following periods.

Secondly, we find that past average price in the neighborhood positively affects the number of seats available in the following period across platforms (note that the effect is not significant for Platform B/C, but the effect is directionally consistent). This likely reflects a natural dynamic where sellers would choose to make their tickets available when the ticket prices are higher.

On the other hand, we also find diverging patterns for primary versus secondary channels. Specifically, we find that the temporal trend ($days\ until\ game$) shows opposite patterns between primary and all other secondary channels. For the primary channel, there are fewer tickets available longer before the game,[10] whereas there are more tickets longer before the game and fewer as it gets closer to the game on secondary channels. This pattern is consistent with the inventory pattern shown in Figure 2.3. There is a noticeable increase in inventory for Game Y on Platform A - primary about 15 days before the game, contributing the overall increasing availability over time. Another variable that has an opposing effect on primary and secondary channels is the non-availability in the past period. On secondary platforms, we see larger number of available tickets in the following period when there was nothing in the neighborhood in the past, but fewer on primary channel. This could be because there could be additional influx of resellers depending on the supply conditions unlike on primary channel.

While zero-inflated negative binomial model results provide insights into how ticket availability evolves based on past availability and prices in the neighborhood, the effects of past supply conditions and time may not be linear. We capture such nonlinearity using GAM with a zero-inflated Poisson model, estimated with GAM function in `mgcv` package in R. For Platform A - primary channel, we find that the linear effects in the zero-inflated Poisson GAM is consistent with that of negative binomial model (no availability_(t-1), Game Y, Capacity). On the other hand, terms for which smooth functions are estimated (availability_(t-1),log$f()$(price_(t-1)), and days until game) show interesting nonlinearities. For example, the effect of availability in the past period on excess zeros is overall negative, such that larger number of past availabilities would predict a lower level of excess zeros in the following period. The estimated

[10] The event takes place when $days\ until\ game = 0$, and the variable takes positive value before the game. E.g., $days\ until\ game = 10$ means 10 days before the game.

smooth partial effects of past availability, shown in the upper left panel of Figure 2.12 shows that the effect is more nuanced than a simple linear negative effect, such that past availability has overall negative effect up to around 82 seats, after which point the effect reverses, and has a positive effect on excess zeros. Moreover, the partial effects of days until game shows a nonlinear pattern with a spike at about 2 weeks before the game for the number of tickets (upper right) and for the fraction of zeros (bottom right), suggesting a higher number of tickets and lower fraction of rows with no availability. This effect likely reflects an unforeseen change in the inventory level shown in Figure 2.3 (upper left pane) where there is a significant jump in the available inventory for Game Y, 16 days before the game. To our knowledge, the jump in the inventory level was driven by team's "buyback" from their brokers. We believe this supply-side shock was unforeseen, and such a shock could prove to be an exogenous shock that could allow us to quantify the impact of supply-side changes on cross-channel supply and demand, shedding light onto the cross-channel structure of the primary and secondary ticket markets.

Table 2.7, Table 2.8, and Table 2.9 and Figure 2.12, Figure 2.13, and Figure 2.14 show the results for Platform A – primary channel, Platform A – resale channel, and Platform B/C, respectively. In interpreting these results, it's important to note that the zero-inflation in the GAM results are coded inversely; that is, in the following results, the outcome for the zero-inflation models the probability of presence (i.e., nonzero), not excess zeros as was the case in zero-inflated negative binomial results.

Table 2.6 Zero-inflated negative binomial model of seat availability

	DV: Number of seats available$_t$		
	Platform A - primary	Platform A - resale	Platform B/C
Count model:			
(Intercept)	-4.9260***	-1.8730***	-2.7095***
	(0.2018)	(0.0512)	(0.1066)
Platform B			0.0227
			(0.0123)
Game Y	0.0094	-0.2467***	-0.1909***
	(0.0133)	(0.0159)	(0.0185)
Availability$_{t-1}$	0.0034***	0.0016***	0.0010**
	(0.0003)	(0.0004)	(0.0004)
log(price$_{t-1}$+1)	1.3579***	0.1682***	0.0726
	(0.0710)	(0.0409)	(0.0370)
1[No availability$_{t-1}$]	-0.2424***	0.3059***	0.2102***
	(0.0532)	(0.0549)	(0.0322)
Days until game	-0.0032***	0.0126***	0.0107***
	(0.0007)	(0.0009)	(0.0008)
Log(theta)	1.6615***	1.3323***	1.2735***
	(0.0314)	(0.0285)	(0.0240)
Section FE	YES	YES	YES[11]
Zero model:			
(Intercept)	0.3282***	0.6058***	1.6991***
	(0.0267)	(0.0261)	(0.0266)
Platform B			-0.0823***
			(0.0196)
Game Y	0.2357***	0.3536***	0.4368***
	(0.0268)	(0.0266)	(0.0274)
Availability$_{t-1}$	-0.0189***	-0.0176***	-0.0154***
	(0.0005)	(0.0006)	(0.0005)
log(price$_{t-1}$+1)	-0.4480***	-0.3898***	0.2462***
	(0.0322)	(0.0248)	(0.0229)
1[No availability$_{t-1}$]	0.8155***	-0.2095**	0.4371***
	(0.0745)	(0.0743)	(0.0365)
Capacity	-0.0220***	0.0175***	0.0119***
	(0.0023)	(0.0025)	(0.0020)
Days until game	0.0028	-0.0072***	-0.0209***
	(0.0014)	(0.0014)	(0.0013)
AIC	116538.2405	108642.9021	138964.4453
Log Likelihood	-58134.1202	-54186.4511	-69351.2227
Num. obs.	43620	43620	86830

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

[11] Due to the large number of zeros on these platforms, it was impossible to estimate a full set of section fixed effects. We collapsed five sections into larger zones – either with adjacent sections of the same price tier or with sections at the same level (e.g., Club seats) of the same price tier. All the collapsed sections were located on the Club level.

For Platform A - primary channel, we find that the linear effects in the zero-inflated Poisson GAM is consistent with that of negative binomial model ($no\ availability_{t-1}$, Game Y, Capacity). On the other hand, terms for which smooth functions are estimated ($availability_{t-1}$, $\log(price_{t-1})$, and $days\ until\ game$) show interesting nonlinearities. For example, the effect of availability in the past period on excess zeros is overall negative, such that larger number of past availabilities would predict a lower level of excess zeros in the following period. The estimated smooth partial effects of past availability, shown in the upper left panel of Figure 2.12 shows that the effect is more nuanced than a simple linear negative effect, such that past availability has overall negative effect up to around 82 seats[12], after which point the effect reverses, and has a positive effect on excess zeros. Moreover, the partial effects of $days\ until\ game$ shows a nonlinear pattern with a spike at about 2 weeks before the game for the number of tickets (upper right) and for the fraction of zeros (bottom right), suggesting a higher number of tickets and lower fraction of rows with no availability. This effect likely reflects an unforeseen change in the inventory level shown in Figure 2.3 (upper left pane) where there is a significant jump in the available inventory for Game Y, 16 days before the game. To our knowledge, the jump in the inventory level was driven by team's "buyback" from their brokers. We believe this supply-side shock was unforeseen, and such a shock could prove to be an exogenous shock that could allow us to quantify the impact of supply-side changes on cross-channel supply and demand, shedding light onto the cross-channel structure of the primary and secondary ticket markets.

[12] The mean of past availability on Platform A - primary for Game X is 28.1 and 29.3 seats for Game Y. The inflection point is around 57 on the mean-centered scale and hence around 90 seats on the original scale.

Table 2.7 Model estimates for zero-inflated Poisson GAM for Platform A - primary

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Count model				
(Intercept)	-2.5934	0.2104	-12.3282	< 0.0001
Game Y	0.0378	0.0105	3.6051	0.0003
1[No availability$_{t-1}$]	-0.3007	0.0467	-6.4432	< 0.0001
Section FE			YES	
Zero model				
(Intercept)	-1.5564	0.0459	-33.9128	< 0.0001
Game Y	-0.1381	0.0231	-5.9709	< 0.0001
1[No availability$_{t-1}$]	-0.6559	0.0704	-9.3164	< 0.0001
Capacity	0.0293	0.0019	15.4384	< 0.0001
B. smooth terms	edf	Ref.df	F-value	p-value
Count model				
s(availability$_{t-1}$)	7.5929	8.3705	141.2233	< 0.0001
s(log(price$_{t-1}$+1))	8.7776	8.9851	1233.5925	< 0.0001
s(days until game)	6.9074	7.9141	130.1352	< 0.0001
Zero model				
s.1(availability$_{t-1}$)	7.5870	8.3828	2378.3005	< 0.0001
s.1(log(price$_{t-1}$+1))	8.9271	8.9980	674.9892	< 0.0001
s.1(days until game)	7.3797	8.3157	78.4180	< 0.0001

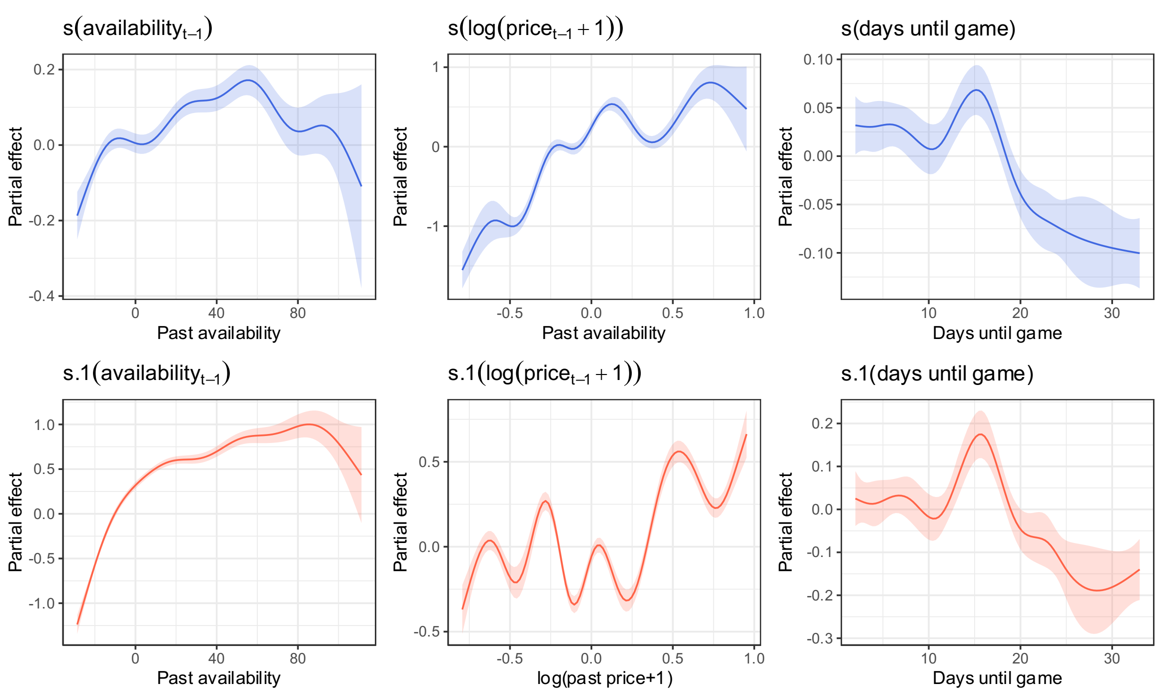

Figure 2.12 Partial effect of past supply conditions and time on availability on Platform A – primary

For Platform A – resale, we find consistent results with the zero-inflated negative binomial model. Zero-inflated Poisson GAM for Platform A – resale reveals nonlinear effects of past availability and pricing in the neighborhood as well as the effect of days until game. While we find that the overall linear effect of past availability is positive such that rows in areas with more tickets in the past will have more tickets in the following period ($\beta_{availability}^{resale,count} = 0.0016, p < 0.001$), the effect is nonlinear, especially when there are over approximately 100 tickets[13] in the neighborhood. This may reflect how individual sellers might choose not to enter the market if it's highly crowded to avoid competition, which could lead to lower prices. We further see that the effect for the past-period prices is also nonlinear, showing that the number of available tickets start decreasing as the past price increases when the price exceeds approximately \$113 for Game X, and \$168 for Game Y. Lastly, we find that the effect of days until game on the number of available tickets is generally increasing, such that there are more tickets longer before the game than close to the game, consistent with previous findings. However, there is a plateau between 16 to 11 days before, with a slightly decreasing availability between 15 to 16 days before the game. While the pattern is subtle, this period coincides with the period when there was a sharp increase in inventory on Platform A – primary, driven by the sudden influx of tickets from ticket buyback.

[13] The mean of past availability on Platform A - resale for Game X is 26.4 and 20.7 seats for Game Y. The inflection point is around 80 on the mean-centered scale and hence around 100 to 105 seats on the original scale.

Table 2.8 Model estimates for zero-inflated Poisson GAM for Platform A - resale

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Count model				
(Intercept)	-1.6854	0.0380	-44.3085	< 0.0001
Game Y	-0.1988	0.0113	-17.6550	< 0.0001
1[No availability$_{t-1}$]	0.0899	0.0390	2.3026	0.0213
Section FE		YES		
Zero model				
(Intercept)	-0.9288	0.0454	-20.4434	< 0.0001
Game Y	-0.3147	0.0215	-14.6530	< 0.0001
1[No availability$_{t-1}$]	0.0905	0.0687	1.3162	0.1881
capacity	0.0022	0.0019	1.1480	0.2510
B. smooth terms	edf	Ref.df	F-value	p-value
Count model				
s(availability$_{t-1}$)	8.7481	8.9718	119.7128	< 0.0001
s(log(price$_{t-1}$+1))	8.2657	8.8615	221.0476	< 0.0001
s(days until game))	5.9094	6.9154	476.0896	< 0.0001
Zero model				
s.1(availability$_{t-1}$)	7.1025	7.9433	1302.1920	< 0.0001
s.1(log(price$_{t-1}$+1))	6.3328	7.5172	293.0086	< 0.0001
s.1(days until game)	3.6983	4.4871	49.7415	< 0.0001

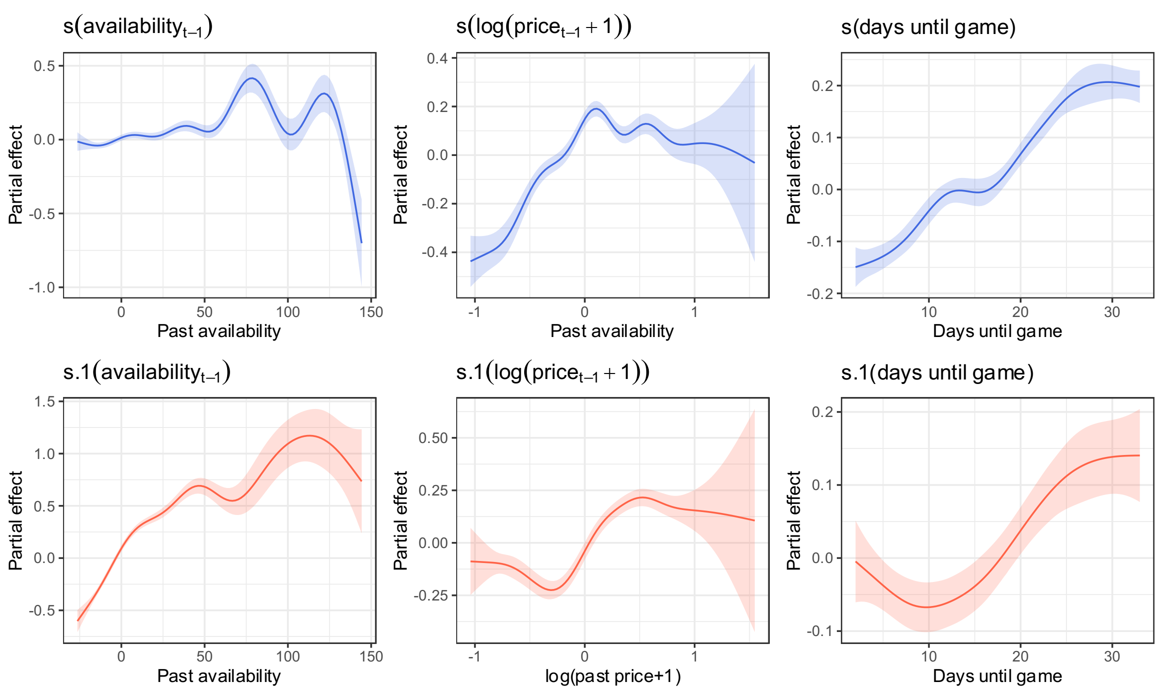

Figure 2.13 Partial effect of past supply conditions and time on availability on Platform A – resale

For Platforms B/C, the nonlinear effects of past availability and prices are very similar to those of Platform A – resale. Meanwhile, we find an interesting pattern in the temporal trend ($days\ until\ game$) where there is a significant drop in the number of available tickets 16 days before the game. Figure 2.15 shows aligned plots for the partial effect of days until game, where the dashed black line represents 16 days prior to the game at which point there was an infusion of tickets on primary channel with ticket buyback.

Table 2.9 Model estimates for zero-inflated Poisson GAM for Platform B/C

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Count model				
(Intercept)	-2.3233	0.0905	-25.6671	< 0.0001
Game Y	-0.1071	0.0127	-8.4374	< 0.0001
Platform B	0.0159	0.0073	2.1685	0.0301
1[no availability$_{t-1}$]	0.0532	0.0246	2.1647	0.0304
Section FE		YES[14]		
Zero model				
(Intercept)	-1.6041	0.0456	-35.1863	< 0.0001
Game Y	-0.4076	0.0290	-14.0388	< 0.0001
Platform B	0.0787	0.0172	4.5682	< 0.0001
1[no availability$_{t-1}$]	-0.1729	0.0443	-3.9044	0.0001
Capacity	-0.0004	0.0018	-0.2067	0.8363
B. smooth terms	edf	Ref.df	F-value	p-value
Count model				
s(availability$_{t-1}$)	8.7336	8.9711	130.7645	< 0.0001
s(log(price$_{t-1}$+1))	8.5239	8.9299	265.8416	< 0.0001
s(days until game)	8.5673	8.9447	365.9796	< 0.0001
Zero model				
s.1(availability$_{t-1}$)	7.7782	8.5281	1478.2490	< 0.0001
s.1(log(price$_{t-1}$+1))	7.5196	8.4588	310.9299	< 0.0001
s.1(days until game)	8.1991	8.8211	239.6523	< 0.0001

[14] Sections were collapsed into larger zones as described in zero-inflated NBD model section.

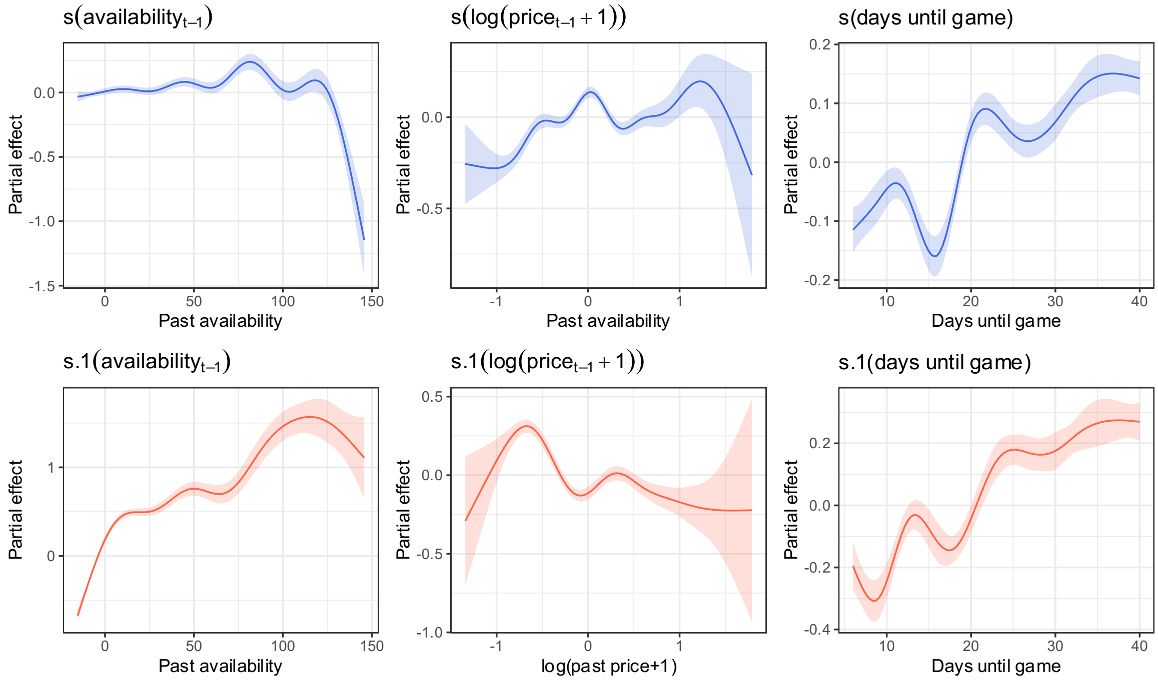

Figure 2.14 Partial effect of past supply conditions and time on availability on Platform B/C

With the use of zero-inflated Poisson GAM, we find suggestive evidence that the sudden change in the level of inventory due to buyback on Platform A – primary not only affects the supply on the primary channel but influences supply other channels, more prominently on Platforms B and C. While we plan to find out further on how and why such buyback took place to further verify that this was an exogenous shock, we believe this was an unexpected shock, which could help us causally identify the cross-channel structure.

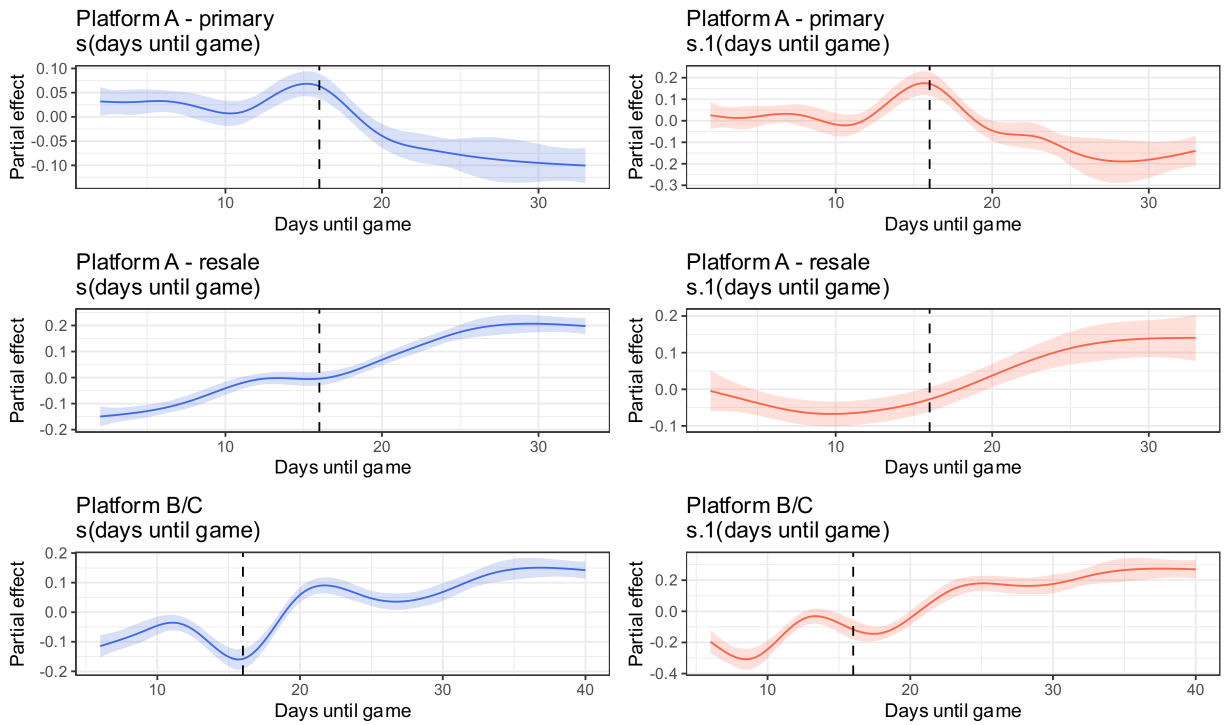

Figure 2.15 Partial effect of days until game on zero-inflated Poisson GAM across channels

2.5.3.2 Prediction

We investigate predictive performances based on predicted aggregate counts. We evaluate the aggregate predictive performance using full predictive density by computing the probability of count for each observation rather than using point predictions. Table 2.10 shows the in- and out-of-sample predictive performance for zero-inflated negative binomial and zero-inflated Poisson models and Figure 2.16 and Figure 2.17 show aggregate distribution of the counts.

We measure the predictive performance using mean absolute percentage error (MAPE), MAPE with top 1% of the counts trimmed, weighted version of MAPE using proportion of counts as weights, and RMSE. Looking at the predictive performance measures, neither model dominates one another. Although the substantive findings from the model are similar, the models have different strengths; zero-inflated negative binomial model allows us to capture the overdispersion in the data, whereas zero-inflated Poisson GAM captures the subtle nonlinearities in the dataset.

Table 2.10 Predictive performance of dynamic availability models

Model	In-sample				
	Channel	MAPE	MAPE [top 1% trimmed]	wMAPE	RMSE
Zero-inflated NBD	Platform A - primary	402.7554	43.98258	0.7469799	598.8967
	Platform A - resale	65.87133	94.06631	0.6007711	655.2412
	Platform B/C	79.55876	78.01434	0.218581	423.8491
Zero-inflated Poisson GAM	Platform A - primary	86.48459	39.73077	0.9818919	861.154
	Platform A - resale	79.03812	70.64524	0.7294229	771.6541
	Platform B/C	73.75404	64.63651	0.3071723	530.9638
	Out-of-sample				
	Channel	MAPE	MAPE [top 1% trimmed]	wMAPE	RMSE
Zero-inflated NBD	Platform A - primary	182.3928	44.63838	1.1720385	117.62177
	Platform A - resale	85.65074	108.82321	1.0910786	120.73038
	Platform B/C	80.33109	71.82259	0.2417923	57.04302
Zero-inflated Poisson GAM	Platform A - primary	42.39753	36.3981	1.4525614	179.568
	Platform A - resale	88.82492	86.74912	1.1677935	135.56008
	Platform B/C	69.08164	71.72857	0.2710193	59.74921

*wMAPE is weighted MAPE with proportion of counts

Visually inspecting the predictive performances (Figure 2.16), both zero-inflated models accurately predict the share of zero availabilities. Comparing the two models, zero-inflated negative binomial model tracks the zero counts better than zero-inflated Poisson GAMs, which tends to overpredict zero counts, both in- and out-of-sample. The most noticeable discrepancy between the predicted and the actual counts is for one or two ticket availabilities. While such two-inflation is observed across all channels, the difference between one and three ticket availability is even more pronounced on secondary channels. Both models significantly overpredict the availability of one ticket and underpredict the availability of two tickets per row. This likely stems from resellers' decisions not to sell single tickets, as sport event tickets are often sold as a pair or more. Further, secondary channels show similar underprediction for four tickets. This may be driven by a similar logic where four tickets can be sold in several possible combinations, including selling two tickets to two different customers. Another region where the model fails to make an adequate count prediction is availability of ten tickets per row, which stands out only on primary channel. Although is not evident why there is a spike at ten tickets per row, it is possible that it is driven by internal distribution policies.

The dynamic availability model and its predictions reveal interesting substantive patterns on how seat availability changes as a function of past availability and prices and provide avenues for further development. First, we could add hierarchical structure to the dynamic count model. Currently, we model Platform A – primary, Platform A – resale, and Platform B/C as separate channels. Given that the channels are not independent of each other, pooling across channels could provide further insights into the temporal and spatial dynamic patterns for ticket availability. We could build in further structures to account for the temporal dependence of the effect of past availability and pricing (e.g., the effect of past availability may be a function of the time between snapshots) and to incorporate additional spatial structures to further explain the effect of seat location, which we currently explain using section fixed effects.

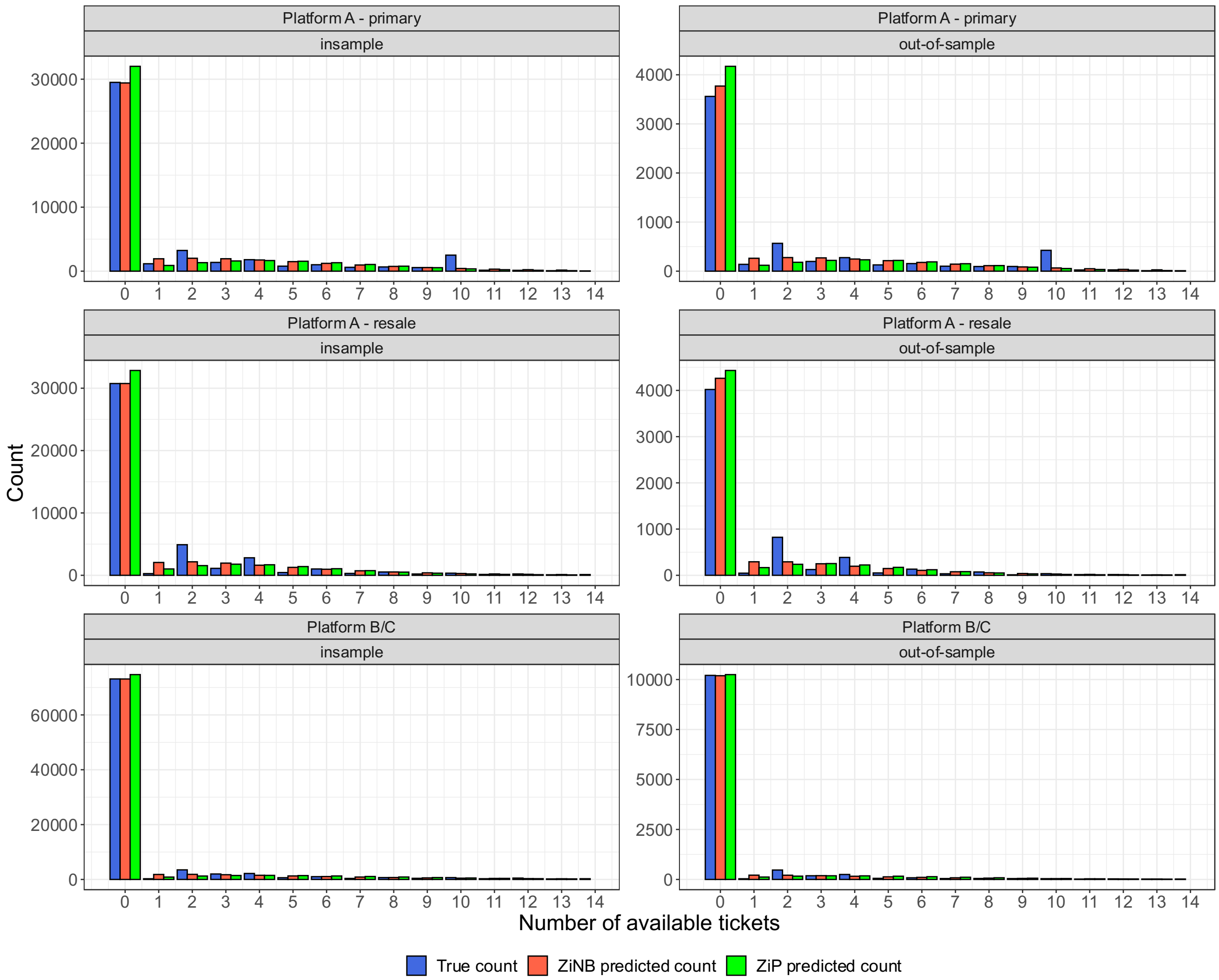

Figure 2.16 Predicted count of the number of available tickets (top 1% trimmed)

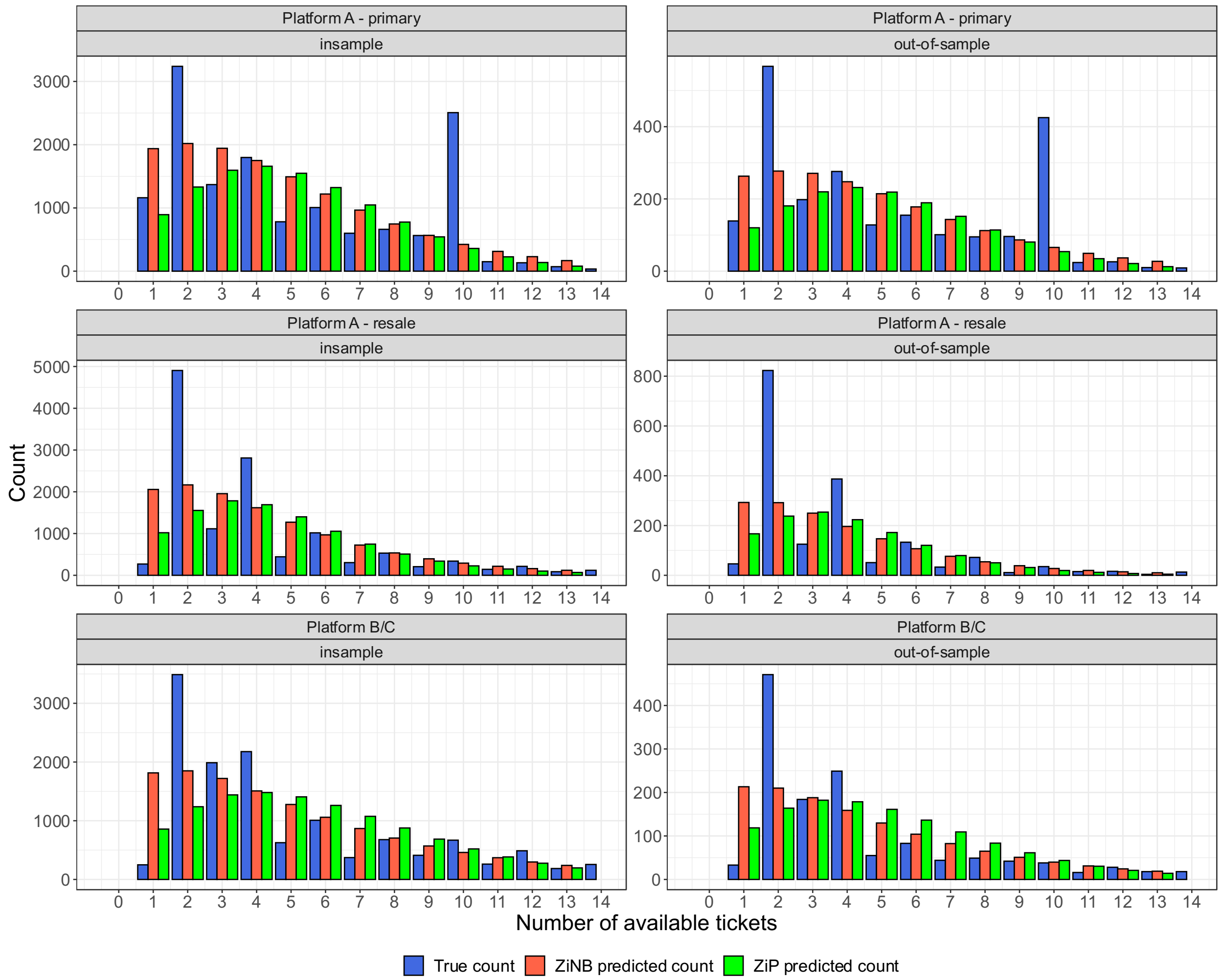

Figure 2.17 Predicted count of the number of available tickets (top 1% trimmed, positive counts only)

Additionally, while the current approach only uses channel-specific supply information to model dynamic availability, it is unlikely that resellers' supply decisions are solely based on within-channel supply conditions as many of the resellers and brokers choose to post their tickets on multiple secondary channels. To account for such complex dependencies, we could further use cross-channel and demand information to model the dynamic availability. Especially with the presence of sudden influx of tickets on Platform A – primary shown in Figure 2.2 and how the effect seems to have cross-channel consequences for the ticket availability, incorporating cross-channel supply information would help us understand the multi-channel structure of the market.

Bibliography

Allison, P. D., & Christakis, N. A. (1994). Logit Models for Sets of Ranked Items. *Sociological Methodology*, *24*(1994), 199–228.

Andrews, R. L., & Srinivasan, T. C. (1995). Studying Consideration Effects in Empirical Choice Models Using Scanner Panel Data. *Journal of Marketing Research*, *32*(1), 30. https://doi.org/10.2307/3152108

Arnold, S. J., Oum, T. H., & Tigert, D. J. (1983). Determinant Attributes in Retail Patronage: Seasonal, Temporal, Regional, and International Comparisons. *Journal of Marketing Research*, *20*(2), 149. https://doi.org/10.2307/3151681

Arslan, H. A., Easley, R. F., Wang, R., & Yılmaz, Ö. (2022). Data-Driven Sports Ticket Pricing for Multiple Sales Channels with Heterogeneous Customers. *Manufacturing and Service Operations Management*, *24*(2), 1241–1260. https://doi.org/10.1287/msom.2021.1005

Berger, J., Humphreys, A., Ludwig, S., Moe, W. W., Netzer, O., & Schweidel, D. A. (2020). Uniting the Tribes: Using Text for Marketing Insight. *Journal of Marketing*, *84*(1), 1–25. https://doi.org/10.1177/0022242919873106

Blanchard, S. J., Dyachenko, T. L., & Kettle, K. L. (2020). Locational Choices: Modeling Consumer Preferences for Proximity to Others in Reserved Seating Venues. *Journal of Marketing Research*, *57*(5), 878–899. https://doi.org/10.1177/0022243720941525

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*(4–5), 993–1022. https://doi.org/10.1016/b978-0-12-411519-4.00006-9

Bradlow, E. T., & Rao, R. (2000). Hierarchical Bayes Assortment Choice Model. *Journal of Marketing Research*, *37*(2), 259–268.

Broniarczyk, S. M., Hoyer, W. D., & McAlister, L. (1998). Consumers' Perceptions of the Assortment Offered in a Grocery Category: The Impact of Item Reduction. *Journal of Marketing Research*, *35*(2), 166–176. https://doi.org/10.1177/002224379803500203

Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing*, *72*(7–9), 1775–1781. https://doi.org/10.1016/j.neucom.2008.06.011

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1). https://doi.org/10.18637/jss.v076.i01

Chernev, A. (2011). Product assortment and consumer choice: An interdisciplinary review. *Foundations and Trends in Marketing*, *6*(1), 1–61. https://doi.org/10.1561/1700000030

Chernev, A., Böckenholt, U., & Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, *25*(2), 333–358. https://doi.org/10.1016/j.jcps.2014.08.002

Chiang, J. (1991). A Simultaneous Approach to the Whether, What and How Much to Buy Questions. *Marketing Science*, *10*(4), 297–315. https://doi.org/10.1287/mksc.10.4.297

Chintagunta, P. K. (1998). Inertia and variety seeking in a model of brand-purchase timing. *Marketing Science*, *17*(3), 253–270. https://doi.org/10.1287/mksc.17.3.253
Chung, J., & Rao, V. R. (2003). A general choice model for bundles with multiple-category products: Application to market segmentation and optimal pricing for bundles. *Journal of Marketing Research*, *40*(2), 115–130. https://doi.org/10.1509/jmkr.40.2.115.19230
Clarkson, J. J., Janiszewski, C., & Cinelli, M. D. (2013). The desire for consumption knowledge. *Journal of Consumer Research*, *39*(6), 1313–1329. https://doi.org/10.1086/668535
Cooper-Martin, E. (1991). Consumers and Movies: Some Findings on Experiential Products. *NA - Advances in Consumer Research*, *18*.
Courty, P. (2003). Some Economics of Ticket Resale. *Journal of Economic Perspectives*, *17*(2), 85–97. https://doi.org/10.1257/089533003765888449
Courty, P., & Davey, L. (2020). The Impact of Variable Pricing, Dynamic Pricing, and Sponsored Secondary Markets in Major League Baseball. *Journal of Sports Economics*, *21*(2), 115–138. https://doi.org/10.1177/1527002519867367
Courty, P., & Pagliero, M. (2012). The impact of price discrimination on revenue: Evidence from the concert industry. *Review of Economics and Statistics*, *94*(1), 359–369. https://doi.org/10.1162/REST_a_00179
Cui, Y., Duenyas, I., & Şahin, Ö. (2014). Should event organizers prevent resale of tickets? *Management Science*, *60*(9), 2160–2179. https://doi.org/10.1287/mnsc.2013.1862
Datta, H., Knox, G., & Bronnenberg, B. J. (2018). Changing their tune: How consumers' adoption of online streaming affects music consumption and discovery. *Marketing Science*, *37*(1), 5–21. https://doi.org/10.1287/mksc.2017.1051
De Los Santos, B., Hortacsu, A., & Wildenbeest, M. R. (2012). Testing models of consumer search using data on web browsing and purchasing behavior. *American Economic Review*, *102*(6), 2955–2980. https://doi.org/10.1257/aer.102.6.2955
Deng, X., Kahn, B. E., Unnava, H. R., & Lee, H. (2016). "Wide" Variety: Effects of Horizontal Versus Vertical Display on Assortment Processing, Perceived Variety, and Choice. *Journal of Marketing Research*, *53*(5), 682–698. https://doi.org/10.1509/jmr.13.0151
Drayer, J., & Shapiro, S. L. (2009). Value Determination in the Secondary Ticket Market: A Quantitative Analysis of the NFL Playoffs. *Sport Marketing Quarterly*, *18*(1), 5–13.
Drayer, J., Shapiro, S. L., & Lee, S. (2012). Dynamic Ticket Pricing in Sport: An Agenda for Research and Practice. *Sport Marketing Quarterly*, *21*(3), 184–194. https://ejwl.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bft&AN=82306237&site=ehost-live
Evers, E. R. K., Inbar, Y., & Zeelenberg, M. (2014). Set-fit effects in choice. *Journal of Experimental Psychology: General*, *143*(2), 504–509. https://doi.org/10.1037/a0033343
Farquhar, P. H., & Rao, V. R. (1976). A Balance Model for Evaluating Subsets of Multiattributed Items. *Management Science*, *22*(5), 528–539.
Feinberg, F. M., Kahn, B. E., & McAlister, L. (1992). Market Share Response When Consumers Seek Variety. *Journal of Marketing Research*, *29*(2), 227. https://doi.org/10.2307/3172572
Feit, E. M., & Bradlow, E. T. (2021). Fusion Modeling. In *Handbook of Market Research* (pp. 147--180).
Forty under 40: Barry Kahn. (2011). *Street and Smith's SportsBusiness Journal*, 37A.
Galak, J., Kruger, J., & Loewenstein, G. (2011). Is variety the spice of life? It all depends on the rate of consumption. *Judgment and Decision Making*, *6*(3), 230–238.
Galak, J., Redden, J. P., & Kruger, J. (2009). Variety amnesia: Recalling past variety can

accelerate recovery from satiation. *Journal of Consumer Research*, *36*(4), 575–584. https://doi.org/10.1086/600066
Geng, X., Wu, R., & Whinston, A. B. (2007). Profiting from partial allowance of ticket resale. *Journal of Marketing*, *71*(2), 184–195. https://doi.org/10.1509/jmkg.71.2.184
Givon, M. (1984). Variety Seeking Through Brand Switching. *Marketing Science*, *3*(1), 1–22. https://doi.org/10.1287/mksc.3.1.1
Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(SUPPL. 1), 5228–5235. https://doi.org/10.1073/pnas.0307752101
Gupta, S. (1988). Impact of Sales Promotions on When, What, and How Much to Buy. *Journal of Marketing Research*, *25*(4), 342. https://doi.org/10.2307/3172945
Hastie, T., & Tibshirani, R. (1987). Generalized additive models: Some applications. *Journal of the American Statistical Association*, *82*(398), 371–386. https://doi.org/10.1080/01621459.1987.10478440
Heckman, J. (1979). Sample Selection Bias as a Specification Error Author (s): James J . Heckman Reviewed work (s): Published by : The Econometric Society Stable URL : http://www.jstor.org/stable/1912352 . *Econometrica*, *47*(1), 153–161. https://www.jstor.org/stable/1912352
Hoch, S. J., Bradlow, E. T., & Wansink, B. (1999). The variety of an assortment. *Marketing Science*, *18*(4), 527–546. https://doi.org/10.1287/mksc.18.4.527
Hoffman, M. D., & Gelman, A. (2014). *The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo*. *15*, 1593–1623.
Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*(6), 995–1006. https://doi.org/10.1037/0022-3514.79.6.995
Kahn, B. E. (1995). Consumer variety-seeking among goods and services. An integrative review. *Journal of Retailing and Consumer Services*, *2*(3), 139–148. https://doi.org/10.1016/0969-6989(95)00038-0
Kahn, B. E. (1998). MARKETING IN THE 21ST CENTURY Dynamic Relationships With Customers : High-Variety Strategies. *Journal of the Academy of Marketing Science*.
Kahn, B. E., & Rafieian, H. (2022). More than just the spice of life: Using variety as a signal for change and diversification. *Consumer Psychology Review*, *5*(1), 87–106. https://doi.org/10.1002/arcp.1078
Kahn, B. E., & Ratner, R. K. (2005). Variety for the sake of variety? Diversification motives in consumer choice. In *Inside Consumption: Consumer Motives, Goals, and Desires* (pp. 102–121). https://doi.org/10.4324/9780203481295
Kahn, B. E., & Wansink, B. (2004). The Influence of Assortment Structure on Perceived Variety and Consumption Quantities. *Journal of Consumer Research*, *30*(4), 519–533. https://doi.org/10.1086/380286
Kuksov, D., & Villas-Boas, J. M. (2010). When more alternatives lead to less choice. *Marketing Science*, *29*(3), 507–524. https://doi.org/10.1287/mksc.1090.0535
Kumar, M., Eckles, D., & Aral, S. (2020). *Scalable bundling via dense product embeddings*. *January*, 1–47. http://arxiv.org/abs/2002.00100
Lattin, J. M., & McAlister, L. (1985). Using a Variety-Seeking Model to Identify Substitute and Complementary Relationships among Competing Products. *Journal of Marketing Research*, *22*(3), 330. https://doi.org/10.2307/3151429

Leslie, P. (2004). Price Discrimination in Broadway Theater. *The RAND Journal of Economics*, *35*(3), 520. https://doi.org/10.2307/1593706

Lewis, M., Wang, Y., & Wu, C. (2019). Season ticket buyer value and secondary market options. *Marketing Science*, *38*(6), 973–993. https://doi.org/10.1287/mksc.2019.1183

Louviere, J. J., & Gaeth, G. J. (1987). Decomposing the Determinants of Retail Facility Choice Using the Method of Hierarchical Information Integration: A Supermarket Illustration. In *Journal of Retailing* (Vol. 63, Issue 1, pp. 25–48). http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=6353501&site=bsi-live

McAlister, L. (1982). A Dynamic Attribute Satiation Model of Variety-Seeking Behavior. *Journal of Consumer Research*, *9*(2), 141. https://doi.org/10.1086/208907

Mogilner, C., Rudnick, T., & Iyengar, S. S. (2008). The mere categorization effect: How the presence of categories increases choosers' perceptions of assortment variety and outcome satisfaction. *Journal of Consumer Research*, *35*(2), 202–215. https://doi.org/10.1086/588698

Natan, O. R. (2022). *Choice Frictions in Large Assortments*.

Nishida, M., & Remer, M. (2018). The Determinants and consequences of search cost heterogeneity: Evidence from local gasoline markets. *Journal of Marketing Research*, *55*(3), 305–320. https://doi.org/10.1509/jmr.15.0146

Orhun, A. Y., Venkataraman, S., & Chintagunta, P. K. (2016). Impact of competition on product decisions: Movie choices of exhibitors. *Marketing Science*, *35*(1), 73–92. https://doi.org/10.1287/mksc.2015.0909

Paul, R. J., & Weinbach, A. P. (2013). Determinants of Dynamic Pricing Premiums in Major League Baseball. *Sport Marketing Quarterly*, *22*(3), 152–165. http://search.ebscohost.com/login.aspx?direct=true&db=buh&AN=90262140&site=ehost-live

Rao, V. R., Russell, G. J., Bhargava, H., Cooke, A., Derdenger, T., Kim, H., Kumar, N., Levin, I., Ma, Y., Mehta, N., Pracejus, J., & Venkatesh, R. (2018). Emerging Trends in Product Bundling: Investigating Consumer Choice and Firm Behavior. *Customer Needs and Solutions*, *5*(1–2), 107–120. https://doi.org/10.1007/s40547-017-0075-x

Rascher, D. A., McEvoy, C. D., Nagel, M. S., & Brown, M. T. (2007). Variable ticket pricing in Major League Baseball. *Journal of Sport Management*, *21*(3), 407–437. https://doi.org/10.1123/jsm.21.3.407

Ratner, R. K., Kahn, B. E., & Kahneman, D. (1999). Choosing less-preferred experiences for the sake of variety. *Journal of Consumer Research*, *26*(1), 1–15. https://doi.org/10.1086/209547

Read, D., Antonides, G., Van den Ouden, L., & Trienekens, H. (2001). Which Is Better: Simultaneous or Sequential Choice? *Organizational Behavior and Human Decision Processes*, *84*(1), 54–70. https://doi.org/10.1006/obhd.2000.2917

Read, D., & Loewenstein, G. (1995). Diversification Bias: Explaining the Discrepancy in Variety Seeking Between Combined and Separated Choices. *Journal of Experimental Psychology: Applied*, *1*(1), 34–49. https://doi.org/10.1037/1076-898X.1.1.34

Salisbury, L. C., & Feinberg, F. M. (2012). All things considered? The role of choice set formation in diversification. *Journal of Marketing Research*, *49*(3), 320–335. https://doi.org/10.1509/jmr.10.0437

Sela, A., Hadar, L., Morgan, S., & Maimaran, M. (2019). Variety-Seeking and Perceived Expertise. *Journal of Consumer Psychology*, *29*(4), 671–679. https://doi.org/10.1002/jcpy.1110

Shaddy, F., & Fishbach, A. (2017). Seller beware: How bundling affects valuation. *Journal of Marketing Research*, *54*(5), 737–751. https://doi.org/10.1509/jmr.15.0277

Shapiro, S. L., & Drayer, J. (2014). An examination of dynamic ticket pricing and secondary market price determinants in Major League Baseball. *Sport Management Review*, *17*(2), 145–159. https://doi.org/10.1016/j.smr.2013.05.002

Simonson, I. (1990). The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research*, *XXVII*(May), 735–757. https://doi.org/10.1017/CBO9780511803475.042

Sweeting, A. (2012). Dynamic pricing behavior in perishable goods markets: Evidence from secondary markets for Major League Baseball tickets. *Journal of Political Economy*, *120*(6), 1133–1172. https://doi.org/10.1086/669254

Tereyağoğlu, N., Fader, P. S., & Veeraraghavan, S. (2017). Pricing Theater Seats: The Value of Price Commitment and Monotone Discounting. *Production and Operations Management*, *26*(6), 1056–1075. https://doi.org/10.1111/poms.12611

Townsend, C., & Kahn, B. E. (2012). A Variety of Views on Variety-Seeking. *Advances in Consumer Research*, *40*, 223–227.

Tversky, A., & Kahneman, D. (1992). Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323. https://doi.org/10.15358/0340-1650-2006-6-331

Van Herpen, E., & Pieters, R. (2002). The variety of an assortment: An extension to the attribute-based approach. *Marketing Science*, *21*(3), 331–341. https://doi.org/10.1287/mksc.21.3.331.144

Veeraraghavan, S., & Vaidyanathan, R. (2012). Measuring seat value in stadiums and theaters. *Production and Operations Management*, *21*(1), 49–68. https://doi.org/10.1111/j.1937-5956.2011.01249.x

Venkatesh, R., & Mahajan, V. (2009). The design and pricing of bundles: A review of normative guidelines and practical approaches. *Handbook of Pricing Research in Marketing*, *15260*(412), 232–257. https://doi.org/10.4337/9781848447448.00020

Xu, J. J., Fader, P. S., & Veeraraghavan, S. (2019). Designing and evaluating dynamic pricing policies for major league baseball tickets. *Manufacturing and Service Operations Management*, *21*(1), 121–138. https://doi.org/10.1287/msom.2018.0760

Zhu, J. (2014). Effect of Resale on Optimal Ticket Pricing : Evidence from Major League Baseball Tickets. *Working Paper*.